

FREE LIBROS

TU BIBLIOTECA VIRTUAL

<http://www.freelibros.com>



Categoría

- [Administración](#)
- [Algebra](#)
- [Análisis Matemático](#)
- [Anatomía](#)
- [Arquitectura](#)
- [Arte](#)
- [Artículos](#)
- [Astronomía](#)
- [Atlas](#)
- [AudioLibros](#)
- [Automatización](#)
- [Base de Datos](#)
- [Biblia](#)
- [Biología](#)
- [Bioquímica](#)
- [Cálculo](#)
- [Circuitos](#)
- [Cirugía](#)
- [Cocina](#)
- [Comic](#)
- [Computer Hoy](#)
- [Contabilidad](#)
- [De Todo](#)
- [Derecho](#)
- [Dermatología](#)
- [Diarios](#)
- [Diccionario](#)
- [Diseño Grafico](#)
- [Diseño Web](#)
- [Documentales](#)
- [Dummies](#)
- [E-Books](#)
- [Ecografía](#)
- [Ecología](#)
- [Economía](#)
- [Ecuaciones diferenciales](#)
- [Educación Primaria](#)
- [Ejemplos](#)
- [Electricidad](#)
- [Electrónica](#)
- [Enciclopedia](#)
- [Estadística](#)
- [Filosofía](#)
- [Física](#)
- [Fisiología](#)
- [Ganar dinero en internet](#)
- [Geología](#)
- [Geometría](#)
- [Ginecología y Obstetricia](#)
- [Guías](#)
- [HackCrack](#)
- [Hidráulica](#)
- [Historia](#)
- [Ingeniería](#)
- [Ingeniería ambiental](#)
- [Ingeniería Civil](#)
- [Ingeniería de Materiales](#)
- [Ingeniería de Minas](#)
- [Ingeniería Industrial](#)
- [Ingeniería Petrolera](#)
- [Ingles](#)
- [Integrales](#)
- [Inv. Operaciones](#)
- [Leer Online](#)
- [Libros](#)
- [Libros Copyleft](#)
- [Libros Unicef](#)
- [Liderazgo y Motivación](#)
- [Linux](#)
- [Logística](#)
- [Maestra Infantil](#)
- [Manga](#)
- [Manual](#)
- [Manualidades](#)
- [Marketing](#)
- [Matemática Discreta](#)
- [Matemáticas](#)
- [Mecánica](#)
- [Medicina](#)
- [Metalurgia](#)
- [Mi Novela Favorita](#)
- [Multimedia](#)
- [Noticias](#)
- [Odontología](#)
- [Ofimática](#)
- [Oftalmología](#)
- [Pediatria](#)
- [Procesos Unitarios](#)
- [Programación](#)
- [Psicología](#)
- [Química](#)
- [Radiología](#)
- [Recetas](#)
- [Redes](#)
- [Religión](#)
- [Revistas](#)
- [Rincón Literario](#)
- [Robótica](#)
- [Romántica](#)
- [Salud](#)
- [Seguridad](#)
- [Sexualidad](#)
- [Sistemas Operativos](#)
- [Sobre Escribir](#)
- [Soldadura](#)
- [Solucionario](#)
- [Termodinámica](#)
- [Tesis](#)
- [Topografía](#)
- [Transferencia de Calor](#)
- [Transferencia de Masa](#)
- [Tutorial](#)
- [TuxInfo](#)
- [VideoTutoriales](#)
- [Windows](#)
- [zoología](#)

PROBABILIDAD
Y
ESTADÍSTICA
PARA INGENIEROS

SEXTA EDICIÓN

WALPOLE



MYERS



MYERS

Prentice
Hall



Ronald E. Walpole
Roanoke College

Raymond H. Myers
Virginia Polytechnic Institute and State University

Sharon L. Myers
Radford University

Probabilidad y estadística para ingenieros

SEXTA EDICIÓN

TRADUCCIÓN:

Ricardo Cruz
Traductor Profesional

REVISIÓN TÉCNICA:

Juan Antonio Torre Marina
Ingeniero Mecánico Electricista
Universidad Anáhuac



México • Argentina • Brasil • Colombia • Costa Rica • Chile • Ecuador
España • Guatemala • Panamá • Perú • Puerto Rico • Uruguay • Venezuela

Autor: Walpole, Ronald E.
 Título, edición: Probabilidad y estadística para ingenieros, 6a. ed.
 PRENTICE-HALL HISPANOAMERICANA, S.A.
 México, 1999

ISBN: 970-17-0264-6

Área: Universitarios

Formato: 19 x 23.5 cms.

Páginas: 752

EDICIÓN EN ESPAÑOL:

EDITOR: PABLO EDUARDO ROIG VÁZQUEZ
 SUPERVISOR DE TRADUCCIÓN: JORGE BONILLA TALAVERA
 SUPERVISOR DE EDICIÓN: ÓSCAR ÁVALOS SALCEDO

EDICIÓN EN INGLÉS:

Executive editor, *Ann Heath*
 Editorial director, *Tim Bozik*
 Editor-in-chief, *Jerome Grant*
 Editorial assistant, *Mindy McClard*
 Executive managing editor, *Kathleen Schiaparelli*
 Managing editor, *Linda Mihalek Behrens*
 Production editor, *Nicholas Romanelli*
 Assistant VP production/manufacturing, *David W. Riccardi*
 Manufacturing manager, *Trudy Pisciotti*
 Manufacturing buyer, *Alan Fischer*
 Creative director, *Paula Maylahn*
 Art director, *Maureen Eide*
 Cover designer, *Jayne Conte*
 Marketing manager, *Melody Marcus*
 Marketing assistant, *Jennifer Pui*
 Composer, *Prepare Inc./Emilcomp s.r.l.*

WALPOLE: PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS, 6a. ed.

Traducido de la sexta edición en inglés de la obra: **Probability and Statistics for Engineers and Scientists**.

All rights reserved. Authorized translation from English language edition published by Prentice-Hall, Inc.
 A Simon & Schuster Company.

Todos los derechos reservados. Traducción autorizada de la edición en inglés publicada por Prentice-Hall, Inc.
 A Simon & Schuster Company.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the publisher.

Prohibida la reproducción total o parcial de esta obra, por cualquier medio o método sin autorización por escrito del editor.

Derechos reservados © 1999 respecto a la primera edición en español publicada por:

PRENTICE-HALL HISPANOAMERICANA, S. A.

Atzacomulco Núm. 500-5º Piso

Col. Industrial Atoto

53519, Naucalpan de Juárez, Edo. de México

ISBN 970-17-0264-6

Miembro de la Cámara Nacional de la Industria Editorial, Reg. Núm. 1524.

Original English Language Edition Published by Prentice-Hall, Inc.

A Simon & Schuster Company.

Copyright © 1998

All rights reserved

ISBN 0-13-840208-6

IMPRESO EN MÉXICO / PRINTED IN MEXICO

Contenido

Prefacio ix

1 *Introducción a la estadística y al análisis de datos 1*

- 1.1 Repaso 1
- 1.2 El papel de la probabilidad 2
- 1.3 Medidas de posición: media de una muestra 4
- 1.4 Medidas de variabilidad 5
- 1.5 Datos discretos y continuos 6
- 1.6 Modelado estadístico, inspección científica y diagnósticos gráficos 7

2 *Probabilidad 9*

- 2.1 Espacio muestral 10
- 2.2 Eventos 13
- 2.3 Conteo de puntos de la muestra 19
- 2.4 Probabilidad de un evento 27
- 2.5 Reglas aditivas 30
- 2.6 Probabilidad condicional 35
- 2.7 Reglas multiplicativas 38
- 2.8 Regla de Bayes 44
- Ejercicios de repaso 49

3 Variables aleatorias y distribuciones de probabilidad 51

- 3.1 Concepto de variable aleatoria 51
- 3.2 Distribuciones discretas de probabilidad 53
- 3.3 Distribuciones continuas de probabilidad 58
- 3.4 Distribuciones empíricas 63
- 3.5 Distribuciones de probabilidad conjunta 69
- Ejercicios de repaso 82

4 Esperanza matemática 84

- 4.1 Media de una variable aleatoria 84
- 4.2 Varianza y covarianza 92
- 4.3 Medias y varianzas de combinaciones lineales de variables aleatorias 101
- 4.4 Teorema de Chebyshev 108
- Ejercicios de repaso 112

5 Algunas distribuciones de probabilidad discreta 114

- 5.1 Introducción 114
- 5.2 Distribución uniforme discreta 114
- 5.3 Distribuciones binomial y multinomial 116
- 5.4 Distribución hipergeométrica 125
- 5.5 Distribuciones binomial negativa y geométrica 132
- 5.6 Distribución de Poisson y proceso de Poisson 135
- Ejercicios de repaso 141

6 Algunas distribuciones continuas de probabilidad 143

- 6.1 Distribución de probabilidad continua 143
- 6.2 Distribución normal 145
- 6.3 Áreas bajo la curva normal 148
- 6.4 Aplicaciones de la distribución normal 153
- 6.5 Aproximación normal a la binomial 160
- 6.6 Distribuciones gamma y exponencial 166
- 6.7 Aplicaciones de las distribuciones exponencial y gamma 170
- 6.8 Distribución ji cuadrada 172
- 6.9 Distribución logarítmica normal 173
- 6.10 Distribución de Weibull 174
- Ejercicios de repaso 178

7 Funciones de variables aleatorias 180

- 7.1 Introducción 180
- 7.2 Transformaciones de variables 180
- 7.3 Momentos y funciones generadoras de momentos 189

8 Distribuciones fundamentales de muestreo y descripciones de datos 198

- 8.1 Muestreo aleatorio 198
- 8.2 Algunas estadísticas importantes 201
- 8.3 Presentaciones de datos y métodos gráficos 208
- 8.4 Distribuciones muestrales 215
- 8.5 Distribuciones muestrales de medias 217
- 8.6 Distribución muestral de S^2 224
- 8.7 Distribución t 228
- 8.8 Distribución F 232
- Ejercicios de repaso 237

9 Problemas de estimación de una y dos muestras 238

- 9.1 Introducción 238
- 9.2 Inferencia estadística 238
- 9.3 Métodos clásicos de estimación 239
- 9.4 Una sola muestra: estimación de la media 243
- 9.5 Error estándar de una estimación puntual 248
- 9.6 Límites de tolerancia 249
- 9.7 Dos muestras: estimación de la diferencia entre dos medias 253
- 9.8 Observaciones pareadas 259
- 9.9 Una sola muestra: estimación de una proporción 264
- 9.10 Dos muestras: estimación de la diferencia entre dos proporciones 268
- 9.11 Una sola muestra: estimación de la varianza 271
- 9.12 Dos muestras: estimación de la razón de dos varianzas 273
- 9.13 Métodos bayesianos de estimación 275
- 9.14 Estimación de probabilidad máxima 282
- Ejercicios de repaso 287

10 Pruebas de hipótesis de una y dos muestras 290

- 10.1 Hipótesis estadísticas: conceptos generales 290
- 10.2 Prueba de una hipótesis estadística 292
- 10.3 Pruebas de una y dos colas 300
- 10.4 Uso de valores P para la toma de decisiones 302
- 10.5 Una sola muestra: pruebas con respecto a una sola media (varianza conocida) 306
- 10.6 Relación con la estimación del intervalo de confianza 309
- 10.7 Una sola muestra: pruebas sobre una sola media (varianza desconocida) 310
- 10.8 Dos muestras: pruebas sobre dos medias 313
- 10.9 Elección del tamaño de la muestra para probar medias 318
- 10.10 Métodos gráficos para comparar medias 323
- 10.11 Una muestra: prueba sobre una sola proporción 330
- 10.12 Dos muestras: pruebas sobre dos proporciones 333
- 10.13 Pruebas referentes a varianzas de una y dos muestras 336

- 10.14 Prueba de la bondad de ajuste 340
- 10.15 Prueba de independencia (datos categóricos) 344
- 10.16 Prueba de homogeneidad 347
- 10.17 Prueba para varias proporciones 348
- 10.18 Estudio de dos muestras 350
 - Ejercicios de repaso 355

11 *Regresión lineal simple y correlación* 358

- 11.1 Introducción a la regresión lineal 358
- 11.2 Regresión lineal simple 361
- 11.3 Propiedades de los estimadores de mínimos cuadrados 365
- 11.4 Inferencias acerca de los coeficientes de regresión 367
- 11.5 Predicción 371
- 11.6 Elección de un modelo de regresión 377
- 11.7 Procedimiento del análisis de varianza 377
- 11.8 Prueba para la linealidad de la regresión: datos con observaciones repetidas 379
- 11.9 Gráficas y transformaciones de datos 387
- 11.10 Estudio de regresión lineal simple 391
- 11.11 Correlación 394
 - Ejercicios de repaso 400

12 *Regresión lineal múltiple* 405

- 12.1 Introducción 405
- 12.2 Estimación de los coeficientes 406
- 12.3 Modelo de regresión lineal con el uso de matrices 410
- 12.4 Propiedades de los estimadores de mínimos cuadrados 418
- 12.5 Inferencias en la regresión lineal múltiple 420
- 12.6 Elección de un modelo de ajuste a través de la prueba de hipótesis 427
- 12.7 Caso especial de ortogonalidad 431
- 12.8 Métodos secuenciales para la selección del modelo 435
- 12.9 Estudio de residuos y violación de suposiciones 441
- 12.10 Validación cruzada, C_p y otros criterios para la selección del modelo 445
 - Ejercicios de repaso 456

13 *Experimentos de un factor: general* 461

- 13.1 Técnica del análisis de varianza 461
- 13.2 La estrategia del diseño experimental 463
- 13.3 Análisis de varianza unilateral: diseño completamente aleatorizado 463
- 13.4 Pruebas de la igualdad de varias varianzas 470

- 13.5 Comparaciones de un solo grado de libertad 475
- 13.6 Comparaciones múltiples 479
- 13.7 Comparación de tratamientos con un control 483
- 13.8 Comparación de un conjunto de tratamientos en bloques 488
- 13.9 Diseño de bloques completos aleatorizados 489
- 13.10 Métodos gráficos y diagnósticos posteriores 498
- 13.11 Cuadrados latinos 499
- 13.12 Modelos de efectos aleatorios 506
- 13.13 Enfoque de regresión para el análisis de varianza 511
- 13.14 Potencia de las pruebas de análisis de varianza 514
- 13.15 Estudio de caso 519
 - Ejercicios de repaso 523

14 Experimentos factoriales 527

- 14.1 Introducción 527
- 14.2 Interacción y el experimento de dos factores 529
- 14.3 Análisis de varianza de dos factores 530
- 14.4 Análisis gráfico en el problema de dos factores 537
- 14.5 Experimentos de tres factores 540
- 14.6 Modelos multifactoriales específicos 544
- 14.7 Experimentos factoriales modelo II y III 549
- 14.8 Elección del tamaño de la muestra 552
 - Ejercicios de repaso 555

15 Experimentos 2^k factoriales y fracciones 559

- 15.1 Introducción 559
- 15.2 Análisis de varianza 560
- 15.3 Experimento 2^k factorial sin réplicas 564
- 15.4 Estudio de caso 565
- 15.5 Experimentos factoriales en bloques incompletos 571
- 15.6 Confusión parcial 577
- 15.7 Experimentos factoriales en un ajuste de regresión 579
- 15.8 Estudio de caso: experimento de purificación de carbón 583
- 15.9 Experimentos factoriales fraccionados 588
- 15.10 Análisis de experimentos factoriales fraccionados 591
- 15.11 Diseños de fracciones superiores y depurados 595
- 15.12 Construcción de diseños de resolución III y IV con 8, 16 y 32 puntos de diseño 596
- 15.13 Otros diseños de resolución III de dos niveles; diseños de Plackett-Burman 597
- 15.14 Diseño robusto de parámetros de Taguchi 599
 - Ejercicios de repaso 607

16 Estadística no paramétrica 609

- 16.1 Pruebas no paramétricas 609
- 16.2 Prueba de signo 610
- 16.3 Prueba de rango con signo 614
- 16.4 Prueba de la suma de rangos 620
- 16.5 Prueba de Kruskal Wallis 623
- 16.6 Pruebas de corridas 625
- 16.7 Límites de tolerancia 629
- 16.8 Coeficiente de correlación de rango 630
- Ejercicios de repaso 634

17 Control estadístico de calidad 635

- 17.1 Introducción 635
- 17.2 Naturaleza de los límites de control 637
- 17.3 Propósitos de la gráfica de control 637
- 17.4 Gráficas de control para variables 638
- 17.5 Gráficas de control para atributos 653
- 17.6 Gráficas de control de cusum 660
- Ejercicios de repaso 664

Bibliografía 667

Apéndice: Tablas estadísticas 671

Respuesta a ejercicios con número impar 723

Índice 737

Prefacio

La sexta edición de *Probabilidad y estadística para ingenieros* destaca la teoría estadística y pone un énfasis apropiado en el análisis de datos, el cálculo estadístico y la mejora de la calidad. Nuestro objetivo es presentar al estudiante los elementos de la probabilidad y las distribuciones de probabilidad, la inferencia estadística, la regresión y correlación lineal, el análisis de varianza, el diseño de experimentos y el control estadístico de la calidad. Este equilibrio entre la teoría y la metodología ha encontrado históricamente aceptación por parte de los revisores y los usuarios del texto. La quinta edición destacó la introducción de impresos de software para realizar los ejemplos e introdujo también los métodos gráficos en el análisis de datos. Ambos temas se refuerzan en la sexta edición, con impresos adicionales incluidos en el texto y también junto con casos de estudio dentro de los ejercicios. Otra ventaja de este texto es la abundancia de ejercicios de la “vida real” que incluyen conjuntos de datos de estudios reales realizados por ingenieros y científicos. Se amplían los ejercicios, en particular para las áreas de pruebas de dos muestras y estimación del intervalo de confianza, análisis de varianza de uno y dos factores, y análisis de regresión. A petición popular, se proporcionan las respuestas para los ejercicios impares.

Un prerrequisito para el uso de este texto es un curso de cálculo diferencial e integral, con derivación parcial e integración múltiple. Si el instructor intenta hacer un tratamiento completo de la regresión lineal múltiple, sería útil un curso de álgebra lineal y matrices.

Cambios importantes en la sexta edición

1. El primer capítulo ha sido ampliado para proporcionar al lector una introducción sólida de las nociones de posición y variabilidad. Se han agregado ejercicios a este capítulo.
2. En casi todos los capítulos aparecen nuevos ejercicios de la "vida real". Se ha puesto énfasis en los ejercicios de repaso.
3. Se introducen distribuciones uniformes continuas y logarítmica normal, y se incluyen aplicaciones y ejercicios con el uso de la distribución logarítmica normal.
4. Se hace énfasis apropiado en el uso de la distribución t para adaptar pruebas y la estimación que incluye dos medias con varianzas no homogéneas.
5. Se omite la sección sobre teoría de la decisión y se hace más énfasis en la estimación de la probabilidad máxima. Se utilizan más ejemplos y ejercicios para reforzar el material del texto sobre probabilidad máxima.
6. Se abarca en forma más completa el uso del álgebra matricial en el análisis de la parte de varianzas de la regresión lineal múltiple.
7. El capítulo sobre experimentos factoriales se complementa con la inclusión de modelos mezclados.
8. En los capítulos de regresión múltiple y análisis de varianzas aparecen ejercicios de "casos de estudio" y se incluyen muestras de impresos de computadora.
9. Los archivos de datos para este libro se pueden descargar del siguiente sitio ftp://w.w.w.prenhall.com/pub/esm/statistics.027/walpole/probability_statistics

Este texto está diseñado para un curso de uno o de dos semestres. Un curso de un semestre que contenga el tratamiento de probabilidad e inferencia estadística se adapta con el material de los capítulos 1 al 10. Sin embargo, existe mucha flexibilidad para el uso de estos capítulos. Después de la nueva introducción, los capítulos 2, 3 y 4 cubren la probabilidad básica y las variables aleatorias. Los capítulos 5 y 6 tratan las distribuciones específicas, sus aplicaciones y las relaciones entre distribuciones. El capítulo 5 trata las distribuciones discretas, incluida la binomial y la hipergeométrica multinomial, binomial negativa, geométrica, uniforme y de Poisson. El capítulo 6 explica la normal (incluye la aproximación normal a la binomial), gamma, logarítmica normal, exponencial, ji cuadrada y otras distribuciones útiles. Es posible que, para un curso más teórico, el instructor pueda elegir no utilizar la totalidad del material en estos dos capítulos. Por otro lado, un curso para ingenieros ciertamente se beneficiaría de la discusión de las áreas de aplicación de estas importantes distribuciones y del uso de los ejercicios. El capítulo 7 es un capítulo corto titulado Funciones de variables aleatorias y es el tema más matemático del libro. Se incluye el uso de funciones generadoras de momento y técnicas para calcular la distribución de funciones de variables aleatorias. El instructor que esté más inclinado a enfatizar la estadística matemática puede preferir incluirlo.

El capítulo 8 contiene tres importantes temas relacionados: muestreo aleatorio, descripción de datos y distribuciones de muestreo fundamentales. Como se mencionó al principio, se pone considerable atención al despliegue de datos y a las gráficas. Además, se colocan las bases para el uso de percentiles y la graficación de la

probabilidad normal en aplicaciones que aparecen en los capítulos siguientes. Se discute la distribución de X y S^2 , y se proporciona la motivación para su uso en los dos capítulos siguientes. A este respecto, se introducen las distribuciones t y F , así como la motivación para su uso en la estimación del intervalo y la prueba de hipótesis.

Los capítulos 9 y 10 contienen material sobre la estimación de una y dos muestras y la prueba de hipótesis, respectivamente. Se incluyen nuevos ejercicios y ejemplos, y se introduce al estudiante a conjuntos de datos de "la vida real". Las ilustraciones incluyen el uso de comparaciones gráficas así como la inferencia formal —impresos de computadora comentados forman parte de la ilustración. El capítulo 9 incluye también material sobre probabilidad máxima y estimación de Bayes. Las pruebas teóricas normales que incluyen las distribuciones t , F y χ^2 se dan en el capítulo 10, como pruebas que implican el uso de datos categóricos. Los capítulos 11 y 12 se ocupan de la regresión lineal simple y la regresión múltiple, respectivamente. Como en el pasado, se incluyen los criterios de selección del modelo de un solo número, procedimientos paso a paso, diagnóstico con gráficas de residuos y una discusión de transformaciones.

El capítulo 13 es una presentación del problema de un factor. Se introduce el análisis de varianza para prueba de medias. Se incluyen los bloques, cuadrados latinos y el modelo de efectos aleatorios. Un extenso estudio de caso ilustra la unión de gráficas estadísticas y la inferencia formal. El capítulo 14 contiene el experimento factorial general, con gráficas e impresos de computadora comentados y resaltados una vez más. El capítulo 15 contiene el material mencionado con anterioridad sobre experimentos 2^k factoriales y fraccionales, incluidos los bloques y la confusión. Se resaltan los métodos modernos de mejora de la calidad por medio de ejemplos y casos de estudio, y se explican la metodología de Taguchi así como modelado de varianza. El capítulo 16 se ocupa de los procedimientos no paramétricos, y el capítulo 17 está dedicado al control estadístico de la calidad.

Los capítulos 11 a 17, a nuestro juicio, proporcionan material suficiente para un segundo semestre. Cada capítulo contiene suficiente flexibilidad y permite la omisión de secciones sin pérdida de continuidad.

Queremos agradecer a los que contribuyeron con la preparación de esta sexta edición. Agradecemos a Linda Seawell, que mecanografió el manuscrito. De nuevo, muchas gracias al Centro de Consulta Estadística del Instituto Politécnico de Virginia: fuente de muchos de los conjuntos de datos de la "vida real" que se utilizan como ejemplos y ejercicios.

Estamos en deuda con aquellos colegas que revisaron la sexta edición y proporcionaron muchas sugerencias útiles para esta edición. Ellos son Melinda McCann, Universidad Estatal de Oklahoma; Bill Beyer, Universidad de Akron; Michael Merscher, Universidad Tecnológica Lawrence; y Kyuil Kim, Universidad de Illinois en Chicago.

También estamos en deuda con el albacea literario del fallecido sir Ronald A. Fisher, F.R.S., Cambridge, y a Oliver & Boyd Ltd., Edimburgo, por su permiso para reimprimir una tabla de su libro *Statistical Methods for Research Workers*; al profesor E. S. Pearson y a los administradores de Biometrika por el permiso para reimprimir de forma condensada las tablas 8 y 18 de *Biometrika Tables for Statisticians*, vol I; a Oliver & Boyd Ltd., por el permiso para reproducir tablas de su libro

Design and Analysis of Industrial Experiments de O. L. Davies; a McGraw-Hill Book Company por el permiso para reproducir las tablas A-25d y A-25e de su libro *Introduction to Statistical Analysis* de W. J. Dixon y R. J. Massey, Jr; a C. Eisenhart, M. W. Hastay y W. A. Wallis por el permiso para reproducir dos tablas de su libro *Techniques of Statistical Analysis*. También deseamos expresar nuestro aprecio por el permiso para reproducir tablas de los *Annals of Mathematical Statistics*, del *Bulletin of the Educational Research at Indiana University*, de una publicación de la American Cyanamid Company, de *Biometrics*, de *Biometrika*, vol 38, y de la *Journal of the American Statistical Association*.

R.H.M. S.L.M.

Introducción a la estadística y al análisis de datos

1.1 Repaso

En la industria estadounidense se dedica una cantidad extraordinaria de atención al *mejoramiento de la calidad* que comenzó en los ochenta y continúa en la presente década. Se ha dicho y escrito mucho del "milagro industrial" japonés que comenzó en la segunda mitad del siglo XX. Fueron capaces de tener éxito donde nuestro país (E.U.) y otros países fallaron —a saber, en la creación de una atmósfera que permite la manufactura de productos de alta calidad. Gran parte del éxito de los japoneses se atribuye al uso de *métodos estadísticos* y al pensamiento estadístico entre el personal gerencial.

El uso de métodos estadísticos en la manufactura, el desarrollo de productos alimenticios, software de computadoras, productos farmacéuticos y muchas otras áreas implica el acopio de información o **datos científicos**. Por supuesto que el acopio de datos no es algo nuevo, ya que se ha realizado por más de mil años. Los datos se han recabado, resumido, reportado y almacenado para su examen cuidadoso. Sin embargo, hay una diferencia profunda entre recabar información científica y la **estadística inferencial**. Es esta última la que ha recibido legítima atención en las décadas recientes. La inferencia estadística produjo ya un número enorme de herramientas analíticas que permiten al ingeniero o al científico comprender mejor los sistemas que generan los datos. Esto refleja la verdadera naturaleza de la ciencia que llamamos estadística inferencial, principalmente el uso de técnicas que nos permiten ir más allá de sólo reportar datos sino, más bien, nos permite obtener conclusiones (o inferencias) sobre el sistema científico. Los estadísticos hacen uso de leyes fundamentales de probabilidad e inferencia estadística para sacar conclusiones sobre los sistemas científicos. La información se colecta en forma de **muestras**, o

agrupaciones de **observaciones**. En el capítulo 2 se introduce el proceso de muestreo y la discusión continúa a lo largo de todo el libro.

Las muestras se reúnen a partir de **poblaciones**, que son agrupaciones de todos los individuos o elementos individuales de un tipo particular. A veces una población significa un sistema científico. Por ejemplo, un fabricante de tarjetas de computadora puede desear eliminar defectos. Un proceso de muestreo puede implicar la recolección de información de 50 tarjetas de computadora probadas aleatoriamente en el proceso. Aquí, la población es todas las tarjetas de computadora producidas por la empresa en un periodo específico. En un experimento con medicamentos, se toma una muestra de pacientes y a cada uno se le administra un medicamento específico para reducir la presión sanguínea. Se enfoca el interés en la obtención de conclusiones sobre la población de los que sufren de hipertensión. Si se hace una mejora en el proceso de producción de las tarjetas de computadora y se reúne una segunda muestra de tarjetas, cualesquiera conclusiones que se obtengan respecto a la efectividad del cambio en el proceso se debe extender a toda la población de tarjetas de computadora que se produzcan bajo el "proceso mejorado".

A menudo, es muy importante el acopio de datos científicos en forma sistemática, cuando la planeación ocupa un lugar importante en la agenda. En ocasiones la planeación está, por necesidad, bastante limitada. Un ingeniero puede necesitar estudiar el efecto de las condiciones del proceso, temperatura, humedad, cantidad de un ingrediente particular, etcétera, sobre la producción del artículo. Él o ella pueden mover de manera sistemática estos **factores** a cualesquiera niveles que se sugieran de acuerdo con cualquier prescripción o **diseño experimental** que se desee. Sin embargo, un científico silvicultor que está interesado en un estudio de los factores que influyen en la densidad de la madera en cierta clase de árbol no necesariamente puede diseñar un experimento. En este caso puede requerir un **estudio observacional** en el que los datos se acopian en el campo pero no se pueden seleccionar de antemano los **niveles de los factores**. Ambos tipos de estudios se prestan a los métodos de la inferencia estadística. En el primero, la calidad de las inferencias dependerá de la planeación apropiada del experimento. En el último, el científico está a merced de lo que pueda recopilar. Por ejemplo, resulta inadecuado si un agrónomo se interesa en estudiar el efecto de la lluvia sobre la producción de plantas y los datos se obtienen durante una sequía.

Se debe comprender la importancia del pensamiento estadístico en los administradores y el uso de la inferencia estadística por el personal científico. Los investigadores obtienen mucho de los datos científicos. Los datos proporcionan una comprensión del fenómeno científico. Los ingenieros de producto y procesos aprenden más en sus esfuerzos fuera de línea para mejorar el proceso. También logran una comprensión valiosa al reunir datos de producción (monitoreo en línea) con una base regular. Esto permite la determinación de las modificaciones necesarias a fin de mantener el proceso en un nivel deseado de calidad.

1.2 El papel de la probabilidad

En este libro, los capítulos 2 a 6 tratan de las nociones fundamentales de la probabilidad. Una obtención esmerada de las bases de estos conceptos permite al lector lograr una mejor comprensión de la inferencia estadística. Sin algo de formalismo en probabilidad el estudiante no puede apreciar la verdadera interpretación del análisis

de datos mediante los métodos estadísticos modernos. Es completamente natural estudiar probabilidad antes de estudiar inferencia estadística. Los elementos de probabilidad nos permiten cuantificar la fuerza o "confianza" en nuestras conclusiones.

Ejemplo 1.1 Suponga que un ingeniero se encuentra con datos de un proceso de producción en el que se muestrean 100 artículos y se encuentra que 10 están defectuosos. Es de esperarse que de vez en cuando habrá artículos defectuosos. Sin embargo, se determina que, a largo plazo, la empresa sólo puede tolerar 5% de artículos defectuosos en el proceso. Ahora bien, los elementos de probabilidad permiten al ingeniero determinar qué tan concluyente es la información muestral con respecto a la naturaleza del proceso. Suponga que averiguamos que *si el proceso es aceptable, es decir, si produce artículos con el 5% defectuosos, hay una probabilidad de 0.00001 de obtener 10 o más artículos defectuosos en una muestra aleatoria de 100 artículos del proceso*. Esta pequeña probabilidad sugiere que el proceso, en realidad, tiene un porcentaje de artículos defectuosos en el largo plazo que excede un 5%. En otras palabras, bajo la condición de un proceso aceptable, la información muestral que se obtuvo casi nunca ocurriría. Claramente, sin embargo, ocurriría con una probabilidad mucho más alta si la tasa de artículos defectuosos del proceso excediese 5% por un monto significativo. ■

De este ejemplo es evidente que los elementos de probabilidad ayudan en la traducción de información muestral en algo concluyente o no concluyente acerca del sistema científico. El ejemplo que sigue proporciona una segunda ilustración. El valor 0.00001, que jugó un gran papel en nuestra conclusión, a menudo se denomina **valor P** . En los capítulos siguientes se trata detenidamente el concepto de **valor- P** .

Ejemplo 1.2 Con frecuencia la naturaleza del estudio científico señalará el papel que juegan la probabilidad y el razonamiento deductivo en la inferencia estadística. El ejercicio 6 al final de la sección 9.8 proporciona datos asociados con un estudio que se llevó a cabo en el Instituto Politécnico y Universidad Estatal de Virginia sobre el desarrollo de una relación entre las raíces de los árboles y la acción de un hongo. Se transfirieron minerales de los hongos a los árboles y azúcares de los árboles al hongo. Se plantaron dos muestras de 10 plántones de roble rojo norteno en un invernadero, una que contenía plántones tratados con nitrógeno y una muestra de plántones sin tratamiento. Todas las demás condiciones ambientales se mantuvieron constantes. Todos los plántones contenían el hongo *Pisolithus tinctoris*. En el capítulo 9 se proporcionan más detalles. Los pesos de los tallos en gramos se registraron al finalizar 140 días. Los datos son como sigue:

Sin nitrógeno	Con nitrógeno
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46



Figura 1.1 Pesos de los tallos.

Es conveniente graficar los datos como se muestra en la figura 1.1. Los valores \circ representan los datos “con nitrógeno” y los valores \times representan los datos “sin nitrógeno”. Ahora bien, el propósito de este experimento es determinar si el uso de nitrógeno tiene influencia en el crecimiento de las raíces. Nótese que la apariencia general de los datos puede sugerir al lector que, en promedio, el uso del nitrógeno aumenta el peso del tallo. Cuatro observaciones con nitrógeno son considerablemente más grandes que cualquiera de las observaciones sin nitrógeno. La mayor parte de las observaciones sin nitrógeno parecen estar por debajo del centro de los datos. La apariencia del conjunto de datos parecería indicar que el nitrógeno es efectivo. Pero, ¿cómo se puede cuantificar esto? ¿Cómo se puede resumir toda la evidencia visual aparente en un significado? Como en el ejemplo anterior, se pueden utilizar los fundamentos de la probabilidad. Las conclusiones se pueden resumir en una declaración de probabilidad o valor P . El problema gira alrededor de la “probabilidad de que datos como estos se puedan observar” *dado que el nitrógeno no tiene efecto*, en otras palabras, *dado que ambas muestras se generaron a partir de la misma población*. Suponga que esta probabilidad es pequeña, digamos 0.03; ésta sería con certeza suficiente evidencia de que el uso del nitrógeno en realidad influye (aparentemente aumenta) el peso promedio del tallo. ■

Dimos dos ejemplos donde los elementos de probabilidad proporcionan un resumen que el científico o el ingeniero puede usar como evidencia sobre la cual fincar una decisión. El puente entre los datos y la conclusión está, por supuesto, basado en los fundamentos de la inferencia estadística, la teoría de la distribución y las distribuciones de muestreo que se tratan e ilustran en los capítulos 7 a 11. Es conveniente en este punto prestar atención a las medidas de posición y variabilidad de la muestra. Tanto el análisis exploratorio o intuitivo de datos como la inferencia estadística formal dependen de estas medidas.

1.3 Medidas de posición: media de una muestra

Las medidas de posición en un conjunto de datos están diseñadas para proporcionar al analista alguna medida cuantitativa de dónde está el centro de los datos en una muestra. En el ejemplo 1.2 ciertamente parece como si el centro de la muestra con nitrógeno excediera al de la muestra sin nitrógeno. Una medida obvia y muy útil es la **media de la muestra**. La media es simplemente un promedio numérico. Suponga que las observaciones en una muestra son x_1, x_2, \dots, x_n . La media de la muestra es

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Hay otras medidas de tendencia central que se explican con detalle en capítulos posteriores. Una medida importante es la **mediana de la muestra**. El propósito de la mediana de la muestra es reflejar la tendencia central de la muestra de manera que

no esté influida por los valores extremos. Dado que las observaciones en una muestra son x_1, x_2, \dots, x_n , acomodados en orden creciente de magnitud, la mediana de la muestra es

$$\begin{aligned}\bar{x} &= x_{(n+1)/2} && \text{si } n \text{ es impar,} \\ &= \frac{x_{n/2} + x_{(n/2)+1}}{2} && \text{si } n \text{ es par.}\end{aligned}$$

Por ejemplo, supongamos que el conjunto de datos es el siguiente: 1.7, 2.2, 3.9, 3.11 y 14.7. La media y la mediana de la muestra son, respectivamente,

$$\bar{x} = 5.12, \quad \tilde{x} = 3.9.$$

Es evidente que la media está influida de manera considerable por la presencia de la observación extrema, 14.7, mientras que la mediana hace énfasis en el verdadero "centro" del conjunto de datos. En el caso del conjunto de datos de dos muestras del ejemplo 1.3, las dos medidas de tendencia central para las muestras individuales son

$$\bar{x} \text{ (sin nitrógeno)} = 0.40 \text{ gramos,}$$

$$\tilde{x} \text{ (sin nitrógeno)} = \frac{0.37 + 0.42}{2} = 0.395 \text{ gramos,}$$

$$\bar{x} \text{ (con nitrógeno)} = 0.57 \text{ gramos,}$$

$$\tilde{x} \text{ (con nitrógeno)} = \frac{0.49 + 0.52}{2} = 0.505 \text{ gramos.}$$

1.4 Medidas de variabilidad

La variabilidad de una muestra juega un papel importante en el análisis de datos. La variabilidad de un proceso y de un producto es un hecho real en los sistemas científicos y de ingeniería: el control o reducción de la variabilidad de un proceso a menudo es una fuente de dificultades mayores. La variabilidad en valores de población y datos de una muestra es un hecho real. Una variabilidad grande entre las observaciones de una muestra a menudo puede "limpiar" cualesquiera efectos que el ingeniero pueda tratar de detectar. Las medidas de posición en una muestra no proporcionan un resumen apropiado de la naturaleza de un conjunto de datos. Es decir, en el ejemplo 1.2 no podemos concluir que el uso del nitrógeno realza el crecimiento sin tomar en cuenta la variabilidad de la muestra. Por ejemplo, compare los dos conjuntos de datos de abajo. Cada uno contiene dos muestras y la diferencia en las medias es aproximadamente la misma para las dos muestras:

Conjunto de datos A: $\times \times \times \times \times \times \quad 0 \times \times 0 \quad 0 \times \times \times 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0$

 \downarrow
 \bar{x}_a
 \downarrow
 \bar{x}_b

Conjunto de datos B: $\times \times \times \times \times \times \quad \times \times \times \times \times \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0$

 \downarrow
 \bar{x}_a
 \downarrow
 \bar{x}_b

El conjunto de datos B parece proporcionar un contraste mucho más claro entre las dos poblaciones de las que se tomaron las muestras. Este no es el caso con el conjunto de datos A. Si el propósito de tal experimento es detectar la diferencia entre las dos poblaciones, la tarea se lleva a cabo en el caso del conjunto de datos B. Sin embargo, en el conjunto de datos A la gran variabilidad dentro de las dos muestras crea dificultad. De hecho, no es claro que haya una diferencia entre las dos poblaciones.

Así como hay muchas medidas de tendencia central o de posición hay muchas medidas de dispersión o variabilidad. Quizá la más simple es el **rango de la muestra** $X_{\max} - X_{\min}$. El rango puede ser muy útil y se discute con amplitud en el capítulo 17 sobre *control estadístico de calidad*. La medida muestral de dispersión que se utiliza más a menudo es la **desviación estándar de la muestra**. Nuevamente denotemos con x_1, x_2, \dots, x_n los valores de la muestra; la desviación estándar de la muestra es

$$s = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{(n-1)}}$$

La cantidad $s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)$ se llama **varianza de la muestra**. Debe quedar claro al lector que la desviación estándar de la muestra es, de hecho, una medida de variabilidad. Una variabilidad grande en un conjunto de datos produce valores relativamente grandes de $(x - \bar{x})^2$ y por ello una varianza de la muestra grande. La cantidad $n-1$ a menudo se denomina **grados de libertad asociados con la varianza** estimada. En este ejemplo simple los grados de libertad representan el número de piezas de información independientes disponibles para calcular la variabilidad. Por ejemplo, suponga que deseamos calcular la varianza de la muestra y la desviación estándar del conjunto de datos (5, 17, 6, 4). El promedio de la muestra es $\bar{x} = 8$. El cálculo de la varianza implica:

$$(5 - 8)^2 + (17 - 8)^2 + (6 - 8)^2 + (4 - 8)^2 = (-3)^2 + (9)^2 + (-2)^2 + (-4)^2.$$

Las cantidades dentro de los paréntesis suman cero. En general, $\sum_{i=1}^n (x_i - \bar{x}) = 0$.

Véase el ejercicio 4. Entonces el cálculo de una varianza de la muestra no implica n **desviaciones cuadradas independientes** de la media \bar{x} . De hecho, como el último valor de $x - \bar{x}$ está determinado por los primeros $n-1$ valores, decimos que estas son $n-1$ "piezas de información" que producen s^2 . Por ello hay $n-1$ grados de libertad en lugar de n grados de libertad para calcular una varianza de la muestra.

1.5 Datos discretos y continuos

La inferencia estadística a través del análisis de estudios observacionales o de experimentos diseñados se utiliza en muchas áreas científicas. Los datos reunidos pueden ser **discretos** o **continuos**, dependiendo del área de aplicación. Por ejemplo, un ingeniero químico se puede interesar en un experimento que lo lleve a condiciones donde se maximice la producción. Aquí, por supuesto, la producción puede ser en porcentaje, o gramos/libra, medida en un continuo. Por otro lado, un toxicólogo que realice un experimento de combinación de fármacos puede encon-

trar datos que son binarios por naturaleza (es decir, el paciente responde o no). Para estas dos situaciones distintas la teoría de la probabilidad que nos permite extraer inferencias estadísticas es completamente diferente. En el capítulo 3 y posteriores se obtienen diferencias claras entre situaciones discretas y continuas.

1.6 Modelado estadístico, inspección científica y diagnósticos gráficos

Muy a menudo el resultado final de un análisis estadístico es la estimación de los parámetros de un **modelo postulado**. Esto es por completo natural para los científicos y los ingenieros, pues con frecuencia tratan con el modelado. Un modelo estadístico no es determinista sino, más bien, debe involucrar algunos aspectos probabilistas. Una forma modelo a menudo es la fundamentación de las **suposiciones** que hace el analista. En nuestro ejemplo 1.2, el científico puede desear extraer algún nivel de distinción entre la población “con nitrógeno” y “sin nitrógeno” a través de la información de la muestra. El análisis puede requerir cierto modelo para los datos, por ejemplo, que las dos muestras provengan de **distribuciones normales** o **gaussianas**. Véase el capítulo 6 para la discusión de una distribución normal.

A veces el modelo postulado puede tener una forma algo más complicada. Considere, por ejemplo, un fabricante de textiles que diseña un experimento donde las muestras de tela se producen de modo que contengan diferentes porcentajes de algodón. Considere los datos siguientes:

Porcentaje del algodón	Resistencia a la tracción
15	7, 7, 9, 8, 10
20	19, 20, 21, 20, 22
25	21, 21, 17, 19, 20
30	8, 7, 8, 9, 10

Se fabrican cinco muestras de tela para cada uno de los cuatro porcentajes de algodón. Ahora bien, en este caso tanto el modelo para el experimento como el tipo de análisis que se utiliza deben tomar en cuenta el objetivo del experimento y los importantes insumos del científico textil. Algunas gráficas simples pueden aclarar la distinción entre las muestras. Véase la figura 1.2; las medias de las mues-

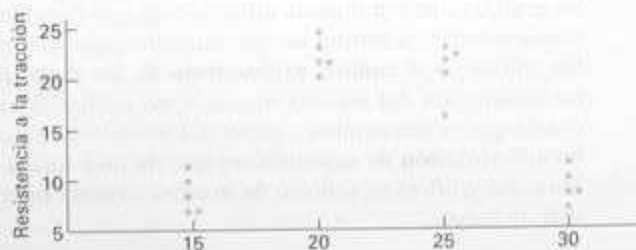


Figura 1.2 Gráfica de la resistencia a la tracción y los porcentajes de algodón.

tras y la variabilidad se describen bien en la gráfica de datos. Un posible objetivo de este experimento es simplemente la determinación de cuáles porcentajes de algodón son en verdad distintos de los otros. En otras palabras, como en el caso de los datos con nitrógeno/sin nitrógeno, ¿para cuáles porcentajes de algodón hay distinciones claras entre las poblaciones o, de manera más específica, entre las medias de las poblaciones? En este caso, quizá un modelo razonable es que cada muestra viene de una distribución normal. Aquí el objetivo es muy semejante al de los datos con nitrógeno/sin nitrógeno, excepto que se involucran más muestras. El formalismo del análisis implica nociones de prueba de hipótesis que se tratan en el capítulo 10. De paso, este formalismo quizá no es necesario a la luz de la gráfica de diagnóstico. Pero, ¿describe el objetivo real del experimento y por consiguiente el enfoque apropiado del análisis de datos? Es probable que el científico anticipe la existencia de una *resistencia a la tracción de la media de la población máxima* en el rango de concentración de algodón en el experimento. Aquí el análisis de los datos debe girar alrededor de un tipo diferente de modelo, uno que postule un tipo de estructura que relacione la resistencia a la tracción de la media de la población con la concentración de algodón. En otras palabras, un modelo se puede escribir como

$$\mu_{t,c} = \beta_0 + \beta_1 C + \beta_2 C^2,$$

donde $\mu_{t,c}$ es la resistencia a la tracción de la media de la población, que varía con la cantidad de algodón en el producto C . La implicación de este modelo es que para un nivel fijo de algodón, hay una población de mediciones de resistencia a la tracción y la media de la población es $\mu_{t,c}$. Este tipo de modelo, que se denomina **modelo de regresión**, se trata en los capítulos 11 y 12. La forma funcional la elige el científico. A veces el análisis de datos puede sugerir que se cambie el modelo. Entonces el analista de datos "considera" un modelo que se puede alterar después de que se haga algún análisis. El uso de un modelo empírico se acompaña por la **teoría de estimación**, donde β_0 , β_1 y β_2 se estiman de los datos. Además, se puede usar entonces la inferencia estadística para determinar lo adecuado del modelo.

Aquí se hacen evidentes varios puntos de las dos ilustraciones de datos: 1) el tipo de modelo que se utiliza para describir los datos a menudo depende del objetivo del experimento y (2) la estructura del modelo debe aprovecharse del insumo científico no estadístico. La selección de un modelo representa una **suposición fundamental** sobre la que se basa la inferencia estadística resultante. Se hará evidente a lo largo del libro qué tan importantes pueden ser las gráficas. A menudo, las gráficas pueden ilustrar información que permite que los resultados de la inferencia estadística formal se comuniquen mejor al científico o al ingeniero. A veces, las gráficas o el **análisis exploratorio de los datos** puede enseñar al analista algo no recuperado del análisis formal. Casi cualquier análisis formal requiere suposiciones que se desarrollan a partir del modelo de datos. Las gráficas pueden resaltar bien la **violación de suposiciones** que de otra forma no se notarían. A lo largo del libro, las gráficas se utilizan de manera extensa para complementar el análisis formal de datos.

Ejercicios

1. Un fabricante de componentes electrónicos se interesa en determinar el tiempo de vida de cierto tipo de batería. La que sigue es una muestra, en horas de vida:

123, 116, 122, 110, 175, 126, 125, 111, 118, 117

(a) Encuentre la media y la mediana de la muestra.
 (b) ¿Cuál característica en este conjunto de datos es la responsable de la diferencia sustancial entre las dos?

2. Un fabricante de neumáticos quiere determinar el diámetro interior de cierto grado de neumático. Idealmente el diámetro sería 570 mm. Los datos son los siguientes:

572, 572, 573, 568, 569, 575, 565, 570

(a) Encuentre la media y la mediana de la muestra.
 (b) Encuentre la varianza, desviación estándar y rango de la muestra.
 (c) Mediante el uso de las estadísticas calculadas en las partes (a) y (b), ¿puede hacer comentarios sobre la calidad de los neumáticos?

3. Se conduce un estudio de los efectos de fumar sobre los patrones de sueño. La medición que se observa es el tiempo, en minutos, que toma quedar dormido. Se obtienen estos datos:

Fumadores: 69.3, 56.0, 22.1, 47.6, 53.2, 48.1,
 52.7, 34.4, 60.2, 43.8, 23.2, 13.8

No fumadores: 28.6, 25.1, 26.4, 34.9, 29.8, 28.4, 38.5, 30.2,
 30.6, 31.8, 41.6, 21.1, 36.0, 37.9, 13.9

(a) Encuentre la media de la muestra para cada grupo.
 (b) Encuentre la desviación estándar de la muestra para cada grupo.
 (c) Grafique los datos como en el caso de los conjuntos de datos A y B de la sección 1.4.
 (d) Comente qué clase de impacto parece tener el fumar sobre el tiempo que se requiere para quedar dormido.

4. Muestre que las n piezas de información en $\sum_{i=1}^n (X_i - \bar{X})^2$ no son independientes; es decir, muestre que

$$\sum_{i=1}^n (X_i - \bar{X}) = 0.$$

5. Dados cinco lanzamientos independientes de una moneda que tienen como resultado *cinco caras*. Resulta que si la moneda es legal, la probabilidad de este resultado es $(\frac{1}{2})^5 = 0.03125$. ¿Produce esto evidencia sólida de que la moneda no es legal? Comente y utilice el concepto de valor P que se discutió en la sección 1.2.

Probabilidad

2.1 Espacio muestral

En el estudio de la estadística tratamos básicamente con la presentación e interpretación de **resultados fortuitos** que ocurren en un estudio planeado o investigación científica. Por ejemplo, podemos registrar el número de accidentes que ocurren mensualmente en la intersección de Driftwood Lane y Royal Oak Drive, con el deseo de justificar la instalación de un semáforo; podemos clasificar los artículos que salen de una línea de montaje como "defectuosos" o "no defectuosos"; o nos podemos interesar en el volumen de gas que se libera en una reacción química cuando se hace variar la concentración de un ácido. Por ello, el estadístico a menudo trata con datos experimentales, conteos o mediciones representativos, o quizá con **datos categóricos** que se pueden clasificar de acuerdo con algún criterio.

Nos referiremos a cualquier registro de información, ya sea numérico o categórico, como una **observación**. Así, los números 2, 0, 1 y 2, que representan el número de accidentes que ocurrieron en cada mes de enero a abril durante el año pasado en la intersección de Driftwood Lane y Royal Oak Drive, constituyen un conjunto de observaciones. De forma similar, los datos categóricos N, D, N, N y D , que representan los artículos defectuosos o no defectuosos cuando se inspeccionan cinco artículos, se registran como observaciones.

Los estadísticos utilizan la palabra *experimento* para describir cualquier proceso que genere un conjunto de datos. Un ejemplo simple de experimento estadístico es el lanzamiento al aire de una moneda. En este experimento sólo hay dos resultados posibles, cara o cruz. Otro experimento puede ser el lanzamiento de un misil y la observación de su velocidad en tiempos específicos. Las opiniones de los votantes

con respecto a un nuevo impuesto sobre ventas también se pueden considerar como observaciones de un experimento. Estamos particularmente interesados en las observaciones que se obtienen por la repetición del experimento varias veces. En la mayor parte de los casos los resultados dependerán del azar y, por tanto, no se pueden predecir con certeza. Si un químico realiza un análisis varias veces bajo las mismas condiciones, obtendrá diferentes medidas, que indican un elemento de probabilidad en el procedimiento experimental. Incluso cuando se lanza al aire una moneda de forma repetida, no podemos tener la certeza de que un lanzamiento dado tendrá como resultado una cara. Sin embargo, conocemos el conjunto completo de posibilidades para cada lanzamiento.

Definición 2.1

El conjunto de todos los resultados posibles de un experimento estadístico se llama **espacio muestral** y se representa con el símbolo S .

Cada resultado en un espacio muestral se llama **elemento** o **miembro** del espacio muestral, o simplemente **punto muestral**. Si el espacio muestral tiene un número finito de elementos, podemos *listar* los miembros separados por comas y encerrarlos en paréntesis. De esta forma el espacio muestral S , de los resultados posibles cuando se lanza al aire una moneda, se puede escribir

$$S = \{H, T\}.$$

donde H y T corresponden a "caras" y "cruces", respectivamente.

Ejemplo 2.1 Considere el experimento de lanzar un dado. Si nos interesamos en el número que muestra en la cara superior, el espacio muestral sería

$$S_1 = \{1, 2, 3, 4, 5, 6\}.$$

Si nos interesamos sólo en si el número es par o impar, el espacio muestral es simplemente

$$S_2 = \{\text{par}, \text{impar}\}. \blacksquare$$

El ejemplo 2.1 ilustra el hecho de que se puede usar más de un espacio muestral para describir los resultados de un experimento. En este caso S_1 proporciona más información que S_2 . Si sabemos cuál elemento en S_1 tiene lugar, podemos decir cuál resultado ocurre en S_2 ; no obstante, el conocimiento de lo que pasa en S_2 no es de ayuda en la determinación de cuál elemento en S_1 ocurre. En general, se desea utilizar un espacio muestral que dé la mayor información acerca de los resultados del experimento.

En algunos experimentos es útil listar los elementos del espacio muestral de forma sistemática mediante un **diagrama de árbol**.

Ejemplo 2.2 Un experimento consiste en lanzar una moneda y después lanzarla una segunda vez si sale cara. Si sale cruz en el primer lanzamiento, entonces se lanza un dado una vez. Para listar los elementos del espacio muestral que proporcione la mayor información, construimos el diagrama de árbol de la figura 2.1. Ahora bien, las diversas trayectorias a lo largo de las ramas del árbol dan los distintos puntos de la muestra. Al comenzar con la rama superior izquierda y movernos a la derecha a

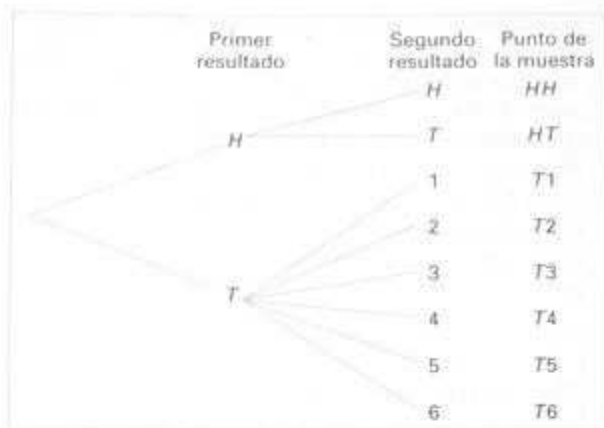


Figura 2.1 Diagrama de árbol para el ejemplo 2.2.

lo largo de la primera trayectoria, obtenemos el punto muestral HH , que indica la posibilidad de que ocurran caras en dos lanzamientos sucesivos de la moneda. Asimismo, el punto muestral $T3$ indica la posibilidad de que la moneda muestre una cruz seguida por un 3 en el lanzamiento del dado. Al seguir a lo largo de todas las trayectorias, vemos que el espacio muestral es

$$S = \{HH, HT, T1, T2, T3, T4, T5, T6\}. \quad \blacksquare$$

Ejemplo 2.3 Suponga que se seleccionan tres artículos de forma aleatoria de un proceso de fabricación. Cada artículo se inspecciona y clasifica como defectuoso, D , o sin defectos, N . Para listar los elementos del espacio muestral que proporcione la mayor información, construimos el diagrama de árbol de la figura 2.2. Ahora bien, las diversas trayectorias a lo largo de las ramas del árbol dan los distintos puntos muestrales. Al comenzar con la primera trayectoria, obtenemos el punto muestral DDD , que indica la posibilidad de que los tres artículos inspeccionados sean defectuosos. Conforme continuamos a lo largo de las demás trayectorias, vemos que el espacio muestral es

$$S = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

Los espacios muestrales con un número grande o infinito de puntos muestrales se describen mejor mediante un **enunciado** o **regla**. Por ejemplo, si los resultados posibles de un experimento son el conjunto de ciudades en el mundo con una población de más de un millón, nuestro espacio muestral se escribe

$$S = \{x | x \text{ es una ciudad con una población de más de un millón}\},$$

que se lee “ S es el conjunto de todas las x tales que x es una ciudad con una población de más de un millón”. La barra vertical se lee “tal que”. De manera similar, si S es el conjunto de todos los puntos (x, y) sobre la frontera o el interior de un círculo de radio 2 con centro en el origen, escribimos

$$S = \{(x, y) | x^2 + y^2 \leq 4\}.$$

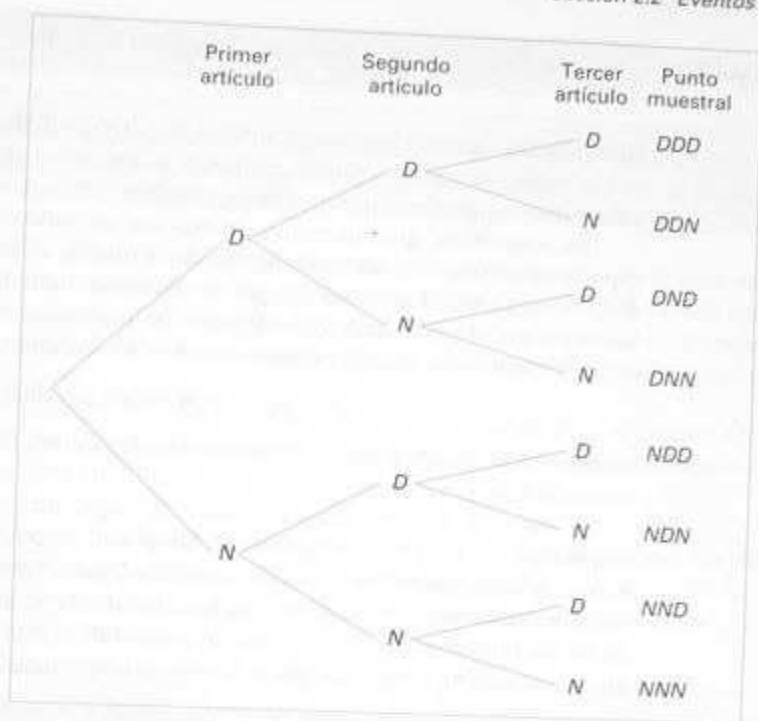


Figura 2.2 Diagrama de árbol para el ejemplo 2.3.

Si describimos el espacio muestral mediante el método de la regla o listando los elementos dependerá del problema específico en cuestión. El método de la regla tiene ventajas prácticas, en particular para los muchos experimentos donde un listado se vuelve una tarea tediosa. ■

2.2 Eventos

Para cualquier experimento dado podemos estar interesados en la ocurrencia de ciertos **eventos** más que en el resultado de un elemento específico en el espacio muestral. Por ejemplo, podemos estar interesados en el evento A en el que el resultado cuando se lanza un dado sea divisible entre 3. Éste ocurrirá si el resultado es un elemento del subconjunto $A = \{3, 6\}$ del espacio muestral S , del ejemplo 2.1. Como ilustración adicional, nos podemos interesar en el evento B de que el número de defectuosos sea mayor que 1 en el ejemplo 2.3. Esto ocurrirá si el resultado es un elemento del subconjunto

$$B = \{DDN, DND, NDD, DDD\}$$

del espacio muestral S .

Para cada evento asignamos una colección de puntos muestrales, que constituye un subconjunto del espacio muestral. Ese subconjunto representa la totalidad de los elementos para los que el evento es cierto.

Definición 2.2

Un **evento** es un subconjunto de un espacio muestral.

Ejemplo 2.4 Dado el espacio muestral $S = \{t | t \geq 0\}$, donde t es la vida en años de cierto componente electrónico, entonces el evento A de que el componente falle antes de que finalice el quinto año es el subconjunto $A = \{t | 0 \leq t < 5\}$.

Es concebible que un evento pueda ser un subconjunto que incluya todo el espacio muestral S , o un subconjunto de S que se denomina **conjunto vacío** y se denota mediante el símbolo \emptyset , que no contiene elemento alguno. Por ejemplo, si hacemos que A sea el evento de detectar un organismo microscópico a simple vista en un experimento biológico, entonces $A = \emptyset$. También, si

$$B = \{x | x \text{ es un factor par de } 7\},$$

entonces B debe ser el conjunto vacío, pues los únicos factores posibles de 7 son los números noes 1 y 7.

Considere un experimento donde se registran los hábitos de fumar de los empleados de una empresa industrial. Un posible espacio muestral podría clasificar a un individuo como no fumador, fumador ligero, fumador moderado o fumador empedernido. Sea el subconjunto de los fumadores un evento. Entonces la totalidad de los no fumadores corresponde a un evento diferente, también subconjunto de S , que se denomina complemento del conjunto de fumadores. ■

Definición 2.3

El **complemento** de un evento A con respecto a S es el subconjunto de todos los elementos de S que no están en A . Denotamos el complemento de A mediante el símbolo A' .

Ejemplo 2.5 Sea R el evento de que una carta roja se seleccione de una baraja ordinaria de 52 cartas, y sea S toda la baraja. Entonces R' es el evento de que la carta seleccionada de la baraja no sea una roja sino una negra. ■

Ejemplo 2.6 Considere el espacio muestral

$$S = \{\text{libro, catalizador, cigarrillo, precipitado, ingeniero, remache}\}.$$

Sea $A = \{\text{catalizador, remache, libro, cigarrillo}\}$. Entonces

$$A' = \{\text{precipitado, ingeniero}\}. \quad \blacksquare$$

Consideremos ahora ciertas operaciones con eventos que tendrán como resultado la formación de nuevos eventos. Estos eventos nuevos serán subconjuntos del mismo espacio muestral como los eventos dados. Suponga que A y B son dos eventos asociados con un experimento. En otras palabras, A y B son subconjuntos del mismo espacio muestral S . Por ejemplo, en el lanzamiento de un dado podemos hacer que A sea el evento de que ocurra un número par y B el evento de que aparezca un número mayor que 3. Entonces los subconjuntos $A = \{2, 4, 6\}$ y $B = \{4, 5, 6\}$ son subconjuntos del mismo espacio muestral

$$S = \{1, 2, 3, 4, 5, 6\}.$$

Nótese que A y B ocurrirán *ambos* en un lanzamiento dado si el resultado es un elemento del subconjunto $\{4, 6\}$, que es precisamente la intersección de A y B .

Definición 2.4

La **intersección** de dos eventos A y B , denotada mediante el símbolo $A \cap B$, es el evento que contiene a todos los elementos que son comunes a A y a B .

Ejemplo 2.7 Sea P el evento de que una persona seleccionada al azar mientras cena en un restaurante de moda sea un contribuyente, y sea Q el evento de que la persona tenga más de 65 años de edad. Entonces el evento $P \cap Q$ es el conjunto de todos los contribuyentes en el restaurante que tienen más de 65 años de edad. ■

Ejemplo 2.8 Sean $M = \{a, e, i, o, u\}$ y $N = \{r, s, t\}$; entonces se sigue que $M \cap N = \emptyset$. Es decir, M y N no tienen elementos en común y, por tanto, no pueden ocurrir ambos de forma simultánea. ■

Para ciertos experimentos estadísticos no es nada extraño definir dos eventos, A y B , que no pueden ocurrir de forma simultánea. Se dice entonces que los eventos A y B son **mutuamente excluyentes**. Expresado de manera más formal, tenemos la definición siguiente:

Definición 2.5

Dos eventos A y B son **mutuamente excluyentes** o **disjuntos** si $A \cap B = \emptyset$; es decir, si A y B no tienen elementos en común.

Ejemplo 2.9 Una compañía de televisión por cable ofrece programas en ocho diferentes canales, tres de los cuales están afiliados con ABC, dos con NBC, y uno con CBS. Los otros dos son un canal educativo y el canal de deportes ESPN. Suponga que una persona que se suscribe a este servicio enciende un televisor sin seleccionar de antemano el canal. Sea A el evento de que el programa pertenezca a la red NBC y B el evento de que pertenezca a la red CBS. Como un programa de televisión no puede pertenecer a más de una red, los eventos A y B no tienen programas en común. Por tanto, la intersección $A \cap B$ no contiene programa alguno y en consecuencia los eventos A y B son mutuamente excluyentes. ■

A menudo nos interesamos en la ocurrencia de al menos uno de dos eventos asociados con un experimento. Así, en el experimento de lanzamiento de un dado, si

$$A = \{2, 4, 6\} \quad \text{y} \quad B = \{4, 5, 6\},$$

podemos interesarnos en que ocurra A o B , o que ocurran A y B . Tal evento, que se llama la **unión** de A y B , ocurrirá si el resultado es un elemento del subconjunto $\{2, 4, 5, 6\}$.

Definición 2.6

La **unión** de dos eventos A y B , que se denota mediante el símbolo $A \cup B$, es el evento que contiene todos los elementos que pertenecen a A o a B o a ambos.

Ejemplo 2.10 Sea $A = \{a, b, c\}$ y $B = \{b, c, d, e\}$; entonces

$$A \cup B = \{a, b, c, d, e\}. \quad \blacksquare$$

Ejemplo 2.11 Sea P el evento de que un empleado seleccionado al azar de una compañía petrolera fume cigarrillos. Sea Q el evento de que el empleado seleccionado ingiera bebidas alcohólicas. Entonces el evento $P \cup Q$ es el conjunto de todos los empleados que beben o fuman, o que hacen ambas cosas. \blacksquare

Ejemplo 2.12 Si $M = \{x | 3 < x < 9\}$ y $N = \{y | 5 < y < 12\}$, entonces

$$M \cup N = \{z | 3 < z < 12\}.$$

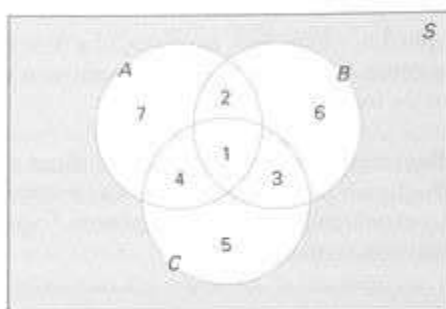


Figura 2.3 Eventos representados por varias regiones.

La relación entre eventos y el correspondiente espacio muestral se puede ilustrar de forma gráfica mediante **diagramas de Venn**. En un diagrama de Venn representamos el espacio muestral como un rectángulo y los eventos con círculos trazados dentro del rectángulo. De esta forma, en la figura 2.3, vemos que

$$A \cap B = \text{regiones 1 y 2,}$$

$$B \cap C = \text{regiones 1 y 3,}$$

$$A \cup C = \text{regiones 1, 2, 3, 4, 5, y 7}$$

$$B' \cap A = \text{regiones 4 y 7,}$$

$$A \cap B \cap C = \text{región 1,}$$

$$(A \cup B) \cap C' = \text{regiones 2, 6, y 7,}$$

y así sucesivamente. En la figura 2.4 vemos que los eventos A , B y C son subconjuntos del espacio muestral S . También es claro que el evento B es un subconjunto del evento A ; el evento $B \cap C$ no tiene elementos y por ello B y C son mutuamente excluyentes; el evento $A \cap C$ tiene al menos un elemento; y el evento $A \cup B = A$. La figura 2.4 puede, por tanto, representar una situación donde seleccionamos una carta al azar de una baraja ordinaria de 52 cartas y observamos si ocurren los siguientes eventos:

A : la carta es roja

B : la carta es la sota, la reina o el rey de diamantes,

C : la carta es un as.

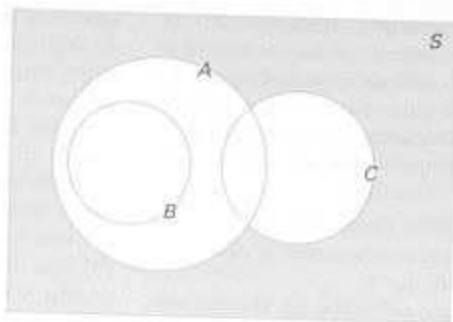


Figura 2.4 Eventos del espacio muestral S .

Claramente, el evento $A \cap C$ consiste sólo en los 2 ases rojos.

Varios resultados que se derivan de las definiciones precedentes, y que se pueden verificar de forma fácil mediante diagramas de Venn, son los que siguen:

1. $A \cap \emptyset = \emptyset$.
2. $A \cup \emptyset = A$.
3. $A \cap A' = \emptyset$.
4. $A \cup A' = S$.
5. $S' = \emptyset$.
6. $\emptyset' = S$.
7. $(A')' = A$.
8. $(A \cap B)' = A' \cup B'$.
9. $(A \cup B)' = A' \cap B'$. ■

Ejercicios

1. Liste los elementos de cada uno de los espacios muestrales siguientes:

(a) el conjunto de enteros entre 1 y 50 divisibles entre 8;

(b) el conjunto $S = \{x \mid x^2 + 4x - 5 = 0\}$;

(c) el conjunto de resultados cuando se lanza una moneda hasta que aparecen una cruz o tres caras;

(d) el conjunto $S = \{x \mid x \text{ es un continente}\}$;

(e) el conjunto $S = \{x \mid 2x - 4 \geq 0 \text{ y } x < 1\}$.

2. Utilice el método de la regla para describir el espacio muestral S , que consiste en todos los puntos del primer cuadrante dentro de un círculo de radio 3 con centro en el origen.

3. ¿Cuáles de los siguientes eventos son iguales?

(a) $A = \{1, 3\}$

(b) $B = \{x \mid x \text{ es un número en un dado}\}$;

(c) $C = \{x \mid x^2 - 4x + 3 = 0\}$;

(d) $D = \{x \mid x \text{ es el número de caras cuando se lanzan seis monedas}\}$.

4. Un experimento implica lanzar un par de dados, 1 verde y 1 rojo, y registrar los números que salen. Si

x es igual al resultado en el dado verde e y es el resultado en el dado rojo, describa el espacio muestral S

(a) mediante la lista de los elementos (x, y) ;

(b) mediante el uso del método de la regla.

5. Un experimento consiste en lanzar un dado y después lanzar una moneda una vez si el número en el dado es par. Si el número en el dado es impar, la moneda se lanza dos veces. Use la notación $4H$, por ejemplo, para denotar el evento de que el dado muestre 4 y después la moneda salga cara, y $3HT$ para denotar el evento de que el dado muestre 3 seguido por una cara y después una cruz en la moneda, construya un diagrama de árbol para mostrar los 18 elementos del espacio muestral S .

6. Se seleccionan dos jurados de 4 suplentes para servir en un juicio por asesinato. Con el uso de la notación A_1A_3 , por ejemplo, para denotar el evento simple de que se seleccionen los suplentes 1 y 3, liste los 6 elementos del espacio muestral S .

7. Se seleccionan al azar cuatro estudiantes de una clase de química y se clasifican como masculino o femenino. Liste los elementos del espacio muestral S_1 con el uso de la letra M para "masculino" y F para "femenino". Defina un segundo espacio muestral S_2 donde los elementos representen el número de mujeres seleccionadas.

8. Para el espacio muestral del ejercicio 4;

- liste los elementos que corresponden al evento A de que la suma sea mayor que 8;
- liste los elementos que corresponden al evento B de que ocurra un 2 en cualquiera de los dos dados;
- liste los elementos que corresponden al evento C de que salga un número mayor que 4 en el dado verde;
- liste los elementos que corresponden al evento $A \cap C$;
- liste los elementos que corresponden al evento $A \cap B$;
- liste los elementos que corresponden al evento $B \cap C$;
- construya un diagrama de Venn para ilustrar las intersecciones y uniones de los eventos A , B y C .

9. Para el espacio muestral del ejercicio 5;

- liste los elementos que corresponden al evento A de que salga en el dado un número menor que 3;
- liste los elementos que corresponden al evento B de que ocurran 2 cruces;
- liste los elementos que corresponden al evento A' ;
- liste los elementos que corresponden al evento $A' \cap B$;
- liste los elementos que corresponden al evento $A \cup B$.

10. Se contrata a una firma de ingenieros para que determine si ciertas vías fluviales en Virginia son seguras para la pesca. Se toman muestras de tres ríos. (a) Liste los elementos de un espacio muestral S , utilice las letras F para "seguro para la pesca" y N para "inseguro para la pesca".

- Liste los elementos de S que corresponden al evento E de que al menos dos de los ríos sean seguros para la pesca.
- Defina un evento que tenga como sus elementos los puntos

$\{FFF, NFF, FFN, FNN\}$.

11. Los currícula de 2 aspirantes masculinos para un puesto de enseñanza de química en una facultad se colocan en la misma fila que los currícula de dos aspirantes mujeres. Están disponibles dos puestos y el primero, con el rango de profesor asistente, se cubre

mediante la selección de 1 de los 4 aspirantes al azar. El segundo puesto, con el rango de instructor, se cubre después mediante la selección aleatoria de uno de los 3 aspirantes restantes. Con el uso de la notación M, F , por ejemplo, para denotar el evento simple de que el primer puesto se cubra con el segundo aspirante hombre y el segundo puesto se cubra después con la primera aspirante mujer;

- liste los elementos de un espacio muestral S ;
- liste los elementos de S que corresponden al evento A de que el puesto de profesor asistente se cubra con un aspirante hombre;
- liste los elementos de S que corresponden al evento B de que exactamente 1 de los 2 puestos se cubra con un aspirante hombre;
- liste los elementos de S que corresponden al evento C de que ningún puesto se cubra con un aspirante hombre;
- liste los elementos de S que corresponden al evento $A \cap B$;
- liste los elementos de S que corresponden al evento $A \cup C$;
- construya un diagrama de Venn para ilustrar las intersecciones y uniones de los eventos A , B y C .

12. Se estudian ejercicio y dieta como posibles sustitutos de la medicación para la presión sanguínea baja. Se utilizarán tres grupos de individuos para estudiar el efecto del ejercicio. El grupo uno es sedentario mientras que el grupo dos camina y el grupo tres nada una hora al día. La mitad de cada uno de los tres grupos de ejercicio tendrá una dieta sin sal. Un grupo adicional de individuos no hará ejercicio ni restringirá su sal, pero tomará la medicación estándar. Use Z para sedentario, W para caminante, S para nadador, Y para sal, N para sin sal, M para medicación, y F para sin medicamentos.

- Muestre todos los elementos del espacio muestral S .
- Dado que A es el conjunto de individuos sin medicamento y B es el conjunto de caminantes, liste los elementos de $A \cup B$.
- Liste los elementos de $A \cap B$.

13. Construya un diagrama de Venn para ilustrar las posibles intersecciones y uniones para los siguientes eventos relativos al espacio muestral que consiste en todos los automóviles construidos en Estados Unidos.

F : cuatro puertas, S : techo corredizo,
 P : dirección asistida

14. Si $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ y $A = \{0, 2, 4, 6, 8\}$, $B = \{1, 3, 5, 7, 9\}$, $C = \{2, 3, 4, 5\}$, y $D = \{1, 6, 7\}$;

liste los elementos de los conjuntos que corresponden a los siguientes eventos:

- (a) $A \cup C$; (b) $A \cap B$;
 (c) C' ; (d) $(C' \cap D) \cup B$;
 (e) $(S \cap C)$; (f) $A \cap C \cap D'$.

15. Considere el espacio muestral

$S = \{\text{cobre, sodio, nitrógeno, potasio, uranio, oxígeno, cinc}\}$

y los eventos

$A = \{\text{cobre, sodio, cinc}\}$,
 $B = \{\text{sodio, nitrógeno, potasio}\}$,
 $C = \{\text{oxígeno}\}$.

Liste los elementos de los conjuntos que corresponden a los eventos siguientes:

- (a) A' ; (b) $A \cup C$;
 (c) $(A \cap B') \cup C'$; (d) $B' \cap C'$;
 (e) $A \cap B \cap C$; (f) $(A' \cup B') \cap (A' \cap C)$.

16. Si $S = \{x \mid 0 < x < 12\}$, $M = \{x \mid 1 < x < 9\}$, y $N = \{x \mid 0 < x < 5\}$, encuentre

- (a) $M \cup N$;
 (b) $M \cap N$;
 (c) $M' \cap N'$.

17. Sean A , B y C eventos relativos al espacio muestral S . Con el uso de diagramas de Venn, sombree las áreas que representan los eventos siguientes:

- (a) $(A \cap B)'$;
 (b) $(A \cup B)'$;
 (c) $(A \cap C) \cup B$.

18. ¿Cuál de los siguientes pares de eventos son mutuamente excluyentes?

- (a) Un golfista que se clasifica en último lugar en la vuelta del hoyo 18 en un torneo de 72 hoyos y pierde el torneo.
 (b) Un jugador de póquer que tiene *flush* (todas las cartas del mismo palo) y 3 de un tipo en la misma mano de 5 cartas.
 (c) Una madre que da a luz a una niña y un par de gemelas el mismo día.
 (d) Un jugador de ajedrez que pierde el último juego y gana el torneo.

19. Suponga que una familia sale de vacaciones de verano en su casa rodante y que M es el evento de que sufrirán problemas mecánicos, T es el evento de que recibirán una infracción por cometer una falta de tránsito, y V es el evento de que llegarán a un lugar para acampar que está lleno. Refiérase al diagrama de Venn de la figura 2.5, ponga en palabras los eventos representados por las regiones siguientes:

- (a) región 5;
 (b) región 3;
 (c) regiones 1 y 2 juntas;
 (d) regiones 4 y 7 juntas;
 (e) regiones 3, 6, 7 y 8 juntas.

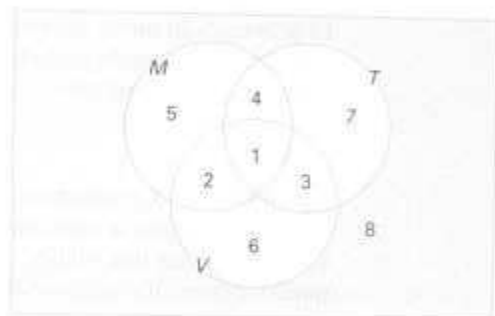


Figura 2.5 Diagrama de Venn para el ejercicio 19.

20. Refiérase al ejercicio 19 y al diagrama de Venn de la figura 2.5, liste los números de las regiones que representan los siguientes eventos:

- (a) La familia no experimentará problemas mecánicos y no cometerá infracciones de tránsito, pero encontrará que el lugar para acampar está lleno.
 (b) La familia experimentará problemas mecánicos y problemas para localizar un lugar para acampar vacío, pero no recibirá una multa por infracción de tránsito.
 (c) La familia tendrá problemas mecánicos o encontrará un lugar para acampar lleno, pero no recibirá una multa por cometer una infracción de tránsito.
 (d) La familia no llegará a un lugar para acampar lleno.

2.3 Conteo de puntos de la muestra

Uno de los problemas que el estadístico debe considerar e intentar evaluar es el elemento de posibilidad asociado con la ocurrencia de ciertos eventos cuando se lleva a cabo un experimento. Estos problemas pertenecen al campo de la probabilidad, un tema que se introducirá en la sección 2.4. En muchos casos debemos

ser capaces de resolver un problema de probabilidad mediante el conteo del número de puntos en el espacio muestral sin listar realmente cada elemento. El principio fundamental del conteo a menudo denominado **regla de multiplicación**, se establece como sigue:

Teorema 2.1

Si una operación se puede llevar a cabo en n_1 formas, y si para cada una de éstas se puede realizar una segunda operación en n_2 formas, entonces las dos operaciones se pueden ejecutar juntas de $n_1 n_2$ formas.

Ejemplo 2.13 ¿Cuántos puntos muestrales hay en el espacio muestral cuando se lanza una vez un par de dados?

SOLUCIÓN

El primer dado puede caer en cualquiera de $n_1 = 6$ maneras. Para cada una de esas 6 maneras el segundo dado también puede caer en $n_2 = 6$ formas. Por tanto, el par de dados puede caer en

$$n_1 n_2 = (6)(6) = 36 \text{ formas posibles. } \blacksquare$$

Ejemplo 2.14 Un urbanista de una nueva subdivisión ofrece a los futuros compradores de una casa la elección del estilo de la fachada entre Tudor, rústica, colonial y tradicional en una planta, dos pisos y desniveles. ¿En cuántas formas diferentes puede ordenar un comprador una de estas casas?

SOLUCIÓN

Como $n_1 = 4$ y $n_2 = 3$ un comprador debe elegir entre

$$n_1 n_2 = (4)(3) = 12 \text{ casas posibles. } \blacksquare$$

Las respuestas a los dos ejemplos anteriores se pueden verificar mediante la construcción de diagramas de árbol y el conteo de las diversas trayectorias a lo largo de las ramas. Así, en el ejemplo 2.14 habrá $n_1 = 4$ ramas que corresponden a los diferentes estilos de las fachadas, y después habrá $n_2 = 3$ ramas que se extienden de cada una de estas 4 ramas para representar los diferentes planes de plantas. Este diagrama de árbol da las $n_1 n_2 = 12$ elecciones de casas dadas por las trayectorias a lo largo de las ramas como se ilustra en la figura 2.6.

La regla de multiplicación del teorema 2.1 se puede extender para cubrir cualquier número de operaciones. Suponga, por ejemplo, que un cliente desea instalar un teléfono Trimline y puede elegir entre $n_1 = 10$ colores decorativos, que supondremos disponibles en cualquiera de $n_2 = 3$ longitudes opcionales de cordón con $n_3 = 2$ tipos de marcado, a saber, de disco o por tonos. Estas tres clasificaciones tienen como resultado

$$n_1 n_2 n_3 = (10)(3)(2) = 60$$

diferentes formas para que un cliente ordene uno de estos teléfonos. La **regla de multiplicación generalizada** que cubre k operaciones se formula en el siguiente teorema.

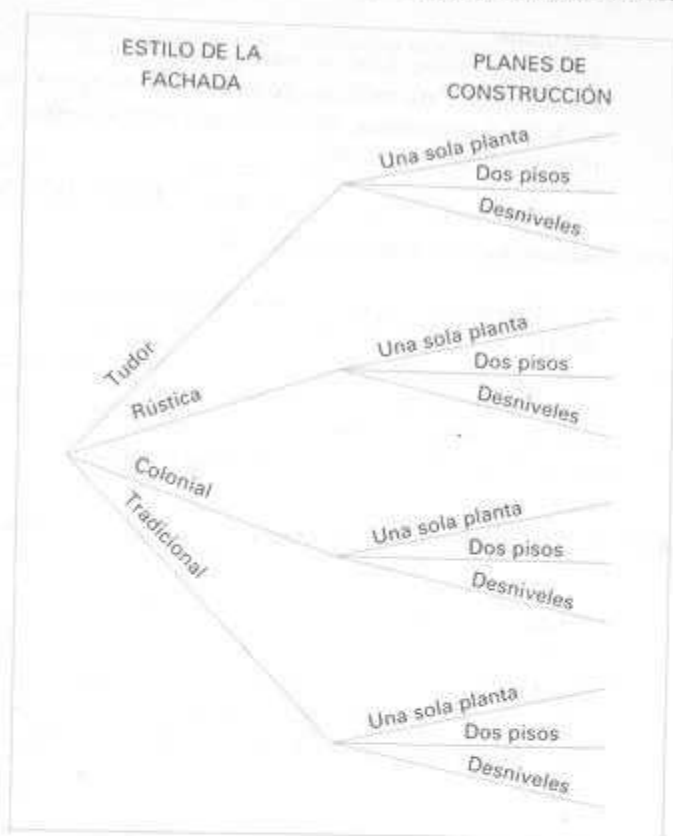


Figura 2.6 Diagrama de árbol para el ejemplo 2.14.

Teorema 2.2

Si una operación se puede ejecutar en n_1 formas, y si para cada una de éstas se puede llevar a cabo una segunda operación en n_2 formas, y para cada una de las primeras dos se puede realizar una tercera operación en n_3 formas, y así sucesivamente, entonces la serie de k operaciones se puede realizar en $n_1 n_2 \dots n_k$ formas.

Ejemplo 2.15 ¿Cuántos almuerzos que consisten en una sopa, emparedado, postre y una bebida son posibles si podemos seleccionar de 4 sopas, 3 tipos de emparedados, 5 postres y 4 bebidas?

SOLUCIÓN

Como $n_1 = 4$, $n_2 = 3$, $n_3 = 5$ y $n_4 = 4$, hay

$$n_1 \times n_2 \times n_3 \times n_4 = 4 \times 3 \times 5 \times 4 = 240$$

diferentes maneras de elegir un almuerzo. ■

Ejemplo 2.16 ¿Cuántos números pares de tres dígitos se pueden formar a partir de los dígitos 1, 2, 5, 6 y 9 si cada dígito se puede usar sólo una vez?

SOLUCIÓN

Como el número debe ser par, tenemos sólo $n_1 = 2$ elecciones para la posición de las unidades. Para cada una de éstas tenemos $n_2 = 4$ elecciones para la posición de las centenas y entonces $n_3 = 3$ elecciones para la posición de las decenas. Por tanto, podemos formar un total de

$$n_1 n_2 n_3 = (2)(4)(3) = 24$$

números pares de tres dígitos. ■

Con frecuencia nos interesamos en un espacio muestral que contiene como elementos a todas las posibles ordenaciones o arreglos de un grupo de objetos. Por ejemplo, podemos querer saber cuántos arreglos diferentes son posibles para sentar a seis personas alrededor de una mesa, o podemos preguntar cuántas ordenaciones diferentes son posibles para sacar dos billetes de lotería de un total de 20. Los diferentes arreglos se llaman **permutaciones**.

Definición 2.7

Una **permutación** es un arreglo de todo o parte de un conjunto de objetos.

Considere las tres letras a, b y c . Las permutaciones posibles son abc, acb, bac, bca, cab y cba . De esta forma vemos que hay 6 arreglos distintos. Con el uso del teorema 2.2 podemos llegar a la respuesta 6 sin realmente listar las diferentes ordenaciones. Hay $n_1 = 3$ elecciones para la primera posición, después $n_2 = 2$ para la segunda, lo que deja sólo $n_3 = 1$ elección para la última posición, lo que da un total de

$$n_1 n_2 n_3 = (3)(2)(1) = 6 \text{ permutaciones.}$$

En general, n objetos distintos se pueden arreglar en

$$n(n-1)(n-2)\cdots(3)(2)(1) \text{ formas.}$$

Representamos este producto mediante el símbolo $n!$, que se lee “ n factorial”. Tres objetos se pueden arreglar en $3! = (3)(2)(1) = 6$ maneras. Por definición, $1! = 1$ y $0! = 1$.

Teorema 2.3

El número de permutaciones de n objetos distintos es $n!$.

El número de permutaciones de las cuatro letras a, b, c y d será $4! = 24$. Consideremos ahora el número de permutaciones que son posibles al tomar las cuatro letras, dos a la vez. Éstas serían $ab, ac, ad, ba, ca, da, bc, cb, bd, db, cd, dc$. De nuevo con el uso del teorema 2.1, tenemos dos posiciones para llenar con $n_1 = 4$ elecciones para la primera y después $n_2 = 3$ elecciones para la segunda para un total de

$$n_1 n_2 = (4)(3) = 12$$

permutaciones. En general, n objetos distintos tomados de r a la vez se pueden arreglar en

$$n(n-1)(n-2)\cdots(n-r+1)$$

formas. Representamos este producto mediante el símbolo

$${}_n P_r = \frac{n!}{(n-r)!}$$

Como resultado tenemos el teorema que sigue.

Teorema 2.4

El número de permutaciones de n objetos distintos tomados de r a la vez es

$${}_n P_r = \frac{n!}{(n-r)!}$$

Ejemplo 2.17 Se sacan dos billetes de lotería de 20 para un primer y un segundo premios. Encuentre el número de puntos muestrales en el espacio S .

SOLUCIÓN

El número total de puntos de la muestra es

$${}_{20}P_2 = \frac{20!}{18!} = (20)(19) = 380. \quad \blacksquare$$

Ejemplo 2.18 ¿De cuántas formas puede una organización local de la Sociedad Americana de Química programar a tres conferencistas para tres reuniones diferentes si todos están disponibles en cualquiera de cinco fechas posibles?

SOLUCIÓN

El número total de programas posibles es

$${}_5P_3 = \frac{5!}{2!} = (5)(4)(3) = 60. \quad \blacksquare$$

Las permutaciones que ocurren al arreglar objetos en un círculo se llaman **permutaciones circulares**. Dos permutaciones circulares no se consideran diferentes a menos que los objetos correspondientes en los dos arreglos estén precedidos o seguidos por un objeto diferente conforme recorramos en la dirección de las manecillas del reloj. Por ejemplo, si cuatro personas juegan bridge, no tenemos una permutación nueva si se mueven una posición en la dirección de las manecillas del reloj. Al considerar a una persona en una posición fija y arreglar a las otras tres de $3!$ formas, encontramos que hay seis arreglos distintos para el juego de bridge.

Teorema 2.5

El número de permutaciones de n objetos distintos arreglados en un círculo es $(n-1)!$

Hasta aquí consideramos permutaciones de objetos distintos. Es decir, todos los objetos fueron por completo diferentes o distinguibles. Obviamente, si las letras b y c son ambas iguales a x , entonces las 6 permutaciones de las letras a, b, c se convierten

en axx , axx , xax , xax , xxa y xxa , de las que sólo 3 son diferentes. Por tanto, con 3 letras, en las que 2 son la misma, tenemos $3!/2! = 3$ permutaciones distintas. Con 4 letras diferentes a , b , c y d tenemos 24 permutaciones distintas. Si hacemos $a = b = x$ y $c = d = y$, podemos listar sólo las siguientes: $xyxy$, $xyxy$, $yxyx$, $yxyx$ y $xyyx$. De esta forma tenemos $4!/(2!2!) = 6$ permutaciones distintas.

Teorema 2.6

El número de permutaciones distintas de n cosas de las que n_1 son de una clase, n_2 de una segunda clase, ..., n_k de una k -ésima clase es

$$\frac{n!}{n_1! n_2! \cdots n_k!}$$

Ejemplo 2.19 ¿De cuántas formas diferentes se pueden arreglar 3 focos rojos, 4 amarillos y 2 azules en una serie de luces navideña con 9 portalámparas?

SOLUCIÓN

El número total de arreglos distintos es

$$\frac{9!}{3! 4! 2!} = 1260. \quad \blacksquare$$

Con frecuencia nos interesa el número de formas de dividir un conjunto de n objetos en r subconjuntos denominados **celdas**. Se consigue una partición si la intersección de todo par posible de los r subconjuntos es el conjunto vacío \emptyset y si la unión de todos los subconjuntos da el conjunto original. El orden de los elementos dentro de una celda no tiene importancia. Considere el conjunto $\{a, e, i, o, u\}$. Las particiones posibles en dos celdas en las que la primera celda contiene cuatro elementos y la segunda un elemento son

$$\{(a, e, i, o), (u)\}, \{(a, i, o, u), (e)\}, \{(e, i, o, u), (a)\}, \{(a, e, o, u), (i)\}, \{(a, e, i, u), (o)\}.$$

Vemos que hay cinco formas de partir un conjunto de cuatro elementos en dos subconjuntos o celdas que contienen cuatro elementos en la primera celda y uno en la segunda.

El número de particiones para esta ilustración se denota con el símbolo

$$\binom{5}{4, 1} = \frac{5!}{4! 1!} = 5,$$

donde el número superior representa el número total de elementos y los números inferiores representan el número de elementos que van en cada celda. Establezcamos esto de forma más general en el siguiente teorema.

Teorema 2.7

El número de formas de partir un conjunto de n objetos en r celdas con n_1 elementos en la primera celda, n_2 elementos en la segunda, y así sucesivamente, es

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

donde $n_1 + n_2 + \cdots + n_r = n$.

Ejemplo 2.20 ¿En cuántas formas se pueden asignar siete científicos a una habitación de hotel triple y a dos dobles?

SOLUCIÓN

El número total de particiones posibles sería

$$\binom{7}{3, 2, 2} = \frac{7!}{3! 2! 2!} = 210. \quad \blacksquare$$

En muchos problemas nos interesamos en el número de formas de seleccionar r objetos de n sin importar el orden. Estas selecciones se llaman **combinaciones**. Una combinación es realmente una partición con dos celdas, una celda contiene los r objetos seleccionados y la otra contiene los $(n - r)$ objetos restantes.

El número de tales combinaciones, denotado por

$$\binom{n}{r, n-r}, \quad \text{por lo general se reduce a } \binom{n}{r},$$

debido a que el número de elementos en la segunda celda debe ser $n - r$.

Teorema 2.8

El número de combinaciones de n objetos distintos tomados de r a la vez es

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Ejemplo 2.21 De cuatro químicos y tres físicos encuentre el número de comités que se pueden formar que consistan en dos químicos y un físico.

SOLUCIÓN

El número de formas de seleccionar a dos químicos de cuatro es $\binom{4}{2} = \frac{4!}{2! 2!} = 6$.

El número de formas de seleccionar un físico de tres es $\binom{3}{1} = \frac{3!}{1! 2!} = 3$.

Al usar la regla de la multiplicación del teorema 1.1 con $n_1 = 6$ y $n_2 = 3$, podemos formar

$$n_1 n_2 = (6)(3) = 18$$

comités con 2 químicos y 1 físico. \blacksquare

Ejercicios

1. A los participantes de una convención se les ofrecen seis recorridos a sitios de interés cada uno de los tres días. ¿De cuántas maneras se puede acomodar una persona para ir a uno de los recorridos planeados por la convención?

2. En un estudio médico los pacientes se clasifican en ocho formas de acuerdo con su tipo sanguíneo AB^+ , AB^- , A^+ , A^- , B^+ , B^- , O^+ u O^- , y también de acuerdo con su presión sanguínea baja, normal o alta.

Encuentre el número de formas en las que se puede clasificar a un paciente.

3. Si un experimento consiste en lanzar un dado y después extraer una letra al azar del alfabeto inglés, ¿cuántos puntos hay en el espacio muestral?

4. Los estudiantes de un colegio de humanidades privado se clasifican como estudiantes de primer año, de segundo, de penúltimo o de último año, y también

de acuerdo con si son hombres o mujeres. Encuentre el número total de clasificaciones posibles para los estudiantes de dicho colegio.

5. Cierta calzado se recibe en cinco diferentes estilos con cada estilo disponible en cuatro colores distintos. Si la tienda desea mostrar pares de estos zapatos que muestren la totalidad de los diversos estilos y colores, ¿cuántos diferentes pares tendría que mostrar?

6. Un estudio en California concluyó que al seguir siete sencillas reglas de salud, la vida de un hombre se puede prolongar 11 años en promedio y la vida de una mujer en siete años. Estas siete reglas son: no fumar, hacer ejercicio, uso moderado del alcohol, dormir siete u ocho horas, mantener el peso apropiado, desayunar y no ingerir alimentos entre comidas. De cuántas formas puede una persona adoptar cinco de estas reglas a seguir:

(a) ¿Si la persona actualmente viola las siete reglas?

(b) ¿Si la persona nunca bebe y siempre desayuna?

7. Un urbanista de un nuevo fraccionamiento ofrece a un futuro comprador de una casa la elección de cuatro diseños, tres diferentes sistemas de calefacción, un garaje o cobertizo, y un patio o un porche cubierto. ¿De cuántos planes diferentes dispone el comprador?

8. Un medicamento para el asma se puede adquirir de cinco diferentes laboratorios en forma de líquido, tabletas o cápsulas, todas en concentración normal o alta. ¿De cuántas formas diferentes puede recetar la medicina un doctor a un paciente que sufre de asma?

9. En un estudio económico de combustibles, cada uno de tres autos de carreras se prueba con cinco marcas diferentes de gasolina en siete lugares que se localizan en diferentes regiones del país. Si se utilizan dos pilotos en el estudio y las pruebas se realizan una vez bajo cada uno de los distintos grupos de condiciones, ¿cuántas pruebas se necesitan?

10. ¿De cuántas formas distintas se puede responder una prueba de falso-verdadero que consta de nueve preguntas?

11. Si una prueba de opción múltiple consiste en cinco preguntas cada una con cuatro respuestas posibles de las que sólo una es correcta,

(a) ¿De cuántas formas diferentes puede elegir un estudiante una respuesta a cada pregunta?

(b) ¿De cuántas maneras puede escoger un estudiante una respuesta a cada pregunta y tener mal todas las respuestas?

12. (a) ¿Cuántas permutaciones distintas se pueden hacer con las letras de la palabra *columns*?

(b) ¿Cuántas de estas permutaciones comienzan con la letra *m*?

13. Un testigo de un accidente de tránsito en el que huye el culpable dice a la policía que el número de placas contenía las letras RLH seguidas de tres dígitos, cuyo primer número es un 5. Si el testigo no puede recordar los últimos dos dígitos, pero tiene la certeza de que los tres eran diferentes, encuentre el número máximo de placas de automóvil que la policía tiene que verificar.

14. (a) ¿De cuántas maneras se pueden formar seis personas para abordar un autobús?

(b) Si tres personas específicas insisten en estar una después de la otra, ¿cuántas maneras son posibles?

(c) Si dos personas específicas rehúsan seguir una a la otra, ¿cuántas maneras son posibles?

15. Un contratista desea construir nueve casas, cada una con diferente diseño. ¿De cuántas formas puede colocar estas casas en una calle si hay seis lotes en un lado de la calle y tres en el lado opuesto?

16. (a) ¿Cuántos números de tres dígitos se pueden formar con los dígitos 0, 1, 2, 3, 4, 5 y 6, si cada dígito se puede usar sólo una vez?

(b) ¿Cuántos de estos números son impares?

(c) ¿Cuántos son mayores que 330?

17. ¿De cuántas maneras se pueden sentar cuatro niños y cinco niñas en una fila si se deben alternar?

18. Cuatro matrimonios compran ocho lugares en la misma fila para un concierto. ¿De cuántas maneras diferentes se pueden sentar:

(a) sin restricciones?

(b) si cada pareja se sienta junta?

(c) si todos los hombres se sientan juntos a la derecha de todas las mujeres?

19. En un concurso regional de ortografía, los ocho finalistas son tres niños y cinco niñas. Encuentre el número de puntos muestrales en el espacio S para el número de ordenamientos posibles al final del concurso para

(a) los ocho finalistas;

(b) las primeras tres posiciones.

20. ¿De cuántas formas se pueden llenar las cinco posiciones iniciales en un equipo de baloncesto con ocho jugadores que pueden jugar cualquiera de las posiciones?

21. Encuentre el número de formas en las que se pueden asignar seis maestros a cuatro secciones de un curso introductorio de psicología si ningún maestro se asigna a más de una sección.

22. Se sacan tres billetes de lotería para el primer, segundo y tercer premios de un grupo de 40 boletos. En

cuentre el número de puntos muestrales en S para dar los tres premios si cada concursante sólo tiene un billete.

23. ¿De cuántas maneras se pueden plantar cinco árboles diferentes en un círculo?

24. ¿De cuántas formas se puede acomodar en un círculo una caravana de ocho carretas que proviene de Arizona?

25. ¿Cuántas permutaciones distintas se pueden hacer con las letras de la palabra *infinito*?

26. ¿De cuántas maneras se pueden colocar tres robles, cuatro pinos y dos arces a lo largo de la línea divisoria de una propiedad si no se distingue entre los árboles del mismo tipo?

27. Un colegio participa en 12 juegos de fútbol durante una temporada. ¿De cuántas formas puede el equipo terminar la temporada con siete ganados, tres perdidos y dos empates?

28. Nueve personas se dirigen a esquiar en tres autos que llevan dos, cuatro y cinco pasajeros, respectivamente. ¿De cuántas maneras es posible transportar a las nueve personas hasta el albergue en todos los autos?

29. ¿Cuántas formas hay para seleccionar a tres candidatos de ocho recién graduados igualmente calificados para las vacantes de una empresa contable?

2.4 Probabilidad de un evento

Quizá fue la sed insaciable del hombre por el juego la que condujo al desarrollo temprano de la teoría de la probabilidad. En un esfuerzo por aumentar sus ganancias pidieron a los matemáticos que les proporcionaran las estrategias óptimas para varios juegos de azar. Algunos de los matemáticos que proporcionaron estas estrategias fueron Pascal, Leibniz, Fermat y James Bernoulli. Como resultado de este primer desarrollo de la teoría de la probabilidad, la inferencia estadística, con todas sus predicciones y generalizaciones, se extiende más allá de los juegos de azar para abarcar muchos otros campos asociados con los eventos aleatorios, como la política, los negocios, la predicción del clima y la investigación científica. Para que estas predicciones y generalizaciones sean razonablemente precisas, es esencial una comprensión de la teoría básica de la probabilidad.

¿Qué queremos decir cuando hacemos afirmaciones como “Juan probablemente ganará el torneo de tenis”, o “tengo una oportunidad de cincuenta por ciento de obtener un número par cuando se lanza un dado”, o “no tengo posibilidad de ganar en la lotería esta noche”, o “la mayoría de nuestros graduados probablemente estarán casados dentro de tres años”? En cada caso expresamos un resultado del cual no estamos seguros, pero debido a la información del pasado o a partir de una comprensión de la estructura del experimento, tenemos algún grado de confianza en la validez de la afirmación.

En el resto de este capítulo consideramos sólo aquellos experimentos para los que el espacio muestral contiene un número finito de elementos. La probabilidad de la ocurrencia de un evento que resulta de tal experimento estadístico se evalúa por medio de un conjunto de números reales denominados **pesos** o **probabilidades** que van de 0 a 1. Para todo punto en el espacio muestral asignamos una probabilidad tal que la suma de todas las probabilidades es 1. Si tenemos razón para creer que es bastante probable que ocurra cierto punto muestral cuando se lleva a cabo el experimento, la probabilidad que se le asigne debe ser cercana a 1. Por otro lado, una probabilidad cercana a cero se asigna a un punto muestral que no es probable que ocurra. En muchos experimentos, como lanzar una moneda o un dado, todos los puntos muestrales tienen la misma oportunidad de ocurrencia y se les asignan probabilidades iguales. Para puntos fuera del espacio muestral, es decir, para eventos simples que no es posible que ocurran, asignamos una probabilidad de cero.

Para encontrar la probabilidad de un evento A , sumamos todas las probabilidades que se asignan a los puntos muestrales en A . Esta suma se denomina probabilidad de A y se denota con $P(A)$.

Definición 2.8

La probabilidad de un evento A es la suma de los pesos de todos los puntos muestrales en A . Por tanto,

$$0 \leq P(A) \leq 1, \quad P(\emptyset) = 0, \quad \text{y} \quad P(S) = 1.$$

Ejemplo 2.22 Se lanza dos veces una moneda. ¿Cuál es la probabilidad de que ocurra al menos una cara?

SOLUCIÓN

El espacio muestral para este experimento es

$$S = \{HH, HT, TH, TT\}.$$

Si la moneda esta balanceada, cada uno de estos resultados tendrá la misma probabilidad de ocurrencia. Por tanto, asignamos una probabilidad de w a cada uno de los puntos muestrales. Entonces $4w = 1$, o $w = 1/4$. Si A representa el evento de que ocurra al menos una cara, entonces

$$A = \{HH, HT, TH\} \quad \text{y} \quad P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}. \quad \blacksquare$$

Ejemplo 2.23 Se carga un dado de forma que sea dos veces más probable que salga un número par que uno non. Si E es el evento de que ocurra un número menor que 4 en un solo lanzamiento del dado, encuentre $P(E)$.

SOLUCIÓN

El espacio muestral es $S = \{1, 2, 3, 4, 5, 6\}$. Asignamos una probabilidad de w a cada número non y una probabilidad de $2w$ a cada número par. Como la suma de las probabilidades debe ser 1, tenemos $9w = 1$ o $w = 1/9$. Por ello se asignan probabilidades de $1/9$ y $2/9$ a cada número non y par, respectivamente. Por tanto,

$$E = \{1, 2, 3\} \quad \text{y} \quad P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}. \quad \blacksquare$$

Ejemplo 2.24 En el ejemplo 2.23 sea A el evento de que salga un número par y sea B el evento de que salga un número divisible entre 3. Encuentre $P(A \cup B)$ y $P(A \cap B)$.

SOLUCIÓN

Para los eventos $A = \{2, 4, 6\}$ y $B = \{3, 6\}$ tenemos

$$A \cup B = \{2, 3, 4, 6\} \quad \text{y} \quad A \cap B = \{6\}.$$

Al asignar una probabilidad de $1/9$ a cada número non y $2/9$ a cada número par, tenemos

$$P(A \cup B) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9} \quad \text{y} \quad P(A \cap B) = \frac{2}{9}. \quad \blacksquare$$

Si el espacio muestral para un experimento contiene N elementos, los cuales tienen la misma probabilidad de ocurrencia, asignamos una probabilidad igual a $1/N$ a cada uno de los N puntos. La probabilidad de cualquier evento A que contenga n de estos N puntos muestrales es entonces la razón del número de elementos en A al número de elementos en S .

Teorema 2.9

Si un experimento puede tener como resultado cualquiera de N diferentes resultados igualmente probables, y si exactamente n de estos resultados corresponden al evento A , entonces la probabilidad del evento A es

$$P(A) = \frac{n}{N}$$

Ejemplo 2.25 Un surtido de dulces contiene seis mentas, cuatro chiclosos y tres chocolates. Si una persona hace una selección aleatoria de uno de los dulces, encuentre la probabilidad de sacar a) una menta o b) un chicloso o un chocolate.

SOLUCIÓN

M , T y C representan los eventos de que la persona seleccione, respectivamente, una menta, un chicloso o un chocolate. El número total de dulces es 13, los cuales tienen la misma probabilidad de ser seleccionados.

- (a) Como seis de los 13 dulces son mentas, la probabilidad del evento M , seleccionar una menta al azar, es

$$P(M) = \frac{6}{13}$$

- (b) Como siete de los 13 dulces son chiclosos o chocolates, se sigue que

$$P(T \cup C) = \frac{7}{13} \quad \blacksquare$$

Ejemplo 2.26 En una mano de póquer que consiste en cinco cartas, encuentre la probabilidad de tener dos ases y tres sotas.

SOLUCIÓN

El número de formas de tener dos ases de cuatro es

$$\binom{4}{2} = \frac{4!}{2!2!} = 6$$

y el número de formas de tener tres sotas de cuatro es

$$\binom{4}{3} = \frac{4!}{3!1!} = 4.$$

Mediante la regla de multiplicación del teorema 2.1, hay $n = (6)(4) = 24$ manos con dos ases y tres sotas. El número total de cinco manos de póquer de cinco cartas, las cuales son igualmente probables, es

$$N = \binom{52}{5} = \frac{52!}{5! 47!} = 2,598,960.$$

Por tanto, la probabilidad del evento C de obtener dos ases y tres sotas en una mano de póquer de cinco cartas es

$$P(C) = \frac{24}{2,598,960} = 0,9 \times 10^{-5}. \blacksquare$$

Si los resultados de un experimento no tienen igual probabilidad de ocurrencia, las probabilidades se deben asignar sobre la base de un conocimiento previo o de evidencia experimental. Por ejemplo, si una moneda no está balanceada, podemos estimar las probabilidades de caras y cruces al lanzar la moneda un número elevado de veces y registrar los resultados. De acuerdo con la definición de **frecuencia relativa** de la probabilidad, las probabilidades verdaderas serían las fracciones de caras y cruces que ocurren a largo plazo.

Para encontrar un valor numérico que represente de manera adecuada la probabilidad de ganar en el tenis, debemos depender de nuestro rendimiento pasado en el juego así como también del de nuestro oponente y, hasta cierto punto, en nuestra creencia de ser capaces de ganar. De manera similar, para encontrar la probabilidad de que un caballo gane una carrera, debemos llegar a una probabilidad que se base en las marcas anteriores de todos los caballos que participan en la carrera, así como de récords de los jockeys que montan los caballos. La intuición sin duda también juega una parte en la determinación del monto de la apuesta que estemos dispuestos a jugar. El uso de la intuición, las creencias personales y otra información indirecta para llegar a probabilidades se denomina como la definición **subjetiva** de probabilidad.

En la mayor parte de las aplicaciones de probabilidad de este libro la interpretación de frecuencia relativa de probabilidad es la que opera. Su fundamento es el experimento estadístico en lugar de la subjetividad. Se le considera más bien como **frecuencia relativa limitante**. Como resultado, muchas aplicaciones de probabilidad en la ciencia y la ingeniería se deben basar en experimentos que se puedan repetir. Nociones menos objetivas de probabilidad se encuentran cuando asignamos probabilidades que se basan en información y opiniones previas. Como ejemplo, "hay una buena oportunidad de que los Leones pierdan el super tazón". Cuando las opiniones y la información previa difieren de individuo a individuo, la probabilidad subjetiva se vuelve el recurso relevante.

2.5 Reglas aditivas

A menudo es más fácil calcular la probabilidad de algún evento a partir del conocimiento de las probabilidades de otros eventos. Esto puede ser cierto si el evento en cuestión se puede representar como la unión de otros dos eventos o como el complemento de algún evento. A continuación se presentan varias leyes importantes que con frecuencia simplifican el cálculo de probabilidades. La primera, que se denomina **regla aditiva**, se aplica a uniones de eventos,

Teorema 2.10

Si A y B son cualesquiera dos eventos, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

PRUEBA Considere el diagrama de Venn de la figura 2.7. $P(A \cup B)$ es la suma de las probabilidades de los puntos muestrales en $A \cup B$. Ahora bien $P(A) + P(B)$ es la suma de todas las probabilidades en A más la suma de todas las probabilidades en B . Por tanto, sumamos dos veces las probabilidades en $(A \cap B)$. Como estas probabilidades se suman a $P(A \cap B)$, debemos restar esta probabilidad una vez para obtener la suma de las probabilidades en $A \cup B$, que es $P(A \cup B)$.

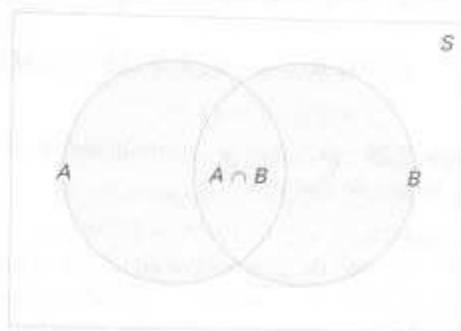


Figura 2.7 Regla aditiva de probabilidad.

Corolario 1

Si A y B son mutuamente excluyentes, entonces

$$P(A \cup B) = P(A) + P(B).$$

El corolario 1 es un resultado inmediato del teorema 2.10, pues si A y B son mutuamente excluyentes, $A \cap B = \emptyset$ y entonces $P(A \cap B) = P(\emptyset) = 0$. En general, escribimos:

Corolario 2

Si $A_1, A_2, A_3, \dots, A_n$ son mutuamente excluyentes, entonces

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Corolario 3

Si A_1, A_2, \dots, A_n es una partición de un espacio muestral S , entonces

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= P(A_1) + P(A_2) + \dots + P(A_n) \\ &= P(S) \\ &= 1. \end{aligned}$$

Como se podría esperar, el teorema 2.10 se extiende de forma análoga.

Teorema 2.11

Para tres eventos A , B y C , $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

Ejemplo 2.27 La probabilidad de que Paula apruebe matemáticas es $2/3$, y la probabilidad de que apruebe inglés es $4/9$. Si la probabilidad de aprobar ambos cursos es $1/4$, ¿cuál es la probabilidad de que Paula apruebe al menos uno de estos cursos?

SOLUCIÓN

Si M es el evento "aprobar matemáticas", y E el evento "aprobar inglés", entonces por la regla de adición tenemos

$$P(M \cup E) = P(M) + P(E) - P(M \cap E) = \frac{2}{3} + \frac{4}{9} - \frac{1}{4} = \frac{31}{36} \quad \blacksquare$$

Ejemplo 2.28 ¿Cuál es la probabilidad de obtener un total de siete u 11 cuando se lanza un par de dados?

SOLUCIÓN

Sea A el evento de que ocurra siete y B el evento de que salga 11. Ahora bien, un total de siete ocurre para seis de los 36 puntos muestrales y un total de 11 ocurre sólo para dos. Como todos los puntos muestrales son igualmente probables, tenemos $P(A) = 1/6$ y $P(B) = 1/18$. Los eventos A y B son mutuamente excluyentes, pues un total de siete y 11 no pueden ocurrir en el mismo lanzamiento. Por tanto,

$$P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{18} = \frac{2}{9}$$

Este resultado también se podría obtener al contar el número total de puntos para el evento $A \cup B$, o sea 8, y escribir

$$P(A \cup B) = \frac{n}{N} = \frac{8}{36} = \frac{2}{9} \quad \blacksquare$$

El teorema 2.10 y sus tres corolarios deben ayudar al lector a ganar una mayor comprensión de la probabilidad y su interpretación. Los corolarios 1 y 2 sugieren el resultado muy intuitivo que trata con la probabilidad de ocurrencia de al menos uno de un número de eventos, sin que puedan ocurrir dos de ellos simultáneamente. La probabilidad de que al menos ocurra uno es la suma de las probabilidades de ocurrencia de los eventos individuales. El tercer corolario simplemente establece que el valor más alto de una probabilidad (uno) se asigna a todo el espacio muestral S .

Ejemplo 2.29 Si las probabilidades de que una persona que compra un automóvil nuevo elija el color verde, blanco, rojo o azul son, respectivamente, 0.09, 0.15, 0.21 y 0.23, ¿cuál es la probabilidad de que un comprador dado adquiera un automóvil nuevo que tenga uno de esos colores?

SOLUCIÓN

Sean G , W , R y B los eventos de que un comprador seleccione, respectivamente un automóvil verde, blanco, rojo o azul. Como estos cuatro eventos son mutuamente excluyentes, la probabilidad es

$$\begin{aligned} P(G \cup W \cup R \cup B) &= P(G) + P(W) + P(R) + P(B) \\ &= 0.09 + 0.15 + 0.21 + 0.23 = 0.68. \end{aligned}$$

A menudo es más difícil calcular la probabilidad de que ocurra un evento que calcular la probabilidad de que el evento no ocurra. Si este es el caso para algún evento A , simplemente encontramos primero $P(A')$ y después, con el uso del teorema 2.11, encontramos $P(A)$ por sustracción. ■

Teorema 2.12

Si A y A' son eventos complementarios, entonces

$$P(A) + P(A') = 1.$$

PRUEBA Como $A \cup A' = S$ y los conjuntos A y A' son disjuntos, entonces

$$1 = P(S) = P(A \cup A') = P(A) + P(A').$$

Ejemplo 2.30 Si las probabilidades de que un mecánico automotriz dé servicio a tres, cuatro, cinco, seis, siete, ocho o más autos en un día de trabajo dado son 0.12, 0.19, 0.28, 0.24, 0.10, y 0.07, respectivamente, ¿cuál es la probabilidad de que dé servicio al menos a cinco autos el siguiente día de trabajo?

SOLUCIÓN

Sea E el evento de que al menos cinco autos reciban servicio. Ahora bien, $P(E) = 1 - P(E')$, donde E' es el evento de que menos de cinco autos reciban servicio. Como

$$P(E') = 0.12 + 0.19 = 0.31,$$

se sigue del teorema 2.12 que

$$P(E) = 1 - 0.31 = 0.69. \quad \blacksquare$$

Ejercicios

1. Encuentre los errores en cada una de las aseveraciones siguientes:

- (a) Las probabilidades de que un vendedor de automóviles venda 0, uno, dos, o tres autos en un día dado de febrero son 0.19, 0.38, 0.29 y 0.15, respectivamente.
 (b) La probabilidad de que llueva mañana es 0.40 y la probabilidad de que no llueva es 0.52.
 (c) Las probabilidades de que una impresora cometa 0, uno, dos, tres, o cuatro errores al imprimir un documento son 0.19, 0.34, -0.25, 0.43 y 0.29, respectivamente.
 (d) Al sacar una carta de una baraja en un solo intento la probabilidad de seleccionar corazones es $1/4$, la probabilidad de seleccionar una carta negra es $1/2$, y la probabilidad de seleccionar una carta negra de corazones es $1/8$.

2. Suponga que todos los elementos de S en el ejercicio 8 de la página 18 tienen la misma probabilidad de ocurrencia, encuentre

- (a) la probabilidad del evento A ;
 (b) la probabilidad del evento C ;
 (c) la probabilidad del evento $A \cap C$.

3. Una caja contiene 500 sobres de los cuales 75 contienen \$100 en efectivo, 150 contienen \$25 y 275 contienen \$10. Se puede comprar un sobre en \$25. ¿Cuál es el espacio muestral para las diferentes cantidades de dinero? Asigne probabilidades a los puntos muestrales y después encuentre la probabilidad de que el primer sobre que se compre contenga menos de \$100.

4. Suponga que en un grupo de último año de facultad de 500 estudiantes se encuentra que 210 fuman, 258 consumen bebidas alcohólicas, 216 comen entre comidas, 122 fuman y consumen bebidas alcohólicas, 83 comen entre comidas y consumen bebidas alcohólicas, 97 fuman y comen entre comidas y 52 tienen estos tres hábitos nocivos para la salud. Si se selecciona al azar un miembro de este grupo, encuentre la probabilidad de que el estudiante

- fume pero no consuma bebidas alcohólicas;
- coma entre comidas y consuma bebidas alcohólicas pero no fume;
- ni fume ni coma entre comidas.

5. La probabilidad de que una industria norteamericana se ubique en Munich es 0.7, la probabilidad de que se ubique en Bruselas es 0.4 y la probabilidad de que se ubique en Munich o Bruselas o en ambas es 0.8. ¿Cuál es la probabilidad de que la industria se ubique

- en ambas ciudades?
- en ninguna de estas ciudades?

6. De experiencias pasadas un agente de bolsa cree que con las condiciones económicas actuales un cliente invertirá en bonos libres de impuestos con una probabilidad de 0.6, invertirá en fondos mutualistas con una probabilidad de 0.3 e invertirá en ambos con una probabilidad de 0.15. Ahora, encuentre la probabilidad de que un cliente invierta

- en bonos libres de impuestos o en fondos mutualistas;
- en ninguno de los instrumentos.

7. Si se elige al azar una letra del alfabeto inglés, encuentre la probabilidad de que la letra

- sea una vocal excluyendo y;
- esté listada en algún lugar antes de la letra j;
- esté listada en algún lugar después de la letra g.

8. Un fabricante de automóviles está preocupado por el posible retiro de su sedán de cuatro puertas de mayor venta. Si hubiera un retiro, hay una probabilidad de 0.25 de que haya un defecto en el sistema de frenos, de 0.18 en la transmisión, de 0.17 en el sistema de combustible y 0.40 en alguna otra área.

- ¿Cuál es la probabilidad de que el defecto esté en los frenos o en el sistema de combustible si la probabilidad de defectos simultáneos en ambos sistemas es 0.2?
- ¿Cuál es la probabilidad de que no haya defectos en los frenos o en el sistema de combustible?

9. Si cada artículo codificado en un catálogo comienza con tres letras distintas seguidas por cuatro dígitos distintos de cero, encuentre la probabilidad de seleccionar aleatoriamente uno de estos artículos codifica-

dos que tenga como primera letra una vocal y el último dígito sea par.

10. Se lanza un par de dados. Encuentre la probabilidad de obtener

- un total de ocho;
- a lo más un total de cinco.

11. Se sacan dos cartas sucesivamente de una baraja sin remplazo. ¿Cuál es la probabilidad de que ambas cartas sean mayores que dos y menores que ocho?

12. Si se toman tres libros al azar de un librero que contiene cinco novelas, tres libros de poemas y un diccionario, ¿cuál es la probabilidad de que

- se seleccione el diccionario?
- se seleccionen dos novelas y un libro de poemas?

13. En una mano de póquer que consiste en cinco cartas, encuentre la probabilidad de tener

- tres ases;
- cuatro cartas de corazones y una de tréboles.

14. En un juego de Yahtzee, donde se lanzan cinco dados de forma simultánea, encuentre la probabilidad de obtener cuatro del mismo tipo.

15. En una clase de 100 estudiantes graduados de preparatoria, 54 estudiaron matemáticas; 69, historia y 35 cursaron matemáticas e historia. Si se selecciona al azar uno de estos estudiantes, encuentre la probabilidad de que

- el estudiante cursó matemáticas o historia;
- el estudiante no llevó alguna de estas materias;
- el estudiante cursó historia pero no matemáticas.

16. La empresa Dom's Pizza Company utiliza pruebas de sabor y el análisis estadístico de los datos antes de comercializar cualquier producto nuevo. Considere un estudio que incluye tres tipos de pastas (delgada, delgada con ajo y orégano, y delgada con trozos de queso). También estudian tres salsas (estándar, una nueva salsa con más ajo y una nueva salsa con albahaca fresca).

- ¿Cuántas combinaciones de pasta y salsa se incluyen?
- ¿Cuál es la probabilidad de que un juez tenga una pasta delgada sencilla con salsa estándar en su primera prueba de sabor?

17. De acuerdo con la Consumer Digest (julio/agosto de 1996), la ubicación probable de las PC en una casa son:

Recámara de adultos	0.03
Recámara de niños	0.15
Otra recámara	0.14
Oficina o estudio	0.40
Otros cuartos	0.28

- (a) ¿Cuál es la probabilidad de que una PC esté en una recámara?
- (b) ¿Cuál es la probabilidad de que no esté en una recámara?

- (c) Suponga que se selecciona una familia al azar entre las familias con una PC; ¿en qué habitación esperaría encontrar una PC?

2.6 Probabilidad condicional

La probabilidad de que un evento B ocurra cuando se sabe que ya ocurrió algún evento A se llama **probabilidad condicional** y se denota por $P(B|A)$. El símbolo $P(B|A)$ por lo general se lee “la probabilidad de que ocurra B dado que ocurrió A ” o simplemente “la probabilidad de B , dado A ”.

Considere el evento B de obtener un cuadrado perfecto cuando se lanza un dado. El dado se construye de modo que los números pares tengan el doble de probabilidad de ocurrencia que los números nones. Con base en el espacio muestral $S = \{1, 2, 3, 4, 5, 6\}$, con probabilidades asignadas de $1/9$ y $2/9$ respectivamente, a los números impares y pares, la probabilidad de que ocurra B es $1/3$. Suponga ahora que se sabe que el lanzamiento del dado tiene como resultado un número mayor que 3. Tenemos ahora un espacio muestral reducido $A = \{4, 5, 6\}$, que es un subconjunto de S . Para encontrar la probabilidad de que ocurra B , en relación con el espacio A , debemos asignar primero nuevas probabilidades a los elementos de A proporcionales a sus probabilidades originales de modo que su suma sea 1. Al asignar una probabilidad de w al número non en A y una probabilidad de $2w$ a los dos números pares, tenemos $5w = 1$ o $w = 1/5$. En relación con el espacio A , encontramos que B contiene sólo el elemento 4. Si denotamos este evento con el símbolo $B|A$, escribimos $B|A = \{4\}$, y de aquí

$$P(B|A) = \frac{2}{5}.$$

Este ejemplo ilustra que los eventos pueden tener probabilidades diferentes cuando se consideran en relación con diferentes espacios muestrales.

También podemos escribir

$$P(B|A) = \frac{2}{5} = \frac{2/9}{5/9} = \frac{P(A \cap B)}{P(A)},$$

donde $P(A \cap B)$ y $P(A)$ se encuentran a partir del espacio muestral original S . En otras palabras, una probabilidad condicional relativa a un subespacio A de S se puede calcular de forma directa de las probabilidades que se asignan a los elementos del espacio muestral original S .

Definición 2.9

La probabilidad condicional de B , dado A , que se denota con $P(B|A)$, se define como

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{si} \quad P(A) > 0.$$

Como ilustración adicional, suponga que nuestro espacio muestral S es la población de adultos en una pequeña ciudad que cumplen con los requisitos para obtener un grado en la facultad. Debemos clasificarlos de acuerdo con su sexo y situación laboral:

	Empleado	Desempleado	Total
Hombre	460	40	500
Mujer	140	260	400
Total	600	300	900

Uno de estos individuos se seleccionará al azar para que realice un viaje a través del país para promover las ventajas de establecer industrias nuevas en la ciudad. Nos interesaremos en los eventos siguientes:

- M : se elige un hombre.
 E : el elegido tiene empleo.

Al utilizar el espacio muestral reducido E , encontramos que

$$P(M|E) = \frac{460}{600} = \frac{23}{30}.$$

Sea $n(A)$ el número de elementos en cualquier conjunto A . Con el uso de esta notación, podemos escribir

$$P(M|E) = \frac{n(E \cap M)}{n(E)} = \frac{n(E \cap M)/n(S)}{n(E)/n(S)} = \frac{P(E \cap M)}{P(E)},$$

donde $P(E \cap M)$ y $P(E)$ se encuentran a partir del espacio muestral original S . Para verificar este resultado, note que

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{y} \quad P(E \cap M) = \frac{460}{900} = \frac{23}{45}.$$

De aquí

$$P(M|E) = \frac{23/45}{2/3} = \frac{23}{30}, \quad \text{como antes.}$$

Ejemplo 2.31 La probabilidad de que un vuelo programado normalmente salga a tiempo es $P(D) = 0.83$; la probabilidad de que llegue a tiempo es $P(A) = 0.82$; y la probabilidad de que salga y llegue a tiempo es $P(D \cap A) = 0.78$. Encuentre la probabilidad de que un avión a) llegue a tiempo, dado que salió a tiempo y b) salió a tiempo, dado que llegó a tiempo.

SOLUCIÓN

(a) La probabilidad de que un avión llegue a tiempo, dado que salió a tiempo es

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) La probabilidad de que un avión saliera a tiempo, dado que llegó a tiempo es

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

En el experimento de lanzamiento de un dado que se discutió en la página 35 notamos que $P(B|A) = 2/5$ mientras que $P(B) = 1/3$. Es decir, $P(B|A) \neq P(B)$, lo que indica que B depende de A . Consideremos ahora un experimento en el que se sacan 2 cartas una después de la otra de una baraja ordinaria, con reemplazo. Los eventos se definen como

- A : la primera carta es un as,
 B : la segunda carta es una espada.

Como la primera carta se reemplaza, nuestro espacio muestral para la primera y segunda cartas consiste en 52 cartas, que contienen cuatro ases y 13 espadas. De aquí

$$P(B|A) = \frac{13}{52} = \frac{1}{4} \quad \text{y} \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

Es decir, $P(B|A) = P(B)$. Cuando esto es cierto, se dice que los eventos A y B son **independientes**. ■

La noción de probabilidad condicional proporciona la capacidad de reevaluar la idea de probabilidad de un evento a la luz de la información adicional; es decir, cuando se sabe que ocurrió otro evento. La probabilidad $P(A|B)$ es una "actualización" de $P(A)$ basada en el conocimiento de que ocurrió el evento B . En el ejemplo 2.31 es importante conocer la probabilidad de que los vuelos lleguen a tiempo. Se nos da la información de que el vuelo no salió a tiempo. Con esta información adicional, la probabilidad más pertinente es $P(A|D')$, esto es, la probabilidad de que llegue a tiempo, dado que no salió a tiempo. En muchas situaciones las conclusiones que se obtienen de observar la probabilidad condicional más importante cambian por completo la situación. En este ejemplo el cálculo de $P(A|D')$ es

$$P(A|D') = \frac{P(A \cap D')}{P(D')} = \frac{0.82 - 0.78}{0.17} = 0.24$$

Como consecuencia, la probabilidad de una llegada a tiempo disminuye considerablemente ante la presencia de la información adicional.

Eventos independientes

Aunque la probabilidad condicional tiene en cuenta la alteración de la probabilidad de un evento a la luz de material adicional, también nos permite comprender mejor el muy importante concepto de **independencia** o, en el contexto actual, el de eventos independientes. En la ilustración del aeropuerto, $P(D|A)$ difiere de $P(D)$. Esto sugiere que la ocurrencia de A influye en D y esto realmente se espera en este caso. Sin embargo, considere la situación donde tenemos los eventos A y B y

$$P(A|B) = P(A).$$

En otras palabras, la ocurrencia de B no tiene impacto en las probabilidades de ocurrencia de A . Aquí la ocurrencia de A es *independiente* de la ocurrencia de B .

La importancia del concepto de independencia no se puede enfatizar con exceso. Juega un papel vital en el material de casi todos los capítulos de este libro y en todas las áreas de la estadística aplicada.

Definición 2.10

Dos eventos A y B son **independientes** si y sólo si

$$P(B|A) = P(B) \quad \text{y} \quad P(A|B) = P(A).$$

De otra forma, A y B son **dependientes**.

La condición $P(B|A) = P(B)$ implica que $P(A|B) = P(A)$, y viceversa. Para los experimentos de extracción de una carta, donde mostramos que $P(B|A) = P(B) = 1/4$, también podemos ver que $P(A|B) = P(A) = 1/13$.

2.7 Reglas multiplicativas

Al multiplicar la fórmula de la definición 2.9 por $P(A)$, obtenemos la siguiente **regla multiplicativa** importante, que nos permite calcular la probabilidad de que ocurran dos eventos.

Teorema 2.13

Si en un experimento pueden ocurrir los eventos A y B , entonces

$$P(A \cap B) = P(A)P(B|A).$$

Así la probabilidad de que ocurran A y B es igual a la probabilidad de que ocurra A multiplicada por la probabilidad de que ocurra B , dado que ocurre A . Como los eventos $A \cap B$ y $B \cap A$ son equivalentes, se sigue del teorema 2.13 que también podemos escribir

$$P(A \cap B) = P(B \cap A) = P(B)P(A|B).$$

En otras palabras, no importa cuál evento se considera como A y cuál como B .

Ejemplo 2.32 Suponga que tenemos una caja de fusibles que contiene 20 unidades, de las cuales cinco están defectuosas. Si se seleccionan dos fusibles al azar y se separan de la caja uno después del otro sin reemplazar el primero, ¿cuál es la probabilidad de que ambos fusibles estén defectuosos?

SOLUCIÓN

Sean A el evento de que el primer fusible esté defectuoso y B el evento de que el segundo esté defectuoso; entonces interpretamos $A \cap B$ como el evento de que ocurra A , y entonces B ocurre después de que ocurre A . La probabilidad de separar primero un fusible defectuoso es $1/4$; entonces la probabilidad de separar un segundo fusible defectuoso de los restantes 4 es $4/19$. Por ello

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{4}{19}\right) = \frac{1}{19} \quad \blacksquare$$

Ejemplo 2.33 Una bolsa contiene cuatro bolas blancas y tres negras, y una segunda bolsa contiene tres blancas y cinco negras. Se saca una bola de la primera bolsa y se coloca sin verla en la segunda. ¿Cuál es la probabilidad de que ahora se saque una bola negra de la segunda bolsa?

SOLUCIÓN

Sean B_1 , B_2 , y W_1 respectivamente, la extracción de una bola negra de la bolsa 1, una negra de la bolsa 2 y una blanca de la bolsa 1. Nos interesa la unión de los eventos mutuamente excluyentes $B_1 \cap B_2$ y $W_1 \cap B_2$. Las diversas posibilidades y sus probabilidades se ilustran en la figura 2.8. Entonces

$$\begin{aligned} P[(B_1 \cap B_2) \cup (W_1 \cap B_2)] &= P(B_1 \cap B_2) + P(W_1 \cap B_2) \\ &= P(B_1)P(B_2|B_1) + P(W_1)P(B_2|W_1) \\ &= \left(\frac{3}{7}\right)\left(\frac{6}{9}\right) + \left(\frac{4}{7}\right)\left(\frac{5}{9}\right) = \left(\frac{38}{63}\right). \end{aligned}$$

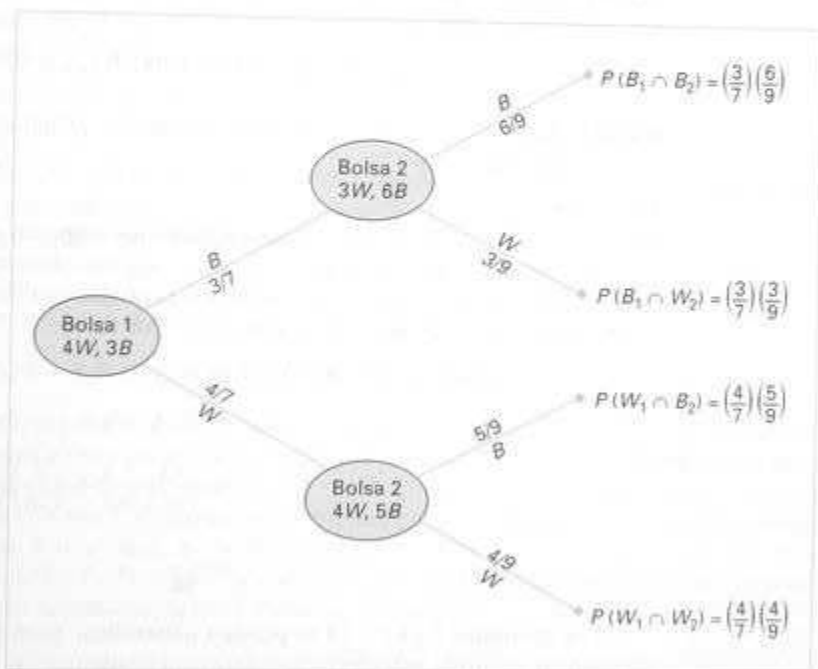


Figura 2.8 Diagrama de árbol para el ejemplo 2.33.

Si, en el ejemplo 2.32, el primer fusible se reemplaza y los fusibles se reacomodan por completo antes de que se extraiga el segundo, entonces la probabilidad de un fusible defectuoso en la segunda selección aún es $1/4$; es decir, $P(B|A) = P(B)$ y los eventos A y B son independientes. Cuando esto es cierto, podemos sustituir $P(B)$ por $P(B|A)$ en el teorema 2.13 para obtener la siguiente regla especial de multiplicación. ■

Teorema 2.14

Dos eventos A y B son independientes si y sólo si

$$P(A \cap B) = P(A)P(B).$$

Por tanto, para obtener la probabilidad de que ocurran dos eventos independientes, simplemente calculamos el producto de sus probabilidades individuales.

Ejemplo 2.34 Una pequeña ciudad tiene un carro de bomberos y una ambulancia disponibles para emergencias. La probabilidad de que el carro de bomberos esté disponible cuando se necesite es 0.98 y la probabilidad de que la ambulancia esté disponible cuando se le requiera es 0.92. En el caso de que resulte un herido de un edificio en llamas, encuentre la probabilidad de que la ambulancia y el carro de bomberos estén disponibles.

SOLUCIÓN

Sean A y B los respectivos eventos de que estén disponibles el carro de bomberos y la ambulancia. Entonces

$$P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016. \quad \blacksquare$$

Ejemplo 2.35 Se lanza dos veces un par de dados. ¿Cuál es la probabilidad de obtener totales de siete y 11?

SOLUCIÓN

Sean $A_1, A_2, B_1,$ y B_2 los eventos independientes respectivos de que ocurra un siete en la primera tirada; ocurra un siete en el segundo lanzamiento, un 11 en el primero y un 11 en el segundo. Nos interesa la probabilidad de la unión de los eventos mutuamente excluyentes $A_1 \cap B_2$ y $B_1 \cap A_2$. Por tanto,

$$\begin{aligned} P[(A_1 \cap B_2) \cup (B_1 \cap A_2)] &= P(A_1 \cap B_2) + P(B_1 \cap A_2) \\ &= P(A_1)P(B_2) + P(B_1)P(A_2) \\ &= \left(\frac{1}{6}\right)\left(\frac{1}{18}\right) + \left(\frac{1}{18}\right)\left(\frac{1}{6}\right) \\ &= \frac{1}{54}. \end{aligned}$$

Los teoremas 2.13 y 2.14 se pueden generalizar para cubrir cualquier número de eventos, como se establece en el teorema siguiente. \blacksquare

Teorema 2.15

Si, en un experimento, pueden ocurrir los eventos $A_1, A_2, A_3, \dots, A_k$, entonces

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_k) &= \\ P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_k|A_1 \cap A_2 \cap \dots \cap A_{k-1}). \end{aligned}$$

Si los eventos $A_1, A_2, A_3, \dots, A_k$ son independientes, entonces

$$P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_k) = P(A_1)P(A_2)P(A_3) \dots P(A_k).$$

Ejemplo 2.36 Se sacan tres cartas una tras otra, sin reemplazo, de una baraja ordinaria. Encuentre la probabilidad de que ocurra el evento $A_1 \cap A_2 \cap A_3$, donde A_1 es el evento de que la primera carta sea un as rojo, A_2 el evento de que la segunda carta sea un 10 o una sota y A_3 el evento de que la tercera carta sea mayor que tres pero menor que siete.

SOLUCIÓN

Primero definimos los eventos

A_1 : la primera carta es un as rojo,

A_2 : la segunda carta es un 10 o una sota,

A_3 : la tercera carta es mayor que tres pero menor que siete.

Entonces

$$P(A_1) = \frac{2}{52}, \quad P(A_2|A_1) = \frac{8}{51}, \quad P(A_3|A_1 \cap A_2) = \frac{12}{50},$$

y de aquí, por el teorema 2.15,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) = \left(\frac{2}{52}\right)\left(\frac{8}{51}\right)\left(\frac{12}{50}\right) = \frac{8}{5525}. \quad \blacksquare$$

Ejemplo 2.37 Se carga una moneda de modo que la cara tenga una posibilidad de ocurrir dos veces mayor que la cruz. Si se lanza tres veces la moneda, ¿cuál es la probabilidad de obtener dos cruces y una cara?

SOLUCIÓN

El espacio muestral para el experimento consiste en los ocho elementos,

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}.$$

Sin embargo, con una moneda no balanceada ya no es posible asignar probabilidades iguales a cada punto de la muestra. Para encontrar las probabilidades, considérese primero el espacio muestral $S_1 = \{H, T\}$, que representa los resultados cuando se lanza una vez la moneda. Si se asignan probabilidades de w y $2w$ para obtener una cruz y una cara, respectivamente, tenemos $3w = 1$ o $w = 1/3$. Por tanto $P(H) = 2/3$ y $P(T) = 1/3$. Sea ahora A el evento de obtener dos cruces y una cara en los tres lanzamientos de la moneda. Entonces,

$$A = \{TTH, THT, HTT\},$$

y como los resultados en cada uno de los tres lanzamientos son independientes, se sigue del teorema 2.15 que

$$P(TTH) = P(T \cap T \cap H) = P(T)P(T)P(H) = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{27}.$$

De manera similar,

$$P(THT) = P(HTT) = \frac{2}{27} \quad \text{y por ello} \quad P(A) = \frac{2}{27} + \frac{2}{27} + \frac{2}{27} = \frac{2}{9}. \quad \blacksquare$$

Ejercicios

1. Si R es el evento de que un convicto cometiera un robo a mano armada y D es el evento de que el convicto promoviera el consumo de drogas, exprese en palabras lo que en probabilidades se indica por

- (a) $P(R|D)$;
 (b) $P(D^c|R)$;
 (c) $P(R^c|D^c)$.

2. Una clase de física avanzada se compone de 10 estudiantes de primer año, tres del último año y 10 graduados. Las calificaciones finales muestran que tres estudiantes de primer año, 10 del último año y cinco de los graduados obtuvieron A en el curso. Si se elige un estudiante al azar de esta clase y se encuentra que es uno de los que obtuvieron A , ¿cuál es la probabilidad de que él o ella sea un estudiante de último año?

3. Una muestra aleatoria de 200 adultos se clasifica abajo por sexo y su nivel de educación.

Educación	Hombre	Mujer
Primaria	38	45
Secundaria	28	50
Facultad	22	17
	77	109

Si se escoge una persona al azar de este grupo, encuentre la probabilidad de que

- (a) la persona sea hombre, dado que la persona tiene educación secundaria;
 (b) la persona no tiene un grado universitario, dado que la persona es mujer.

4. En un experimento para estudiar la relación de la hipertensión arterial y los hábitos de fumar, se reúnen los siguientes datos para 180 individuos:

	No fumadores	Fumadores moderados	Fumadores empedernidos
Con hipertensión	21	36	30
Sin hipertensión	48	26	19

Si se selecciona uno de estos individuos al azar, encuentre la probabilidad de que la persona

- (a) sufre hipertensión, dado que la persona es un fumador empedernido;
 (b) sea un no fumador, dado que la persona no sufre de hipertensión.

5. En el último año de una clase de graduados de preparatoria con 100 alumnos, 42 cursaron matemáticas; 68, psicología; 54, historia; 22 matemáticas e historia; 25, matemáticas y psicología, siete historia pero ni matemáticas ni psicología; 10, cursaron las tres ma-

terias y ocho no tomaron alguna de las tres. Si se selecciona un estudiante al azar, encuentre la probabilidad de que

- (a) una persona inscrita en psicología cursa las tres materias;
 (b) una persona que no se inscribió en psicología cursa historia y matemáticas.

6. Un fabricante de una vacuna para la gripe se interesa en la calidad de su suero. Tres diferentes departamentos procesan los lotes de suero y tienen tasas de rechazo de 0.10, 0.08 y 0.12, respectivamente. Las inspecciones de los tres departamentos son secuenciales e independientes.

- (a) ¿Cuál es la probabilidad de que un lote de suero sobreviva a la primera inspección departamental pero sea rechazado por el segundo departamento?
 (b) ¿Cuál es la probabilidad de que un lote de suero sea rechazado por el tercer departamento?

7. En *USA Today* (5 de septiembre de 1996) se listaron como sigue los resultados de una encuesta sobre el uso de ropa para dormir mientras se viaja:

	Hombre	Mujer
Ropa interior	0.220	0.024
Camisón	0.002	0.180
Nada	0.160	0.018
Pijamas	0.102	0.073
Camiseta	0.046	0.088
Otros	0.084	0.003

- (a) ¿Cuál es la probabilidad de que un viajero sea una mujer que duerme desnuda?
 (b) ¿Cuál es la probabilidad de que un viajero sea hombre?
 (c) Suponga que el viajero es hombre, ¿cuál es la probabilidad de que un viajero sea mujer si duerme en pijama?
 (d) ¿Cuál es la probabilidad de que un viajero sea hombre si duerme en pijama o en camiseta?

8. La probabilidad de que un automóvil al que se llena el tanque de gasolina también necesite un cambio de aceite es 0.25, la probabilidad de que necesite un nuevo filtro de aceite es 0.40 y la probabilidad de que necesite cambio de aceite y filtro es 0.14.

- (a) Si se tiene que cambiar el aceite, ¿cuál es la probabilidad de que se necesite un nuevo filtro?
 (b) Si necesita un filtro de aceite nuevo, ¿cuál es la probabilidad de que se tenga que cambiar el aceite?

9. La probabilidad de que un hombre casado vea cierto programa de televisión es 0.4 y la probabilidad de que una mujer casada vea el programa es 0.5. La probabilidad de que un hombre vea el programa, dado que su esposa lo hace, es 0.7. Encuentre la probabilidad de que

- (a) un matrimonio vea el programa;
- (b) una esposa vea el programa dado que su esposo lo ve;
- (c) al menos una persona de un matrimonio vea el programa.

10. Para matrimonios que viven en cierto suburbio la probabilidad de que el esposo vote en un referéndum es 0.21, la probabilidad de que su esposa vote es 0.28 y la probabilidad de que ambos voten es 0.15. ¿Cuál es la probabilidad de que

- (a) al menos un miembro de un matrimonio vote?
- (b) una esposa vote, dado que su esposo votará?
- (c) un esposo vote, dado que su esposa no vote?

11. La probabilidad de que un vehículo que entra a las Cavernas Luray tenga placas de Canadá es 0.12, la probabilidad de que sea una casa rodante es 0.28 y la probabilidad de que sea una casa rodante con placas de Canadá es 0.09. ¿Cuál es la probabilidad de que

- (a) una casa rodante que entra a las Cavernas Luray tenga placas de Canadá?
- (b) un vehículo con placas de Canadá que entra a las Cavernas Luray sea una casa rodante?
- (c) un vehículo que entra a las Cavernas Luray no tenga placas de Canadá o que no sea una casa rodante?

12. La probabilidad de que el jefe de familia esté en casa cuando un representante de MCI llame es 0.4. Dado que el jefe de familia está en casa, la probabilidad de que ocurra un cambio de compañía para las llamadas de larga distancia es 0.3. Encuentre la probabilidad de que el jefe de familia esté en casa y cambie a MCI para el servicio de llamadas de larga distancia.

13. La probabilidad de que un doctor diagnostique de manera correcta una enfermedad particular es 0.7. Dado que el doctor hace un diagnóstico incorrecto, la probabilidad de que el paciente presente una demanda es 0.9. ¿Cuál es la probabilidad de que el doctor haga un diagnóstico incorrecto y el paciente lo demande?

14. En 1970, 11% de los estadounidenses completaron cuatro años de profesional, de los cuales 43% eran mujeres. En 1990, 22% de los estadounidenses completaron cuatro años de profesional, de los cuales 53% fueron mujeres. (*Time*, 19 de enero de 1996.)

(a) Dado que una persona completo cuatro años de profesional en 1970, ¿cuál es la probabilidad de que la persona sea mujer?

(b) ¿Cuál es la probabilidad de que una mujer termine cuatro años de profesional en 1990?

(c) ¿Cuál es la probabilidad de que en 1990 un hombre no haya terminado la carrera profesional?

15. Un agente de bienes raíces tiene ocho llaves maestras para abrir varias casas nuevas. Sólo una llave maestra abrirá cualquiera de las casas. Si 40% de estas casas por lo general se dejan abiertas, ¿cuál es la probabilidad de que el agente pueda entrar en una casa específica si selecciona tres llaves maestras al azar antes de salir de la oficina?

16. Antes de la distribución de cierto software estadístico se prueba la precisión de cada cuarto disco compacto (CD). El proceso de prueba consiste en correr cuatro programas independientes y verificar los resultados. La tasa de falla para los cuatro programas de prueba son 0.01, 0.03, 0.02 y 0.01, respectivamente.

(a) ¿Cuál es la probabilidad de que un CD que se pruebe falle cualquier prueba?

(b) Dado que se prueba un CD, ¿cuál es la probabilidad de que falle el programa dos o tres?

(c) En una muestra de 100, ¿cuántos CD esperaría que se rechazaran?

(d) Dado que un CD está defectuoso, ¿cuál es la probabilidad de que se pruebe?

17. Una ciudad tiene dos carros de bomberos que operan de forma independiente. La probabilidad de que un carro específico esté disponible cuando se le necesite es 0.96.

(a) ¿Cuál es la probabilidad de que ninguno esté disponible cuando se les necesite?

(b) ¿Cuál es la probabilidad de que un carro de bomberos esté disponible cuando se le necesite?

18. La probabilidad de que Tom viva 20 años más es 0.7, y la probabilidad de que Nancy viva 20 años más es 0.9. Si suponemos independencia para ambos, ¿cuál es la probabilidad de que ninguno viva 20 años más?

19. Un neceser contiene dos frascos de aspirina y tres frascos de tabletas para la tiroides. Un segundo bolso grande contiene tres frascos de aspirinas, dos frascos de tabletas para la tiroides y un frasco de tabletas laxantes. Si se saca un frasco de tabletas al azar de cada equipaje, encuentre la probabilidad de que

(a) ambos frascos contengan tabletas para la tiroides;

- (b) ningún frasco contenga tabletas para la tiroides;
 (c) los dos frascos contengan tabletas diferentes.

20. La probabilidad de que una persona que visita a su dentista necesite rayos X es 0.6, la probabilidad de que una persona que necesite una placa de rayos X también tenga una amalgama es 0.3, y la probabilidad de que una persona que tenga una placa de rayos X y una amalgama también tenga una extracción dental es 0.1. ¿Cuál es la probabilidad de que una persona que visita a su dentista tenga una placa de rayos X, una amalgama y un diente extraído?

21. Encuentre la posibilidad de seleccionar aleatoriamente 4 litros de leche en buenas condiciones sucesivamente de un refrigerador que contiene 20 litros de los cuales cinco están echados a perder, mediante el uso de

- (a) la primera fórmula del teorema 2.15 en la página 40;
 (b) las fórmulas de los teoremas 2.8 y 2.9 en las páginas 25 y 29, respectivamente.

2.8 Regla de Bayes

Regresemos a la ilustración de la sección 2.6, donde un individuo se selecciona al azar de los adultos de una pequeña ciudad para viajar por el país y promover las ventajas de establecer industrias nuevas en la ciudad. Suponga que ahora damos la información adicional de que 36 de los empleados y 12 de los desempleados son miembros del Club Rotario. Deseamos encontrar la probabilidad del evento A de que el individuo seleccionado sea miembro del Club Rotario. Con referencia a la figura 2.9, podemos escribir A como la unión de los dos eventos mutuamente excluyentes $E \cap A$ y $E' \cap A$. De aquí

$$A = (E \cap A) \cup (E' \cap A),$$

y por el corolario 1 del teorema 2.10, y además el teorema 2.13, podemos escribir

$$\begin{aligned} P(A) &= P[(E \cap A) \cup (E' \cap A)] \\ &= P(E \cap A) + P(E' \cap A) \\ &= P(E)P(A|E) + P(E')P(A|E'). \end{aligned}$$

Los datos de la sección 2.6, junto con los datos adicionales dados arriba para el conjunto A , nos permiten calcular

$$P(E) = \frac{600}{900} = \frac{2}{3}, \quad P(A|E) = \frac{36}{600} = \frac{3}{50},$$

y

$$P(E') = \frac{1}{3}, \quad P(A|E') = \frac{12}{300} = \frac{1}{25}.$$



Figura 2.9 Diagrama de Venn para los eventos A , E y E' .

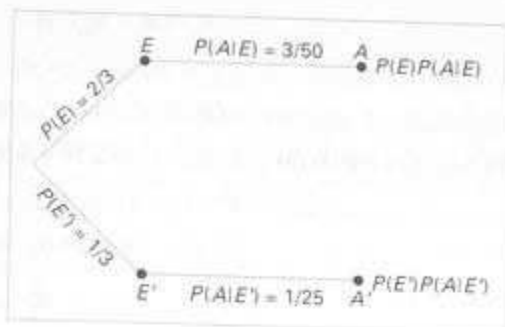


Figura 2.10 Diagrama de árbol para los datos de la página 44.

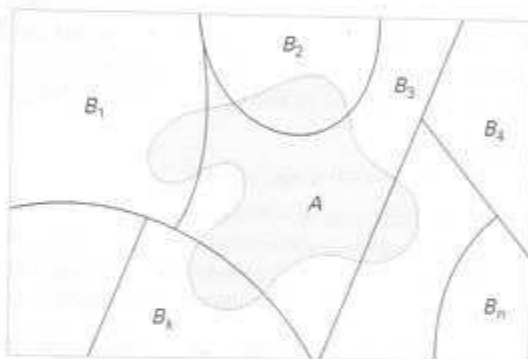


Figura 2.11 Partición del espacio muestral S .

Si mostramos estas probabilidades por medio del diagrama de árbol de la figura 2.10, donde la primera rama da la probabilidad $P(E)P(A|E)$ y la segunda rama da la probabilidad $P(E')P(A|E')$, se sigue que

$$P(A) = \left(\frac{2}{3}\right)\left(\frac{3}{50}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{25}\right) = \frac{4}{75}.$$

Una generalización de la ilustración precedente al caso donde el espacio muestral se parte en k subconjuntos que cubren el siguiente teorema, que algunas veces se denomina **teorema de probabilidad total** o **regla de eliminación**.

Teorema 2.16

Si los eventos B_1, B_2, \dots, B_k constituyen una partición del espacio muestral S tal que $P(B_i) \neq 0$ para $i = 1, 2, \dots, k$, entonces para cualquier evento A de S ,

$$P(A) = \sum_{i=1}^k P(B_i \cap A) = \sum_{i=1}^k P(B_i)P(A|B_i).$$

PRUEBA Considere el diagrama de Venn de la figura 2.11. Se observa que el evento A es la unión de los eventos mutuamente excluyentes

$$B_1 \cap A, B_2 \cap A, \dots, B_k \cap A;$$

es decir,

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A).$$

Con el uso del corolario 2 del teorema 2.10 y además el teorema 2.13, tenemos

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_k \cap A) \\ &= \sum_{i=1}^k P(B_i \cap A) = \sum_{i=1}^k P(B_i)P(A|B_i). \end{aligned}$$

Ejemplo 2.38 En cierta planta de montaje, tres máquinas, B_1 , B_2 y B_3 , montan 30%, 45% y 25% de los productos, respectivamente. Se sabe de la experiencia pasada que 2%, 3% y 2% de los productos ensamblados por cada máquina, respectivamente, tienen defectos. Ahora, suponga que se selecciona de forma aleatoria un producto terminado. ¿Cuál es la probabilidad de que esté defectuoso?

SOLUCIÓN

Considere los eventos siguientes:

A : el producto está defectuoso.

B_1 : el producto está ensamblado por la máquina B_1 .

B_2 : el producto está ensamblado por la máquina B_2 .

B_3 : el producto está ensamblado por la máquina B_3 .

Al aplicar la regla de eliminación, podemos escribir

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

Con referencia al diagrama de árbol de la figura 2.12, encontramos que las tres ramas dan las probabilidades

$$P(B_1)P(A|B_1) = (0.3)(0.02) = 0.006,$$

$$P(B_2)P(A|B_2) = (0.45)(0.03) = 0.0135,$$

$$P(B_3)P(A|B_3) = (0.25)(0.02) = 0.005.$$

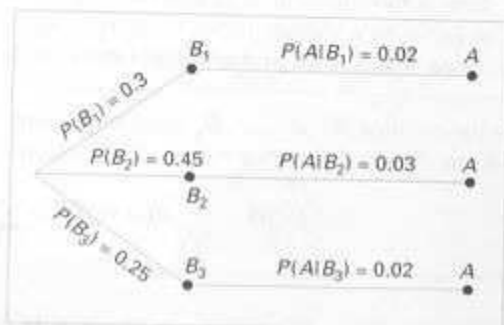


Figura 2.12 Diagrama de árbol para el ejemplo 2.38

y de aquí

$$P(A) = 0.006 + 0.0135 + 0.005 = 0.0245. \quad \blacksquare$$

En lugar de preguntar por $P(A)$, por la regla de eliminación, suponga que consideramos ahora el problema de encontrar la probabilidad condicional $P(B_r|A)$ en el ejemplo 2.38. En otras palabras, suponga que se seleccionó un producto de forma aleatoria y es defectuoso. ¿Cuál es la probabilidad de que este producto fuera hecho por la máquina B_r ? Preguntas de este tipo se pueden contestar mediante el uso del teorema siguiente, que se llama **regla de Bayes**:

Teorema 2.17

(Regla de Bayes) Si los eventos B_1, B_2, \dots, B_k constituyen una partición del espacio muestral S donde $P(B_i) \neq 0$ para $i = 1, 2, \dots, k$, entonces para cualquier evento A en S tal que $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^k P(B_i)P(A|B_i)} \quad \text{para } r = 1, 2, \dots, k.$$

PRUEBA Por la definición de probabilidad condicional,

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)}$$

y con el uso del teorema 2.16 en el denominador, tenemos

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)}.$$

Al aplicar el teorema 2.13 al numerador y al denominador, obtenemos la forma alterna,

$$P(B_r|A) = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^k P(B_i)P(A|B_i)},$$

que completa la demostración.

Ejemplo 2.39 Con referencia al ejemplo 2.38, si se elige al azar un producto y se encuentra que es defectuoso, ¿cuál es la probabilidad de que esté ensamblado por la máquina B_3 ?

SOLUCIÓN

Con el uso de la regla de Bayes para escribir

$$P(B_3|A) = \frac{P(B_3)P(A|B_3)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3)}$$

y después sustituir las probabilidades calculadas en el ejemplo 2.38, tenemos

$$P(B_3|A) = \frac{0.005}{0.006 + 0.0135 + 0.005} = \frac{0.005}{0.0245} = \frac{10}{49}$$

En vista del hecho de que se seleccionó un producto defectuoso, este resultado sugiere que probablemente no fue hecho con la máquina B_3 . ■

Ejercicios

1. En cierta región del país se sabe por experiencia del pasado que la probabilidad de seleccionar un adulto mayor de 40 años de edad con cáncer es 0.05. Si la probabilidad de que un doctor diagnostique de forma correcta que una persona con cáncer tiene la enfermedad es 0.78 y la probabilidad de que diagnostique de forma incorrecta que una persona sin cáncer como si tuviera la enfermedad es 0.06, ¿cuál es la probabilidad de que a una persona se le diagnostique cáncer?

2. La policía planea reforzar los límites de velocidad mediante el uso de un sistema de radar en cuatro diferentes puntos dentro de la ciudad. Las trampas de radar en cada uno de los sitios L_1 , L_2 , L_3 y L_4 operan 40%, 30%, 20% y 30% del tiempo, y si una persona que maneja a gran velocidad cuando va a su trabajo tiene las probabilidades de 0.2, 0.1, 0.5 y 0.2, respectivamente, de pasar por esos lugares, ¿cuál es la probabilidad de que reciba una multa por conducir con exceso de velocidad?

3. Refiérase al ejercicio 1, ¿cuál es la probabilidad de que una persona a la que se le diagnostica cáncer realmente tenga la enfermedad?

4. Si en el ejercicio 2 la persona es multada por conducir con exceso de velocidad en su camino al trabajo, ¿cuál es la probabilidad de que pase por el sistema de radar que se ubica en L_2 ?

5. Suponga que los cuatro inspectores de una fábrica de película colocan la fecha de caducidad en cada paquete de película al final de la línea de montaje. John, que coloca la fecha de caducidad en 20% de los paquetes, no la pone una vez en cada 200 paquetes; Tom, que la coloca en 60% de los paquetes, no la coloca una vez en cada 100 paquetes; Jeff, quien la coloca en el 15% de los paquetes, no lo hace una vez en cada 90 paquetes; y Pat, que fecha 5% de los paquetes, no lo hace una vez en cada 200 paquetes. Si un consumidor se queja de que su paquete de película no mues-

tra la fecha de caducidad, ¿cuál es la probabilidad de que haya sido inspeccionado por John?

6. Una compañía telefónica regional opera tres estaciones relevadoras idénticas en diferentes sitios. Durante un periodo de un año, el número de desperfectos reportados por cada estación y las causas se muestran abajo.

	Estaciones: A	B	C
Problemas con el suministro de electricidad	2	1	1
Desperfecto de la computadora	4	3	2
Fallas del equipo eléctrico	5	4	2
Fallas ocasionadas por otros errores humanos	7	7	5

Suponga que se reporta una falla y que se encuentra que fue ocasionada por otros errores humanos, ¿Cuál es la probabilidad de que provenga de la estación C?

7. La contaminación de los ríos en Estados Unidos es un problema de hace varios años. Considere los eventos siguientes:

A = {El río está contaminado.}

B = {Una prueba en una muestra de agua detecta contaminación.}

C = {Se permite la pesca.}

Suponga: $P(A) = 0.3$, $P(B|A) = 0.75$,

$P(B|A') = 0.20$, $P(C|A \cap B) = 0.20$,

$P(C|A' \cap B) = 0.15$,

$P(C|A \cap B') = 0.80$,

$P(C|A' \cap B') = 0.90$.

(a) Encuentre $P(A \cap B \cap C)$.

(b) Encuentre $P(B' \cap C)$.

(c) Encuentre $P(C)$.

(d) Encuentre la probabilidad de que el río esté contaminado, dado que se permite la pesca y que la prueba de la muestra no detecte contaminación.

8. Una cadena de tiendas de pintura produce y vende pintura de látex y semiesmaltada. Con base en las ventas de largo plazo, la probabilidad de que un cliente compre látex es 0.75. De los que compran pintura de látex, 60% también compran rodillos. Pero 30%

de los compradores de pintura semiesmaltada compran rodillos. Un comprador que se selecciona al azar compra un rodillo y una lata de pintura. ¿Cuál es la probabilidad de que la pintura sea de látex?

Ejercicios de repaso

1. Un suero de la verdad tiene la propiedad de que 90% de los sospechosos culpables se juzgan de forma adecuada mientras que, por supuesto, 10% de los sospechosos culpables resultan erróneamente inocentes. Por otro lado, a los sospechosos inocentes se les juzga de manera errónea 1% de las veces. Si el sospechoso se selecciona de un grupo de sospechosos de los que sólo 5% alguna vez han cometido un crimen, y el suero indica que es culpable, ¿cuál es la probabilidad de que sea inocente?

2. Una alergista afirma que 50% de los pacientes que examina son alérgicos a algún tipo de hierba. ¿Cuál es la probabilidad de que

(a) exactamente tres de sus cuatro próximos pacientes sean alérgicos a hierbas?

(b) ninguno de sus siguientes cuatro pacientes sean alérgicos a hierbas?

3. Mediante la comparación de las regiones apropiadas en un diagrama de Venn, verifique que

(a) $(A \cap B) \cup (A \cap B') = A$;

(b) $A' \cup (B' \cup C) = (A \cap B') \cup (A' \cup C)$.

4. Las probabilidades de que una estación de servicio bombee gasolina en 0, uno, dos, tres, cuatro, cinco o más autos durante cierto periodo de 30 minutos son 0.03, 0.18, 0.24, 0.28, 0.10 y 0.17, respectivamente. Encuentre la probabilidad de que en este periodo de 30 minutos

(a) más de dos autos reciban gasolina;

(b) a lo más cuatro autos reciban gasolina;

(c) cuatro o más autos reciban gasolina.

5. ¿Cuántas manos de bridge que contengan cuatro espadas, seis diamantes, una de bastos y dos corazones son posibles?

6. Si la probabilidad de que una persona cometa un error en su declaración de impuestos es 0.1, encuentre la probabilidad de que

(a) cuatro personas no relacionadas cometan cada una un error;

(b) el señor Jones y la señora Clark cometan un error, y el señor Roberts y la señora Williams no cometan errores.

7. Una empresa industrial grande usa tres hoteles locales para proporcionar hospedaje nocturno a sus clientes. Por experiencia pasada se sabe que a 20% de los clientes se les asignan habitaciones en el Ramada Inn, al 50% en el Sheraton y al 30% en el Lakeview Motor Lodge. Si hay una falla en la plomería en 5% de las habitaciones del Ramada Inn, en 4% de las habitaciones del Sheraton y en 8% de las habitaciones del Lakeview Motor Lodge, ¿cuál es la probabilidad de que

(a) a un cliente se le asigne una habitación con fallas en la plomería?

(b) a una persona con una habitación que tiene problemas de plomería se le haya asignado acomodo en el Lakeview Motor Lodge?

8. De un grupo de cuatro hombres y cinco mujeres, ¿cuántos comités de tres miembros son posibles

(a) sin restricciones?

(b) con un hombre y dos mujeres?

(c) con dos hombres y una mujer si cierto hombre debe estar en el comité?

9. La probabilidad de que un paciente se recupere de una delicada operación de corazón es 0.8. ¿Cuál es la probabilidad de que

(a) exactamente dos de los siguientes tres pacientes que tienen esta operación sobrevivan?

(b) los siguientes tres pacientes que tengan esta operación sobrevivan?

10. En cierta prisión federal se sabe que $\frac{2}{3}$ de los reclusos son menores de 25 años de edad. También se sabe que $\frac{3}{5}$ de los reos son hombres y que $\frac{5}{8}$ son mujeres de 25 años de edad o mayores. ¿Cuál es la probabilidad de que un prisionero seleccionado al azar de esta prisión sea mujer y de al menos 25 años de edad?

11. De cuatro manzanas rojas, cinco verdes y seis amarillas, ¿cuántas selecciones de nueve manzanas son posibles si se deben seleccionar tres de cada color?
12. De una caja que contiene seis bolas negras y cuatro verdes se extraen tres bolas sucesivamente, cada bola se reemplaza en la caja antes de que se extraiga la siguiente. ¿Cuál es la probabilidad de que
- las tres sean del mismo color?
 - cada color esté representado?
13. Un cargamento de 12 televisores contiene tres defectuosos. ¿De cuántas formas un hotel puede comprar cinco de estas unidades y recibir al menos dos defectuosas?
14. Se examinaron los planes de estudio de ingeniería eléctrica, química, industrial y mecánica. Se encuentra que algunos estudiantes no cursan estadística, algunos cursan un semestre y otros cursan dos semestres. Considere los eventos siguientes:
- A: Se cursa algo de estadística
 B: Ingenieros eléctricos e industriales
 C: Ingenieros químicos
- Utilice diagramas de Venn y sombree las áreas que representan los eventos siguientes:
- $(A \cap B)'$;
 - $(A \cup B)'$;
 - $(A \cap C) \cup B$.
15. Cierta agencia federal emplea a tres empresas consultoras (A , B y C) con probabilidades de 0.110,

0.35 y 0.25, respectivamente. De la experiencia pasada se sabe que las probabilidades de excesos en costos de las empresas son 0.05, 0.03 y 0.15, respectivamente. Suponga que la agencia experimenta un exceso en los costos.

- ¿Cuál es la probabilidad de que la empresa consultora implicada sea la compañía C ?
- ¿Cuál es la probabilidad de que sea la compañía A ?

16. Un fabricante estudia los efectos de la temperatura de cocción, tiempo de cocción y tipo de aceite para la cocción para hacer papas fritas. Se utilizan tres diferentes temperaturas, cuatro diferentes tiempos de cocción y tres diferentes aceites.

- ¿Cuál es el número total de combinaciones a estudiar?
- ¿Cuántas combinaciones se utilizarán para cada tipo de aceite?
- Discuta por qué las permutaciones no son un problema en este ejercicio.

17. Considere la situación del ejercicio 16 y suponga que el fabricante puede probar sólo dos combinaciones en un día.

- ¿Cuál es la probabilidad de que se elija cualquier conjunto dado de dos corridas?
- ¿Cuál es la probabilidad de que se utilice la temperatura más alta en cualquiera de estas dos combinaciones?



VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD

3.1 Concepto de variable aleatoria

La estadística se ocupa de realizar inferencias acerca de poblaciones y sus características. Se llevan a cabo experimentos cuyos resultados se encuentran sujetos al azar. La prueba de un número de componentes electrónicos es un ejemplo de **experimento estadístico**, término que se utiliza para describir cualquier proceso mediante el cual se generan varias observaciones al azar. A menudo es importante asignar una descripción numérica al resultado. Por ejemplo, el espacio muestral que da una descripción detallada de cada posible resultado cuando se prueban tres componentes electrónicos se puede escribir como

$$S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\},$$

donde N denota "no defectuoso" y D denota "defectuoso". Naturalmente, estamos interesados en el número de defectuosos que ocurren. De esta forma a cada punto en el espacio muestral se le *asignará un valor numérico* de 0, 1, 2 o 3. Estos valores son, por supuesto, cantidades aleatorias *determinadas por el resultado del experimento*. Se pueden ver como valores que toma la *variable aleatoria* X , el número de artículos defectuosos cuando se prueban tres componentes electrónicos.

Definición 3.1

Una **variable aleatoria** es una función que asocia un número real con cada elemento del espacio muestral.

Utilizaremos una letra mayúscula, digamos X , para denotar una variable aleatoria y su correspondiente minúscula, x en este caso, para uno de sus valores. En el ejemplo anterior de prueba de componentes electrónicos, notamos que la variable aleatoria X toma el valor 2 para todos los elementos en el subconjunto

$$E = \{DDN, DND, NDD\}$$

del espacio muestral S . Es decir, cada valor posible de X representa un evento que es un subconjunto del espacio muestral para el experimento dado.

Ejemplo 3.1 Se sacan dos bolas de manera sucesiva sin reemplazo de una urna que contiene cuatro bolas rojas y tres negras. Los posibles resultados y los valores y de la variable aleatoria Y , donde Y es el número de bolas rojas, son ■

Espacio muestral	y
RR	2
RB	1
BR	1
BB	0

Ejemplo 3.2 El empleado de un almacén regresa tres cascos de seguridad al azar a tres empleados de un taller siderúrgico que ya los habían probado. Si Smith, Jones y Brown, en ese orden, reciben uno de los tres cascos, liste los puntos muestrales para los posibles órdenes de regreso de los cascos y encuentre el valor m de la variable aleatoria M que representa el número de asociaciones correctas.

SOLUCIÓN

Si S , J y B representan los cascos de Smith, Jones y Brown, respectivamente, entonces los posibles arreglos en los que se pueden regresar los cascos y el número de asociaciones correctas son

Espacio muestral	m
SJB	3
SBJ	1
JSB	1
JBS	0
BSJ	0
BJS	1

En cada uno de los dos ejemplos anteriores, el espacio muestral contiene un número finito de elementos. Por otro lado, cuando se lanza un dado hasta que ocurre un cinco, obtenemos un espacio muestral con una secuencia interminable de elementos,

$$S = \{F, NF, NNF, NNNF, \dots\},$$

donde F y N representan, respectivamente, la ocurrencia y no ocurrencia de un 5. Pero incluso en este experimento el número de elementos puede ser igual a todos los números enteros de modo que hay un primer elemento, un segundo, un tercero y así sucesivamente, y en este sentido se pueden contar. ■

Definición 3.2

Si un espacio muestral contiene un número finito de posibilidades o una serie interminable con tantos elementos como números enteros existen, se llama **espacio muestral discreto**.

Los resultados de algunos experimentos estadísticos no pueden ser ni finitos ni contables. Tal es el caso, por ejemplo, cuando se lleva a cabo una investigación para medir las distancias que recorre cierta marca de automóvil en una ruta de prueba preestablecida con cinco litros de gasolina. Supongamos que la distancia es una variable que se mide con algún grado de precisión, entonces claramente tenemos un número infinito de posibles distancias en el espacio muestral que no se pueden igualar a todos los números enteros. También, si se registrara el tiempo requerido para que tenga lugar una reacción química, una vez más los posibles intervalos de tiempo que forman nuestro espacio muestral son infinitos en número e incontables. Vemos ahora que no todos los espacios muestrales necesitan ser discretos.

Definición 3.3

Si un espacio muestral contiene un número infinito de posibilidades igual al número de puntos en un segmento de línea, se llama **espacio muestral continuo**.

Una variable aleatoria se llama **variable aleatoria discreta** si se puede contar su conjunto de resultados posibles. Como los posibles valores de Y en el ejemplo 3.1 son 0, 1 y 2 y los valores posibles de M en el ejemplo 3.2 son 0, 1 y 3, se sigue que Y y M son variables aleatorias discretas. Cuando una variable aleatoria puede tomar valores en una escala continua, se le denomina **variable aleatoria continua**. A menudo los posibles valores de una variable aleatoria continua son precisamente los mismos valores que contiene el espacio muestral continuo. Tal es el caso cuando la variable aleatoria representa la medición de la distancia que cierta marca de automóvil recorre en una pista de prueba con cinco litros de gasolina.

En la mayor parte de los problemas prácticos, las variables aleatorias continuas representan datos *medidos*, como son todos los posibles pesos, alturas, temperaturas, distancias o periodos de vida, mientras que las variables aleatorias discretas representan datos *contados*, como el número de artículos defectuosos en una muestra de k artículos o el número de accidentes de carretera por año en un estado dado. Nótese que las variables aleatorias Y y M de los ejemplos 3.1 y 3.2 representan datos contados, Y el número de bolas rojas y M el número de asignaciones correctas de los cascos.

3.2 Distribuciones discretas de probabilidad

Una variable aleatoria discreta toma cada uno de sus valores con cierta probabilidad. En el caso de lanzar una moneda tres veces, la variable X que representa el número de caras, toma el valor 2 con probabilidad $3/8$, pues tres de los ocho puntos muestrales igualmente probables tienen como resultado dos caras y una cruz. Si se suponen pesos iguales para los eventos simples del ejemplo 3.2, la probabilidad de que ningún empleado obtenga de vuelta su casco correcto, es decir, la probabilidad de que M tome el valor cero, es $1/3$. Los valores posibles m de M y sus probabilidades son

m	0	1	3
$P(M = m)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Note que los valores de m agotan todos los casos posibles y por ello las probabilidades suman 1.

Con frecuencia es conveniente representar todas las probabilidades de una variable aleatoria X mediante una fórmula. Tal fórmula necesariamente sería una función de los valores numéricos x que denotaremos con $f(x)$, $g(x)$, $r(x)$, y así sucesivamente. Por tanto, escribimos $f(x) = P(X = x)$; es decir, $f(3) = P(X = 3)$. El conjunto de pares ordenados $(x, f(x))$ se llama **función de probabilidad** o **distribución de probabilidad** de la variable aleatoria discreta X .

Definición 3.4

El conjunto de pares ordenados $(x, f(x))$ es una **función de probabilidad**, **función masa de probabilidad** o **distribución de probabilidad** de la variable aleatoria discreta X si, para cada resultado posible x ,

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

Ejemplo 3.3 Un embarque de ocho microcomputadoras similares para una tienda contiene tres que están defectuosas. Si una escuela hace una compra al azar de dos de estas computadoras, encuentre la distribución de probabilidad para el número de defectuosas.

SOLUCIÓN

Sea X una variable aleatoria cuyos valores x son los números posibles de computadoras defectuosas que compra la escuela. Entonces x puede ser cualquiera de los números 0, 1 y 2. Ahora bien,

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} = \frac{10}{28}$$

$$f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} = \frac{15}{28}$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} = \frac{3}{28}$$

Entonces la distribución de probabilidad de X es ■

x	0	1	2
$f(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Ejemplo 3.4 Si una agencia de autos vende 50% de su inventario de cierto vehículo extranjero equipado con bolsas de aire, encuentre una fórmula para la distribución de probabilidad del número de autos con bolsas de aire entre los siguientes cuatro vehículos que venda la agencia.

SOLUCIÓN

Como la probabilidad de vender un automóvil con bolsas de aire es 0.5, los $2^4 = 16$ puntos del espacio muestral tienen la misma probabilidad de ocurrencia. Por tanto, el denominador para todas las probabilidades, y también para nuestra función, es 16. Para obtener el número de maneras de vender tres modelos con bolsas de aire, necesitamos considerar el número de formas de dividir cuatro resultados en dos celdas con tres modelos con bolsas de aire asignadas a una celda, y el modelo sin bolsas de aire asignado a la otra. Esto se puede hacer de $\binom{4}{3} = 4$ formas. En general, el evento de vender x modelos con bolsas de aire y $4 - x$ modelos sin bolsas de aire puede ocurrir de $\binom{4}{x}$ formas, donde x puede ser 0, 1, 2, 3 o 4. Entonces, la distribución de probabilidad $f(x) = P(X = x)$ es

$$f(x) = \frac{\binom{4}{x}}{16} \quad \text{para } x = 0, 1, 2, 3, 4.$$

Hay muchos problemas donde deseamos calcular la probabilidad de que el valor observado de una variable aleatoria X sea menor o igual que algún número real x . Al escribir $F(x) = P(X \leq x)$ para cualquier número real x , definimos a $F(x)$ como la **distribución acumulada** de la variable aleatoria X . ■

Definición 3.5

La **distribución acumulada** $F(x)$ de una variable aleatoria discreta X con distribución de probabilidad $f(x)$ es

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad \text{para } -\infty < x < \infty,$$

Para la variable aleatoria M , el número de asociaciones correctas en el ejemplo 3.2, tenemos

$$F(2.4) = P(M \leq 2.4) = f(0) + f(1) = \binom{1}{3} + \binom{1}{2} = \binom{2}{6}.$$

La distribución acumulada de M es

$$F(m) = \begin{cases} 0 & \text{para } m < 0 \\ \frac{1}{3} & \text{para } 0 \leq m < 1 \\ \frac{2}{6} & \text{para } 1 \leq m < 3 \\ 1 & \text{para } m \geq 3. \end{cases}$$

Se debe notar en particular el hecho de que la distribución acumulada se define no sólo para los valores que toma la variable aleatoria dada sino para todos los números reales.

Ejemplo 3.5 Encuentre la distribución acumulada de la variable aleatoria X del ejemplo 3.4. Mediante el uso de $F(x)$, verifique que $f(2) = 3/8$.

SOLUCIÓN

El cálculo directo de la distribución de probabilidad del ejemplo 3.4 da $f(0) = 1/16$, $f(1) = 1/4$, $f(2) = 3/8$, $f(3) = 1/4$, y $f(4) = 1/16$. Por tanto,

$$F(0) = f(0) = \frac{1}{16},$$

$$F(1) = f(0) + f(1) = \frac{5}{16},$$

$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16},$$

$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16},$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1.$$

De aquí

$$F(x) = \begin{cases} 0 & \text{para } x < 0 \\ \frac{1}{16} & \text{para } 0 \leq x < 1 \\ \frac{5}{16} & \text{para } 1 \leq x < 2 \\ \frac{11}{16} & \text{para } 2 \leq x < 3 \\ \frac{15}{16} & \text{para } 3 \leq x < 4 \\ 1 & \text{para } x \geq 4. \end{cases}$$

Entonces,

$$f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}. \quad \blacksquare$$

A menudo es útil ver una distribución de probabilidad en forma gráfica. Se pueden graficar los puntos $(x, f(x))$ del ejemplo 3.4 para obtener la figura 3.1. Al unir los puntos al eje x , ya sea con una línea punteada o con una sólida, obtenemos lo que por lo general se denomina como **gráfica de barras**. La figura 3.1 permite ver de forma fácil qué valores de X tienen más probabilidad de ocurrencia, y también indica en este caso una situación perfectamente simétrica.

En lugar de graficar los puntos $(x, f(x))$, más a menudo construimos rectángulos, como en la figura 3.2. Aquí los rectángulos se construyen de modo que sus bases, de igual ancho, se centren en cada valor x y sus alturas sean iguales a las probabilidades dadas por $f(x)$. Las bases se construyen de forma tal que no dejen espacios entre los rectángulos. La figura 3.2 se llama **histograma de probabilidad**.

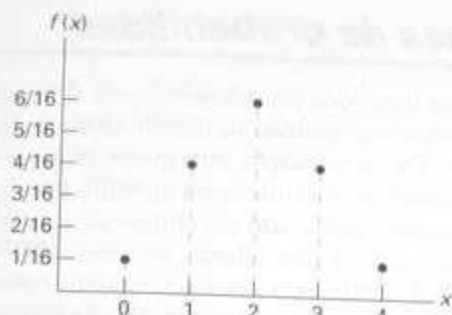


Figura 3.1 Gráfica de barras.

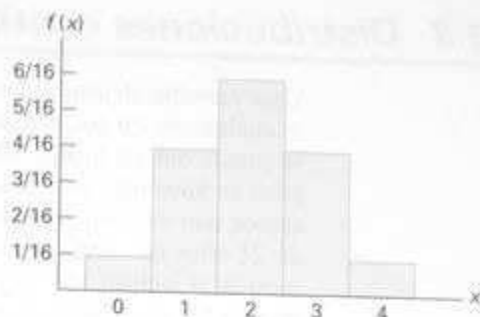


Figura 3.2 Histograma de probabilidad.



Figura 3.3 Distribución acumulada discreta.

Como cada base en la figura 3.2 tiene ancho unitario $P(X = x)$ es igual al área del rectángulo centrado en x . Incluso si las bases no fueran de ancho unitario, podríamos ajustar las alturas de los rectángulos para que las áreas aún fueran iguales a las probabilidades de X al tomar cualquiera de sus valores x . Este concepto de utilizar áreas para representar probabilidades es necesario para nuestra consideración de la distribución de probabilidad de una variable aleatoria continua.

La gráfica de la distribución acumulada del ejemplo 3.2, que aparece como una función escalonada en la figura 3.3, se obtiene al graficar los puntos $(x, F(x))$.

Ciertas distribuciones de probabilidad se aplican a más de una situación física. La distribución de probabilidad del ejemplo 3.4, por ejemplo, también se aplica a la variable aleatoria Y , donde Y es el número de caras cuando se lanza 4 veces una moneda, o a la variable aleatoria W , donde W es el número de cartas rojas que resultan cuando se sacan 4 cartas sucesivamente al azar de una baraja con el reemplazo de cada carta y barajando antes de sacar la siguiente. En el capítulo 5 se considerarán distribuciones discretas especiales que se pueden aplicar a muchas situaciones experimentales diferentes.

3.3 Distribuciones continuas de probabilidad

Una variable aleatoria continua tiene una probabilidad cero de tomar *exactamente* cualquiera de sus valores. En consecuencia, su distribución de probabilidad no se puede dar en forma tabular. En un principio esto puede parecer sorprendente, pero se hace más plausible cuando consideramos un ejemplo particular. Consideremos una variable aleatoria cuyos valores son las alturas de toda la gente mayor de 21 años de edad. Entre cualesquiera dos valores, digamos 163.5 y 164.5 centímetros, o incluso 163.99 y 164.01 centímetros, hay un número infinito de alturas, una de las cuales es 164 centímetros. La probabilidad de seleccionar una persona al azar que mida exactamente 164 centímetros de estatura y no sea una del conjunto infinitamente grande de estaturas tan cercanas a 164 centímetros que humanamente no se pueda medir la diferencia es remota, por ello asignamos una probabilidad cero a tal evento. Éste no es el caso, sin embargo, si nos referimos a la probabilidad de seleccionar una persona que al menos mida 163 centímetros pero no más de 165 centímetros de estatura. Tratamos ahora con un intervalo en lugar de un valor puntual de nuestra variable aleatoria.

Trataremos el cálculo de probabilidades para varios intervalos de variables aleatorias continuas como $P(a < X < b)$, $P(W > c)$, etcétera. Nótese que cuando X es continua,

$$P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b).$$

Es decir, no importa si incluimos o no un extremo del intervalo. Esto no es cierto, sin embargo, cuando X es discreta.

Aunque la distribución de probabilidad de una variable aleatoria continua no se puede representar de forma tabular, se puede establecer como una fórmula. Dicha fórmula necesariamente será función de los valores numéricos de la variable continua X y como tal se representará mediante la notación funcional $f(x)$. Al tratar con variables continuas, $f(x)$ por lo general se llama **función de densidad de probabilidad**, o simplemente **función de densidad** de X . Como X se define sobre un espacio muestral continuo, es posible que $f(x)$ tenga un número finito de discontinuidades. Sin embargo, la mayor parte de las funciones de densidad que tienen aplicaciones prácticas en el análisis de datos estadísticos son continuas y sus gráficas pueden tomar cualquiera de varias formas, algunas de las cuales se muestran en la figura 3.4. Como se utilizarán áreas para representar probabilidades y éstas son valores numéricos positivos, la función de densidad debe estar completamente por arriba del eje x .

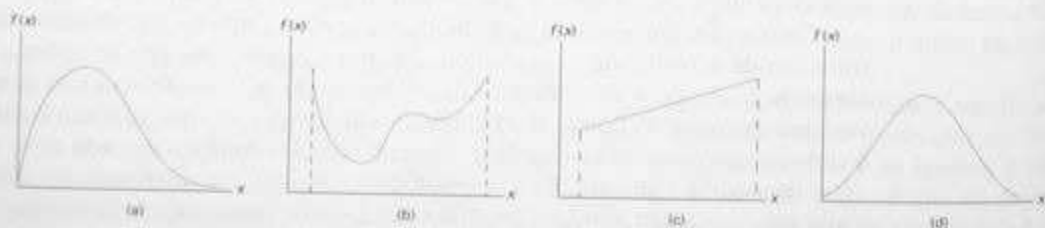
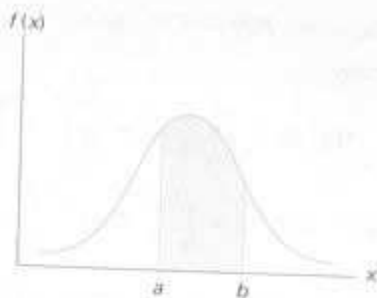


Figura 3.4 Funciones de densidad típicas.

Figura 3.5 $P(a < X < b)$.

Una función de densidad de probabilidad se construye de modo que el área bajo su curva limitada por el eje x sea igual a 1 cuando se calcula en el rango de X para el que se define $f(x)$. Si este rango de X es un intervalo finito, siempre es posible extender el intervalo para incluir a todo el conjunto de números reales al definir $f(x)$ como cero en todos los puntos de las partes extendidas del intervalo. En la figura 3.5, la probabilidad de que X tome un valor entre a y b es igual al área sombreada bajo la función de densidad entre las ordenadas en $x = a$ y $x = b$, y del cálculo integral es dada por

$$P(a < X < b) = \int_a^b f(x) dx.$$

Definición 3.6

La función $f(x)$ es una **función de densidad de probabilidad** para la variable aleatoria continua X , definida en el conjunto de números reales R , si

1. $f(x) \geq 0$, para toda $x \in R$.
2. $\int_{-\infty}^{\infty} f(x) dx = 1$.
3. $P(a < X < b) = \int_a^b f(x) dx$.

Ejemplo 3.6 Suponga que el error en la temperatura de reacción, en $^{\circ}\text{C}$, para un experimento de laboratorio controlado es una variable aleatoria continua X que tiene la función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Verifique la condición 2 de la definición 3.6.

(b) Encuentre $P(0 < X \leq 1)$.**SOLUCIÓN**

$$(a) \int_{-\infty}^{\infty} f(x) dx = \int_{-1}^2 \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_{-1}^2 = \frac{8}{9} + \frac{1}{9} = 1.$$

$$(b) P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_0^1 = \frac{1}{9}. \quad \blacksquare$$

Definición 3.7

La **distribución acumulada** $F(x)$ de una variable aleatoria continua X con función de densidad $f(x)$ es

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad \text{para } -\infty < x < \infty.$$

Como consecuencia inmediata de la definición 3.7 se pueden escribir los dos resultados,

$$P(a < X < b) = F(b) - F(a) \quad \text{y} \quad f(x) = \frac{dF(x)}{dx},$$

si existe la derivada.

Ejemplo 3.7 Para la función de densidad del ejemplo 3.6 encuentre $F(x)$, y utilícela para evaluar $P(0 < X \leq 1)$.

SOLUCIÓN

Para $-1 < x < 2$,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-1}^x \frac{t^2}{3} dt = \frac{t^3}{9} \Big|_{-1}^x = \frac{x^3 + 1}{9}.$$

Por tanto,

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{x^3 + 1}{9}, & -1 < x < 2 \\ 1, & x \geq 2. \end{cases}$$

La distribución acumulada $F(x)$ se expresa de forma gráfica en la figura 3.6. Ahora bien,

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

que concuerda con el resultado que se obtuvo al utilizar la función de densidad en el ejemplo 3.6. \blacksquare

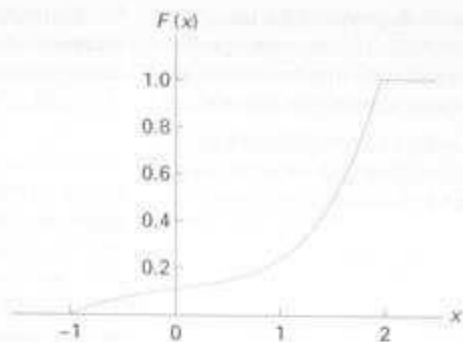


Figura 3.6 Distribución acumulada continua.

Ejercicios

1. Clasifique las siguientes variables aleatorias como discretas o continuas:

X : el número de accidentes automovilísticos por año en Virginia. ☹

Y : el tiempo para jugar 18 hoyos de golf. ☺

M : la cantidad de leche que una vaca particular produce anualmente. ☺

N : el número de huevos que pone mensualmente una gallina. ☹

P : el número de permisos de construcción que se emiten cada mes en una ciudad. ☹

Q : el peso del grano producido por acre. ☺

2. Un embarque de cinco automóviles extranjeros contiene dos que tienen ligeras manchas de pintura. Si una agencia recibe tres de estos automóviles al azar, liste los elementos del espacio muestral S con las letras B y N para “manchado” y “no manchado”, respectivamente; luego a cada punto muestral asigne un valor x de la variable aleatoria X que representa el número de automóviles que la agencia compra con manchas de pintura.

3. Sea W la variable aleatoria que da el número de caras menos el número de cruces en tres lanzamientos de una moneda. Liste los elementos del espacio muestral S para los tres lanzamientos de la moneda y asigne un valor w de W a cada punto muestral.

4. Se lanza una moneda hasta que ocurren 3 caras sucesivamente. Liste sólo aquellos elementos del espacio muestral que requieren seis o menos lanzamientos. ¿Es un espacio muestral discreto? Explique.

5. Determine el valor c de modo que cada una de las funciones siguientes puedan servir como distribución de probabilidad de la variable aleatoria discreta X :

$$(a) f(x) = c(x^2 + 4) \quad \text{para } x = 0, 1, 2, 3;$$

$$(b) f(x) = c \binom{2}{x} \binom{3}{3-x} \quad \text{para } x = 0, 1, 2.$$

6. La vida útil, en días, para frascos de cierta medicina de prescripción es una variable aleatoria que tiene la función de densidad

$$f(x) = \begin{cases} \frac{20,000}{(x+100)^3}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la probabilidad de que un frasco de esta medicina tenga una vida útil de

(a) al menos 200 días;

(b) cualquier duración entre 80 a 120 días.

7. El número total de horas, medidas en unidades de 100 horas, que una familia utiliza una aspiradora en un periodo de un año es una variable aleatoria continua X que tiene la función de densidad

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la probabilidad de que en un periodo de un año, una familia utilice su aspiradora

(a) menos de 120 horas;

(b) entre 50 y 100 horas.

8. Encuentre la distribución de probabilidad de la variable aleatoria W del ejercicio 3, suponga que la moneda está cargada de modo que una cara tenga el doble de probabilidad de ocurrencia que una cruz.

9. La proporción de personas que responden a cierta encuesta enviada por correo es una variable aleatoria continua X que tiene la función de densidad

$$f(x) = \begin{cases} \frac{2(x+2)}{5}, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Muestre que $P(0 < X < 1) = 1$.

(b) Encuentre la probabilidad de que más de $1/4$ pero menos de $1/2$ de las personas contactadas respondan a este tipo de encuesta.

10. Encuentre una fórmula para la distribución de probabilidad de la variable aleatoria X que representa el resultado cuando se lanza una vez un solo dado.

11. Un embarque de siete televisores contiene dos unidades defectuosas. Un hotel hace una compra al azar de tres de los televisores. Si x es el número de unidades defectuosas que compra el hotel, encuentre la distribución de probabilidad de X . Exprese los resultados de forma gráfica como histograma de probabilidad.

12. Una empresa de inversiones ofrece a sus clientes bonos municipales que vencen después de diferentes años. Dado que la distribución acumulada de T , el número de años de vencimiento para un bono que se selecciona al azar es,

$$F(t) = \begin{cases} 0, & t < 1 \\ \frac{1}{4}, & 1 \leq t < 3 \\ \frac{1}{2}, & 3 \leq t < 5 \\ \frac{3}{4}, & 5 \leq t < 7 \\ 1, & t \geq 7. \end{cases}$$

encuentre

(a) $P(T = 5)$;

(b) $P(T > 3)$;

(c) $P(1.4 < T < 6)$.

13. La distribución de probabilidad de X , el número de imperfecciones por 10 metros de una tela sintética en rollos continuos de ancho uniforme, está dada por

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

Construya la distribución acumulada de X .

14. El tiempo de espera, en horas, entre corredores sucesivos detectados por un radar es una variable aleatoria continua con distribución acumulada

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-2x}, & x > 0, \end{cases}$$

Encuentre la probabilidad de esperar menos de 12 minutos entre corredores sucesivos

(a) con el uso de la distribución acumulada de X ;

(b) con el uso de la función de densidad de probabilidad de X .

15. Encuentre la distribución acumulada de la variable aleatoria X que representa el número de unidades defectuosas en el ejercicio 11. Con el uso de $F(x)$, encuentre

(a) $P(X = 1)$;

(b) $P(0 < X \leq 2)$.

16. Construya una gráfica de la distribución acumulada del ejercicio 15.

→ 17. Una variable aleatoria continua X que puede tomar valores entre $x = 1$ y $x = 3$ tiene una función de densidad dada por $f(x) = 1/2$.

(a) Muestre que el área bajo la curva es igual a 1.

(b) Encuentre $P(2 < X < 2.5)$.

(c) Encuentre $P(X \leq 1.6)$.

18. Una variable aleatoria continua X que puede tomar valores entre $x = 2$ y $x = 5$ tiene una función de densidad dada por $f(x) = 2(1+x)/27$. Encuentre

(a) $P(X < 4)$;

(b) $P(3 < X < 4)$.

19. Para la función de densidad del ejercicio 17, encuentre $F(x)$. Utilícela para evaluar $P(2 < X < 2.5)$.

20. Para la función de densidad del ejercicio 18, encuentre $F(x)$, y utilícela para evaluar $P(3 \leq X < 4)$.

→ 21. Considere la función de densidad

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Evalúe k .

(b) Encuentre $F(x)$ y utilícela para evaluar

$$P(0.3 < X < 0.6).$$

22. Se sacan tres cartas sucesivamente de una baraja sin reemplazo. Encuentre la distribución de probabilidad para el número de espadas.

23. Encuentre la distribución acumulada de la variable aleatoria W del ejercicio 8. Con el uso de $F(w)$, encuentre

(a) $P(W > 0)$;

(b) $P(-1 \leq W < 3)$.

24. Encuentre la distribución de probabilidad para el número de CD de jazz cuando se seleccionan cuatro CD al azar de una colección que consiste en cinco CD de jazz, dos de música clásica y tres de rock. Exprese sus resultados mediante una fórmula.
25. Se seleccionan tres monedas sin reemplazo de una caja que contiene cuatro de diez centavos y dos de cinco centavos. Encuentre la distribución de pro-

babilidad para el total T de las tres monedas. Exprese la distribución de probabilidad de forma gráfica como un histograma de probabilidad.

26. Se sacan tres bolas sucesivamente de una caja que contiene cuatro bolas negras y dos verdes, cada bola se reemplaza en la caja antes de sacar la siguiente. Encuentre la distribución de probabilidad para el número de bolas verdes.

3.4 Distribuciones empíricas

En las secciones anteriores el lector fue expuesto a conceptos acerca de distribuciones discretas y continuas. Los métodos de cálculo de probabilidades dependen del conocimiento de la función de masa de probabilidad o función de densidad de probabilidad. La función de probabilidad para el caso discreto y la función de densidad para el caso continuo son formas de caracterizar la distribución de probabilidad para una población o sistema.

Obviamente, el usuario de los métodos estadísticos no puede generar suficiente información o datos experimentales para caracterizar por completo la distribución. Pero a menudo se utilizan conjuntos de datos para aprender acerca de ciertas propiedades de la distribución. Los científicos e ingenieros están acostumbrados a tratar con conjuntos de datos. La importancia de caracterizar o *resumir* la naturaleza de las recopilaciones de datos debe ser obvia. Con frecuencia un resumen de un conjunto de datos por medio de una representación gráfica puede proporcionar más información del sistema del que se tomaron los datos.

En capítulos posteriores se explora con detalle el muestreo de distribuciones y la representación de datos a fin de reforzar la **inferencia estadística** acerca de sistemas científicos. En esta sección sólo introducimos algunas representaciones simples pero a menudo efectivas que complementan el estudio de las distribuciones estadísticas.

Por lo general, en un experimento que incluye una variable aleatoria continua la función de densidad $f(x)$ se desconoce y se supone su forma. Para que la elección de $f(x)$ sea razonablemente válida se necesita en su selección un buen juicio que se base en toda la información disponible. Los datos estadísticos, generados en gran cantidad, pueden ser muy útiles para estudiar el comportamiento de la distribución si se presentan en una representación tabular y gráfica combinada denominada **diagrama de tronco y hojas**.

Para ilustrar la construcción de un diagrama de tronco y hojas, considere los datos de la tabla 3.1, que especifican la "vida" de 40 baterías de coche similares que se registran al décimo de año más cercano. Se garantiza que las baterías duran tres

Tabla 3.1 Vida de las baterías de automóvil

2.2	4.1	3.5	4.5	3.2	3.7	3.0	2.6
3.4	1.6	3.1	3.3	3.8	3.1	4.7	3.7
2.5	4.3	3.4	3.6	2.9	3.3	3.9	3.1
3.3	3.1	3.7	4.4	3.2	4.1	1.9	3.4
4.7	3.8	3.2	2.6	3.9	3.0	4.2	3.5

Tabla 3.2 Diagrama de tronco y hojas de la vida de las baterías

Tronco	Hoja	Frecuencia
1	69	2
2	25669	5
3	001111222333444556777889	25
4	11234577	8

años. Primero, se divide cada observación en dos partes que consisten en un tronco y una hoja de modo que el tronco represente el dígito entero y la hoja correspondiente a la parte decimal del número. En otras palabras, para el número 3.7 el dígito 3 se designa como el tronco y el dígito 7 como la hoja. Los cuatro troncos 1, 2, 3 y 4 para nuestros datos se listan de forma consecutiva en el lado izquierdo de una línea vertical en la tabla 3.2; las hojas se registran a la derecha de la línea en oposición al valor del tronco apropiado. De esta forma la hoja 6 del número 1.6 se registra en el lado opuesto del tronco 1; la hoja 5 del número 2.5 se registra en el lado opuesto del tronco 2; etcétera. El número de hojas que se registran en el lado opuesto de cada tronco se resume en la columna de frecuencia.

El diagrama de tronco y hoja de la tabla 3.2 contiene sólo cuatro troncos y en consecuencia no proporciona una representación adecuada de la distribución. Para remediar este problema, necesitamos aumentar el número de troncos en nuestro diagrama. Una forma simple de llevar a cabo esto es escribir cada valor de tronco dos veces a la izquierda de la línea vertical y registrar las hojas 0, 1, 2, 3 y 4 del lado opuesto del valor del tronco apropiado donde aparece por primera vez; y las hojas 5, 6, 7, 8 y 9 del lado opuesto al mismo valor del tronco donde aparezca por segunda vez. Este diagrama de doble tronco y hojas modificado se ilustra en la tabla 3.3, donde los troncos que corresponden a las hojas 0 a 4 se codifican con el símbolo * y el tronco que corresponde a las hojas 5 a 9 con el símbolo ·.

Se puede lograr un aumento en el número de troncos al escribir cada valor de tronco cinco veces a la izquierda de una línea vertical, donde podemos ahora codificar los troncos con *a* para las hojas 0 y 1, *b* para las hojas 2 y 3, *c* para las hojas 4 y 5, *d* para las hojas 6 y 7, y *e* para las hojas 8 y 9. Para los datos de la tabla 3.1 utilizaremos entonces los troncos 1*d*, 1*e*, 2*a*, 2*b*, 2*c*, 2*d* y 2*e* para construir un diagrama de cinco troncos y hojas.

En cualquier problema dado, debemos decidir los valores apropiados de los troncos. Esta decisión se toma de forma un poco arbitraria, aunque nos guiamos por

Tabla 3.3 Diagrama de doble tronco y hojas para la vida de las baterías

Tronco	Hoja	Frecuencia
1	69	2
2*	2	1
2·	5669	4
3*	001111222333444	15
3·	5567778899	10
4*	13234	5
4·	577	3

el tamaño de nuestra muestra. Por lo general, elegimos entre cinco y 20 troncos. Entre más pequeño es el número de datos disponibles, menor es la cantidad de troncos a seleccionar. Por ejemplo, si los datos consisten en números del 1 al 21 que representan la cantidad de personas en la fila de una cafetería para 40 días hábiles que se seleccionan de manera aleatoria y elegimos un diagrama de doble tronco y hojas, los troncos serían 0^* , 0^- , 1^* , 1^- y 2^* por lo que a la observación más pequeña 1 le corresponde el tronco 0^* y la hoja 1, al número 18 le corresponde el tronco 1^- y la hoja 8, y a la observación mayor 21 le corresponde el tronco 2^* y la hoja 1. Por otro lado, si los datos consisten en cantidades de \$8800 a \$9600 que representan las mejores ventas de 100 automóviles nuevos de cierto distribuidor y elegimos un diagrama de un solo tronco y hojas, los troncos serían 88, 89, 89, ..., y 96 y las hojas contendrían dos dígitos cada una. A un carro que se vende en \$9385 le correspondería un valor de tronco de 93 y la hoja de dos dígitos 85. Las hojas de dígitos múltiples que pertenecen al mismo tronco por lo general se separan por comas en el diagrama de tronco y hojas. Los puntos decimales de los datos por lo general se ignoran cuando todos los dígitos a la derecha del punto decimal representan la hoja. Tal es el caso de las tablas 3.2 y 3.3. Sin embargo, si los datos consisten en números que van de 21.8 a 74.9, debemos elegir los dígitos 2, 3, 4, 5, 6 y 7 como nuestros troncos de modo que un número como el 48.3 tenga un valor de tronco de 4 y una hoja de 8.3.

El diagrama de tronco y hojas representa una forma efectiva de resumir datos. Otra forma es a través del uso de la **distribución de frecuencias** donde los datos, que se agrupan en clases o intervalos diferentes, se pueden construir mediante el conteo de las hojas que pertenecen a cada tronco y al notar que cada tronco define un intervalo de clase. En la tabla 3.3 el tronco 1 con 2 hojas define el intervalo 1.0-1.9 que contiene 2 observaciones; el tronco 2 con 5 hojas define el intervalo 2.0-2.9 que contiene 5 observaciones; el tronco 3 con 25 hojas define el intervalo 3.0-3.9 con 25 observaciones; y el tronco 4 con 8 hojas define el intervalo 4.0-4.9 que contiene 8 observaciones. Para el diagrama de doble tronco y hojas de la tabla 3.1 los troncos definen los siete intervalos de clase 1.5-1.9, 2.0-2.4, 2.5-2.9, 3.0-3.4, 3.5-3.9, 4.0-4.4 y 4.5-4.9 con frecuencias 2, 1, 4, 15, 10, 5 y 3, respectivamente. Al dividir cada frecuencia de clase entre el número total de observaciones, obtenemos la proporción del conjunto de observaciones en cada una de las clases. Una tabla que lista las frecuencias relativas se denomina **distribución de frecuencias relativas**. La distribución de frecuencias relativas para los datos de la tabla 3.1, que muestra los puntos medios de cada intervalo de clase se da en la tabla 3.4.

Tabla 3.4 Distribución de frecuencias relativas de la vida de las baterías

Intervalo de clase	Marca de clase	Frecuencia, f	Frecuencia relativa
1.5-1.9	1.7	2	0.050
2.0-2.4	2.2	1	0.025
2.5-2.9	2.7	4	0.100
3.0-3.4	3.2	15	0.375
3.5-3.9	3.7	10	0.250
4.0-4.4	4.2	5	0.125
4.5-4.9	4.7	3	0.075

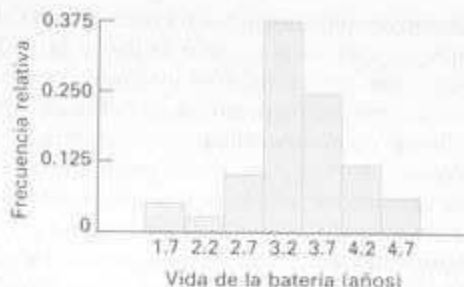


Figura 3.7 Histograma de frecuencias relativas.

La información que proporciona una distribución de frecuencias relativas en forma tabular es más fácil de comprender si se presenta de forma gráfica. Con el uso de los puntos medios de cada intervalo y las frecuencias relativas correspondientes, construimos un **histograma de frecuencias relativas** (figura 3.7) exactamente de la misma manera que construimos el histograma de probabilidad de la sección 3.2.

En la sección 3.2 sugerimos que las alturas de los rectángulos se ajustaran de modo que las áreas representasen probabilidades. Una vez hecho esto, el eje vertical se puede omitir. Si deseamos estimar la distribución de probabilidad $f(x)$ de una variable aleatoria continua X mediante una curva suave como en la figura 3.8, es importante que los rectángulos del histograma de frecuencias relativas se ajuste de modo que el área total sea igual a 1.

La probabilidad de que una batería dure entre 3.45 y 4.45 años cuando se selecciona al azar de la línea de producción infinita de tales baterías está dada por el área sombreada bajo la curva. Nuestra probabilidad estimada que se basa en las vidas registradas de las 40 baterías sería la suma de las áreas que contienen los rectángulos entre 3.45 y 4.45.

Aunque dibujamos una estimación de la forma de $f(x)$ en la figura 3.8, aún no tenemos conocimiento de su fórmula o ecuación y por tanto no podemos encontrar el área sombreada. Como ayuda para comprender el método de estimación de la fórmula para $f(x)$, recordemos algo de geometría analítica elemental. Las parábolas, hipérbolas, círculos, elipses, etcétera, tienen formas bien conocidas de ecuaciones, y en cada caso reconoceríamos sus gráficas. Al pensar de manera inversa, si sólo tenemos sus gráficas pero reconocemos su forma, entonces no es difícil estimar las constantes desconocidas o parámetros y llegar a la ecuación exacta. Por ejemplo, si



Figura 3.8 Estimación de la función de densidad de probabilidad.

la curva parece tener la forma de una parábola, entonces sabemos que tiene una ecuación de la forma $f(x) = ax^2 + bx + c$, donde a , b y c son parámetros que se pueden determinar mediante diversos procedimientos de estimación.

Muchas distribuciones continuas se pueden representar de manera gráfica mediante la curva en forma de campana característica de la figura 3.8. La ecuación de la función de densidad de probabilidad $f(x)$ en este caso se conoce tanto como la de una parábola o la de un círculo y depende sólo de la determinación de dos parámetros. Una vez que estos parámetros se estiman a partir de los datos podemos escribir la ecuación estimada y, después, con el uso de las tablas apropiadas, encontrar cualesquier probabilidades que elijamos.

Se dice que una distribución es **simétrica** si se puede doblar a lo largo de un eje vertical de modo que los lados coincidan. Una distribución que carece de simetría con respecto a un eje vertical se dice que es **asimétrica** o **sesgada**. La distribución que se ilustra en la figura 3.9(a) se dice que está sesgada a la derecha, pues tiene una cola derecha larga y una cola izquierda mucho más corta. En la figura 3.9(b) vemos que la distribución es simétrica, mientras que en la figura 3.9(c) está sesgada a la izquierda.

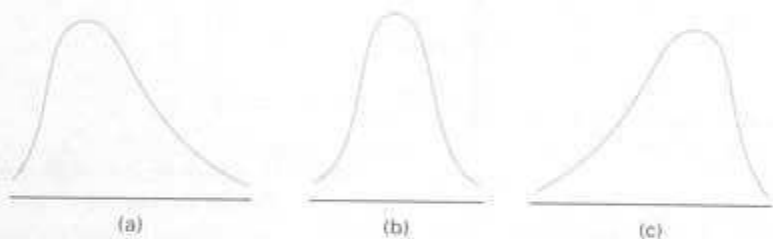


Figura 3.9 Asimetría de los datos.

Al girar un diagrama de tronco y hojas en una dirección contraria a las manecillas del reloj en un ángulo de 90° , observamos que las columnas de hojas que resultan forman una imagen similar a un histograma. En consecuencia, si nuestro propósito principal al ver los datos es determinar la forma general de la función de densidad, rara vez será necesario construir un histograma de frecuencias relativas. En el capítulo 6 consideraremos la mayor parte de las funciones de densidad importantes que se utilizan en ingeniería y en investigaciones científicas.

La distribución acumulada de X , donde X representa la vida de la batería del coche, se puede estimar de forma geométrica mediante el uso de los datos de la tabla 3.4. Para construir dicha gráfica, primero acomodamos nuestros datos como en la tabla 3.5, en una distribución de frecuencias **relativas acumuladas**, y después graficamos la frecuencia relativa acumulada menor que cada límite de clase contra el correspondiente límite de clase como en la figura 3.10. Estimamos $F(x)$ al dibujar una curva suave a través de los puntos.

Los puntos de percentil, decil y cuartil se pueden leer rápidamente en la distribución acumulada. En la figura 3.10 las líneas punteadas indican que el percentil veinticinco o primer cuartil y el séptimo decil son aproximadamente 3.05 y 3.70 años,

Tabla 3.5 Distribución de frecuencias relativas acumuladas de la vida de las baterías

Límites de clase	Frecuencia relativa acumulada
Menor que 1.45	0.000
Menor que 1.95	0.050
Menor que 2.45	0.075
Menor que 2.95	0.175
Menor que 3.45	0.550
Menor que 3.95	0.800
Menor que 4.45	0.925
Menor que 4.95	1.000

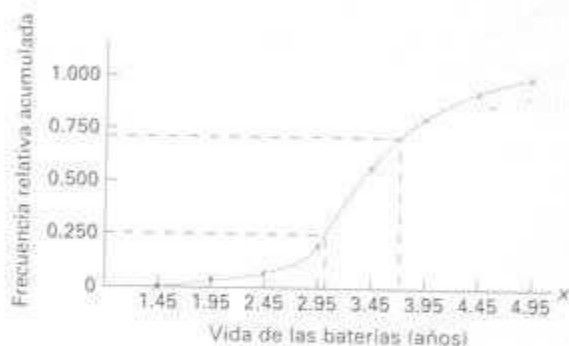


Figura 3.10 Distribución continua acumulada.

respectivamente. Esto significa que 25% o una cuarta parte de todas las baterías de este tipo se espera que duren menos de 3.05 años, mientras que se puede esperar que 70% de tales baterías duren menos de 3.70 años.

Ejercicios

1. Los resultados siguientes representan las calificaciones del examen final de un curso de estadística elemental:

23	60	79	32	57	74	52	70	82	36
80	77	81	95	41	65	92	85	55	76
52	10	64	75	78	25	80	98	81	67
41	71	83	54	64	72	88	62	74	43
60	78	89	76	84	48	84	90	15	79
34	67	17	82	69	74	63	80	85	61

(a) Construya un diagrama de tronco y hojas para las calificaciones del examen en el que los troncos sean 1, 2, 3, ..., 9.

(b) Determine una distribución de frecuencias relativas.

(c) Construya un histograma de frecuencias relativas, dibuje una estimación de la gráfica de $f(x)$, y discuta la asimetría de la distribución.

(d) Construya una distribución de frecuencias relativas acumuladas.

(e) Dibuje una estimación de la gráfica de $F(x)$.

(f) Estime el primer cuartil y el séptimo decil.

2. Los siguientes datos representan la duración de la vida en años, medida al décimo más cercano, de 30 bombas de combustible similares:

2.0	3.0	0.3	3.3	1.3	0.4
0.2	6.0	5.5	6.5	0.2	2.3
1.5	4.0	5.9	1.8	4.7	0.7
4.5	0.3	1.5	0.5	2.5	5.0
1.0	6.0	5.6	6.0	1.2	0.2

(a) Construya un diagrama de tronco y hojas para las vidas de las bombas de combustible, utilice el dígito a la izquierda del punto decimal como el tronco para cada observación.

(b) Determine una distribución de frecuencias relativas.

(c) Construya una distribución de frecuencias relativas acumuladas.

(d) Dibuje una estimación de la gráfica de $F(x)$.

(e) Estime el valor bajo el cual caen dos tercios de los valores.

3. Los siguientes datos representan la duración de la vida, en segundos, de 50 moscas sometidas a un nuevo atomizador en un experimento de laboratorio controlado:

17	20	10	9	23	13	12	19	18	24
12	14	6	9	13	6	7	10	13	7
16	18	8	13	3	32	9	7	10	11
13	7	18	7	10	4	27	19	16	8
7	10	5	14	15	10	9	6	7	15

(a) Construya un diagrama de doble tronco y hojas para la vida de las moscas, utilice los troncos 0^* , 0^- , 1^* , 1^- , 2^* , 2^- y 3^* de modo que los troncos codificados con los símbolos * y - estén asociados, respectivamente, con las hojas 0 a 4 y 5 a 9.

(b) Determine una distribución de frecuencias relativas.

(c) Construya un histograma de frecuencias relativas, dibuje una estimación de la gráfica de $f(x)$, y discuta la asimetría de la distribución.

(d) Construya una distribución de frecuencias relativas acumulada.

(e) Dibuje una estimación de la gráfica de $F(x)$.

(f) Estime el percentil 75.

4. Construya un diagrama de tronco y hojas para los datos de la tabla 3.1, escriba cada tronco cinco veces y después codifique los troncos como se describe en la página 64.

5. El contenido de nicotina, en miligramos, de 40 cigarrillos de cierta marca se registraron como sigue:

1.09	1.92	2.31	1.79	2.28
1.74	1.47	1.97	0.85	1.24
1.58	2.03	1.70	2.17	2.55
2.11	1.86	1.90	1.68	1.51
1.64	0.72	1.69	1.85	1.82
1.79	2.46	1.88	2.08	1.67
1.37	1.93	1.40	1.64	2.09
1.75	1.63	2.37	1.75	1.69

(a) Construya un diagrama de tronco y hojas para los datos en que los troncos son los dígitos a la izquierda del punto decimal, cada uno repetido cinco veces de modo que las hojas de doble dígito de 00 a 19 se asocien con los troncos codificados con la letra a , las hojas 20 a 39 se asocien con los troncos codificados con la letra b , etcétera. De esta forma, un número como el 1.29 tiene un valor de tronco $1b$ y una hoja igual a 29.

(b) Determine una distribución de frecuencias relativas.

3.5 Distribuciones de probabilidad conjunta

Nuestro estudio de variables aleatorias y sus distribuciones de probabilidad en las secciones anteriores se restringe a espacios muestrales unidimensionales, en los que registramos los resultados de un experimento como valores que toma una sola variable aleatoria. Habrá situaciones, sin embargo, donde podemos encontrar que es deseable registrar los resultados simultáneos de diversas variables aleatorias. Por ejemplo, podemos medir la cantidad de precipitado P y volumen V de gas liberado en un experimento químico controlado, que dan lugar a un espacio muestral bidimensional que consiste en los resultados (p, v) , o podemos interesarnos en la dureza H y en la resistencia a la tensión T de cobre estirado en frío que conducen a los resultados (h, t) . En un estudio para determinar la probabilidad de éxito en la facultad, que se basa en los datos de preparatoria, se puede utilizar un espacio muestral tridimensional y registrar para cada individuo su calificación de la prueba de aptitud, su clasificación en la clase de preparatoria y el promedio en puntos al final del primer año en licenciatura.

Si X y Y son dos variables aleatorias discretas, la distribución de probabilidad para sus ocurrencias simultáneas se puede representar mediante una función con valores $f(x, y)$ para cualquier par de valores (x, y) dentro del rango de las variables aleatorias X y Y . Se acostumbra referirse a esta función como la **distribución de probabilidad conjunta** de X y Y .

De aquí, en el caso discreto,

$$f(x, y) = P(X = x, Y = y);$$

es decir, los valores $f(x, y)$ dan la probabilidad de que ocurran los resultados x y y al mismo tiempo. Por ejemplo, si se le va a dar servicio a un televisor y X representa la edad de la unidad al año más próximo y Y representa el número de bulbos defectuosos en el televisor, entonces $f(5, 3)$ es la probabilidad de que el televisor tenga cinco años y necesite tres bulbos nuevos.

Definición 3.8

La función $f(x, y)$ es una **distribución de probabilidad conjunta** o **función de masa de probabilidad** de las variables aleatorias discretas X y Y si

1. $f(x, y) \geq 0$ para toda (x, y) ,
2. $\sum_x \sum_y f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$,

Para cualquier región A en el plano xy , $P[(X, Y) \in A] = \sum_A f(x, y)$.

Ejemplo 3.8 Se seleccionan al azar dos repuestos para un bolígrafo de una caja que contiene tres repuestos azules, dos rojos y 3 verdes. Si X es el número de repuestos azules y Y es el número de repuestos rojos que se seleccionan, encuentre (a) la función de probabilidad conjunta $f(x, y)$ y (b) $P[(X, Y) \in A]$, donde A es la región $\{(x, y) | x + y \leq 1\}$.

SOLUCIÓN

- (a) Los posibles pares de valores (x, y) son $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(0, 2)$ y $(2, 0)$. Ahora bien, $f(0, 1)$, por ejemplo, representa la probabilidad de que se seleccionen un repuesto rojo y uno verde. El número total de formas igualmente probables de seleccionar cualesquiera dos repuestos de los ocho es $\binom{8}{2} = 28$. El número de formas de seleccionar uno rojo de dos repuestos rojos y uno verde es $\binom{2}{1} \binom{3}{1} = 6$. De aquí $f(0, 1) = 6/28 = 3/14$. Cálculos similares dan las probabilidades para los otros casos, que se presentan en la tabla 3.6. Nótese que las probabilidades suman 1. En el capítulo 4 quedará claro que la distribución de probabilidad conjunta de la tabla 3.6 se puede representar mediante la fórmula

Tabla 3.6 Distribución de probabilidad conjunta para el ejemplo 3.8

$f(x, y)$	x			Totales por renglón
	0	1	2	
0	$\frac{3}{28}$	$\frac{6}{28}$	$\frac{3}{28}$	$\frac{12}{28}$
$y = 1$	$\frac{3}{14}$	$\frac{3}{14}$		$\frac{3}{7}$
2	$\frac{1}{28}$			$\frac{1}{28}$
Totales por columna	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

$$f(x, y) = \frac{\binom{3}{x} \binom{2}{y} \binom{3}{2-x-y}}{\binom{8}{2}}$$

para $x = 0, 1, 2$; $y = 0, 1, 2$; $0 \leq x + y \leq 2$.

$$\begin{aligned} \text{(b) } P[(X, Y) \in A] &= P(X + Y \leq 1) \\ &= f(0, 0) + f(0, 1) + f(1, 0) \\ &= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} \\ &= \frac{9}{14} \end{aligned}$$

Cuando X y Y son variables aleatorias continuas, la **función de densidad conjunta** $f(x, y)$ es una superficie sobre el plano xy , y $P[(X, Y) \in A]$, donde A es cualquier región en el plano xy , es igual al volumen del cilindro recto limitado por la base A y la superficie. ■

Definición 3.9

La función $f(x, y)$ es una **función de densidad conjunta** de las variables aleatorias continuas X y Y si

1. $f(x, y) \geq 0$ para toda (x, y) ,
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$,
3. $P[(X, Y) \in A] = \iint_A f(x, y) dx dy$.

para cualquier región A en el plano xy .

Ejemplo 3.9 Una fábrica de dulces distribuye cajas de chocolates con un surtido de cremas, chicolosos y nueces cubiertas con chocolate claro y oscuro. Para una caja seleccionada al azar, sean X y Y , respectivamente, las proporciones de chocolates claro y oscuro que son cremas y suponga que la función de densidad conjunta es

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y), & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Verifique la condición 2 de la definición 3.9.

(b) Encuentre $P[(X, Y) \in A]$, donde A está en la región $\{(x, y) | 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$.

SOLUCIÓN

$$\begin{aligned} \text{(a)} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy &= \int_0^1 \int_0^1 \frac{2}{5}(2x + 3y) \, dx \, dy \\ &= \int_0^1 \left. \frac{2x^2}{5} + \frac{6xy}{5} \right|_{x=0}^{x=1} dy \\ &= \int_0^1 \left(\frac{2}{5} + \frac{6y}{5} \right) dy = \left. \frac{2y}{5} + \frac{3y^2}{5} \right|_0^1 \\ &= \frac{2}{5} + \frac{3}{5} = 1. \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad P[(X, Y) \in A] &= P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) \\ &= \int_{1/4}^{1/2} \int_0^{1/2} \frac{2}{5}(2x + 3y) \, dx \, dy \\ &= \int_{1/4}^{1/2} \left. \frac{2x^2}{5} + \frac{6xy}{5} \right|_{x=0}^{x=1/2} dy \\ &= \int_{1/4}^{1/2} \left(\frac{1}{10} + \frac{3y}{5} \right) dy = \left. \frac{y}{10} + \frac{3y^2}{10} \right|_{1/4}^{1/2} \\ &= \frac{1}{10} \left[\left(\frac{1}{2} + \frac{3}{4} \right) - \left(\frac{1}{4} + \frac{3}{16} \right) \right] = \frac{13}{160}. \end{aligned}$$

Dada la distribución de probabilidad conjunta $f(x, y)$ de las variables aleatorias discretas X y Y , la distribución de probabilidad $g(x)$ de X sola se obtiene al sumar $f(x, y)$ sobre los valores de Y . De manera similar, la distribución de probabilidad $h(y)$ de Y sola se obtiene al sumar $f(x, y)$ sobre los valores de X . Definimos $g(x)$ y $h(y)$ como **distribuciones marginales** de X y Y , respectivamente. Cuando X y Y son variables aleatorias continuas, las sumatorias se reemplazan por integrales. Ahora podemos establecer la definición general siguiente. ■

Definición 3.10

Las **distribuciones marginales** de X sola y Y sola son

$$g(x) = \sum_y f(x, y) \quad \text{y} \quad h(y) = \sum_x f(x, y)$$

para el caso discreto, y

$$g(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \text{y} \quad h(y) = \int_{-\infty}^{\infty} f(x, y) \, dx$$

para el caso continuo.

El término *marginal* se utiliza aquí debido a que, en el caso discreto, los valores de $g(x)$ y $h(y)$ son exactamente los totales marginales de las columnas y renglones respectivos cuando los valores de $f(x, y)$ se muestran en una tabla rectangular.

Ejemplo 3.10 Muestre que los totales de columnas y renglones de la tabla 3.6 dan las distribuciones marginales de X sola y Y sola

SOLUCIÓN

Para la variable aleatoria X , vemos que

$$\begin{aligned} P(X=0) &= g(0) = \sum_{y=0}^2 f(0, y) = f(0, 0) + f(0, 1) + f(0, 2) \\ &= \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14} \end{aligned}$$

$$\begin{aligned} P(X=1) &= g(1) = \sum_{y=0}^2 f(1, y) = f(1, 0) + f(1, 1) + f(1, 2) \\ &= \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28} \end{aligned}$$

y

$$\begin{aligned} P(X=2) &= g(2) = \sum_{y=0}^2 f(2, y) = f(2, 0) + f(2, 1) + f(2, 2) \\ &= \frac{3}{28} + 0 + 0 = \frac{3}{28} \end{aligned}$$

que son exactamente los totales de columna de la tabla 3.6. De manera similar podemos mostrar que los valores de $h(y)$ están dados por los totales de los renglones. En forma tabular, estas distribuciones marginales se pueden escribir como sigue: ■

x	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Ejemplo 3.11 Encuentre $g(x)$ y $h(y)$ para la función de densidad conjunta del ejemplo 3.9.

SOLUCIÓN

Por definición,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 \frac{2}{5}(2x + 3y) dy = \frac{4xy}{5} + \frac{6y^2}{10} \Big|_{y=0}^{y=1} = \frac{4x + 3}{5}$$

para $0 \leq x \leq 1$ y $g(x) = 0$ en cualquier otro caso. De manera similar,

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^1 \frac{2}{5}(2x + 3y) dx = \frac{2(1 + 3y)}{5}$$

para $0 \leq y \leq 1$ y $h(y) = 0$ en cualquier otro caso.

El hecho de que las distribuciones marginales $g(x)$ y $h(y)$ sean en realidad las distribuciones de probabilidad de las variables individuales X y Y solas se puede verificar al mostrar que se satisfacen las condiciones de la definición 3.4 o de la definición 3.6. Por ejemplo, en el caso continuo

$$\int_{-\infty}^{\infty} g(x) dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1$$

y

$$\begin{aligned} P(a < X < b) &= P(a < X < b, -\infty < Y < \infty) \\ &= \int_a^b \int_{-\infty}^{\infty} f(x, y) dy dx = \int_a^b g(x) dx. \end{aligned}$$

En la sección 3.1 establecimos que el valor x de la variable aleatoria X representa un evento que es un subconjunto del espacio muestral. Si utilizamos la definición de probabilidad condicional que se establece en el capítulo 2,

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0,$$

donde A y B son ahora los eventos definidos por $X = x$ y $Y = y$, respectivamente, entonces

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x, y)}{g(x)}, \quad g(x) > 0,$$

donde X y Y son variables aleatorias discretas. ■

No es difícil mostrar que la función $f(x, y)/g(x)$, que es estrictamente una función de y con x fija, satisface todas las condiciones de una distribución de probabilidad. Esto también es cierto cuando $f(x, y)$ y $g(x)$ son la densidad conjunta y la distribución marginal de variables aleatorias continuas. Como resultado es en extremo importante que hagamos uso del tipo especial de distribución de la forma $f(x, y)/g(x)$ a fin de ser capaces de calcular de manera efectiva probabilidades condicionales. Este tipo de distribución se llama **distribución de probabilidad condicional**; la definición condicional es la que sigue.

Definición 3.11

Sean X y Y dos variables aleatorias, discretas o continuas. La **distribución condicional** de la variable aleatoria Y , dado que $X = x$, es

$$f(y|x) = \frac{f(x, y)}{g(x)}, \quad g(x) > 0.$$

De manera similar, la distribución condicional de la variable aleatoria X , dado que $Y = y$, es

$$f(x|y) = \frac{f(x, y)}{h(y)}, \quad h(y) > 0.$$

Si se desea encontrar la probabilidad de que la variable aleatoria discreta X caiga entre a y b cuando se sabe que la variable discreta $Y = y$, evaluamos

$$P(a < X < b | Y = y) = \sum_x f(x|y),$$

donde la sumatoria se extiende a todos los valores de X entre a y b . Cuando X y Y son continuas, evaluamos

$$P(a < X < b | Y = y) = \int_a^b f(x|y) dx.$$

Ejemplo 3.12 Con referencia al ejemplo 3.8, encuentre la distribución condicional de X , dado que $Y = 1$, y utilícela para determinar $P(X = 0 | Y = 1)$.

SOLUCIÓN

Necesitamos encontrar $f(x|y)$, donde $y = 1$. Primero, encontramos que

$$h(1) = \sum_{x=0}^2 f(x, 1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Ahora bien,

$$f(x|1) = \frac{f(x, 1)}{h(1)} = \frac{7}{3} f(x, 1), \quad x = 0, 1, 2.$$

Por tanto,

$$f(0|1) = \frac{7}{3} f(0, 1) = \left(\frac{7}{3}\right) \left(\frac{3}{14}\right) = \frac{1}{2},$$

$$f(1|1) = \frac{7}{3} f(1, 1) = \left(\frac{7}{3}\right) \left(\frac{3}{14}\right) = \frac{1}{2},$$

$$f(2|1) = \frac{7}{3} f(2, 1) = \left(\frac{7}{3}\right) (0) = 0,$$

y la distribución condicional de X , dado que $Y = 1$, es

x	0	1	2
$f(x 1)$	$\frac{1}{2}$	$\frac{1}{2}$	0

Finalmente,

$$P(X = 0 | Y = 1) = f(0|1) = \frac{1}{2}.$$

Por tanto, si se sabe que uno de los dos repuestos seleccionados es rojo, tenemos una probabilidad igual a $1/2$ de que el otro repuesto no sea azul. ■

Ejemplo 3.13 La densidad conjunta para las variables aleatorias (X, Y) , donde X es el cambio de temperatura unitario y Y es la proporción de desplazamiento espectral que produce cierta partícula atómica es

$$f(x, y) = \begin{cases} 10xy^2, & 0 < x < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Encuentre las densidades marginales $g(x)$, $h(y)$ y la densidad condicional $f(y|x)$.
- (b) Encuentre la probabilidad de que el espectro se desplace más de la mitad de las observaciones totales, dado que la temperatura aumenta a 0.25 de unidad.

SOLUCIÓN

(a) Por definición,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_x^1 10xy^2 dy = \frac{10}{3} xy^3 \Big|_{y=x}^{y=1} = \frac{10}{3} x(1 - x^3), \quad 0 < x < 1$$

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_0^y 10xy^2 dx = 5x^2 y^2 \Big|_{x=0}^{x=y} = 5y^4, \quad 0 < y < 1$$

Ahora bien

$$f(y|x) = \frac{f(x, y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1-x^3)} = \frac{3y^2}{(1-x^3)}, \quad 0 < x < y < 1.$$

(b) Por tanto,

$$P\left(Y > \frac{1}{2} \mid X = 0.25\right) = \int_{1/2}^1 f(y|x=0.25) dy = \int_{1/2}^1 \frac{3y^2}{(1-0.25^3)} dy = \frac{8}{9}. \quad \blacksquare$$

Ejemplo 3.14 Dada la función de densidad conjunta

$$f(x, y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

encuentre $g(x)$, $h(y)$, $f(x|y)$, y evalúe $P(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3})$.

SOLUCIÓN

Por definición,

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 \frac{x(1+3y^2)}{4} dy \\ &= \frac{xy}{4} + \frac{xy^3}{4} \Big|_{y=0}^{y=1} = \frac{x}{2}, \quad 0 < x < 2 \end{aligned}$$

y

$$\begin{aligned} h(y) &= \int_{-\infty}^{\infty} f(x, y) dx = \int_0^2 \frac{x(1+3y^2)}{4} dx \\ &= \frac{x^2}{8} + \frac{3x^2 y^2}{8} \Big|_{x=0}^{x=2} = \frac{1+3y^2}{2}, \quad 0 < y < 1. \end{aligned}$$

Por tanto,

$$f(x|y) = \frac{f(x, y)}{h(y)} = \frac{x(1 + 3y^2)/4}{(1 + 3y^2)/2} = \frac{x}{2}, \quad 0 < x < 2$$

y

$$P\left(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3}\right) = \int_{1/4}^{1/2} \frac{x}{2} dx = \frac{3}{64}. \quad \blacksquare$$

Independencia estadística

Si $f(x|y)$ no depende de y , como es el caso en el ejemplo 3.14, entonces $f(x|y) = g(x)$ y $f(x, y) = g(x)h(y)$. La demostración se tiene al sustituir

$$f(x, y) = f(x|y)h(y)$$

en la distribución marginal de X . Es decir,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-\infty}^{\infty} f(x|y)h(y) dy.$$

Si $f(x|y)$ no depende de y , podemos escribir

$$g(x) = f(x|y) \int_{-\infty}^{\infty} h(y) dy.$$

Ahora bien

$$\int_{-\infty}^{\infty} h(y) dy = 1,$$

pues $h(y)$ es la función de densidad de probabilidad de Y . Por tanto,

$$g(x) = f(x|y) \quad \text{y entonces} \quad f(x, y) = g(x)h(y).$$

Debe tener sentido para el lector que si $f(x|y)$ no depende de y , entonces por supuesto el resultado de la variable aleatoria Y no tiene impacto en el resultado de la variable aleatoria X . En otras palabras, decimos que X y Y son variables aleatorias independientes. Ofrecemos ahora la siguiente definición formal de independencia estadística.

Definición 3.12

Sean X y Y dos variables aleatorias, discretas o continuas, con distribución de probabilidad conjunta $f(x, y)$ y distribuciones marginales $g(x)$ y $h(y)$, respectivamente. Se dice que las variables aleatorias X y Y son **estadísticamente independientes** si y sólo si

$$f(x, y) = g(x)h(y)$$

para toda (x, y) dentro de sus rangos.

Las variables aleatorias continuas del ejemplo 3.14 son estadísticamente independientes, pues el producto de las dos distribuciones marginales da la función de densidad conjunta. Obviamente este no es el caso, sin embargo, para las variables continuas del ejemplo 3.13. La comprobación de la independencia estadística de variables aleatorias discretas requiere una investigación más profunda, pues es posible que el producto de las distribuciones marginales sea igual a la distribución de probabilidad conjunta para algunas de las combinaciones de (x, y) pero no para todas. Si puede encontrar algún punto (x, y) para el que $f(x, y)$ se define de modo que $f(x, y) \neq g(x)h(y)$, las variables discretas X y Y no son estadísticamente independientes.

Ejemplo 3.15 Muestre que las variables aleatorias del ejemplo 3.8 no son estadísticamente independientes.

SOLUCIÓN

Consideremos el punto $(0, 1)$. De la tabla 3.6 encontramos que las tres probabilidades $f(0, 1)$, $g(0)$ y $h(1)$ son

$$f(0, 1) = \frac{3}{14},$$

$$g(0) = \sum_{y=0}^2 f(0, y) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$h(1) = \sum_{x=0}^2 f(x, 1) = \frac{3}{14} + \frac{3}{14} + 0 = \frac{3}{7}.$$

Claramente,

$$f(0, 1) \neq g(0)h(1),$$

y por tanto X y Y no son estadísticamente independientes. ■

Todas las definiciones anteriores respecto a dos variables aleatorias se pueden generalizar al caso de n variables aleatorias. Sea $f(x_1, x_2, \dots, x_n)$ la función de probabilidad conjunta de las variables aleatorias x_1, x_2, \dots, x_n . La distribución marginal de X_1 , por ejemplo, es

$$g(x_1) = \sum_{x_2} \cdots \sum_{x_n} f(x_1, x_2, \dots, x_n)$$

para el caso discreto, y

$$g(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_2 dx_3 \cdots dx_n$$

para el caso continuo. Ahora podemos obtener **distribuciones marginales conjuntas** como $\phi(x_1, x_2)$, donde

$$\phi(x_1, x_2) = \begin{cases} \sum_{x_3} \cdots \sum_{x_n} f(x_1, x_2, \dots, x_n) & \text{(caso discreto)} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_3 dx_4 \cdots dx_n & \text{(caso continuo).} \end{cases}$$

Se pueden considerar numerosas distribuciones condicionales. Por ejemplo, la **distribución condicional conjunta** de X_1, X_2 y X_3 , dado que $X_4 = x_4, X_5 = x_5, \dots, X_n = x_n$, se escribe como

$$f(x_1, x_2, x_3 | x_4, x_5, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n)}{g(x_4, x_5, \dots, x_n)},$$

donde $g(x_4, x_5, \dots, x_n)$ es la distribución marginal conjunta de las variables aleatorias X_4, X_5, \dots, X_n .

Una generalización de la definición 3.12 conduce a la siguiente definición para la independencia estadística mutua de las variables X_1, X_2, \dots, X_n .

Definición 3.13

Sean X_1, X_2, \dots, X_n n variables aleatorias, discretas o continuas, con distribución de probabilidad conjunta $f(x_1, x_2, \dots, x_n)$ y distribuciones marginales $f_1(x_1), f_2(x_2), \dots, f_n(x_n)$, respectivamente. Se dice que las variables aleatorias X_1, X_2, \dots, X_n son **estadísticamente independientes** mutuamente si y sólo si

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2)\cdots f_n(x_n)$$

para toda (x_1, x_2, \dots, x_n) dentro de sus rangos.

Ejemplo 3.16 Suponga que el tiempo de vida, en años, de cierto producto alimenticio perecedero empacado en cajas de cartón es una variable aleatoria cuya función de densidad de probabilidad está dada por

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Sean X_1, X_2 , y X_3 los tiempos de vida para tres de estas cajas que se seleccionan de forma independiente y encuentre $P(X_1 < 2, 1 < X_2 < 3, X_3 > 2)$.

SOLUCIÓN

Como las cajas se seleccionan de forma independiente, podemos suponer que las variables aleatorias X_1, X_2 y X_3 son estadísticamente independientes y que tienen la densidad de probabilidad conjunta

$$f(x_1, x_2, x_3) = f(x_1)f(x_2)f(x_3) = e^{-x_1}e^{-x_2}e^{-x_3} = e^{-x_1-x_2-x_3}$$

para $x_1 > 0, x_2 > 0, x_3 > 0$, y $f(x_1, x_2, x_3) = 0$ en cualquier otro caso. De aquí

$$\begin{aligned} P(X_1 < 2, 1 < X_2 < 3, X_3 > 2) &= \int_2^{\infty} \int_1^3 \int_0^2 e^{-x_1-x_2-x_3} dx_1 dx_2 dx_3 \\ &= (1 - e^{-2})(e^{-1} - e^{-3})e^{-2} = 0.0376. \end{aligned}$$

Ejercicios

1. Determine el valor de c tal que las siguientes funciones representen distribuciones de probabilidad conjunta de las variables aleatorias X y Y :

(a) $f(x, y) = cxy$, para $x = 1, 2, 3$; $y = 1, 2, 3$.

(b) $f(x, y) = c|x - y|$, para $x = -2, 0, 2$; $y = -2, 3$.

2. Si la distribución de probabilidad conjunta de X y Y está dada por

$$f(x, y) = \frac{(x+y)}{30}, \quad \text{para } x = 0, 1, 2, 3;$$

$$y = 0, 1, 2.$$

encuentre

- (a) $P(X \leq 2, Y = 1)$; (b) $P(X > 2, Y \leq 1)$;
 (c) $P(X > Y)$; (d) $P(X + Y = 4)$.

3. De un costal de frutas que contiene tres naranjas, dos manzanas y tres plátanos se selecciona una muestra aleatoria de cuatro frutas. Si X es el número de naranjas y Y el de manzanas en la muestra, encuentre

- (a) la distribución de probabilidad conjunta de X y Y ;
 (b) $P\{(X, Y) \in A\}$, donde A es la región dada por $\{(x, y) | x + y \leq 2\}$.

4. Una vinatería opera instalaciones para atención en el automóvil y para atender a quien llega caminando. En un día seleccionado al azar, sean X y Y , respectivamente, las proporciones del tiempo que se utiliza cada instalación, y suponga que la función de densidad conjunta de estas variables aleatorias es:

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y), & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{en cualquier otra parte.} \end{cases}$$

- (a) Encuentre la densidad marginal de X .
 (b) Encuentre la densidad marginal de Y .
 (c) Encuentre la probabilidad de que las instalaciones para atención en carro estén ocupadas menos de la mitad del tiempo.

5. Una compañía dulcera distribuye cajas de chocolates con un surtido de cremas, chicolosos y envinados. Suponga que el peso de cada caja es un kilogramo, pero que los pesos individuales de las cremas, chicolosos y envinados varían de una caja a otra. Para una caja seleccionada al azar, sean X y Y los pesos de las cremas y los chicolosos, respectivamente, y suponga que la función de densidad conjunta de estas variables es

$$f(x, y) = \begin{cases} 24xy, & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ & x + y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Encuentre la probabilidad de que en una caja dada los envinados representen más de $1/2$ del peso.
 (b) Encuentre la densidad marginal para el peso de las cremas.
 (c) Encuentre la probabilidad de que el peso de los chicolosos en una caja sea menor de $1/8$ de kilogramo si se sabe que las cremas constituyen $3/4$ del peso.

6. Sean X y Y la duración de la vida, en años, de dos componentes en un sistema electrónico. Si la función de densidad conjunta de estas variables es

$$f(x, y) = \begin{cases} e^{-(x+y)}, & x > 0, y > 0 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

encuentre $P(0 < X < 1 | Y = 2)$.

7. Sea X el tiempo de reacción, en segundos, a cierto estimulante y Y la temperatura ($^{\circ}\text{F}$) a la que cierta reacción comienza a suceder. Suponga que dos variables aleatorias X y Y tienen la densidad conjunta

$$f(x, y) = \begin{cases} 4xy, & 0 < x < 1; 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre

- (a) $P(0 \leq X \leq \frac{1}{2} \text{ y } \frac{1}{4} \leq Y \leq \frac{1}{2})$;
 (b) $P(X < Y)$.

8. Se supone que cada rueda trasera de un aeroplano experimental se llena a una presión de 40 psi. Sea X la presión real del aire para la rueda derecha y Y la presión real de aire de la rueda izquierda. Suponga que X y Y son variables aleatorias con la densidad conjunta

$$f(x, y) = \begin{cases} k(x^2 + y^2), & 30 \leq x < 50; 30 \leq y < 50 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Encuentre k .
 (b) Encuentre $P(30 \leq X \leq 40 \text{ y } 40 \leq Y < 50)$.
 (c) Encuentre la probabilidad de que ambas ruedas no estén completamente llenas.

9. Sea X el diámetro de un cable eléctrico blindado y Y el diámetro del molde cerámico que hace el cable. X y Y tienen una escala tal que están entre 0 y 1. Suponga que X y Y tienen la densidad conjunta

$$f(x, y) = \begin{cases} \frac{1}{y}, & 0 < x < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre $P(X + Y > 1/2)$.

10. Con referencia al ejercicio 2, encuentre
 (a) la distribución marginal de X ;
 (b) la distribución marginal de Y .

11. La cantidad de queroseno, en miles de litros, en un tanque al principio de cualquier día es una cantidad aleatoria Y de la que una cantidad aleatoria X se vende durante el día. Suponga que el tanque no se reabastece durante el día por lo que $x \leq y$, y supon-

ga que la función de densidad conjunta de estas variables es

$$f(x, y) = \begin{cases} 2, & 0 < x < y, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Determine si X y Y son independientes.
 (b) Encuentre $P(1/4 < X < 1/2 | Y = 3/4)$.

12. Refiérase al ejercicio 3, encuentre

- (a) $f(y|2)$ para todos los valores de y ;
 (b) $P(y = 0 | X = 2)$.

13. Sea X el número de veces que falla cierta máquina de control numérico: 1, 2 o 3 veces en un día dado. Sea Y el número de veces que se llama a un técnico para una emergencia. Su distribución de probabilidad conjunta está dada como.

$f(x, y)$	x	1	2	3
	1	0.05	0.05	0.1
y	2	0.05	0.1	0.35
	3	0	0.2	0.1

- (a) Evalúe la distribución marginal de X .
 (b) Evalúe la distribución marginal de Y .
 (c) Encuentre $P(Y = 3 | X = 2)$.

14. Suponga que X y Y tienen la siguiente distribución de probabilidad conjunta:

$f(x, y)$	x	2	4
	1	0.10	0.15
y	3	0.20	0.30
	5	0.10	0.15

- (a) Encuentre la distribución marginal de X .
 (b) Encuentre la distribución marginal de Y .

15. Considere un experimento que consiste en dos lanzamientos de un dado balanceado. Si X es el número de cuatros y Y es el número de cincos que se obtienen en los dos lanzamientos del dado, encuentre

- (a) la distribución de probabilidad conjunta de X y Y ;
 (b) $P\{(X, Y) \in A\}$, donde A es la región $\{(x, y) | 2x + y < 3\}$.

16. Sea X el número de caras y Y el número de caras menos el número de cruces cuando se lanzan tres monedas. Encuentre la distribución de probabilidad conjunta de X y Y .

17. Se sacan tres cartas sin reemplazo de las 12 cartas mayores (sotas, reinas y reyes) de una baraja ordinaria de 52 cartas. Sea X el número de reyes que se seleccionan y Y el número de sotas. Encuentre

- (a) la distribución de probabilidad conjunta de X y Y ;
 (b) $P\{(X, Y) \in A\}$, donde A es la región dada por $\{(x, y) | x + y \geq 2\}$.

18. Se lanza dos veces una moneda. Sea Z el número de caras en el primer lanzamiento y W el número total de caras en los dos lanzamientos. Si la moneda no está balanceada y una cara tiene una probabilidad de ocurrencia del 40%, encuentre

- (a) la distribución de probabilidad conjunta de W y Z ;
 (b) la distribución marginal de W ;
 (c) la distribución marginal de Z ;
 (d) la probabilidad de que ocurra al menos una cara.

19. Dada la función de densidad conjunta

$$f(x, y) = \begin{cases} \frac{6-x-y}{8}, & 0 < x < 2, 2 < y < 4 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

encuentre $P(1 < Y < 3 | X = 2)$.

20. Determine si las dos variables aleatorias del ejercicio 13 son dependientes o independientes.

21. Determine si las dos variables aleatorias del ejercicio 14 son dependientes o independientes.

22. La función de densidad conjunta de las variables aleatorias X y Y es

$$f(x, y) = \begin{cases} 6x, & 0 < x < 1, 0 < y < 1 - x \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Muestre que X y Y no son independientes.
 (b) Encuentre $P(X > 0.3 | Y = 0.5)$.

23. Si X, Y y Z tienen la función de densidad de probabilidad conjunta

$$f(x, y, z) = \begin{cases} kxy^2z, & 0 < x < 1, 0 < y < 1, \\ & 0 < z < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Encuentre k .
 (b) Encuentre $P(X < 1/4, Y > 1/2, 1 < Z < 2)$.

24. Determine si las dos variables aleatorias del ejercicio 7 son dependientes o independientes.

25. Determine si las dos variables aleatorias del ejercicio 8 son dependientes o independientes.

26. La función de densidad de probabilidad conjunta de las variables aleatorias X , Y y Z es

$$f(x, y, z) = \begin{cases} 4xyz^2, & 0 < x < 1, 0 < y < 1, \\ 9, & 0 < z < 3 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Ejercicios de repaso

1. Una compañía tabacalera produce mezclas de tabaco, cada mezcla contiene varias partes de tabaco turco, tabaco de la región y otros. Las proporciones de turco y de la región en una mezcla son variables aleatorias con función de densidad conjunta (X = turco y Y = de la región)

$$f(x, y) = \begin{cases} 24xy, & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ x + y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre la probabilidad de que en una caja dada el tabaco turco represente más de la mitad de la mezcla.

(b) Encuentre la función de densidad marginal para la proporción de tabaco de la región.

(c) Encuentre la probabilidad de que la parte de tabaco turco sea menos de $1/8$ si se sabe que la mezcla contiene $3/4$ de tabaco de la región.

2. Una compañía de seguros ofrece a sus asegurados varias opciones diferentes de pago de la prima. Para un asegurado seleccionado al azar, sea X el número de meses entre pagos sucesivos. La función de distribución acumulada de X es

$$F(x) = \begin{cases} 0 & \text{si } x < 1 \\ 0.4 & \text{si } 1 \leq x < 3 \\ 0.6 & \text{si } 3 \leq x < 5 \\ 0.8 & \text{si } 5 \leq x < 7 \\ 1.0 & \text{si } x \geq 7. \end{cases}$$

(a) ¿Cuál es la función masa de probabilidad de X ?

(b) Calcule $P(4 < X \leq 7)$.

3. Dos componentes electrónicos de un sistema de misiles trabajan en armonía para el éxito de todo el sistema. Sean X y Y la vida en horas de los dos sistemas. La densidad conjunta de X y Y es

$$f(x, y) = \begin{cases} ye^{-y(1+x)}, & x \geq 0, y \geq 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre

(a) la función de densidad marginal conjunta de Y y Z ,

(b) la densidad marginal de Y ;

(c) $P(\frac{1}{2} < X < \frac{3}{4}, Y > \frac{1}{2}, 1 < Z < 2)$;

(d) $P(0 < X < \frac{1}{2} | Y = \frac{1}{2}, Z = 2)$.

(a) Determine las funciones de densidad marginal para ambas variables aleatorias.

(b) ¿Cuál es la probabilidad de que ambos componentes duren más de dos horas?

4. Una instalación de servicio opera con dos líneas. En un día seleccionado al azar, sea X la proporción de tiempo que la primera línea está en uso en tanto que Y es la proporción de tiempo en que la segunda línea está en uso. Suponga que la función de densidad de probabilidad conjunta para (X, Y) es

$$f(x, y) = \begin{cases} \frac{2}{3}(x^2 + y^2), & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Calcule la probabilidad de que ninguna línea esté ocupada más de la mitad del tiempo.

(b) Encuentre la probabilidad de que la primera línea esté ocupada más del 75% del tiempo.

5. Sea el número de llamadas telefónicas que recibe un conmutador durante un intervalo de cinco minutos una variable aleatoria X con función de probabilidad

$$f(x) = \frac{e^{-2} 2^x}{x!} \quad \text{para } x = 0, 1, 2, \dots$$

(a) Determine la probabilidad de que X sea igual a 0, 1, 2, 3, 4, 5 y 6.

(b) Grafique la función de masa de probabilidad para estos valores de x .

(c) Determine la función de distribución acumulada para estos valores de X .

6. Considere las variables aleatorias X y Y con función de densidad conjunta

$$f(x, y) = \begin{cases} x + y, & 0 \leq x, y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Determine las distribuciones marginales de X y Y .

(b) Encuentre $P(X > 0.5, Y > 0.5)$.

7. Un proceso industrial produce artículos que se pueden clasificar como defectuosos o no defectuosos. La probabilidad de que un artículo esté defectuoso es 0.1. Se lleva a cabo un experimento en el cual se sacan cinco artículos al azar del proceso. Sea la variable aleatoria X el número de defectuosos en esta muestra de cinco. ¿Cuál es la función de masa de probabilidad de X ?

8. Considere la siguiente función de densidad de probabilidad conjunta de las variables aleatorias X y Y :

$$f(x, y) = \begin{cases} \frac{3x - y}{9}, & 1 < x < 3, 1 < y < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Encuentre las distribuciones marginales de X y Y .
 (b) ¿Son independientes X y Y ?
 (c) Encuentre $P(X > 2)$.

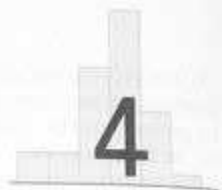
9. La duración en horas de un componente eléctrico es una variable aleatoria con función de distribución acumulada

$$F(x) = \begin{cases} 1 - e^{-x/50}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

- (a) Determine la función de densidad de probabilidad.
 (b) Determine la probabilidad de que la duración de tal componente exceda 70 horas.

10. Una fábrica produce pantalones. Un grupo de 10 trabajadores los “revisan”. Los trabajadores inspeccionan pantalones que se toman al azar de la línea de producción. A cada inspector se le asigna un número entre el 1 y el 10. Un comprador selecciona un pantalón para comprarlo. Sea la variable aleatoria X el número del inspector.

- (a) Determine una función de masa de probabilidad razonable para X .
 (b) Grafique la función de distribución acumulada para X .



Esperanza matemática

4.1 Media de una variable aleatoria

Si se lanzan 16 veces dos monedas y X es el número de caras que ocurre por lanzamiento, entonces los valores de X pueden ser 0, 1 y 2. Suponga que en el experimento salen cero caras, una cara y dos caras un total de cuatro, siete y cinco veces, respectivamente. El número promedio de caras por lanzamiento de las dos monedas es entonces

$$\frac{(0)(4) + (1)(7) + (2)(5)}{16} = 1.06.$$

Este es un valor promedio y no es necesariamente un posible resultado del experimento. Por ejemplo, el ingreso mensual promedio de un vendedor probablemente no es igual a alguno de sus cheques de pago mensual.

Reestructuremos ahora nuestro cálculo del número promedio de caras de modo que tengamos la siguiente forma equivalente:

$$(0)\left(\frac{4}{16}\right) + (1)\left(\frac{7}{16}\right) + (2)\left(\frac{5}{16}\right) = 1.06.$$

Los números $4/16$, $7/16$, y $5/16$ son las fracciones de los lanzamientos totales que tienen como resultado cero, una y dos caras, respectivamente. Estas fracciones también son la frecuencias relativas de los diferentes valores de X en nuestro experimento. En efecto, entonces, podemos calcular la media o promedio de un conjunto de datos mediante el conocimiento de los distintos valores que ocurren y sus frecuencias relativas, sin un conocimiento del número total de observaciones en nuestro conjunto de datos. Por tanto, si $4/16$ o $1/4$ de los lanzamientos tiene como

resultado cero caras, 7/16 de los lanzamientos tienen como resultado una cara y 5/16 de éstos tienen dos caras, el número medio de caras por lanzamiento será 1.06 sin importar si el número total de lanzamientos fue 16, 1000 o incluso 10,000.

Utilicemos ahora este método de frecuencias relativas para calcular el número promedio de caras por lanzamiento que podríamos esperar en el largo plazo. Nos referiremos a este valor promedio como la **media de la variable aleatoria X** o la **media de la distribución de probabilidad de X** y la denotamos por μ , o simplemente como μ cuando esté claro a qué variable aleatoria nos referimos. También es común entre los estadísticos referirse a esta media como la esperanza matemática o el valor esperado de la variable aleatoria X y denotarla como $E(X)$.

Suponga que se lanzan monedas legales, encontramos que el espacio muestral para nuestro experimento es

$$S = \{HH, HT, TH, TT\}.$$

Como los 4 puntos muestrales son igualmente probables, se sigue que

$$P(X = 0) = P(TT) = \frac{1}{4}, \quad P(X = 1) = P(TH) + P(HT) = \frac{1}{2},$$

y

$$P(X = 2) = P(HH) = \frac{1}{4},$$

donde un elemento típico, digamos TH , indica que el primer lanzamiento tuvo como resultado una cruz seguida de una cara en el segundo lanzamiento. Ahora bien, estas probabilidades son justo las frecuencias relativas para los eventos dados en el largo plazo. Por tanto,

$$\mu = E(X) = (0)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{2}\right) + (2)\left(\frac{1}{4}\right) = 1.$$

Este resultado significa que una persona que lance dos monedas una y otra vez, en promedio, obtendrá una cara por lanzamiento.

El método descrito antes para calcular el número esperado de caras en el lanzamiento de dos monedas sugiere que la media o valor esperado de cualquier variable aleatoria discreta se puede obtener al multiplicar cada uno de los valores x_1, x_2, \dots, x_n de la variable aleatoria X por su correspondiente probabilidad $f(x_1), f(x_2), \dots, f(x_n)$ y sumar los productos. Esto es cierto, sin embargo, sólo si la variable aleatoria es discreta. En el caso de variables aleatorias continuas, la definición de un valor esperado es esencialmente la misma pero con integrales que reemplazan a las sumatorias.

Definición 4.1

Sea X una variable aleatoria con distribución de probabilidad $f(x)$. La **media o valor esperado** de X es

$$\mu = E(X) = \sum_x xf(x)$$

si X es discreta, y

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x) dx$$

si X es continua.

Ejemplo 4.1 Un inspector de calidad muestrea un lote que contiene siete componentes; el lote contiene cuatro componentes buenos y tres defectuosos. El inspector toma una muestra de tres componentes. Encuentre el valor esperado del número de componentes buenos en esta muestra.

SOLUCIÓN

Sea X el número de componentes buenos en la muestra. La distribución de probabilidad de X es

$$f(x) = \frac{\binom{4}{x} \binom{3}{3-x}}{\binom{7}{3}}, \quad x = 0, 1, 2, 3.$$

Unos simples cálculos dan $f(0) = 1/35$, $f(1) = 12/35$, $f(2) = 18/35$, $f(3) = 4/35$. Por tanto,

$$\mu = E(X) = (0) \left(\frac{1}{35}\right) + (1) \left(\frac{12}{35}\right) + (2) \left(\frac{18}{35}\right) + (3) \left(\frac{4}{35}\right) = \frac{12}{7} = 1.7.$$

De esta manera, si se selecciona al azar una muestra de tamaño tres una y otra vez de un lote de cuatro componentes buenos y tres defectuosos, contendría, en promedio, 1.7 componentes buenos. ■

Ejemplo 4.2 En un juego de azar se pagarán \$5 a una persona si sólo salen caras o cruces cuando se lanzan tres monedas, y ella pagará \$3 si salen una o dos caras. ¿Cuál es su ganancia esperada?

SOLUCIÓN

El espacio muestral para los posibles resultados cuando se lanzan de manera simultánea tres monedas, o de manera equivalente si se lanza tres veces una moneda, es

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}.$$

Se podría argumentar que cada una de estas posibilidades es igualmente probable y que ocurre con probabilidad de $1/8$. Un método alternativo sería aplicar la regla de la multiplicación de probabilidad para eventos independientes a cada elemento de S . Por ejemplo,

$$P(HHT) = P(H)P(H)P(T) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}.$$

La variable aleatoria de interés es Y , el monto que el jugador puede ganar; y los valores posibles de Y son \$5 si ocurre el evento $E_1 = \{HHH, TTT\}$ y $-\$3$ si ocurre el evento

$$E_2 = \{HHT, HTH, THH, HTT, THT, TTH\}$$

evento. Como E_1 y E_2 ocurren con probabilidades $1/4$ y $3/4$, respectivamente, se sigue que

$$\mu = E(Y) = (5) \left(\frac{1}{4}\right) + (-3) \left(\frac{3}{4}\right) = -1.$$

En este juego la persona perderá, en promedio, \$1 por lanzamiento de las tres monedas. Un juego se considera "equitativo" si el jugador, en promedio, queda empatao. Por tanto, una ganancia esperada de cero define un juego equitativo. ■

Los ejemplos 4.1 y 4.2 se diseñan para permitir al lector obtener una mejor comprensión de lo que queremos decir por valor esperado de una variable aleatoria. En ambos casos las variables aleatorias son discretas. Seguimos con un ejemplo de variable aleatoria continua donde un ingeniero se interesa en la *vida media* de cierto tipo de dispositivo electrónico. Esta es una ilustración de problema de *tiempo de falla* que a menudo se presenta en la práctica. El valor esperado de la vida del dispositivo es un parámetro importante para su evaluación.

Ejemplo 4.3 Sea X la variable aleatoria que denota la vida en horas de cierto dispositivo electrónico. La función de densidad de probabilidad es

$$f(x) = \begin{cases} \frac{20,000}{x^3}, & x > 100 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la vida esperada de este tipo de dispositivo.

SOLUCIÓN

Con el uso de la definición 4.1, tenemos

$$\mu = E(X) = \int_{100}^{\infty} x \frac{20,000}{x^3} dx = \int_{100}^{\infty} \frac{20,000}{x^2} dx = 200.$$

Por tanto, podemos esperar que este tipo de dispositivo dure, *en promedio*, 200 horas. ■

Consideremos ahora una nueva variable aleatoria $g(X)$, que depende de X ; es decir, cada valor de $g(X)$ está determinado al conocer los valores de X . Por ejemplo, $g(X)$ podría ser X^2 o $3X - 1$, de modo que siempre que X tome el valor 2, $g(X)$ toma el valor $g(2)$. En particular, si X es una variable aleatoria discreta con distribución de probabilidad $f(x)$, $x = -1, 0, 1, 2$, y $g(X) = X^2$ entonces

$$P[g(X) = 0] = P(X = 0) = f(0),$$

$$P[g(X) = 1] = P(X = -1) + P(X = 1) = f(-1) + f(1),$$

$$P[g(X) = 4] = P(X = 2) = f(2).$$

de modo que la distribución de probabilidad de $g(X)$ se puede escribir como

$g(x)$	0	1	4
$P[g(X) = g(x)]$	$f(0)$	$f(-1) + f(1)$	$f(2)$

Por definición del valor esperado de una variable aleatoria, obtenemos

$$\begin{aligned} \mu_{g(X)} &= E[g(x)] = 0f(0) + 1[f(-1) + f(1)] + 4f(2) \\ &= (-1)^2 f(-1) + (0)^2 f(0) + (1)^2 f(1) + (2)^2 f(2) = \sum_x g(x)f(x). \end{aligned}$$

valor esperado de la variable aleatoria $g(X)$ es

$$\mu_{g(X)} = E[g(X)] = \sum g(x)f(x)$$

si X es discreta, y

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx$$

si X es continua.

Ejemplo 4.4 Suponga que el número de autos X que pasa por un lavado entre 4:00 p.m. y 5:00 p.m. en cualquier viernes soleado tiene la siguiente función de probabilidad:

x	4	5	6	7	8	9
$P(X = x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Sea $g(X) = 2X - 1$ la cantidad de dinero en dólares, que el administrador le paga al dependiente. Encuentre las ganancias esperadas del dependiente en particular.

SOLUCIÓN

Por el teorema 4.1, el dependiente puede esperar recibir

$$\begin{aligned} E[g(X)] &= E(2X - 1) \\ &= \sum_{x=4}^9 (2x - 1)f(x) \\ &= (7)\left(\frac{1}{12}\right) + (9)\left(\frac{1}{12}\right) + (11)\left(\frac{1}{4}\right) + (13)\left(\frac{1}{4}\right) + (15)\left(\frac{1}{6}\right) \\ &= \$12.67. \quad \blacksquare \end{aligned}$$

Ejemplo 4.5 Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de $g(X) = 4X + 3$.

SOLUCIÓN

Por el teorema 4.1, tenemos

$$E(4X + 3) = \int_{-1}^2 \frac{(4x + 3)x^2}{3} dx = \frac{1}{3} \int_{-1}^2 (4x^3 + 3x^2) dx = 8.$$

Debemos extender ahora nuestro concepto de esperanza matemática al caso de dos variables aleatorias X y Y con distribución de probabilidad conjunta $f(x, y)$. ■

Definición 4.2

Sean X y Y variables aleatorias con distribución de probabilidad conjunta $f(x, y)$. La media o valor esperado de la variable aleatoria $g(X, Y)$ es

$$\mu_{g(X, Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

si X y Y son discretas, y

$$\mu_{g(X, Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

si X y Y son continuas.

La generalización de la definición 4.2 para el cálculo de la esperanza matemática de funciones de diversas variables aleatorias es clara.

Ejemplo 4.6 Sean X y Y variables aleatorias con distribución de probabilidad conjunta que se indica en la tabla 3.6 de la página 71. Encuentre el valor esperado de $g(X, Y) = XY$.

SOLUCIÓN

Por la definición 4.2, escribimos

$$\begin{aligned} E(XY) &= \sum_{x=0}^2 \sum_{y=0}^2 xyf(x, y) = (0)(0)f(0, 0) + (0)(1)f(0, 1) + (0)(2)f(0, 2) \\ &\quad + (1)(0)f(1, 0) + (1)(1)f(1, 1) \\ &\quad + (2)(0)f(2, 0) \\ &= f(1, 1) = \frac{3}{14} \quad \blacksquare \end{aligned}$$

Ejemplo 4.7 Encuentre $E\left(\frac{Y}{X}\right)$ para la función de densidad

$$f(x, y) = \begin{cases} \frac{x(1 + 3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

SOLUCIÓN

Tenemos

$$E\left(\frac{Y}{X}\right) = \int_0^1 \int_0^2 \frac{y(1+3y^2)}{4} dx dy = \int_0^1 \frac{y+y^3}{2} dy = \frac{5}{8}.$$

Nótese que si $g(X, Y) = X$ en la definición 4.2, tenemos:

$$E(X) = \begin{cases} \sum_x \sum_y x f(x, y) = \sum_x x g(x) & \text{(caso discreto)} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy = \int_{-\infty}^{\infty} x g(x) dx & \text{(caso continuo),} \end{cases}$$

donde $g(x)$ es la distribución marginal de X . Por tanto, para calcular $E(X)$ en un espacio bidimensional, se puede usar la distribución de probabilidad conjunta de X y Y o la distribución marginal de X .

De manera similar, definimos

$$E(X) = \begin{cases} \sum_x \sum_y y f(x, y) = \sum_y y h(y) & \text{(caso discreto)} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \int_{-\infty}^{\infty} y h(y) dy & \text{(caso continuo),} \end{cases}$$

donde $h(y)$ es la distribución marginal de la variable aleatoria Y . ■

Ejercicios

1. Suponga que dos variables (X, Y) se distribuyen de manera uniforme en un círculo de radio a . La función densidad de probabilidad conjunta es

$$f(x, y) = \begin{cases} \frac{1}{\pi a^2}, & x^2 + y^2 \leq a^2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de X , μ_x .

2. La distribución de probabilidad de la variable aleatoria discreta X es:

$$f(x) = \left(\frac{3}{4}\right) \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3.$$

Encuentre la media de X .

3. Encuentre la media de la variable aleatoria T que representa el total de las tres monedas del ejercicio 25 de la página 63.

4. Una moneda está cargada de modo que la probabilidad de ocurrencia de una cara es tres veces mayor que la de una cruz. Encuentre el número esperado de cruces cuando esta moneda se lanza dos veces.

5. La distribución de probabilidad de X , el número de imperfecciones por cada 10 metros de una tela sintética en rollos continuos de ancho uniforme, está dada en el ejercicio 13 de la página 62 como

x	0	1	2	3	4
$f(x)$	0.041	0.37	0.16	0.05	0.01

Encuentre el número promedio de imperfecciones en 10 metros de esta tela.

6. A un empleado de un lavado de autos se le paga de acuerdo con el número de autos que lava. Suponga que las probabilidades son $1/12$, $1/12$, $1/4$, $1/4$, $1/6$, y $1/6$, respectivamente, de que el empleado reciba \$7, \$9, \$11, \$13, \$15 o \$17 entre 4:00 p.m. y 5:00 p.m. en cualquier viernes soleado. Encuentre las ganancias esperadas del empleado para este periodo particular.

7. Al invertir en unas acciones particulares, una persona puede tener una ganancia en un año de \$4000 con probabilidad de 0.3 o tener una pérdida de \$1000 con probabilidad de 0.7. ¿Cuál es la ganancia esperada de esta persona?

8. Suponga que un distribuidor de joyas antiguas se interesa en la compra de un collar de oro para el que las probabilidades son 0.22, 0.36, 0.28 y 0.14, respectivamente, de que pueda venderlo con una ganancia de \$250, una ganancia de \$150, venderlo al costo o venderlo con una pérdida de \$150. ¿Cuál es su ganancia esperada?

9. En un juego de apuestas se le paga a una mujer \$3 si saca una sota o una reina y \$5 si saca un rey o un as de una baraja ordinaria de 52 cartas. Si saca cualquier otra carta, pierde. ¿Cuánto debe pagar si el juego es justo?

10. Dos expertos en calidad de neumáticos examinan lotes de éstos y asignan tasas de calidad a cada neumático en una escala de tres puntos. Sea X la calificación dada por el experto A y Y la del experto B . La siguiente tabla da la distribución conjunta para X y Y .

	Y	1	2	3
X	1	0.1	0.05	0.02
	2	0.1	0.35	0.05
	3	0.03	0.1	0.2

Encuentre μ_x y μ_y .

11. Un piloto privado desea asegurar su aeroplano por \$200,000. La compañía de seguros estima que puede ocurrir una pérdida total con probabilidad 0.002, una pérdida de 50% con probabilidad 0.01 y una pérdida de 25% con probabilidad 0.1. Si se ignoran todas las demás pérdidas parciales, ¿qué prima debe cobrar cada año la compañía de seguros para tener una utilidad promedio de \$500?

12. Si la utilidad de un distribuidor, en unidades de \$5000, para un automóvil nuevo se puede ver como una variable aleatoria X que tiene la función de densidad

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

encuentre la utilidad promedio por automóvil.

13. La función de densidad de las mediciones codificadas del diámetro de paso de los hilos de un encaje es

$$f(x) = \begin{cases} \frac{4}{\pi(1+x^2)}, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de X .

14. ¿Qué proporción de individuos se puede esperar que respondan a cierta encuesta que se envía por correo si la proporción X tiene la función de densidad

$$f(x) = \begin{cases} \frac{2(x+2)}{5}, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso?} \end{cases}$$

15. La función de densidad de la variable aleatoria continua X , el número total de horas, en unidades de 100 horas, que una familia utiliza una aspiradora en un periodo de un año, se da en el ejercicio 7 de la página 61 como

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2-x, & 1 \leq x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el número promedio de horas por año que las familias utilizan sus aspiradoras.

16. Suponga que inspecciona un lote de 1000 focos, entre los cuales 20 son defectuosos. Elija dos focos al azar del lote sin reemplazo. Sean

$$X_1 = \begin{cases} 1, & \text{si el primer foco es defectuoso} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{si el segundo foco es defectuoso} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Encuentre la probabilidad de que ninguno de los focos sea defectuoso. [PISTA: Calcule $P(X_1 + X_2 = 1)$.]

17. Sea X una variable aleatoria con la siguiente distribución de probabilidad:

x	-3	6	9
$f(x)$	1/6	1/2	1/3

Encuentre $\mu_g(X)$, donde $g(X) = (2X + 1)^2$.

18. Encuentre el valor esperado de la variable aleatoria $g(X) = X^2$, donde X tiene la distribución de probabilidad del ejercicio 2.

19. Una empresa industrial grande compra varios procesadores de palabras nuevos al final de cada año, el número exacto depende de la frecuencia de reparaciones en el año anterior. Suponga que el número de procesadores de palabras, X , que se compran cada año tiene la distribución de probabilidad siguiente:

x	0	1	2	3
$f(x)$	1/10	3/10	2/5	1/5

Si el costo del modelo que se desea permanece fijo a \$1200 a lo largo de este año y se obtiene un descuent-

to de $50X^2$ dólares en cualquier compra, ¿cuánto espera gastar esta empresa en nuevos procesadores de palabras al final de este año?

20. Una variable aleatoria continua X tiene la función de densidad

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de $g(X) = e^{2X/3}$.

21. ¿Cuál es la ganancia promedio por automóvil del distribuidor si la ganancia en cada uno está dada por $g(X) = X^2$, donde X es una variable aleatoria que tiene la función de densidad del ejercicio 12?

22. El periodo de hospitalización, en días, para pacientes que siguen el tratamiento para cierto tipo de trastorno renal es una variable aleatoria $Y = X + 4$, donde X tiene la función de densidad

$$f(x) = \begin{cases} \frac{32}{(x+4)^3}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el número promedio de días que una persona permanece hospitalizada para seguir el tratamiento para este desorden.

23. Suponga que X y Y tienen la siguiente función de probabilidad conjunta:

$f(x, y)$	$x: 2$	4
$y: 1$	0.10	0.15
3	0.20	0.30
5	0.10	0.15

- (a) Encuentre el valor esperado de $g(X, Y) = XY^2$.
 (b) Encuentre μ_X y μ_Y .

24. Con referencia a las variables aleatorias cuya distribución de probabilidad conjunta se da en el ejercicio 3 de la página 80,

- (a) encuentre $E(X^2Y - 2XY)$;
 (b) encuentre $\mu_X - \mu_Y$.

25. Refiérase a las variables aleatorias cuya distribución de probabilidad conjunta se da en el ejercicio 17 de la página 81, encuentre la media para el número total de sotas y reyes cuando se sacan tres cartas sin reemplazo de las 12 cartas mayores de una baraja ordinaria de 52 cartas.

26. Sean X y Y variables aleatorias con función de densidad conjunta:

$$f(x, y) = \begin{cases} 4xy, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de $Z = \sqrt{X^2 + Y^2}$.

4.2 Varianza y Covarianza

La media o valor esperado de una variable aleatoria X es de especial importancia en estadística, pues describe el lugar donde se centra la distribución de probabilidad. Por sí misma, sin embargo, la media no da una descripción adecuada de la forma de la distribución. Necesitamos caracterizar la variabilidad en la distribución. En la figura 4.1 tenemos los histogramas de dos distribuciones discretas de probabilidad con la misma media $\mu = 2$ que difieren de forma considerable en la variabilidad o dispersión de sus observaciones alrededor de la media.

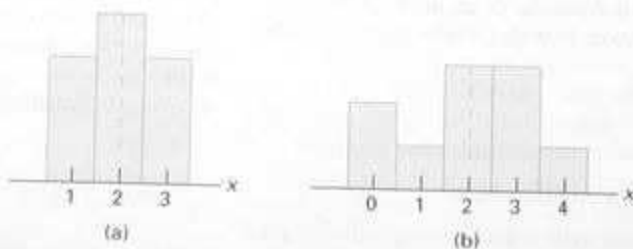


Figura 4.1 Distribuciones con medias iguales y dispersiones diferentes.

La medida de variabilidad más importante de una variable aleatoria X se obtiene al hacer $g(X) = (X - \mu)^2$ en el teorema 4.1. Debido a su importancia en estadística, se le denomina **varianza de la variable aleatoria X** o **varianza de la distribución de probabilidad de X** y se denota por $\text{Var}(X)$ o con el símbolo σ_x^2 , o simplemente σ^2 cuando queda claro a qué variable aleatoria nos referimos.

Definición 4.3

Sea X una variable aleatoria con distribución de probabilidad $f(x)$ y media μ . La varianza de X es

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x)$$

si X es discreta, y

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

si X es continua. La raíz cuadrada positiva de la varianza, σ , se llama **desviación estándar** de X .

La cantidad $x - \mu$ en la definición 4.3 se llama **desviación de una observación** respecto a su media. Como estas desviaciones se elevan al cuadrado y después se promedian, σ^2 será mucho menor para un conjunto de valores x que sean cercanos a μ que para un conjunto de valores que varíe de forma considerable de μ .

Ejemplo 4.8 Sea la variable aleatoria X el número de automóviles que se utilizan con propósitos de negocios en un día de trabajo dado. La distribución de probabilidad para la compañía A [Figura 4.1(a)] es

x	1	2	3
$f(x)$	0.3	0.4	0.3

y para la compañía B [Figura 4.1(b)] es

x	0	1	2	3	4
$f(x)$	0.2	0.1	0.3	0.3	0.1

Muestre que la varianza de la distribución de probabilidad para la compañía B es mayor que la de la compañía A.

SOLUCIÓN

Para la compañía A, encontramos que

$$\mu = E(X) = (1)(0.3) + (2)(0.4) + (3)(0.3) = 2.0$$

y entonces

$$\begin{aligned} \sigma^2 &= \sum_{x=1}^3 (x - 2)^2 f(x) = (1 - 2)^2(0.3) + (2 - 2)^2(0.4) + (3 - 2)^2(0.3) \\ &= 0.6. \end{aligned}$$

Para la compañía B, tenemos

$$\begin{aligned}\mu &= E(X) = (0)(0.2) + (1)(0.1) + (2)(0.3) + (3)(0.3) + (4)(0.1) \\ &= 2.0\end{aligned}$$

por lo que

$$\begin{aligned}\sigma^2 &= \sum_{x=0}^4 (x - 2)^2 f(x) \\ &= (0 - 2)^2(0.2) + (1 - 2)^2(0.1) + (2 - 2)^2(0.3) + (3 - 2)^2(0.3) \\ &\quad + (4 - 2)^2(0.1) = 1.6.\end{aligned}$$

De forma clara, la varianza del número de automóviles que se utilizan con propósitos de negocios es mayor para la compañía B que para la A. ■

Una fórmula alternativa que se prefiere para encontrar σ^2 , que a menudo simplifica los cálculos, se da en el siguiente teorema.

Teorema 4.2

La varianza de una variable aleatoria X es

$$\sigma^2 = E(X^2) - \mu^2.$$

PRUEBA Para el caso discreto podemos escribir

$$\begin{aligned}\sigma^2 &= \sum_x (x - \mu)^2 f(x) = \sum_x (x^2 - 2\mu x + \mu^2) f(x) \\ &= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x).\end{aligned}$$

Como $\mu = \sum_x x f(x)$ por definición, y $\sum_x f(x) = 1$ para cualquier distribución de probabilidad discreta, se sigue que

$$\sigma^2 = \sum_x x^2 f(x) - \mu^2 = E(X^2) - \mu^2.$$

Para el caso continuo la demostración es la misma paso a paso, si se reemplazan las sumatorias por integrales.

Ejemplo 4.9 Sea la variable aleatoria X el número de partes defectuosas de una máquina cuando se muestrean tres partes de una línea de producción y se prueban. La siguiente es la distribución de probabilidad de X .

x	0	1	2	3
$f(x)$	0.51	0.38	0.10	0.01

Con el uso del teorema 4.2, calcule σ^2 .

SOLUCIÓN

Primero, calculamos

$$\mu = (0)(0.51) + (1)(0.38) + (2)(0.10) + (3)(0.01) = 0.61.$$

Ahora bien

$$E(X^2) = (0)(0.51) + (1)(0.38) + (4)(0.10) + (9)(0.01) = 0.87.$$

Por tanto,

$$\sigma^2 = 0.87 - (0.61)^2 = 0.4979. \quad \blacksquare$$

Ejemplo 4.10 La demanda semanal de Pepsi, en miles de litros, de una cadena local de tiendas, es una variable aleatoria continua X que tiene la densidad de probabilidad

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la media y la varianza de X .

SOLUCIÓN

$$\mu = E(X) = 2 \int_1^2 x(x-1) dx = \frac{5}{3},$$

y

$$E(X^2) = 2 \int_1^2 x^2(x-1) dx = \frac{17}{6}.$$

Por tanto,

$$\sigma^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}. \quad \blacksquare$$

Hasta este momento la varianza o la desviación estándar sólo tienen significado cuando comparamos dos o más distribuciones que tienen las mismas unidades de medida. Por tanto, podemos comparar las varianzas de las distribuciones de contenidos, medidos en litros, para dos compañías que embotellan jugo de naranja y el valor más grande indicaría la compañía cuyo producto es más variable o menos uniforme. No tendría significado comparar la varianza de una distribución de alturas con la varianza de una distribución de calificaciones de aptitud. En la sección 4.4 mostramos cómo se puede usar la desviación estándar para describir una sola distribución de observaciones.

Extenderemos ahora nuestro concepto de varianza de una variable aleatoria X para incluir también variables aleatorias relacionadas con X . Para la variable aleatoria $g(X)$, la varianza se denotará con $\sigma_{g(X)}^2$ y se calculará mediante el siguiente teorema.

Teorema 4.3

Sea X una variable aleatoria con distribución de probabilidad $f(x)$. La varian-za de la variable aleatoria $g(X)$ es

$$\sigma_{g(X)}^2 = E\{[g(X) - \mu_{g(X)}]^2\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x)$$

si X es discreta, y

$$\sigma_{g(X)}^2 = E\{[g(X) - \mu_{g(X)}]^2\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x) dx$$

si X es continua.

PRUEBA Como $g(X)$ es en sí misma una variable aleatoria con media $\mu_{g(X)}$ como se define en el teorema 4.1, se sigue de la definición 4.3 que

$$\sigma_{g(X)}^2 = E\{[g(X) - \mu_{g(X)}]^2\}.$$

Ahora bien, cuando se aplica el teorema 4.1 nuevamente a la variable aleatoria $[g(X) - \mu_{g(X)}]^2$, la demostración queda completa.

Ejemplo 4.11 Calcule la varianza de $g(X) = 2X + 3$, donde X es una variable aleatoria con distribución de probabilidad

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

SOLUCIÓN

Primero encontramos la media de la variable aleatoria $2X + 3$. De acuerdo con el teorema 4.1,

$$\mu_{2X+3} = E(2X + 3) = \sum_{x=0}^3 (2x + 3)f(x) = 6.$$

Ahora, con el uso del teorema 4.3, tenemos

$$\begin{aligned}\sigma_{2X+3}^2 &= E\{[(2X + 3) - \mu_{2X+3}]^2\} = E\{[2X + 3 - 6]^2\} \\ &= E(4X^2 - 12X + 9) = \sum_{x=0}^3 (4x^2 - 12x + 9)f(x) = 4. \quad \blacksquare\end{aligned}$$

Ejemplo 4.12 Sea X una variable aleatoria que tiene la función de densidad dada en el ejemplo 4.5 de la página 88. Encuentre la varianza de la variable aleatoria $g(X) = 4X + 3$.

SOLUCIÓN

En el ejemplo 4.5 encontramos que $\mu_{4X+3} = 8$. Ahora, con el uso del teorema 4.3,

$$\begin{aligned}\sigma_{4X+3}^2 &= E\{[(4X + 3) - 8]^2\} = E\{(4X - 5)^2\} \\ &= \int_{-1}^2 (4x - 5)^2 \frac{x^2}{3} dx = \frac{1}{3} \int_{-1}^2 (16x^4 - 40x^3 + 25x^2) dx = \frac{51}{5}.\end{aligned}$$

Si $g(X, Y) = (X - \mu_X)(Y - \mu_Y)$, donde $\mu_X = E(X)$ y $\mu_Y = E(Y)$, la definición 4.2 da un valor esperado que se llama covarianza de X y Y , que denotamos como σ_{XY} o $\text{cov}(X, Y)$. ■

Definición 4.4

Sean X y Y variables aleatorias con distribución de probabilidad conjunta $f(x, y)$. La covarianza de X y Y es

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_Y)f(x, y)$$

si X y Y son discretas, y

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)f(x, y) dx dy$$

si X y Y son continuas.

La covarianza entre dos variables aleatorias es una medida de la naturaleza de la asociación entre las dos. Si valores grandes de X a menudo tienen como resultado valores grandes de Y o valores pequeños de X tienen como resultado valores pequeños de Y , $X - \mu_X$ positiva a menudo tendrá como resultado $Y - \mu_Y$ positiva y $X - \mu_X$ negativa a menudo tendrá como resultado $Y - \mu_Y$ negativa. De esta forma el producto $(X - \mu_X)(Y - \mu_Y)$ tenderá a ser positivo. Por otro lado, si valores grandes de X tienen como resultado valores pequeños de Y , entonces el producto $(X - \mu_X)(Y - \mu_Y)$ tenderá a ser negativo. Así, el *signo* de la covarianza indica si la relación entre dos variables aleatorias dependientes es positiva o negativa. Cuando X y Y son estadísticamente independientes se puede mostrar que la covarianza es cero (véase el teorema 4.10, corolario 1). Lo opuesto, sin embargo, por lo general no es cierto. Dos variables pueden tener covarianza cero e incluso así no ser estadísticamente independientes.

La fórmula alternativa que se prefiere para σ_{XY} se establece en el siguiente teorema.

Teorema 4.4

La covarianza de dos variables aleatorias X y Y con medias μ_X y μ_Y , respectivamente, está dada por

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

PRUEBA Para el caso discreto podemos escribir

$$\begin{aligned} \sigma_{XY} &= \sum_x \sum_y (x - \mu_X)(y - \mu_Y)f(x, y) \\ &= \sum_x \sum_y (xy - \mu_X y - \mu_Y x + \mu_X \mu_Y)f(x, y) \\ &= \sum_x \sum_y xyf(x, y) - \mu_X \sum_x \sum_y yf(x, y) \end{aligned}$$

$$-\mu_Y \sum_x \sum_y xf(x, y) + \mu_X \mu_Y \sum_x \sum_y f(x, y).$$

Como $\mu_X = \sum_x \sum_y xf(x, y)$ y $\mu_Y = \sum_x \sum_y yf(x, y)$ por definición, y además $\sum_x \sum_y f(x, y) = 1$ para cualquier distribución discreta conjunta, se sigue que

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y = E(XY) - \mu_X \mu_Y.$$

Para el caso continuo la prueba es idéntica pero con las sumatorias reemplazadas por integrales.

Ejemplo 4.13 El número de repuestos azules X y el número de repuestos rojos Y , cuando se seleccionan al azar dos repuestos para bolígrafo de cierta caja, se describió en el ejemplo 3.8 de la página 70 mediante la siguiente distribución de probabilidad conjunta:

$f(x, y)$	$x: 0$	1	2	$h(y)$
$y: 0$	$\frac{2}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
1	$\frac{3}{14}$	$\frac{3}{14}$		$\frac{3}{7}$
2	$\frac{1}{28}$			$\frac{1}{28}$
$g(x)$	$\frac{5}{14}$	$\frac{12}{14}$	$\frac{3}{28}$	1

Encuentre la covarianza de X y Y .

SOLUCIÓN

Del ejemplo 4.6, vemos que $E(XY) = 3/14$. Ahora bien

$$\begin{aligned} \mu_X = E(X) &= \sum_{x=0}^2 \sum_{y=0}^2 xf(x, y) = \sum_{x=0}^2 xg(x) = (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right) \\ &= \frac{3}{4} \end{aligned}$$

y

$$\begin{aligned} \mu_Y = E(Y) &= \sum_{x=0}^2 \sum_{y=0}^2 yf(x, y) = \sum_{y=0}^2 yh(y) = (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right) \\ &= \frac{1}{2}. \end{aligned}$$

Por tanto

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y = \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) = -\frac{9}{56}. \quad \blacksquare$$

Ejemplo 4.14 La fracción X de corredores y la fracción Y de corredoras que compiten en la maratón se describen mediante la función de densidad conjunta

$$f(x, y) = \begin{cases} 8xy, & 0 \leq x \leq 1, 0 \leq y \leq x \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la covarianza de X y Y .

SOLUCIÓN

Debemos calcular primero las funciones de densidad marginal. Éstas son

$$g(x) = \begin{cases} 4x^3, & 0 \leq x \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

y

$$h(y) = \begin{cases} 4y(1 - y^2), & 0 \leq y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

De las funciones de densidad marginal dadas arriba, calculamos

$$\mu_X = E(X) = \int_0^1 4x^4 dx = \frac{4}{5}, \quad \mu_Y = E(Y) = \int_0^1 4y^2(1 - y^2) dy = \frac{8}{15}.$$

De las funciones de densidad conjunta dadas, tenemos

$$E(XY) = \int_0^1 \int_y^1 8x^2y^2 dx dy = \frac{4}{9}.$$

Entonces

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{4}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right) = \frac{4}{225}.$$

Aunque la covarianza entre dos variables aleatorias proporciona información con respecto a la naturaleza de la relación, la magnitud de σ_{XY} *no indica nada respecto a la fuerza de la relación* debido a que σ_{XY} depende de la escala. Su magnitud dependerá de las unidades que se miden para X y Y . Hay una versión de la covarianza libre de la escala que se denomina **coeficiente de correlación** que se utiliza con amplitud en estadística. ■

Definición 4.5

Sean X y Y variables aleatorias con covarianza σ_{XY} y desviaciones estándar σ_X y σ_Y , respectivamente. El coeficiente de correlación X y Y es

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Debe quedar claro al lector que ρ_{XY} es independiente de las unidades de X y Y . El coeficiente de correlación satisface la desigualdad $-1 \leq \rho_{XY} \leq 1$. Toma un valor de cero cuando $\sigma_{XY} = 0$. Donde hay una dependencia lineal exacta, digamos $Y = a + bX$, $\rho_{XY} = 1$ si $b > 0$ y $\rho_{XY} = -1$ si $b < 0$. (Véase el ejercicio 16.) El coeficiente de correlación es tema de más discusión en el capítulo 12, donde se trata la regresión lineal.

Ejercicios

- Use la definición 4.3 de la página 93 para encontrar la varianza de la variable aleatoria X del ejercicio 7 de la página 90.
- Sea X una variable aleatoria con la siguiente distribución de probabilidad:

x	-2	3	5
$f(x)$	0.3	0.2	0.5

Encuentre la desviación estándar de X .

- La variable aleatoria X , que representa el número de errores por 100 líneas de código de programación, tiene la siguiente distribución de probabilidad:

x	2	3	4	5	6
$f(x)$	0.01	0.25	0.4	0.3	0.04

Aplique el teorema 4.2 y encuentre la varianza de X .

- Suponga que las probabilidades son 0.4, 0.3, 0.2 y 0.1, respectivamente, de que 0, 1, 2 o 3 fallas de energía eléctrica afecten cierta subdivisión en cualquier año dado. Encuentre la media y la varianza de la variable aleatoria X que representa el número de fallas de energía que afectan esta subdivisión.
- La utilidad de un distribuidor, en unidades de \$5000, para un automóvil nuevo es una variable aleatoria X que tiene la función de densidad que se da en el ejercicio 12 de la página 91. Encuentre la varianza de X .
- La proporción de personas que responden cierta encuesta que se manda por correo es una variable aleatoria X que tiene la función de densidad que se da en el ejercicio 14 de la página 91. Encuentre la varianza de X .
- El número total de horas, en unidades de 100 horas, que una familia utiliza una aspiradora en un pe-

riodo de un año es una variable aleatoria X que tiene la función de densidad dada en el ejercicio 15 de la página 91. Encuentre la varianza de X .

- Con referencia al ejercicio 14 de la página 91, encuentre $\sigma_{g(X)}$ para la función $g(X) = 3X^2 + 4$.
- Encuentre la desviación estándar de la variable aleatoria $g(X) = (2X + 1)^2$ del ejercicio 17 en la página 91.
- Utilice los resultados del ejercicio 21 de la página 92, y encuentre la varianza de $g(X) = X^2$, donde X es una variable aleatoria que tiene la función de densidad que se da en el ejercicio 12 de la página 91.
- El tiempo, en minutos, para que un aeroplano obtenga vía libre para aterrizar en cierto aeropuerto es una variable aleatoria $Y = 3X - 2$, donde X tiene la función de densidad

$$f(x) = \begin{cases} \frac{1}{4}e^{-x/4}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la media y la varianza de la variable aleatoria Y .

- Encuentre la covarianza de las variables aleatorias X y Y del ejercicio 3 de la página 100.
- Encuentre la covarianza de las variables aleatorias X y Y del ejercicio 13 de la página 81.
- Encuentre la covarianza de las variables aleatorias X y Y del ejercicio 8 de la página 80.
- Refiérase a las variables aleatorias cuya función de densidad conjunta está dada en el ejercicio 4 de la página 80, encuentre la covarianza de X y Y .
- Dada una variable aleatoria X , con desviación estándar σ , y una variable aleatoria $Y = a + bx$, muestre que si $b < 0$, el coeficiente de correlación $\rho_{XY} = -1$, y si $b > 0$, $\rho_{XY} = 1$.

4.3 Medias y varianzas de combinaciones lineales de variables aleatorias

Desarrollaremos ahora algunas propiedades útiles que simplificarán los cálculos de las medias y varianzas de variables aleatorias que aparecen en capítulos posteriores. Estas propiedades nos permitirán tratar con las esperanzas matemáticas en relación con otros parámetros que ya se conocen o que son fáciles de calcular. Todos los resultados que presentamos aquí son válidos para variables aleatorias continuas y discretas. Las demostraciones se dan sólo para el caso continuo. Comenzamos con un teorema y dos corolarios que deben ser, de forma intuitiva, razonables para el lector.

Teorema 4.5

Si a y b son constantes, entonces

$$E(aX + b) = aE(X) + b.$$

PRUEBA Por definición de un valor esperado,

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b)f(x) dx = a \int_{-\infty}^{\infty} xf(x) dx + b \int_{-\infty}^{\infty} f(x) dx.$$

La primera integral de la derecha es $E(X)$ y la segunda integral es igual a 1. Por tanto, tenemos

$$E(aX + b) = aE(X) + b.$$

Corolario 1

Al hacer $a = 0$, vemos que $E(b) = b$.

Corolario 2

Al hacer $b = 0$, vemos que $E(aX) = aE(X)$.

Ejemplo 4.15 Mediante la aplicación del teorema 4.5 a la variable aleatoria discreta $g(X) = 2X - 1$, resuelva de nuevo el ejemplo 4.4.

SOLUCIÓN

De acuerdo con el teorema 4.5, podemos escribir

$$E(2X - 1) = 2E(X) - 1,$$

Ahora bien

$$\begin{aligned} \mu &= E(X) = \sum_{x=4}^9 xf(x) \\ &= (4)\left(\frac{1}{12}\right) + (5)\left(\frac{1}{12}\right) + (6)\left(\frac{1}{4}\right) + (7)\left(\frac{1}{4}\right) + (8)\left(\frac{1}{6}\right) + (9)\left(\frac{1}{6}\right) = \frac{41}{6} \end{aligned}$$

Por tanto,

$$\mu_{2X-1} = (2)\left(\frac{41}{6}\right) - 1 = \$12.67,$$

como antes. ■

Ejemplo 4.16 Para resolver de nuevo el ejemplo 4.5, aplíquese el teorema 4.5 a la variable aleatoria continua $g(X) = 4X + 3$.

SOLUCIÓN

En el ejemplo 4.5 podemos utilizar el teorema 4.5 para escribir

$$E(4X + 3) = 4E(X) + 3.$$

Ahora bien

$$E(X) = \int_{-1}^2 x \left(\frac{x^2}{3}\right) dx = \int_{-1}^2 \frac{x^3}{3} dx = \frac{5}{4}.$$

Por tanto,

$$E(4X + 3) = (4)\left(\frac{5}{4}\right) + 3 = 8,$$

como antes. ■

Teorema 4.6

El valor esperado de la suma o diferencia de dos o más funciones de una variable aleatoria X es la suma o diferencia de los valores esperados de las funciones. Es decir,

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)].$$

PRUEBA Por definición, *

$$\begin{aligned} E[g(X) \pm h(X)] &= \int_{-\infty}^{\infty} [g(x) \pm h(x)]f(x) dx \\ &= \int_{-\infty}^{\infty} g(x)f(x) dx \pm \int_{-\infty}^{\infty} h(x)f(x) dx \\ &= E[g(X)] \pm E[h(X)]. \end{aligned}$$

Ejemplo 4.17 Sea X una variable aleatoria con la distribución de probabilidad que sigue:

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

Encuentre el valor esperado de $Y = (X - 1)^2$.

SOLUCIÓN

Al aplicar el teorema 4.6 a la función $Y = (X - 1)^2$, podemos escribir

$$E[(X - 1)^2] = E(X^2 - 2X + 1) = E(X^2) - 2E(X) + E(1).$$

Del corolario 1 del teorema 4.5, $E(1) = 1$, y por cálculo directo

$$E(X) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (2)(0) + (3)\left(\frac{1}{6}\right) = 1$$

y

$$E(X^2) = (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (4)(0) + (9)\left(\frac{1}{6}\right) = 2.$$

De aquí

$$E[(X - 1)^2] = 2 - (2)(1) + 1 = 1. \quad \blacksquare$$

Ejemplo 4.18 La demanda semanal de cierta bebida, en miles de litros, en una cadena de tiendas es una variable aleatoria continua $g(X) = X^2 + X - 2$, donde X tiene la función de densidad

$$f(x) = \begin{cases} 2(x - 1), & 1 < x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado para la demanda semanal de la bebida.

SOLUCIÓN

Por el teorema 4.6, escribimos

$$E(X^2 + X - 2) = E(X^2) + E(X) - E(2).$$

Del corolario 1 del teorema 4.5, $E(2) = 2$, y por integración directa

$$E(X) = \int_1^2 2x(x - 1) dx = 2 \int_1^2 (x^2 - x) dx = \frac{5}{3},$$

y

$$E(X^2) = \int_1^2 2x^2(x - 1) dx = 2 \int_1^2 (x^3 - x^2) dx = \frac{17}{6}.$$

Ahora bien

$$E(X^2 + X - 2) = \frac{17}{6} + \frac{5}{3} - 2 = \frac{5}{2},$$

por lo que la demanda semanal promedio de la bebida en esta cadena de tiendas es de 2500 litros. \blacksquare

Suponga que tenemos dos variables aleatorias X y Y con distribución de probabilidad conjunta $f(x, y)$. Dos propiedades adicionales que serán muy útiles en los capítulos siguientes incluyen los valores esperados de la suma, diferencia y producto de estas dos variables aleatorias. Primero, sin embargo, demos un teorema sobre

el valor esperado de la suma o diferencia de funciones de las variables dadas. Éste, por supuesto, es sólo una extensión del teorema 4.6.

Teorema 4.7

El valor esperado de la suma o diferencia de dos o más funciones de las variables aleatorias X y Y es la suma o diferencia de los valores esperados de las funciones. Es decir,

$$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)].$$

PRUEBA Por la definición 4.2,

$$\begin{aligned} E[g(X, Y) \pm h(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [g(x, y) \pm h(x, y)]f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y) dx dy \\ &\quad \pm \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y)f(x, y) dx dy \\ &= E[g(X, Y)] \pm E[h(X, Y)]. \end{aligned}$$

Corolario 1

Al hacer $g(X, Y) = g(X)$ y $h(X, Y) = h(Y)$, vemos que

$$E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)].$$

Corolario 2

Al hacer $g(X, Y) = X$ y $h(X, Y) = Y$, vemos que

$$E(X \pm Y) = E(X) \pm E(Y).$$

Si X representa la producción diaria de algún artículo de la máquina A y Y la producción diaria de la misma clase de artículo de la máquina B , entonces $X + Y$ representa el número total de artículos que ambas máquinas producen diariamente. El segundo corolario del teorema 4.7 establece que la producción promedio diaria para ambas máquinas es igual a la suma de la producción promedio diaria de cada máquina.

Teorema 4.8

Sean X y Y dos variables aleatorias independientes. Entonces

$$E(XY) = E(X)E(Y).$$

PRUEBA Por la definición 4.2

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx dy.$$

Como X y Y son independientes, podemos escribir

$$f(x, y) = g(x)h(y),$$

donde $g(x)$ y $h(y)$ son las distribuciones marginales de X y Y , respectivamente. De aquí

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyg(x)h(y) dx dy = \int_{-\infty}^{\infty} xg(x) dx \int_{-\infty}^{\infty} yh(y) dy \\ &= E(X)E(Y). \end{aligned}$$

El teorema 4.8 se puede ilustrar para variables discretas al lanzar un dado verde y uno rojo. Representétese con la variable aleatoria X el resultado del dado verde y con la variable aleatoria Y el resultado del dado rojo. Entonces XY representa el producto de los números que ocurren en el par de dados. A largo plazo, el promedio de los productos de los números es igual al producto del número promedio que ocurre en el dado verde y el número promedio que ocurre en el dado rojo.

Ejemplo 4.19 En la producción de microchips de arseniuro de galio, se sabe que la proporción entre galio y arseniuro es independiente de la producción de un alto porcentaje de obleas manejables, que son los principales componentes de los microchips. Denótese con X la proporción de galio y arseniuro y con Y el porcentaje de microobleas manejables producidas durante un periodo de 1 hora. X y Y son variables aleatorias independientes con la siguiente densidad conjunta

$$f(x, y) = \begin{cases} \frac{x(1 + 3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Ilustre que $E(XY) = E(X)E(Y)$, como sugiere el teorema 4.8.

SOLUCIÓN

Por definición,

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^2 xyf(x, y) dx dy = \int_0^1 \int_0^2 \frac{x^2y(1 + 3y^2)}{4} dx dy \\ &= \int_0^1 \frac{x^3y(1 + 3y^2)}{12} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{2y(1 + 3y^2)}{3} dy = \frac{5}{6}. \end{aligned}$$

$$\begin{aligned} E(X) &= \int_0^1 \int_0^2 xf(x, y) dx dy = \int_0^1 \int_0^2 \frac{x^2(1 + 3y^2)}{4} dx dy \\ &= \int_0^1 \frac{x^3(1 + 3y^2)}{12} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{2(1 + 3y^2)}{3} dy = \frac{4}{3}. \end{aligned}$$

$$E(Y) = \int_0^1 \int_0^2 yf(x, y) dx dy = \int_0^1 \int_0^2 \frac{xy(1 + 3y^2)}{4} dx dy$$

$$= \int_0^1 \frac{x^2 y (1 + 3y^2)}{8} \Big|_{x=0}^{x=2} dy = \int_0^1 \frac{y(1 + 3y^2)}{2} dy = \frac{5}{8}.$$

De aquí

$$E(X)E(Y) = \left(\frac{4}{3}\right)\left(\frac{5}{8}\right) = \frac{5}{6} = E(XY).$$

Concluimos esta sección con la demostración de dos teoremas que son útiles para calcular varianzas o desviaciones estándar. ■

Teorema 4.9

Si a y b son constantes, entonces

$$\sigma_{aX+b}^2 = a^2 \sigma_X^2 = a^2 \sigma^2.$$

PRUEBA Por definición

$$\sigma_{aX+b}^2 = E\{[(aX + b) - \mu_{aX+b}]^2\}.$$

Ahora bien

$$\mu_{aX+b} = E(aX + b) = a\mu + b$$

por el teorema 4.5. Por tanto,

$$\sigma_{aX+b}^2 = E\{(aX + b - a\mu - b)^2\} = a^2 E\{(X - \mu)^2\} = a^2 \sigma^2$$

Corolario 1

Al hacer $a = 1$, vemos que

$$\sigma_{X+b}^2 = \sigma_X^2 = \sigma^2.$$

Corolario 2

Al hacer $b = 0$, vemos que

$$\sigma_{aX}^2 = a^2 \sigma_X^2 = a^2 \sigma^2.$$

El corolario 1 establece que la varianza permanece sin cambio si se suma o se resta una constante a la variable aleatoria. La suma o resta de una constante simplemente corre los valores de X a la derecha o a la izquierda pero no cambia su variabilidad. Sin embargo, si una variable aleatoria se multiplica por una constante o se divide entre una constante, entonces el corolario 2 establece que la varianza se multiplica por el cuadrado de la constante o se divide entre el cuadrado de la constante.

Teorema 4.10

Si X y Y son variables aleatorias con distribución de probabilidad conjunta $f(x, y)$, entonces

$$\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab\sigma_{XY}.$$

PRUEBA Por definición

$$\sigma_{aX+bY}^2 = E\{[(aX + bY) - \mu_{aX+bY}]^2\}.$$

Ahora bien

$$\mu_{aX+bY} = E(aX + bY) = aE(X) + bE(Y) = a\mu_X + b\mu_Y,$$

mediante el uso del corolario 2 del teorema 4.7 seguido por el corolario 2 del teorema 4.5. Por tanto,

$$\begin{aligned}\sigma_{aX+bY}^2 &= E\{[(aX + bY) - (a\mu_X + b\mu_Y)]^2\} \\ &= E\{[a(X - \mu_X) + b(Y - \mu_Y)]^2\} \\ &= a^2E\{(X - \mu_X)^2\} + b^2E\{(Y - \mu_Y)^2\} + 2abE\{(X - \mu_X)(Y - \mu_Y)\} \\ &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}.\end{aligned}$$

Corolario 1

Si X y Y son variables aleatorias independientes, entonces

$$\sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2.$$

El resultado que se establece en el corolario 1 se obtiene a partir del teorema 4.10 al demostrar que la covarianza de las variables independientes X y Y es cero. De aquí, por el teorema 4.4,

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y = 0,$$

pues $E(XY) = E(X)E(Y)$ para variables independientes.

Corolario 2

Si X y Y son variables aleatorias independientes, entonces

$$\sigma_{aX-bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2.$$

El corolario 2 se obtiene al reemplazar b por $-b$ en el corolario 1. Al generalizar a una combinación lineal de n variables aleatorias independientes, escribimos

Corolario 3

Si X_1, X_2, \dots, X_n son variables aleatorias independientes, entonces

$$\sigma_{a_1X_1 + a_2X_2 + \dots + a_nX_n}^2 = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + \dots + a_n^2\sigma_{X_n}^2.$$

Ejemplo 4.20 Si X y Y son variables aleatorias con varianzas $\sigma_X^2 = 2$, $\sigma_Y^2 = 4$, y covarianza $\sigma_{XY} = -2$, entonces encuentre la varianza de la variable aleatoria $Z = 3X - 4Y + 8$.

SOLUCIÓN

$$\begin{aligned}\sigma_Z^2 &= \sigma_{3X-4Y+8}^2 \\ &= \sigma_{3X-4Y}^2 && \text{(por el teorema 4.9)} \\ &= 9\sigma_X^2 + 16\sigma_Y^2 - 24\sigma_{XY} && \text{(por el teorema 4.10)}\end{aligned}$$

$$\begin{aligned}
 &= (9)(2) + (16)(4) - (24)(-2) \\
 &= 130.
 \end{aligned}$$

Ejemplo 4.21 Denótese con X y Y la cantidad de dos tipos diferentes de impurezas en un lote de cierto producto químico. Suponga que X y Y son variables aleatorias independientes con varianzas $\sigma_x^2 = 2$ y $\sigma_y^2 = 3$. Encuentre la varianza de la variable aleatoria $Z = 3X - 2Y + 5$.

SOLUCIÓN

$$\begin{aligned}
 \sigma_z^2 &= \sigma_{3x-2y+5}^2 \\
 &= \sigma_{3x-2y}^2 && \text{(por el teorema 4.9, corolario 1)} \\
 &= 9\sigma_x^2 + 4\sigma_y^2 && \text{(por el teorema 4.10, corolario 2)} \\
 &= (9)(2) + (4)(3) \\
 &= 30. \quad \blacksquare
 \end{aligned}$$

4.4 Teorema de Chebyshev

En la sección 4.2 establecimos que la varianza de una variable aleatoria nos dice algo acerca de la variabilidad de las observaciones alrededor de la media. Si una variable aleatoria tiene una varianza o desviación estándar pequeña, esperaríamos que la mayor parte de los valores se agruparan alrededor de la media. Por tanto, la probabilidad de que una variable aleatoria tome un valor dentro de cierto intervalo alrededor de la media es mayor que para una variable aleatoria similar con una desviación estándar mayor. Si pensamos en la probabilidad en términos de un área, esperaríamos una distribución continua con una desviación estándar pequeña que tenga la mayor parte de su área cercana a μ , como en la figura 4.2(a). Sin embargo, un valor grande de σ indica una variabilidad mayor y, por tanto, esperaríamos que el área estuviera más extendida, como en la figura 4.2(b).

Podemos argumentar lo mismo para una distribución discreta. El área en el histograma de probabilidad de la figura 4.3(b) se extiende mucho más que la de la figura 4.3(a), lo que indica una distribución más variable de mediciones o resultados.

El matemático ruso P. L. Chebyshev (1821-1894) descubrió que la fracción del área entre dos valores simétricos cualesquiera alrededor de la media está relacionada

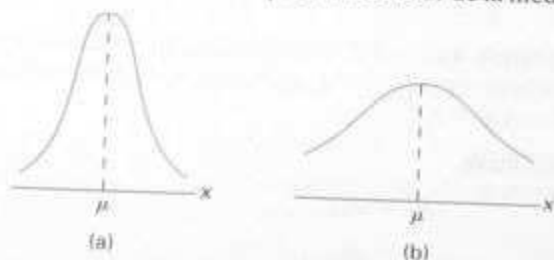


Figura 4.2 Variabilidad de observaciones continuas alrededor de la media.

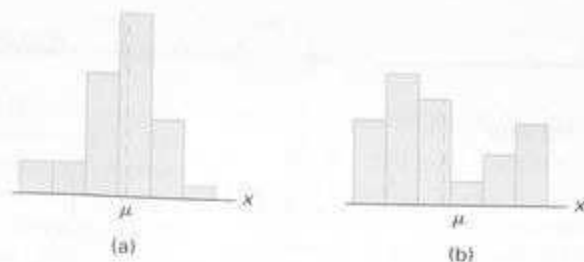


Figura 4.3 Variabilidad de observaciones discretas alrededor de la media.

con la desviación estándar. Como el área bajo una curva de distribución de probabilidad o en un histograma de probabilidad suma 1, el área entre cualesquiera dos números es la probabilidad de que la variable aleatoria tome un valor entre estos números.

El siguiente teorema, debido a Chebyshev, da una estimación conservadora de la probabilidad de que una variable aleatoria tome un valor dentro de k desviaciones estándar de su media para cualquier número real k . Proporcionaremos la demostración sólo para el caso continuo y se deja el caso discreto como ejercicio.

Teorema 4.11

(Teorema de Chebyshev) La probabilidad de que cualquier variable aleatoria X tome un valor dentro de k desviaciones estándar de la media es **al menos** $1 - 1/k^2$. Es decir,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

PRUEBA Por nuestra definición anterior de varianza de X podemos escribir

$$\begin{aligned} \sigma^2 &= E[(X - \mu)^2] \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x) dx \\ &\geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x) dx, \end{aligned}$$

debido a que la segunda de las tres integrales es no negativa. Ahora bien, como $|x - \mu| \geq k\sigma$ para cualquier $x \geq \mu + k\sigma$ o $x \leq \mu - k\sigma$, tenemos que $(x - \mu)^2 \geq k^2\sigma^2$ en ambas integrales restantes. Se sigue que

$$\sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} k^2\sigma^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} k^2\sigma^2 f(x) dx$$

y que

$$\int_{-\infty}^{\mu - k\sigma} f(x) dx + \int_{\mu + k\sigma}^{\infty} f(x) dx \leq \frac{1}{k^2}.$$

De aquí

$$P(\mu - k\sigma < X < \mu + k\sigma) = \int_{\mu - k\sigma}^{\mu + k\sigma} f(x) dx \geq 1 - \frac{1}{k^2},$$

con lo que el teorema queda establecido.

Para $k = 2$ el teorema establece que la variable aleatoria X tiene una probabilidad de al menos $1 - 1/2^2 = 3/4$ de caer dentro de dos desviaciones estándar de la media. Es decir, tres cuartos o más de las observaciones de cualquier distribución yacen en el intervalo $\mu \pm 2\sigma$. De manera similar, el teorema dice que al menos ocho novenos de las observaciones de cualquier distribución caen en el intervalo $\mu \pm 3\sigma$.

Ejemplo 4.22 Una variable aleatoria X tiene una media $\mu = 8$, varianza $\sigma^2 = 9$, y distribución de probabilidad desconocida. Encuentre

(a) $P(-4 < X < 20)$,

(b) $P(|X - 8| \geq 6)$.

SOLUCIÓN

(a) $P(-4 < X < 20) = P[8 - (4)(3) < X < 8 + (4)(3)] \geq \frac{15}{16}$.

(b) $P(|X - 8| \geq 6) = 1 - P(|X - 8| < 6)$
 $= 1 - P(-6 < X - 8 < 6)$
 $= 1 - P[8 - (2)(3) < X < 8 + (2)(3)]$
 $\leq \frac{1}{4}$. ■

El teorema de Chebyshev tiene validez para cualquier distribución de observaciones y, por esta razón, los resultados por lo general son débiles. El valor que el teorema proporciona es sólo un límite inferior. Es decir, sabemos que la probabilidad de una variable aleatoria que cae dentro de dos desviaciones estándar de la media *no puede ser menor* que $3/4$, pero nunca sabemos cuánto podría ser en realidad. Sólo cuando se conoce la distribución de probabilidad podemos determinar probabilidades exactas. Por esta razón llamamos al teorema resultado de *distribución libre*. Cuando se supongan distribuciones específicas en los capítulos siguientes, los resultados serán menos conservadores. El uso del teorema de Chebyshev se relega a situaciones donde se desconoce la forma de la distribución.

Ejercicios

1. Refiérase al ejercicio 3 de la página 100, encuentre la media y la varianza de la variable aleatoria discreta $Z = 3X - 2$, donde X representa el número de errores por 100 líneas de código.

2. Con el uso de los teoremas 4.5 y 4.9, encuentre la media y la varianza de la variable aleatoria $Z = 5X + 3$, donde X tiene la distribución de probabilidad del ejercicio 4 en la página 100.

3. Suponga que una tienda de abarrotes compre cinco envases de leche descremada al precio de mayoreo de \$1.20 por envase y la revende a \$1.65 por envase. Después de la fecha de caducidad, la leche que no se vende se retira de los anaqueles y el tendero recibe un crédito del distribuidor igual a tres cuartos del precio de mayoreo. Si la distribución de probabilidad de la variable aleatoria X , el número de envases que se venden de este lote, es

x	0	1	2	3	4	5
$f(x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$

encuentre la utilidad esperada.

4. Repita el ejercicio 11 de la página 100, con la aplicación de los teoremas 4.5 y 4.9.

5. Sea X una variable aleatoria con la siguiente distribución de probabilidad:

x	-3	6	9
$f(x)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Encuentre $E(X)$ y $E(X^2)$ y después, con el uso de estos valores, evalúe $E[(2X + 1)^2]$.

6. El tiempo total, medido en unidades de 100 horas, que un adolescente utiliza su estéreo en un periodo de un año es una variable continua X que tiene la función de densidad

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Utilice el teorema 4.6 para evaluar la media de la variable aleatoria $Y = 60X^2 + 39X$, donde Y es igual al número de kilowatt hora que gasta al año.

7. Si una variable X se define de modo que

$$E[(X - 1)^2] = 10, \quad E[(X - 2)^2] = 6,$$

encuentre μ y σ^2 .

8. Suponga que X y Y son variables aleatorias independientes que tienen la distribución de probabilidad conjunta

$f(x, y)$	$x:$	2	4
	1	0.10	0.15
y	3	0.20	0.30
	5	0.10	0.15

Encuentre

(a) $E(2X - 3Y)$;

(b) $E(XY)$.

9. Use el teorema 4.7 para evaluar $E(2XY^2 - X^2Y)$ para la distribución de probabilidad conjunta que se muestra en la tabla 3.6.

10. Se crean setenta nuevos trabajos en una planta de ensamblaje automatizada, pero 1000 aspirantes solicitan los 70 puestos. Para seleccionar a los 70 mejores entre los aspirantes, la armadora aplica un examen que cubre habilidad mecánica, destreza manual y capacidad matemática. La calificación media de este examen resulta 60, y las calificaciones tienen una desviación estándar de 6. ¿Una persona que tiene una calificación de 84 puede obtener uno de los trabajos? [PISTA: Utilice el teorema de Chebyshev.] Suponga que la distribución es simétrica alrededor de la media.

11. Una empresa eléctrica fabrica un foco de 100 watts que, de acuerdo con las especificaciones escritas en la caja, tiene una vida media de 900 horas con una desviación estándar de 50 horas. A lo más, ¿qué porcentaje de los focos no duran al menos 700 horas? Suponga que la distribución es simétrica alrededor de la media.

12. Una compañía local fabrica cable telefónico. La longitud promedio del cable es de 52 pulgadas con una desviación estándar de 6.5 pulgadas. A lo más, ¿qué porcentaje del cable telefónico de esta compañía excede 71.5 pulgadas? Suponga que la distribución es simétrica alrededor de la media.

13. Suponga que lanza un dado de 10 lados (0, 1, 2, ..., 9) 500 veces. Con el uso del teorema de Chebyshev, calcule la probabilidad de que la media de la muestra, \bar{X} , esté entre 4 y 5.

14. Si X y Y son variables aleatorias independientes con varianzas $\sigma_X^2 = 5$ y $\sigma_Y^2 = 3$, encuentre la varianza de la variable aleatoria $Z = -2X + 4Y - 3$.

15. Repita el ejercicio 14 si X y Y no son independientes y $\sigma_{XY} = 1$.

16. Una variable aleatoria X tiene una media $\mu = 12$, una varianza $\sigma^2 = 9$, y una distribución de probabilidad desconocida. Con el uso del teorema de Chebyshev, estime

(a) $P(6 < X < 18)$;

(b) $P(3 < X < 21)$.

17. Una variable aleatoria X tiene una media $\mu = 10$ y una varianza $\sigma^2 = 4$. Con el uso del teorema de Chebyshev, encuentre

(a) $P(|X - 10| \geq 3)$;

(b) $P(|X - 10| < 3)$;

(c) $P(5 < X < 15)$;

(d) el valor de la constante c tal que

$$P(|X - 10| \geq c) \leq 0.04.$$

18. Calcule $P(\mu - 2\sigma < X < \mu + 2\sigma)$, donde X tiene la función de densidad

$$f(x) = \begin{cases} 6x(1-x), & 0 < x < 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

y compare con el resultado dado por el teorema de Chebyshev.

19. Sea X el número que ocurre cuando se lanza un dado rojo y Y el número que sale cuando se lanza un dado verde. Encuentre

(a) $E(X + Y)$;

(b) $E(X - Y)$;

(c) $E(XY)$.

20. Suponga que X y Y son variables aleatorias independientes con densidades de probabilidad

$$g(x) = \begin{cases} 8/x^3, & x > 2 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

y

$$h(y) = \begin{cases} 2y, & 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre el valor esperado de $Z = XY$.

21. Si la función de densidad conjunta de X y Y está dada por

$$f(x, y) = \begin{cases} \frac{1}{2}(x + 2y), & 0 < x < 1, 1 < y < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

encuentre el valor esperado de $g(X, Y) = (X/Y^5) + X^2Y$.

22. Sea X el número que ocurre cuando se lanza un dado verde y Y el número que ocurre cuando se lanza un dado rojo. Encuentre la varianza de la variable aleatoria

(a) $2X - Y$;

(b) $X + 3Y - 5$.

23. Suponga que X es una variable aleatoria normal con media μ y desviación estándar σ . Demuestre que el teorema de Chebyshev es válido para $k = 2$ y $k = 3$.

24. Considere una variable aleatoria X con función de densidad

$$f(x) = \begin{cases} \frac{1}{5}, & 0 \leq x \leq 5 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre $\mu = E(X)$ y $\sigma^2 = E(X - \mu)^2$.

(b) Demuestre que el teorema de Chebyshev es válido para $k = 2$ y $k = 3$.

Ejercicios de repaso

1. Demuestre el teorema de Chebyshev cuando X es una variable aleatoria discreta.

2. Encuentre la covarianza de las variables aleatorias X y Y que tienen la función de densidad de probabilidad conjunta

$$f(x, y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

3. Refiérase a las variables aleatorias cuya función de densidad de probabilidad conjunta está dada en el

ejercicio 11 de la página 80, encuentre la cantidad promedio de queroseno que queda en el tanque al final del día.

4. Suponga que la duración X en minutos de un tipo particular de conversación telefónica es una variable aleatoria con función de densidad de probabilidad

$$f(x) = \frac{1}{3}e^{-x/5}, \quad x > 0$$

$$= 0, \quad \text{en cualquier otro caso.}$$

Determine

(a) La duración media $[E(X)]$ de este tipo de conversación telefónica.

(b) Encuentre la varianza y la desviación estándar de X .

(c) Encuentre $E(X + 5)^2$.

5. Refiérase a las variables aleatorias cuya función de densidad conjunta está dada en el ejercicio 5 de la página 80, encuentre la covarianza entre el peso de las cremas y el peso de los chiclosos en estas cajas de chocolates.

6. Refiérase a las variables aleatorias cuya función de densidad conjunta está dada en el ejercicio 4 de la página 80,

(a) encuentre μ_X y μ_Y ;

(b) encuentre $E[(X + Y)/2]$.

7. Suponga que se sabe que la vida X en horas de un compresor particular tiene la función de densidad

$$f(x) = \frac{1}{900}e^{-x/900}, \quad X > 0$$

$$= 0, \quad \text{en cualquier otro caso.}$$

(a) Encuentre la vida media del compresor.

(b) Encuentre $E(X^2)$.

(c) Encuentre la varianza y la desviación estándar de la variable aleatoria X .

8. Refiérase a las variables aleatorias cuya función de densidad conjunta está dada en el ejercicio 5 de la página 80, encuentre el peso esperado de la suma de las cremas y los chiclosos si se compra una caja de estos chocolates.

9. Muestre que $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$.

10. Considere la función de densidad del ejercicio 7 anterior. Demuestre que el teorema de Chebyshev es válido para $k = 2$ y $k = 3$.

11. Considere la función de densidad conjunta

$$g(x) = 8/x^3, \quad x > 2$$

$$= 0, \quad \text{en cualquier otro caso.}$$

$$h(y) = 2y, \quad 0 < y < 1$$

$$= 0, \quad \text{en cualquier otro caso.}$$

Calcule el coeficiente de correlación ρ_{XY} .

12. Considere las variables aleatorias X y Y del ejercicio 15 de la página 111. Calcule ρ_{XY} .

13. La ganancia de un distribuidor en unidades de \$5000 en un automóvil nuevo es una variable aleatoria X que tiene la función de densidad

$$f(x) = \begin{cases} 2(1-x), & 0 \leq x \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre la varianza de la ganancia del distribuidor.

(b) Demuestre que la desigualdad de Chebyshev es válida para $k = 2$ con la función de densidad anterior.

(c) ¿Cuál es la probabilidad de que la ganancia exceda \$500?

14. Considere el ejercicio 10 de la página 91. ¿Se puede decir que las calificaciones dadas por los dos expertos son independientes? Explique por qué sí o por qué no.

Algunas distribuciones de probabilidad discreta

5.1 Introducción

El comportamiento de una variable aleatoria queda descrito por su distribución de probabilidad discreta sin importar si ésta se representa de forma gráfica mediante un histograma, en forma tabular o con una fórmula. A menudo, las observaciones que se generan en diferentes experimentos estadísticos tienen el mismo tipo general de comportamiento. En consecuencia, las variables aleatorias discretas asociadas con estos experimentos se pueden describir esencialmente con la misma distribución de probabilidad y por tanto se pueden representar mediante una sola fórmula. De hecho, se necesita sólo un puñado de distribuciones de probabilidad importantes para describir muchas de las variables aleatorias discretas que se encuentran en la práctica.

5.2 Distribución uniforme discreta

La más simple de todas las distribuciones de probabilidad discreta es una donde la variable aleatoria toma cada uno de sus valores con una probabilidad idéntica. Tal distribución de probabilidad se denomina **distribución uniforme discreta**.

Distribución uniforme discreta

Si la variable aleatoria X toma los valores x_1, x_2, \dots, x_k , con idénticas probabilidades, entonces la distribución uniforme discreta está dada por

$$f(x; k) = \frac{1}{k}, \quad x = x_1, x_2, \dots, x_k.$$

Utilizamos la notación $f(x; k)$ en lugar de $f(x)$ para indicar que la distribución uniforme depende del parámetro k .

Ejemplo 5.1 Cuando se selecciona un foco al azar de una caja que contiene un foco de 40 watts, uno de 60, uno de 75 y uno de 100, cada elemento del espacio muestral $S = \{40, 60, 75, 100\}$ ocurre con probabilidad $1/4$. Por tanto, tenemos una distribución uniforme, con

$$f(x; 4) = \frac{1}{4}, \quad x = 40, 60, 75, 100. \quad \blacksquare$$

Ejemplo 5.2 Cuando se lanza un dado, cada elemento del espacio muestral $S = \{1, 2, 3, 4, 5, 6\}$ ocurre con probabilidad $1/6$. Por tanto, tenemos una distribución uniforme, con

$$f(x; 6) = \frac{1}{6}, \quad x = 1, 2, 3, 4, 5, 6.$$

La representación gráfica de la distribución uniforme por medio de un histograma siempre resulta ser un conjunto de rectángulos con alturas iguales. El histograma para el ejemplo 5.2 se muestra en la figura 5.1. \blacksquare

Teorema 5.1

La media y la varianza de la distribución uniforme discreta $f(x; k)$ son

$$\mu = \frac{\sum_{i=1}^k x_i}{k} \quad \text{y} \quad \sigma^2 = \frac{\sum_{i=1}^k (x_i - \mu)^2}{k}.$$

PRUEBA Por definición

$$\mu = E(X) = \sum_{i=1}^k x_i f(x_i; k) = \sum_{i=1}^k \frac{x_i}{k} = \frac{\sum_{i=1}^k x_i}{k}.$$



Figura 5.1 Histograma para el lanzamiento de un dado.

También, por definición,

$$\sigma^2 = E[(X - \mu)^2] = \sum_{i=1}^k (x_i - \mu)^2 f(x_i; k) = \sum_{i=1}^k \frac{(x_i - \mu)^2}{k} = \frac{\sum_{i=1}^k (x_i - \mu)^2}{k}$$

Ejemplo 5.3 Con referencia al ejemplo 5.2, encontramos que

$$\mu = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

y

$$\sigma^2 = \frac{(1 - 3.5)^2 + (2 - 3.5)^2 + \dots + (6 - 3.5)^2}{6} = \frac{35}{12} \quad \blacksquare$$

5.3 Distribuciones binomial y multinomial

Un experimento a menudo consiste en pruebas repetidas, cada una con dos posibles resultados que se pueden etiquetar como **éxito** o **fracaso**. La aplicación más obvia tiene que ver con la prueba de artículos a medida que salen de una línea de montaje, donde cada prueba o experimento puede indicar si un artículo está defectuoso o no. Podemos elegir definir cualquiera de los resultados como éxito. Se puede considerar la extracción sucesiva de cartas de una baraja ordinaria y cada prueba se etiqueta como éxito o fracaso, dependiendo de si la carta es de corazones o no. Si cada carta se reemplaza y el paquete se baraja antes de la siguiente extracción, los dos experimentos recién descritos tienen propiedades similares, pues los ensayos que se repiten son independientes y la probabilidad de éxito permanece constante entre cada uno de ellos. El proceso se denomina **proceso de Bernoulli**. Cada ensayo se llama **experimento de Bernoulli**. Observe en el ejemplo de extracción de cartas que las probabilidades de éxito para los ensayos que se repiten cambia si las cartas no se reemplazan. Es decir, la probabilidad de seleccionar una carta de corazones en la primera extracción es $1/4$, pero en la segunda es una probabilidad condicional que tiene un valor de $13/51$ o $12/51$, lo cual depende de si aparece una de corazones en la primera extracción; este, entonces, ya no se considerará como un conjunto de experimentos de Bernoulli.

Proceso de Bernoulli

Si se habla con exactitud, el proceso de Bernoulli debe tener las siguientes propiedades:

1. El experimento consiste en n pruebas que se repiten.
2. Cada prueba produce un resultado que se puede clasificar como éxito o fracaso.
3. La probabilidad de un éxito, que se denota con p , permanece constante en cada prueba.
4. Las pruebas que se repiten son independientes.

Considere el conjunto de experimentos de Bernoulli en el que se seleccionan tres artículos al azar de un proceso de ensamblaje, se inspeccionan y se clasifican como defectuosos o no defectuosos. Un artículo defectuoso se designa como un éxito. El número de éxitos es una variable aleatoria X que toma valores integrales de cero a 3. Los ocho resultados posibles y los valores correspondientes de X son

Resultado	x
NNN	0
NDN	1
NND	1
DNN	1
NDD	2
DND	2
DDN	2
DDD	3

Como los artículos se seleccionan de forma independiente de un proceso que supondremos produce 25% de artículos defectuosos,

$$P(NDN) = P(N)P(D)P(N) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{9}{64}.$$

Cálculos similares dan las probabilidades para los demás resultados posibles. La distribución de probabilidad de X es por tanto:

x	0	1	2	3
$f(x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

El número X de éxitos en n experimentos de Bernoulli se denomina **variable aleatoria binomial**. La distribución de probabilidad de esta variable aleatoria discreta se llama **distribución binomial**, y sus valores se denotarán como $b(x; n, p)$, pues dependen del número de pruebas y de la probabilidad de éxito en una prueba dada. De esta forma, para la distribución de probabilidad de X , el número de defectuosos es,

$$P(X = 2) = f(2) = b(2; 3, \frac{1}{4}) = \frac{9}{64}.$$

Generalicemos ahora la ilustración anterior para obtener una fórmula para $b(x; n, p)$. Es decir, deseamos encontrar una fórmula que dé la probabilidad de x éxitos en n pruebas para un experimento binomial. Primero, considere la probabilidad de x éxitos y $n - x$ fracasos en un orden específico. Como las pruebas son independientes, podemos multiplicar todas las probabilidades que corresponden a los diferentes resultados. Cada éxito ocurre con probabilidad p y cada fracaso con probabilidad $q = 1 - p$. Por tanto, la probabilidad para el orden específico es $p^x q^{n-x}$. Debemos determinar ahora el número total de puntos muestrales en el experimento que tiene x éxitos y $n - x$ fracasos. Este número es igual al número de particiones de n resultados en dos grupos con x en un grupo y $n - x$ en el otro

y se escribe $\binom{n}{x}$. Como estas particiones son mutuamente excluyentes, sumamos las probabilidades de todas las diferentes particiones para obtener la fórmula general o simplemente multiplicamos $p^x q^{n-x}$ por $\binom{n}{x}$.

Distribución binomial

Un experimento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$. Entonces la distribución de probabilidad de la variable aleatoria binomial X , el número de éxitos en n pruebas independientes, es

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

Nótese que cuando $n = 3$ y $p = 1/4$, la distribución de probabilidad de X , el número de defectuosos, se puede escribir como

$$b\left(x; 3, \frac{1}{4}\right) = \binom{3}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3,$$

en lugar de la forma tabular anterior.

Ejemplo 5.4 La probabilidad de que cierta clase de componente sobreviva a una prueba de choque dada es $3/4$. Encuentre la probabilidad de que sobrevivan exactamente dos de los siguientes cuatro componentes que se prueben.

SOLUCIÓN

Suponga que las pruebas son independientes y como $p = 3/4$ para cada una de las cuatro pruebas, obtenemos

$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^2 = \frac{4!}{2!2!} \cdot \frac{3^2}{4^4} = \frac{27}{128}.$$

La distribución binomial deriva su nombre del hecho de que los $n + 1$ términos en la expansión binomial de $(q + p)^n$ corresponden a los diversos valores de $b(x; n, p)$ para $x = 0, 1, 2, \dots, n$. Es decir,

$$\begin{aligned} (q + p)^n &= \binom{n}{0} q^n + \binom{n}{1} p q^{n-1} + \binom{n}{2} p^2 q^{n-2} + \dots + \binom{n}{n} p^n \\ &= b(0; n, p) + b(1; n, p) + b(2; n, p) + \dots + b(n; n, p). \end{aligned}$$

Como $p + q = 1$ vemos que

$$\sum_{x=0}^n b(x; n, p) = 1,$$

condición que debe ser válida para cualquier distribución de probabilidad.

Con frecuencia, nos interesamos en problemas donde se necesita encontrar $P(X < r)$ o $P(a \leq X \leq b)$. Por fortuna, disponemos de las sumas binomiales

$$B(r; n, p) = \sum_{x=0}^r b(x; n, p)$$

y se dan en la tabla A.1 del apéndice para $n = 1, 2, \dots, 20$, y para valores seleccionados de p entre 0.1 y 0.9. Ilustramos el uso de la tabla A.1 con el siguiente ejemplo. ■

Ejemplo 5.5 La probabilidad de que un paciente se recupere de una rara enfermedad sanguínea es 0.4. Si se sabe que 15 personas contraen esta enfermedad, ¿cuál es la probabilidad de que (a) sobrevivan al menos 10, (b) sobrevivan de 3 a 8 y (c) sobrevivan exactamente 5?

SOLUCIÓN

Sea X el número de personas que sobreviven.

$$\begin{aligned} \text{(a) } P(X \geq 10) &= 1 - P(X < 10) = 1 - \sum_{x=0}^9 b(x; 15, 0.4) = 1 - 0.9662 \\ &= 0.0338. \end{aligned}$$

$$\begin{aligned} \text{(b) } P(3 \leq X \leq 8) &= \sum_{x=3}^8 b(x; 15, 0.4) \\ &= \sum_{x=0}^8 b(x; 15, 0.4) - \sum_{x=0}^2 b(x; 15, 0.4) = 0.9050 - 0.0271 \\ &= 0.8779. \end{aligned}$$

$$\begin{aligned} \text{(c) } P(X = 5) &= b(5; 15, 0.4) = \sum_{x=0}^5 b(x; 15, 0.4) - \sum_{x=0}^4 b(x; 15, 0.4) \\ &= 0.4032 - 0.2173 = 0.1859. \quad \blacksquare \end{aligned}$$

Áreas de aplicación

De los ejemplos 5.4 y 5.5 debe quedar claro que la distribución binomial encuentra aplicaciones en muchos campos científicos. Un ingeniero industrial está profundamente interesado en la “proporción de defectuosos” en un proceso industrial. A menudo, las mediciones de control de calidad y los esquemas de muestreo para procesos se basan en la distribución binomial. Ésta se aplica en cualquier situación industrial donde el resultado de un proceso es dicotómico y los resultados del proceso son independientes, y la probabilidad de éxito es constante de una prueba a otra. La distribución binomial también se utiliza de manera extensa en aplicaciones médicas y militares. En ambos casos, un resultado de éxito o fracaso es importante. Por ejemplo, “cura” o “no cura” es importante en el trabajo farmacéutico, mientras “dar en el blanco” o “errar” a menudo es la interpretación del resultado de lanzar un misil guiado.

Como la distribución de probabilidad de cualquier variable aleatoria binomial depende sólo de los valores que toman los parámetros n , p y q , parecería razonable suponer que la media y la varianza de una variable aleatoria binomial también dependen de los valores que toman estos parámetros. En realidad, esto es cierto, y en el teorema 5.2 derivamos las fórmulas generales como funciones de n , p y q que se

pueden utilizar para calcular la media y la varianza de cualquier variable aleatoria binomial.

Teorema 5.2

La media y la varianza de la distribución binomial $b(x; n, p)$ son

$$\mu = np \quad \text{y} \quad \sigma^2 = npq.$$

PRUEBA Representemos el resultado de la j -ésima prueba mediante la variable aleatoria de Bernoulli I_j , que toma los valores 0 y 1 con probabilidades q y p , respectivamente.

Por tanto, en un experimento binomial el número de éxitos se puede escribir como la suma de las n variables indicadoras independientes. De aquí,

$$X = I_1 + I_2 + \cdots + I_n.$$

La media de cualquier I_j es $E(I_j) = (0 \cdot q) + (1 \cdot p) = p$. Por tanto, con el uso del corolario 2 del teorema 4.7, la media de la distribución binomial es

$$\mu = E(X) = E(I_1) + E(I_2) + \cdots + E(I_n) = \underbrace{p + p + \cdots + p}_{n \text{ términos}} = np.$$

La varianza de cualquier I_j es

$$\sigma_{I_j}^2 = E[(I_j - p)^2] = E(I_j^2) - p^2 = (0)^2q + (1)^2p - p^2 = p(1 - p) = pq.$$

Al extender el corolario 2 del teorema 4.10 al caso de n variables independientes, la varianza de la distribución binomial es

$$\sigma_X^2 = \sigma_{I_1}^2 + \sigma_{I_2}^2 + \cdots + \sigma_{I_n}^2 = \underbrace{pq + pq + \cdots + pq}_{n \text{ términos}} = npq.$$

Ejemplo 5.6 Encuentre la media y la varianza de la variable aleatoria binomial del ejemplo 5.5, y después utilice el teorema de Chebyshev para interpretar el intervalo $\mu \pm 2\sigma$.

SOLUCIÓN

Como el ejemplo 5.5 fue un experimento binomial con $n = 15$ y $p = 0.4$, por el teorema 5.2, tenemos

$$\mu = (15)(0.4) = 6 \quad \text{y} \quad \sigma^2 = (15)(0.4)(0.6) = 3.6.$$

Al tomar la raíz cuadrada de 3.6, encontramos que $\sigma = 1.897$. Por tanto el intervalo que se requiere es $6 \pm (2)(1.897)$, o de 2.206 a 9.794. El teorema de Chebyshev afirma que el número de recuperaciones entre 15 pacientes sujetos a la enfermedad mencionada tiene una probabilidad de al menos $3/4$ de caer entre 2.206 y 9.794 o, como los datos son discretos, entre 3 y 9 inclusive. ■

Como el lector se puede dar cuenta ahora, en muchas aplicaciones hay más de dos resultados posibles. Para tomar un ejemplo del campo de la genética, el color de las crías de los conejillos de indias pueden ser rojos, negros o blancos. A menudo la dicotomía "defectuoso" o "no defectuoso" en situaciones de ingeniería es ciertamente una gran simplificación. En realidad, con frecuencia hay más de dos categorías que caracterizan artículos o partes que salen de una línea de producción.

Experimentos multinomiales

El experimento binomial se convierte en un **experimento multinomial** si cada prueba tiene más de dos resultados posibles. Por ello la clasificación de un producto fabricado como ligero, pesado o aceptable y el registro de los accidentes en cierto crucero de acuerdo con el día de la semana constituyen experimentos multinomiales. Extraer una carta de una baraja con *reemplazo* también es un experimento multinomial si los cuatro palos son los resultados de interés.

En general, si una prueba dada puede tener como consecuencia cualquiera de los k resultados posibles E_1, E_2, \dots, E_k con probabilidades p_1, p_2, \dots, p_k , entonces la **distribución multinomial** dará la probabilidad de que E_1 ocurra x_1 veces; E_2 ocurra x_2 veces; ...; y E_k ocurra x_k veces en n pruebas independientes, donde

$$x_1 + x_2 + \dots + x_k = n.$$

Denotaremos esta distribución de probabilidad conjunta como

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n).$$

Claramente, $p_1 + p_2 + \dots + p_k = 1$, pues el resultado de cada prueba debe ser uno de los k resultados posibles.

Para derivar la fórmula general, procedemos como en el caso binomial. Como las pruebas son independientes, cualquier orden especificado que produzca x_1 resultados para E_1 , x_2 para E_2 , ..., x_k para E_k ocurrirá con probabilidad $p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$. El número total de órdenes que den resultados similares para las n pruebas es igual al número de particiones de n pruebas en k grupos con x_1 en el primer grupo; x_2 en el segundo grupo; ...; y x_k en el k -ésimo grupo. Esto se puede hacer en

$$\binom{n}{x_1, x_2, \dots, x_k} = \frac{n!}{x_1! x_2! \dots x_k!}$$

formas. Como todas las particiones son mutuamente excluyentes y ocurren con igual probabilidad, obtenemos la distribución multinomial al multiplicar la probabilidad para un orden específico por el número total de particiones.

Distribución multinomial

Si una prueba dada puede conducir a los k resultados E_1, E_2, \dots, E_k con probabilidades p_1, p_2, \dots, p_k , entonces la distribución de probabilidad de las variables aleatorias X_1, X_2, \dots, X_k , que representan el número de ocurrencias para E_1, E_2, \dots, E_k en n pruebas independientes es

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n) = \binom{n}{x_1, x_2, \dots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

con

$$\sum_{i=1}^k x_i = n \quad \text{y} \quad \sum_{i=1}^k p_i = 1.$$

La distribución multinomial deriva su nombre del hecho que los términos de la expansión multinomial de $(p_1 + p_2 + \cdots + p_k)^n$ corresponden a todos los posibles valores de

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n).$$

Ejemplo 5.7 Si se lanza seis veces un par de dados, ¿cuál es la probabilidad de obtener un total de siete u 11 dos veces, un par igual una vez y cualquier otra combinación tres veces?

SOLUCIÓN

Listamos los siguientes eventos posibles,

E_1 : ocurre un total de 7 u 11,

E_2 : ocurre un par igual,

E_3 : no ocurre ni un par ni un total de 7 u 11.

Las probabilidades correspondientes para una prueba dada son $p_1 = 2/9$, $p_2 = 1/6$, y $p_3 = 11/18$. Estos valores permanecen constantes para todas las seis pruebas. Al usar la distribución multinomial con $x_1 = 2$, $x_2 = 1$, y $x_3 = 3$, encontramos que la probabilidad que se requiere es

$$\begin{aligned} f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) &= \binom{6}{2, 1, 3} \left(\frac{2}{9}\right)^2 \left(\frac{1}{6}\right)^1 \left(\frac{11}{18}\right)^3 \\ &= \frac{6!}{2! 1! 3!} \cdot \frac{2^2}{9^2} \cdot \frac{1}{6} \cdot \frac{11^3}{18^3} = 0.1127. \quad \blacksquare \end{aligned}$$

Ejercicios

1. Se selecciona un empleado de un equipo de 10 para supervisar cierto proyecto mediante la selección de una etiqueta al azar de una caja que contiene 10 etiquetas numeradas del 1 al 10. Encuentre la fórmula para la distribución de probabilidad de X que representa el número en la etiqueta que se saca. ¿Cuál es la probabilidad de que el número que se extrae sea menor que 4?

2. Se dan dos altavoces idénticos a doce personas para que escuchen diferencias, si hay alguna. Suponga que estas personas responden sólo adivinando. Encuentre la probabilidad de que tres personas afirmen escuchar una diferencia entre los dos altavoces.

3. Encuentre la media y la varianza de la variable aleatoria X del ejercicio 1.

4. En cierto distrito de la ciudad la necesidad de dinero para comprar drogas se establece como la razón del 75% de los robos. Encuentre la probabilidad de que entre los siguientes cinco casos de robo que se reporten en este distrito,

- (a) exactamente dos resulten de la necesidad de dinero para comprar drogas;
- (b) al menos tres resulten de la necesidad de dinero para comprar drogas.

5. De acuerdo con la *Chemical Engineering Progress* (noviembre de 1990), aproximadamente 30% de todas las fallas de operación de tuberías en plantas químicas son ocasionadas por errores del operador.

- (a) ¿Cuál es la probabilidad de que de las siguientes 20 fallas al menos 10 se deban a error del operador?
- (b) ¿Cuál es la probabilidad de que no más de cuatro de 20 fallas se deban a error del operador?
- (c) Suponga, para una planta particular, que de la muestra aleatoria de 20 de tales fallas exactamente cinco sean errores de operación. ¿Considera que la cifra de 30% anterior se aplique a esta planta? Comente.

6. De acuerdo con una investigación de la Administrative Management Society, la mitad de las compañías estadounidenses dan a sus empleados cuatro semanas de vacaciones después de 15 años de trabajar en la compañía. Encuentre la probabilidad de que entre seis compañías investigadas al azar, el número que da a sus empleados cuatro semanas de vacaciones después de 15 años de servicio es

- (a) cualquiera entre dos y cinco;
- (b) menor que tres.

7. Un prominente médico afirma que 70% de las personas con cáncer de pulmón son fumadores empedernidos. Si su aseveración es correcta:

- (a) encuentre la probabilidad de que de 10 de tales pacientes admitidos recientemente en un hospital, menos de la mitad sean fumadores empedernidos.
- (b) encuentre la probabilidad de que de 20 de tales pacientes que recientemente hayan ingresado a un hospital, menos de la mitad sean fumadores empedernidos.

8. De acuerdo con un estudio publicado por un grupo de sociólogos de la Universidad de Massachusetts, aproximadamente 60% de los consumidores de Valium en el estado de Massachusetts tomaron Valium por primera vez debido a problemas psicológicos. Encuentre la probabilidad de que entre los siguientes ocho consumidores entrevistados de este estado:

- (a) exactamente tres comenzaron a tomar Valium por problemas psicológicos;
- (b) al menos cinco comenzaron a consumir Valium por problemas que no fueron psicológicos.

9. Al probar cierta clase de neumático para camión en un terreno escabroso, se encuentra que 25% de los camiones no completaban la prueba sin ponchaduras. De los siguientes 15 camiones probados, encuentre la probabilidad de que

- (a) de tres a seis tengan ponchaduras;
- (b) menos de cuatro tengan ponchaduras;
- (c) más de cinco tengan ponchaduras.

10. De acuerdo con un reporte publicado en la revista *Parade*, una encuesta a nivel nacional de la Universidad de Michigan a estudiantes universitarios de último año revela que casi el 70% desaprobaban el consumo diario de marihuana. Si se seleccionan 12 estudiantes al azar y se les pide su opinión, encuentre la probabilidad de que el número de los que desaprobaban fumar marihuana todos los días sea

- (a) cualquier valor de siete a nueve;
- (b) a lo más cinco;
- (c) no menos de ocho.

11. La probabilidad de que un paciente se recupere de una delicada operación de corazón es 0.9. ¿Cuál es la probabilidad de que exactamente cinco de los siguientes siete pacientes intervenidos sobrevivan?

12. Un ingeniero de control de tráfico reporta que 75% de los vehículos que pasan por un punto de verificación son del interior del estado. ¿Cuál es la probabilidad de que menos de cuatro de los siguientes nueve vehículos sean de otro estado?

13. Un estudio examinó las actitudes nacionales acerca de los antidepresivos. El estudio reveló que aproximadamente 70% cree que los "antidepresivos en realidad no curan nada, sólo encubren el problema real". De acuerdo con este estudio, ¿cuál es la probabilidad de que al menos tres de las siguientes cinco personas seleccionadas al azar sean de esta opinión?

14. Se sabe que el porcentaje de victorias para que el equipo de baloncesto Toros de Chicago pasara a las finales en la temporada 1996-97 fue 87.7. Redondee 87.7 a 90 a fin de utilizar la tabla A.1.

- (a) ¿Cuál es la probabilidad de que los Toros ganen los primeros cuatro de los siete de la serie final?
- (b) ¿Cuál es la probabilidad de que los Toros ganen toda la serie final?
- (c) ¿Qué suposición importante se realiza para contestar las partes a y b?

15. Se sabe que 40% de los ratones inoculados con un suero quedan protegidos de cierta enfermedad. Si se inoculan cinco ratones, encuentre la probabilidad de que

- (a) ninguno contraiga la enfermedad;

(b) menos de dos contraigan la enfermedad;

(c) más de tres contraigan la enfermedad.

16. Suponga que los motores de un aeroplano operan de forma independiente y fallan con probabilidad igual a 0.4. Suponga que un avión tiene un vuelo seguro si funcionan al menos la mitad de sus motores; determine si un avión de cuatro motores o uno de dos tiene la probabilidad más alta de un vuelo exitoso.

17. Si X representa el número de personas del ejercicio 13 que creen que los antidepresivos no curan sino que sólo encubren el problema real, encuentre la media y la varianza de X cuando se seleccionan al azar cinco personas y después utilice el teorema de Chebyshev para interpretar el intervalo $\mu \pm 2\sigma$.

18. (a) ¿En el ejercicio 9 cuántos de los 15 camiones esperaría que tuviera ponchaduras?

(b) De acuerdo con el teorema de Chebyshev, ¿hay una probabilidad de al menos $\frac{3}{4}$ de que el número de camiones entre los siguientes 15 que tengan ponchaduras caiga en cuál intervalo?

19. Un estudiante que maneja hacia su escuela encuentra un semáforo. Este semáforo permanece verde por 35 segundos, ámbar cinco segundos y rojo 60 segundos. Suponga que el estudiante va a la escuela toda la semana entre 8:00 y 8:30. Sea X_1 el número de veces que encuentra una luz verde, X_2 el número de veces que encuentra una luz ámbar y X_3 el número de veces que encuentra una luz roja. Encuentre la distribución conjunta de X_1 , X_2 y X_3 .

20. De acuerdo con el periódico *USA Today* (18 de marzo de 1997) de cuatro millones de trabajadores en la fuerza de trabajo, 5.8% resultó positivo en una prueba de drogas. De los que resultaron positivos, 22.5% fueron usuarios de cocaína y 54.4% de marihuana.

(a) ¿Cuál es la probabilidad de que de 10 trabajadores que resultaron positivos, dos sean usuarios de cocaína, cinco de marihuana y tres de otras drogas?

(b) ¿Cuál es la probabilidad de que de 10 trabajadores que resultaron positivos, todos sean usuarios de marihuana?

(c) ¿Cuál es la probabilidad de que de 10 trabajadores que resultaron positivos, ninguno sea usuario de cocaína?

21. La superficie de un tablero circular para dardos tiene un pequeño círculo central llamado ojo de buey y 20 regiones en forma de rebanada de pastel numeradas del 1 al 20. Cada una de las regiones en forma de rebanada de pastel está dividida en tres partes de modo que una persona que lanza un dardo que cae en

un número específico obtiene un puntaje igual al valor del número, el doble del número o el triple de éste, según en cuál de las tres partes cae el dardo. Si una persona atina al ojo de buey con probabilidad 0.01, atina un doble con probabilidad 0.10, un triple con probabilidad 0.05 y no le atina al tablero con probabilidad 0.02, ¿cuál es la probabilidad de que siete lanzamientos tengan como resultado ningún centro, ningún triple, un doble dos veces y una falla completa?

22. De acuerdo con la teoría genética, cierta cruce de conejillos de indias tendrá crías rojas, negras y blancas con la relación 8:4:4. Encuentre la probabilidad de que entre ocho crías cinco sean rojas, dos negras y una blanca.

23. Las probabilidades de que un delegado a cierta convención llegue por aire, autobús, automóvil o tren son 0.4, 0.2, 0.3 y 0.1, respectivamente. ¿Cuál es la probabilidad de que entre nueve delegados a esta convención seleccionados al azar, tres lleguen por aire, tres por autobús, uno en automóvil y dos en tren?

24. Un ingeniero de seguridad afirma que sólo 40% de todos los trabajadores utilizan cascos de seguridad cuando almuerzan en el lugar de trabajo. Suponga que esta afirmación es cierta, encuentre la probabilidad de que cuatro de seis trabajadores elegidos al azar utilicen sus cascos mientras almuerzan en el lugar de trabajo.

25. Suponga que para un embarque muy grande de chips de circuitos integrados, la probabilidad de falla para cualquier chip es 0.10. Suponga que se cumplen las suposiciones en que se basan las distribuciones binomiales y encuentre la probabilidad de que a lo más tres chips fallen en una muestra aleatoria de 20.

26. Suponga que seis de 10 accidentes automovilísticos se deben principalmente a una violación del límite de velocidad, encuentre la probabilidad de que entre ocho accidentes automovilísticos seis se deban principalmente a una violación del límite de velocidad

(a) mediante el uso de la fórmula para la distribución binomial;

(b) con el uso de la tabla binomial.

27. Si la probabilidad de que una luz fluorescente tenga una vida útil de al menos 800 horas es 0.9, encuentre las probabilidades de que entre 20 de tales luces

(a) exactamente 18 tengan una vida útil de al menos 800 horas;

(b) al menos 15 tengan una vida útil de al menos 800 horas;

(c) al menos dos *no* tengan una vida útil de al menos 800 horas.

28. Un fabricante sabe que en promedio 20% de los tostadores eléctricos que fabrica requerirán reparaciones dentro de un año después de su venta. Cuando se seleccionan al azar 20 tostadores, encuentre los números x y y apropiados tales que

- (a) la probabilidad de que al menos x de ellos requieran reparaciones sea menor que 0.5;
 (b) la probabilidad de que al menos y de ellos *no* requieran reparaciones sea mayor que 0.8.

5.4 Distribución hipergeométrica

La manera más simple de ver la diferencia entre la distribución binomial de la sección 5.3 y la distribución hipergeométrica está en la forma en que se realiza el muestreo. Los tipos de aplicaciones de la distribución hipergeométrica son muy similares a los de la binomial. Nos interesamos en el cálculo de probabilidades para el número de observaciones que caen en una categoría particular. Pero en el caso de la binomial, se requiere la independencia entre las pruebas. Como resultado, si se aplica la binomial a, digamos, tomar muestras de un lote de artículos (barajas, lotes de artículos producidos), el muestreo se debe efectuar **con reemplazo** de cada artículo después de que se observe. Por otro lado, la distribución hipergeométrica no requiere independencia y se basa en el muestreo que se realiza **sin reemplazo**.

Las aplicaciones de la distribución hipergeométrica se encuentran en muchas áreas, con gran uso en muestreo de aceptación, pruebas electrónicas y garantía de calidad. Obviamente, para muchos de estos campos el muestreo se realiza a expensas del artículo que se prueba. Es decir, el artículo se destruye y por ello no se puede reemplazar en la muestra. Así, el muestreo sin reemplazo es necesario. Utilizamos un ejemplo simple con barajas para ilustración.

Si deseamos encontrar la probabilidad de observar tres cartas rojas en cinco extracciones de una baraja ordinaria de 52 cartas, la distribución binomial de la sección 5.3 no se aplica a menos que cada carta se reemplace y que el paquete se baraje antes de que se extraiga la siguiente carta. Para resolver el problema de muestrear sin reemplazo, replanteamos el problema. Si se sacan cinco cartas al azar, nos interesamos en la probabilidad de seleccionar tres cartas rojas de las 26 disponibles y dos negras de las 26 cartas negras de que dispone la baraja. Hay $\binom{26}{3}$ formas de seleccionar tres cartas rojas, y para cada una de estas formas podemos elegir dos cartas negras de $\binom{26}{2}$ maneras. Por tanto, el número total de formas de seleccionar tres cartas rojas y dos negras en cinco extracciones es el producto $\binom{26}{3}\binom{26}{2}$. El número total de formas de seleccionar cualesquiera cinco cartas de las 52 disponibles es $\binom{52}{5}$. Por ello, la probabilidad de seleccionar cinco cartas sin reemplazo de las cuales tres sean rojas y dos negras está dada por

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = \frac{(26!/3!23!)(26!/2!24!)}{(52!/5!47!)} = 0.3251.$$

En general, nos interesa la probabilidad de seleccionar x éxitos de los k artículos considerados como éxito y $n - x$ fracasos de los $N - k$ artículos que se consideran fracasos cuando se selecciona una muestra aleatoria de tamaño n de N artículos. Esto se conoce como **experimento hipergeométrico**; es decir, uno que posee las siguientes dos propiedades:

1. Se selecciona sin reemplazo una muestra aleatoria de tamaño n de N artículos.
2. k de los N artículos se pueden clasificar como éxitos y $N - k$ se clasifican como fracasos.

El número X de éxitos de un experimento hipergeométrico se denomina **variable aleatoria hipergeométrica**. En consecuencia, la distribución de probabilidad de la variable hipergeométrica se llama **distribución hipergeométrica**, y sus valores se denotan como $h(x; N, n, k)$, debido a que dependen del número de éxitos k en el conjunto N del que seleccionamos n artículos.

Ejemplo 5.8 Se selecciona al azar un comité de cinco personas entre tres químicos y cinco físicos. Encuentre la distribución de probabilidad para el número de químicos en el comité.

SOLUCIÓN

Sea la variable aleatoria X el número de químicos en el comité. Se satisfacen las dos propiedades de un experimento hipergeométrico. Por ello

$$P(X = 0) = h(0; 8, 5, 3) = \frac{\binom{3}{0} \binom{5}{5}}{\binom{8}{5}} = \frac{1}{56}$$

$$P(X = 1) = h(1; 8, 5, 3) = \frac{\binom{3}{1} \binom{5}{4}}{\binom{8}{5}} = \frac{15}{56}$$

$$P(X = 2) = h(2; 8, 5, 3) = \frac{\binom{3}{2} \binom{5}{3}}{\binom{8}{5}} = \frac{30}{56}$$

$$P(X = 3) = h(3; 8, 5, 3) = \frac{\binom{3}{3} \binom{5}{2}}{\binom{8}{5}} = \frac{10}{56}$$

En forma tabular la distribución hipergeométrica de X es como sigue:

x	0	1	2	3
$h(x; 8, 5, 3)$	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{30}{56}$	$\frac{10}{56}$

La distribución de probabilidad se establece mediante la fórmula

$$h(x; 8, 5, 3) = \frac{\binom{3}{x} \binom{5}{5-x}}{\binom{8}{5}}, \quad x = 0, 1, 2, 3.$$

Generalicemos el ejemplo 5.8 para encontrar una fórmula para $h(x; N, n, k)$. El número total de muestras de tamaño n que se eligen de N artículos es $\binom{N}{n}$. Se supone que estas muestras tienen igual probabilidad. Hay $\binom{k}{x}$ formas de seleccionar x éxitos de los k disponibles y por cada una de estas formas podemos elegir $n - x$ fracasos en $\binom{N-k}{n-x}$ formas. De esta manera, el número total de muestras favorables entre las $\binom{N}{n}$ muestras posibles está dado por $\binom{k}{x} \binom{N-k}{n-x}$. De aquí, tenemos la siguiente definición. ■

Distribución hipergeométrica

La distribución de probabilidad de la variable aleatoria hipergeométrica X , el número de éxitos en una muestra aleatoria de tamaño n que se selecciona de N artículos de los que k se denominan **éxito** y $N - k$ **fracaso**, es

$$h(x; N, n, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, 2, \dots, n.$$

Ejemplo 5.9 Lotes de 40 componentes cada uno se denominan aceptables si no contienen más de tres defectuosos. El procedimiento para muestrear el lote es la selección de cinco componentes al azar y rechazar el lote si se encuentra un componente defectuoso. ¿Cuál es la probabilidad de que se encuentre exactamente un defectuoso en la muestra si hay tres defectuosos en todo el lote?

SOLUCIÓN

Si se utiliza la distribución hipergeométrica con $n = 5$, $N = 40$, $k = 3$ y $x = 1$ encontramos que la probabilidad de obtener un defectuoso es

$$h(1; 40, 5, 3) = \frac{\binom{3}{1} \binom{37}{4}}{\binom{40}{5}} = 0.3011. \quad \blacksquare$$

Teorema 5.3

La media y la varianza de la distribución hipergeométrica $h(x; N, n, k)$ son

$$\mu = \frac{nk}{N} \quad \text{y} \quad \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right).$$

A continuación demostramos el resultado sólo para la media.

PRUEBA Para encontrar la media de la distribución hipergeométrica, escribimos

$$\begin{aligned} E(X) &= \sum_{x=0}^n x \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} = k \sum_{x=1}^n \frac{(k-1)!}{(x-1)!(k-x)!} \cdot \frac{\binom{N-k}{n-x}}{\binom{N}{n}} \\ &= k \sum_{x=1}^n \frac{\binom{k-1}{x-1} \binom{N-k}{n-x}}{\binom{N}{n}}. \end{aligned}$$

Al hacer $y = x - 1$, encontramos que se convierte en

$$E(X) = k \sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{N-k}{n-1-y}}{\binom{N}{n}}.$$

Al escribir

$$\binom{N-k}{n-1-y} = \binom{(N-1)-(k-1)}{n-1-y}$$

y

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)}{n(n-1)},$$

obtenemos

$$E(X) = \frac{nk}{N} \sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{(N-1)-(k-1)}{n-1-y}}{\binom{N-1}{n-1}} = \frac{nk}{N},$$

debido a que la sumatoria representa el total de todas las probabilidades en un experimento hipergeométrico cuando se seleccionan al azar $N-1$ artículos de $N-1$, de los que $k-1$ se consideran éxitos.

Ejemplo 5.10 Encuentre la media y la varianza de la variable aleatoria del ejemplo 5.9, y después utilice el teorema de Chebyshev para interpretar el intervalo $\mu \pm 2\sigma$.

SOLUCIÓN

Como el ejemplo 5.9 fue un experimento hipergeométrico con $n = 40$, $n = 5$, y $k = 3$, entonces por el teorema 5.3 tenemos

$$\mu = \frac{(5)(3)}{40} = \frac{3}{8} = 0.375$$

y

$$\sigma^2 = \left(\frac{40-5}{39}\right)(5)\left(\frac{3}{40}\right)\left(1 - \frac{3}{40}\right) = 0.3113.$$

Al obtener la raíz cuadrada de 0.3113, encontramos que $\sigma = 0.558$. De aquí, el intervalo que se requiere es $0.375 \pm (2)(0.558)$, o de -0.741 a 1.491 . El teorema de Chebyshev establece que el número de componentes defectuosos que se obtienen cuando se seleccionan al azar cinco de un lote de 40 componentes de los que tres son defectuosos tiene una probabilidad de al menos $3/4$ de caer entre -0.741 y 1.491 . Es decir, al menos tres cuartos de las veces los cinco componentes incluirán menos de dos defectuosos. ■

Relación con la distribución binomial

En este capítulo discutimos varias distribuciones discretas importantes que tienen amplia aplicabilidad. Muchas de estas distribuciones se relacionan bien entre ellas. El estudiante principiante debe tener una clara comprensión de estas relaciones. Hay una relación interesante entre la distribución hipergeométrica y la binomial. Como se podría esperar, si n es pequeño comparado con N , la naturaleza de los N artículos cambia muy poco en cada prueba. Así la cantidad k/n juega el papel del parámetro binomial p . Como consecuencia, la distribución binomial se puede ver como una versión de población grande de las distribuciones hipergeométricas. La media y la varianza entonces se obtienen de las fórmulas

$$\mu = np = \frac{nk}{N}, \quad \sigma^2 = npq = n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right).$$

Al comparar estas fórmulas con las del teorema 5.3, vemos que la media es la misma mientras que la varianza difiere por un factor de corrección $(N-n)/(N-1)$, que es insignificante cuando n es pequeña en relación con N .

Ejemplo 5.11 Un fabricante de neumáticos para automóvil reporta que entre un cargamento de 5000 que se manda a un distribuidor local, 1000 están ligeramente manchados. Si se compran al azar 10 de estos neumáticos al distribuidor, ¿cuál es la probabilidad de que exactamente tres estén manchados?

SOLUCIÓN

Como $N = 5000$ es grande en relación con la muestra de tamaño $n = 10$, aproximaremos la probabilidad que se desea con el uso de la distribución binomial. La probabilidad de obtener un neumático manchado es 0.2. Por tanto, la probabilidad de obtener exactamente tres manchados es

$$h(3; 5000, 10, 1000) \approx b(3; 10, 0.2)$$

$$\begin{aligned}
 &= \sum_{x=0}^3 b(x; 10, 0.2) - \sum_{x=0}^2 b(x; 10, 0.2) \\
 &= 0.8791 - 0.6778 = 0.2013.
 \end{aligned}$$

La distribución hipergeométrica se puede extender para tratar el caso donde los N artículos se pueden dividir en k celdas A_1, A_2, \dots, A_k con a_1 elementos en la primera celda, a_2 en la segunda, \dots , a_k elementos en la k -ésima celda. Nos interesamos ahora en la probabilidad de que una muestra aleatoria de tamaño n dé x_1 elementos de A_1 , x_2 elementos de A_2 , \dots , x_k y de A_k . Representemos esta probabilidad por

$$f(x_1, x_2, \dots, x_k; a_1, a_2, \dots, a_k, N, n).$$

Para obtener una fórmula general, notamos que el número total de muestras de tamaño n que se pueden elegir a partir de N artículos es aún $\binom{N}{n}$. Hay $\binom{a_1}{x_1}$ formas de seleccionar x_1 artículos de los que hay en A_1 , y para cada uno de éstos podemos elegir x_2 de los de A_2 en $\binom{a_2}{x_2}$ formas. Por tanto, podemos seleccionar x_1 artículos de A_1 y x_2 de A_2 en $\binom{a_1}{x_1} \binom{a_2}{x_2}$ formas. Si continuamos de esta forma, podemos seleccionar todos los n artículos que consisten en x_1 de A_1 , x_2 de A_2 , \dots , y x_k de A_k en

$$\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k} \text{ formas.}$$

La distribución de probabilidad que se requiere se define ahora como sigue. ■

Distribución hipergeométrica multivariada

Si N artículos se pueden dividir en las k celdas A_1, A_2, \dots, A_k con a_1, a_2, \dots, a_k elementos, respectivamente, entonces la distribución de probabilidad de las variables aleatorias X_1, X_2, \dots, X_k , que representan el número de elementos que se seleccionan de A_1, A_2, \dots, A_k en una muestra aleatoria de tamaño n , es

$$f(x_1, x_2, \dots, x_k; a_1, a_2, \dots, a_k, N, n) = \frac{\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k}}{\binom{N}{n}}$$

$$\text{con } \sum_{i=1}^k x_i = n \text{ y } \sum_{i=1}^k a_i = N.$$

Ejemplo 5.12 Un grupo de 10 individuos se usa para un estudio biológico. El grupo contiene tres personas con sangre tipo O, cuatro con sangre tipo A y tres con tipo B. ¿Cuál es la probabilidad de que una muestra aleatoria de cinco contenga una persona con sangre tipo O, dos personas con tipo A y dos personas con tipo B?

SOLUCIÓN

Al usar la extensión de la distribución hipergeométrica con $x_1 = 1$, $x_2 = 2$, $x_3 = 2$, $a_1 = 3$, $a_2 = 4$, $a_3 = 3$, $N = 10$ y $n = 5$, encontramos que la probabilidad que se desea es

$$f(1, 2, 2; 3, 4, 3, 10, 5) = \frac{\binom{3}{1} \binom{4}{2} \binom{3}{2}}{\binom{10}{5}} = \frac{3}{14} \quad \blacksquare$$

Ejercicios

1. Si se reparten siete cartas de una baraja ordinaria de 52 cartas, ¿cuál es la probabilidad de que
(a) exactamente dos de ellas sean mayores?
(b) al menos una de ellas sea una reina?

2. Para evitar la detección en la aduana, un viajero coloca seis tabletas de narcótico en una botella que contiene nueve píldoras de vitamina similares en apariencia. Si el oficial de la aduana selecciona tres de las tabletas al azar para su análisis, ¿cuál es la probabilidad de que el viajero sea arrestado por posesión ilegal de narcóticos?

3. El dueño de una casa planta seis bulbos seleccionados al azar de una caja que contiene 5 bulbos de tulipán y cuatro de narciso. ¿Cuál es la probabilidad de que plante dos bulbos de narciso y cuatro de tulipán?

4. De un lote de 10 misiles, se seleccionan cuatro al azar y se lanzan. Si el lote contiene tres misiles defectuosos que no explotarán, ¿cuál es la probabilidad de que
(a) los cuatro exploten?
(b) a lo más dos fallen?

5. Se selecciona al azar un comité de tres personas a partir de cuatro doctores y dos enfermeras. Escriba una fórmula para la distribución de probabilidad de la variable aleatoria X que representa el número de doctores en el comité. Encuentre $P(2 \leq X \leq 3)$.

6. ¿Cuál es la probabilidad de que una mesera se rehusa a servir bebidas alcohólicas a sólo dos menores si verifica al azar las identificaciones de cinco estudiantes de entre nueve estudiantes de los cuales cuatro no tienen la edad legal?

7. Una compañía está interesada en evaluar su procedimiento de inspección actual en embarques de 50 artículos idénticos. El procedimiento es tomar una muestra de cinco y pasar el embarque si no se encuentran más de dos defectuosos. ¿Qué proporción del 20% de embarques defectuosos se aceptará?

8. Una compañía fabricante utiliza un esquema de aceptación de producción de artículos antes de que se embarquen. El plan tiene dos etapas. Se preparan cajas de 25 artículos para su embarque y se prueba una muestra de tres en busca de defectuosos. Si se encuentra alguno defectuoso, toda la caja se regresa para verificar el 100%. Si no se encuentran defectuosos, la caja se embarca.

(a) ¿Cuál es la probabilidad de que una caja que contiene tres defectuosos se embarque?
(b) ¿Cuál es la probabilidad de que una caja que contenga sólo un artículo defectuoso se regrese para su revisión?

9. Suponga que la compañía fabricante del ejercicio 8 decide cambiar su esquema de aceptación. Bajo el nuevo esquema un inspector toma un artículo al azar, lo inspecciona y después lo reemplaza en la caja; un segundo inspector hace lo mismo. Finalmente, un tercer inspector lleva a cabo el mismo procedimiento. La caja no se embarca si cualquiera de los tres encuentra un defectuoso. Responda el ejercicio 8 bajo este nuevo plan.

10. En el ejercicio 4, ¿cuántos misiles defectuosos se pueden incluir entre los cuatro que se seleccionan? Utilice el teorema de Chebyshev para describir la variabilidad del número de misiles defectuosos que se incluyen cuando se seleccionan cuatro de varios lotes cada uno de tamaño 10 con tres misiles defectuosos.

11. Si a una persona se le reparten varias veces 13 cartas de una baraja ordinaria de 52 cartas, ¿cuántas cartas de corazones por mano puede esperar? ¿Entre cuáles dos valores esperaría que cayera el número de corazones al menos 75% de las veces?

12. Se estima que 4000 de los 10,000 residentes de una ciudad que votan están en contra de un nuevo impuesto sobre ventas. Si se seleccionan al azar 15 votantes y se les pide su opinión, ¿cuál es la probabilidad de que a lo más siete estén a favor del nuevo impuesto?

13. Una ciudad vecina considera una petición de anexión de 1200 residencias contra una subdivisión del condado. Si los ocupantes de la mitad de las residencias objetan la anexión, ¿cuál es la probabilidad de que en una muestra aleatoria de 10 al menos tres estén a favor de la petición de anexión?
14. Entre 150 empleados de IRS en una ciudad grande, sólo 30 son mujeres. Si se eligen al azar 10 de los aspirantes para que proporcionen asistencia libre de impuestos a los residentes de esta ciudad, utilice la aproximación binomial a la hipergeométrica para encontrar la probabilidad de que el menos se seleccionen tres mujeres.
15. Una encuesta a nivel nacional de la Universidad de Michigan a 17,000 estudiantes de último año revela que casi el 70% desapruueba el consumo diario de marihuana. Si se seleccionan al azar 18 de estos estudiantes y se les pide su opinión, ¿cuál es la probabilidad de que más de nueve pero menos de 14 desaprueben el consumo de marihuana?
16. Encuentre la probabilidad de que cuando se le reparta una mano de bridge de 13 cartas tenga cinco de espadas, dos de corazones, tres de diamantes y tres de bastos.
17. Un club de estudiantes extranjeros tiene como miembros a dos canadienses, tres japoneses, cinco italianos y dos alemanes. Si se selecciona al azar un comité de cuatro, encuentre la probabilidad de que
(a) todas las nacionalidades estén representadas;
(b) todas las nacionalidades estén representadas excepto los italianos.
18. Una urna contiene tres bolas verdes, dos azules, y cuatro rojas. En una muestra aleatoria de cinco, encuentre la probabilidad de que se seleccionen bolas azules y al menos una roja.
19. Estudios de población de biología y el ambiente a menudo etiquetan y sueltan a sujetos a fin de estimar el tamaño y el grado de ciertas características en la población. Se capturan diez animales de cierta población que se piensa extinta (o cerca de la extinción), se etiquetan y se liberan en cierta región. Después de un periodo se selecciona en la región una muestra aleatoria de 15 animales del tipo. ¿Cuál es la probabilidad de que cinco de estos seleccionados sean animales etiquetados si hay 25 animales de este tipo en la región?
20. Una compañía grande tiene un sistema de inspección para los lotes de compresores pequeños que se compran a los vendedores. Un lote típico contiene 15 compresores. En el sistema de inspección se selecciona una muestra aleatoria de cinco y todos se prueban. Suponga que en el lote de 15 hay dos compresores defectuosos.
(a) ¿Cuál es la probabilidad de que para una muestra dada haya un compresor defectuoso?
(b) ¿Cuál es la probabilidad de que la inspección descubra ambos compresores defectuosos?
21. Una fuerza de tarea gubernamental sospecha que algunas fábricas violan los reglamentos contra la contaminación ambiental con respecto a la descarga de cierto tipo de producto. Veinte empresas están bajo sospecha pero no todas se pueden inspeccionar. Suponga que tres de las empresas violan los reglamentos.
(a) ¿Cuál es la probabilidad de que la inspección de cinco empresas no encuentre ninguna violación?
(b) ¿Cuál es la probabilidad de que el plan anterior encuentre dos que violan el reglamento?

5.5 Distribuciones binomial negativa y geométrica

Consideremos un experimento donde las propiedades son las mismas que las que se indican para un experimento binomial, con la excepción de que las pruebas se repetirán hasta que ocurra un número *fijo* de éxitos. Por tanto, en lugar de encontrar la probabilidad de x éxitos en n pruebas, donde n es fija, nos interesamos ahora en la probabilidad de que ocurra el k -ésimo éxito en la x -ésima prueba. Los experimentos de este tipo se llaman **experimentos binomiales negativos**.

Como ejemplo, considere el uso de un medicamento que se sabe que es efectivo en 60% de los casos en que se utiliza. El uso del medicamento se considerará un éxito si es efectivo al proporcionar algún grado de alivio al paciente. Nos interesa encontrar la probabilidad de que el quinto paciente que experimente alivio sea el

séptimo paciente en recibir el medicamento en una semana dada. Designamos éxito con S y fracaso con F , un orden posible para alcanzar el resultado que se desea es $SFSSSFS$, que ocurre con probabilidad

$$(0.6)(0.4)(0.6)(0.6)(0.6)(0.4)(0.6) = (0.6)^5(0.4)^2.$$

Podríamos listar todos los posibles órdenes mediante el reacomodo de la F y las S excepto para el último resultado, que debe ser el quinto éxito. El número total de órdenes posibles es igual al número de particiones de las primeras seis pruebas en dos grupos con dos fracasos asignados a un grupo y los cuatro éxitos asignados al otro grupo. Esto se puede realizar de $\binom{6}{4} = 15$ formas mutuamente excluyentes.

De aquí, si X representa el resultado en el que ocurre el quinto éxito, entonces

$$P(X = 7) = \binom{6}{4}(0.6)^5(0.4)^2 = 0.1866.$$

El número X de pruebas que produce k éxitos en un experimento binomial negativo se llama **variable aleatoria binomial negativa** y su distribución de probabilidad se llama **distribución binomial negativa**. Como sus probabilidades dependen del número de éxitos que se desean y la probabilidad de un éxito en una prueba dada, las denotaremos con el símbolo $b^*(x; k, p)$. Para obtener la fórmula general para $b^*(x; k, p)$, considere la probabilidad de un éxito en la x -ésima prueba precedido por $k - 1$ éxitos y $x - k$ fracasos en un orden específico. Como las pruebas son independientes, podemos multiplicar todas las probabilidades que corresponden a cada resultado que se desea. Cada éxito ocurre con probabilidad p y cada fracaso con probabilidad $q = 1 - p$. Por tanto, la probabilidad para el orden específico, que termina en un éxito, es

$$p^{k-1}q^{x-k}p = p^kq^{x-k}.$$

El número total de puntos muestrales en el experimento que termina en un éxito, después de la ocurrencia de $k - 1$ éxitos y $x - k$ fracasos en cualquier orden, es igual al número de particiones de $x - 1$ pruebas en dos grupos con $k - 1$ éxitos que corresponden a un grupo y $x - k$ fracasos que corresponden al otro grupo.

Este número se especifica con el término $\binom{x-1}{k-1}$, cada uno es mutuamente excluyente y ocurre con igual probabilidad p^kq^{x-k} . Obtenemos la fórmula general al multiplicar p^kq^{x-k} por $\binom{x-1}{k-1}$:

Distribución binomial negativa

Si pruebas independientes repetidas pueden tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X , el número de la prueba en la que ocurre el k -ésimo éxito, es

$$b^*(x; k, p) = \binom{x-1}{k-1}p^kq^{x-k}, \quad x = k, k+1, k+2, \dots$$

Ejemplo 5.13 Encuentre la probabilidad de que una persona que lanza tres monedas obtenga sólo caras o sólo cruces por segunda vez en el quinto lanzamiento.

SOLUCIÓN

Al utilizar la distribución binomial negativa con $x = 5$, $k = 2$, y $p = 1/4$, tenemos

$$b^*(5; 2, \frac{1}{4}) = \binom{4}{1} \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = \frac{4!}{1!3!} \cdot \frac{3^3}{4^5} = \frac{27}{256}.$$

La distribución binomial negativa deriva su nombre del hecho de que cada término de la expansión de $p^k(1-q)^{-k}$ corresponde a los valores de $b^*(x; k, p)$ para $x = k, k+1, k+2, \dots$

Si consideramos el caso especial de la distribución binomial negativa donde $k = 1$, tenemos una distribución de probabilidad para el número de pruebas que se requieren para un solo éxito. Un ejemplo sería lanzar una moneda hasta que salga una cara. Nos podemos interesar en la probabilidad de que ocurra la primera cara en el cuarto lanzamiento. La distribución binomial negativa se reduce a la forma

$$b^*(x; 1, p) = pq^{x-1}, \quad x = 1, 2, 3, \dots$$

Como los términos sucesivos constituyen una progresión geométrica, se acostumbra referirse a este caso especial como la **distribución geométrica** y denotar sus valores con $g(x; p)$. ■

Distribución Geométrica

Si pruebas independientes repetidas pueden tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X , el número de la prueba en el que ocurre el primer éxito, es

$$g(x; p) = pq^{x-1}, \quad x = 1, 2, 3, \dots$$

Ejemplo 5.14 Se sabe que en cierto proceso de fabricación, en promedio, uno de cada 100 artículos está defectuoso. ¿Cuál es la probabilidad de que el quinto artículo que se inspecciona sea el primer defectuoso que se encuentra?

SOLUCIÓN

Al usar la distribución geométrica con $x = 5$ y $p = 0.01$, tenemos

$$g(5; 0.01) = (0.01)(0.99)^4 = 0.0096. \quad \blacksquare$$

Ejemplo 5.15 En "tiempo ocupado" un conmutador telefónico está muy cerca de su capacidad, por lo que los usuarios tienen dificultad al hacer sus llamadas. Puede ser de interés conocer el número de intentos necesario a fin de conseguir un enlace telefónico. Suponga que $p = 0.05$ es la probabilidad de conseguir un enlace durante el tiempo ocupado. Nos interesa conocer la probabilidad de que se necesiten cinco intentos para una llamada exitosa.

SOLUCIÓN

El uso de la distribución geométrica con $x = 5$ y $p = 0.05$ da

$$P(X = x) = g(5; 0.05) = (0.05)(0.95)^4 = 0.041.$$

Muy a menudo, en aplicaciones que tienen que ver con la distribución hipergeométrica, la media y la varianza son importantes. Es así que, en el ejemplo 5.15 el número *esperado* de llamadas necesario para lograr un enlace es muy importante. A continuación se establecen, sin demostración, la media y la varianza de la distribución geométrica. ■

Teorema 5.4

La media y la varianza de una variable aleatoria que sigue la distribución geométrica son

$$\mu = \frac{1}{p}, \quad \sigma^2 = \frac{1-p}{p^2}.$$

Aplicaciones de distribuciones binomial negativa y geométrica

Las áreas de aplicación para las distribuciones binomial negativa y geométrica se vuelven obvias cuando se enfocan los ejemplos de esta sección y los ejercicios que se dedican a estas distribuciones al final de la sección 5.6. En el caso de la distribución geométrica, el ejemplo 5.15 describe una situación donde los ingenieros o administradores intentan determinar cuán ineficiente es un sistema de conmutación telefónica durante periodos ocupados. Claramente en este caso, las pruebas ocurren antes de que un éxito represente un costo. Si hay una alta probabilidad de hacer varios intentos antes del enlace, entonces se deben hacer planes para rediseñar el sistema.

Las aplicaciones de la binomial negativa son similares en naturaleza. Los intentos son costosos en algún sentido y *ocurren en sucesión*. Una alta probabilidad de que se requiera un número "grande" de intentos para experimentar un número fijo de éxitos no es benéfica para el científico o el ingeniero. Considere los escenarios de los ejercicios 14 y 15 de los ejercicios de repaso. En el ejercicio 15 el perforador define cierto nivel de éxitos a partir de los sitios de perforación secuenciales en busca de petróleo. Si sólo se llevan seis intentos al momento en que se experimenta el segundo éxito, las utilidades parecen dominar de forma considerable la inversión requerida por la perforación.

5.6 Distribución de Poisson y proceso de Poisson

Los experimentos que dan valores numéricos de una variable aleatoria X , el número de resultados que ocurren durante un intervalo dado o en una región específica, se llaman **experimentos de Poisson**. El intervalo dado puede ser de cualquier longitud, como un minuto, un día, una semana, un mes o incluso un año. Por ello un experimento de Poisson puede generar observaciones para la variable aleatoria X que representa el número de llamadas telefónicas por hora que recibe una oficina, el número de días que la escuela permanece cerrada debido a la nieve durante el invierno o el número de juegos suspendidos debido a la lluvia durante la temporada de béisbol. La región específica podría ser un segmento de línea, un área o quizá

una pieza de material. En tales casos X puede representar el número de ratas de campo por acre, el número de bacterias en un cultivo dado o el número de errores mecanográficos por página.

Un experimento de Poisson se deriva del **proceso de Poisson** y posee las siguientes propiedades:

Propiedades del proceso de Poisson

1. El número de resultados que ocurren en un intervalo o región específica es independiente del número que ocurre en cualquier otro intervalo o región del espacio disjunto. De esta forma vemos que el proceso de Poisson no tiene memoria.
2. La probabilidad de que ocurra un solo resultado durante un intervalo muy corto o en una región pequeña es proporcional a la longitud del intervalo o al tamaño de la región y no depende del número de resultados que ocurren fuera de este intervalo o región.
3. La probabilidad de que ocurra más de un resultado en tal intervalo corto o que caiga en tal región pequeña es insignificante.

El número X de resultados que ocurren durante un experimento de Poisson se llama **variable aleatoria de Poisson** y su distribución de probabilidad se llama **distribución de Poisson**. El número medio de resultados se calcula de $\mu = \lambda t$, donde t es el "tiempo" o "región" específico de interés. Como sus probabilidades dependen de λ , la tasa de ocurrencia de los resultados, las denotaremos con el símbolo $P(x; \lambda t)$. La derivación de la fórmula para $p(x; \lambda t)$, que se basa en las tres propiedades de un proceso de Poisson que se listan arriba, está fuera del alcance de este libro. El siguiente concepto se utiliza para calcular probabilidades de Poisson.

Distribución de Poisson

La distribución de probabilidad de la variable aleatoria de Poisson X , que representa el número de resultados que ocurren en un intervalo dado o región específica que se denota con t , es

$$p(x; \lambda t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}, \quad x = 0, 1, 2, \dots,$$

donde λ es el número promedio de resultados por unidad de tiempo o región y $e = 2.71828 \dots$

La tabla A.2 contiene la suma de la probabilidad de Poisson

$$P(r; \lambda t) = \sum_{x=0}^r p(x; \lambda t)$$

para algunos valores selectos de λt que van de 0.1 a 18. Ilustramos el uso de esta tabla con los siguientes dos ejemplos.

Ejemplo 5.16 Durante un experimento de laboratorio el número promedio de partículas radiactivas que pasan a través de un contador en un milisegundo es cuatro. ¿Cuál es la probabilidad de que seis partículas entren al contador en un milisegundo dado?

SOLUCIÓN

Al usar la distribución de Poisson con $x = 6$ y $\lambda t = 4$, encontramos en la tabla A.2 que

$$p(6; 4) = \frac{e^{-4}4^6}{6!} = \sum_{x=0}^6 p(x; 4) - \sum_{x=0}^5 p(x; 4) = 0.8893 - 0.7851 = 0.1042. \quad \blacksquare$$

Ejemplo 5.17 El número promedio de camiones tanque que llega cada día a cierta ciudad portuaria es 10. Las instalaciones en el puerto pueden manejar a lo más 15 camiones tanque por día. ¿Cuál es la probabilidad de que en un día dado los camiones se tengan que regresar?

SOLUCIÓN

Sea X el número de camiones tanque que llegan cada día. Entonces, con el uso de la tabla A.2, tenemos

$$\begin{aligned} P(X > 15) &= 1 - P(X \leq 15) = 1 - \sum_{x=0}^{15} p(x; 10) = 1 - 0.9513 \\ &= 0.0487. \end{aligned}$$

Como la distribución binomial, la distribución de Poisson se utiliza para control de calidad, seguro de calidad y muestreo de aceptación. Además, ciertas distribuciones continuas importantes que se usan en la teoría de confiabilidad y teoría de colas dependen del proceso de Poisson. Algunas de estas distribuciones se discuten y se desarrollan en el capítulo 6. \blacksquare

Teorema 5.5

La media y la varianza de la distribución de Poisson $p(x; \lambda t)$ tienen el valor λt .

PRUEBA Para verificar que la media es en realidad λt , sea $\mu = \lambda t$. Podemos escribir

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu} \mu^x}{x!} = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu} \mu^x}{x!} = \mu \sum_{x=1}^{\infty} \frac{e^{-\mu} \mu^{x-1}}{(x-1)!}$$

Ahora bien, sea $y = x - 1$ lo que da

$$E(X) = \mu \sum_{y=0}^{\infty} \frac{e^{-\mu} \mu^y}{y!} = \mu,$$

puesto que

$$\sum_{y=0}^{\infty} \frac{e^{-\mu} \mu^y}{y!} = \sum_{y=0}^{\infty} p(y; \mu) = 1.$$

La varianza de la distribución de Poisson se obtiene al encontrar primero

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\mu} \mu^x}{x!} = \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\mu} \mu^x}{x!}$$

$$= \mu^2 \sum_{x=2}^{\infty} \frac{e^{-\mu} \mu^{x-2}}{(x-2)!}.$$

Al hacer $y = x - 2$, tenemos:

$$E[X(X-1)] = \mu^2 \sum_{y=0}^{\infty} \frac{e^{-\mu} \mu^y}{y!} = \mu^2.$$

De aquí

$$\sigma^2 = E[X(X-1)] + \mu - \mu^2 = \mu^2 + \mu - \mu^2 = \mu = \lambda t.$$

En el ejemplo 5.16, donde $\lambda t = 4$, también tenemos $\sigma^2 = 4$ y por ellos $\sigma = 2$. Con el uso del teorema de Chebyshev, podemos establecer que nuestra variable aleatoria tiene una probabilidad de al menos $3/4$ de caer en el intervalo $\mu \pm 2\sigma = 4 \pm (2)(2)$, o de 0 a 8. Por tanto, concluimos que al menos tres cuartos de las veces el número de partículas radiactivas que entran al contador estará en cualquier valor entre 0 y 8 durante un milisegundo dado.

Distribución de Poisson como forma limitante de la binomial

Debe ser evidente de los tres principios del proceso de Poisson que la distribución de Poisson se relaciona con la distribución binomial. Aunque la de Poisson por lo general encuentra aplicaciones en problemas de espacio y tiempo como se ilustra con los ejemplos 5.16 y 5.17, se puede ver como una forma limitante de la distribución binomial. En el caso de la binomial, si n es bastante grande y p es pequeña, las condiciones comienzan a simular las implicaciones de *espacio continuo o región temporal* del proceso de Poisson. La independencia entre las pruebas de Bernoulli en el caso binomial es consistente con la propiedad 2 del proceso de Poisson. Si se hace al parámetro p cercano a cero se relaciona con la propiedad 3. En realidad, derivaremos ahora la distribución de Poisson como forma limitante de la distribución binomial cuando $n \rightarrow \infty$, $p \rightarrow 0$, y np permanece constante. De aquí, si n es grande y p cercana a 0, se puede usar la distribución de Poisson, con $\mu = np$, para aproximar probabilidades binomiales. Si p es cercana a 1, aún podemos utilizar la distribución de Poisson para aproximar probabilidades binomiales mediante el intercambio de lo que definimos como éxito y fracaso, cambiamos con ello p a un valor cercano a 0.

Teorema 5.6

Sea X una variable aleatoria binomial con distribución de probabilidad $b(x; n, p)$. Cuando $n \rightarrow \infty$, $p \rightarrow 0$, y $\mu = np$ permanece constante,

$$b(x; n, p) \rightarrow p(x; \mu).$$

PRUEBA La distribución binomial se puede escribir como

$$\begin{aligned} b(x; n, p) &= \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{n(n-1) \cdots (n-x+1)}{x!} p^x (1-p)^{n-x}. \end{aligned}$$

Al sustituir $p = \mu/n$, tenemos:

$$\begin{aligned} b(x; n, p) &= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\mu}{n}\right)^x \left(1 - \frac{\mu}{n}\right)^{n-x} \\ &= 1\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right) \frac{\mu^x}{x!} \left(1 - \frac{\mu}{n}\right)^n \left(1 - \frac{\mu}{n}\right)^{-x}. \end{aligned}$$

Conforme $n \rightarrow \infty$ mientras x y μ permanecen constantes,

$$\lim_{n \rightarrow \infty} 1\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right) = 1, \quad \lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{n}\right)^n = e^{-\mu},$$

y de la definición del número e ,

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{n}\right)^n = \lim_{n \rightarrow \infty} \left[1 + \frac{1}{(-n)/\mu}\right]^{-n/\mu} = e^{-\mu}.$$

De aquí, bajo las condiciones limitantes dadas,

$$b(x; n, p) \rightarrow \frac{e^{-\mu} \mu^x}{x!}, \quad x = 0, 1, 2, \dots$$

Ejemplo 5.18 En un proceso de fabricación donde se manufacturan productos de vidrio ocurren defectos o burbujas, lo que deja ocasionalmente a la pieza indeseable para su venta. Se sabe que, en promedio, uno de cada 1000 de estos artículos que se producen tiene una o más burbujas. ¿Cuál es la probabilidad de que una muestra aleatoria de 8000 tenga menos de siete artículos con burbujas?

SOLUCIÓN

Éste es en esencia un experimento binomial con $n = 8000$ y $p = 0.001$. Como p es muy cercana a cero y n es bastante grande, haremos la aproximación con la distribución de Poisson utilizando

$$\mu = (8000)(0.001) = 8.$$

De aquí, si X representa el número de burbujas, tenemos

$$P(X < 7) = \sum_{x=0}^6 b(x; 8000, 0.001) \approx \sum_{x=0}^6 p(x; 8) = 0.3134.$$

Ejercicios

1. La probabilidad de que una persona, que vive en cierta ciudad, tenga un perro se estima en 0.3. Encuentre la probabilidad de que la décima persona entrevistada al azar en esta ciudad sea la quinta que tiene un perro.

2. Un científico inocula varios ratones, uno a la vez, con el germen de una enfermedad hasta que encuentre dos que contraen la enfermedad. Si la probabilidad de contraer la enfermedad es $1/6$, ¿cuál es la probabilidad de que se requieran ocho ratones?

3. El estudio de un inventario determina que, en promedio, las demandas de un artículo particular en un almacén se realizan cinco veces al día. ¿Cuál es la probabilidad de que en un día dado se pida este artículo

- (a) más de cinco veces?
(b) ninguna vez?

4. Encuentre la probabilidad de que una persona que lanza una moneda obtenga

- (a) la tercera cara en el séptimo lanzamiento;

(b) la primera cara en el cuarto lanzamiento.

5. Tres personas lanzan una moneda y el disparaje paga los cafés. Si todas las monedas tienen el mismo resultado, se lanzan de nuevo. Encuentre la probabilidad de que se necesiten menos de cuatro lanzamientos.

6. De acuerdo con un estudio publicado por un grupo de sociólogos de la Universidad de Massachusetts, cerca de dos tercios de los 20 millones de personas que en este país consumen Valium son mujeres. Suponga que esta cifra es una estimación válida, encuentre la probabilidad de que en un día dado la quinta prescripción de Valium que escribe un doctor es

- (a) la primera que prescribe Valium para una mujer;
 (b) la tercera que prescribe Valium para una mujer.

7. La probabilidad de que un estudiante para piloto apruebe el examen escrito para una licencia de piloto privado es 0.7. Encuentre la probabilidad de que el estudiante apruebe el examen

- (a) en el tercer intento;
 (b) antes del cuarto intento.

8. En promedio en cierta intersección ocurren tres accidentes de tránsito por mes. ¿Cuál es la probabilidad de que para cualquier mes dado en esta intersección

- (a) ocurran exactamente cinco accidentes?
 (b) ocurran menos de tres accidentes?
 (c) ocurran al menos dos accidentes?

9. Una secretaria comete dos errores por página, en promedio. ¿Cuál es la probabilidad de que en la siguiente página cometa

- (a) cuatro o más errores?
 (b) ningún error.

10. Cierta área del este de Estados Unidos resulta, en promedio, afectada por seis huracanes al año. Encuentre la probabilidad de que para cierto año esta área resulte afectada por

- (a) menos de cuatro huracanes;
 (b) cualquier cantidad entre seis a ocho huracanes.

11. Suponga que la probabilidad de que una persona dada crea un chisme acerca de las transgresiones de cierta actriz famosa es 0.8. ¿Cuál es la probabilidad de que

- (a) la sexta persona en escuchar este chisme sea la cuarta en creerlo?
 (b) la tercera persona en escuchar este chisme sea la primera en creerlo?

12. El número promedio de ratas de campo por acre en un campo de cinco acres de trigo se estima en 12. Encuentre la probabilidad de que se encuentren menos de siete ratas de campo

(a) en un acre dado;

(b) en dos de los siguientes tres acres que se inspeccionen.

13. El chef de un restaurante prepara una ensalada revuelta que contiene, en promedio, cinco vegetales. Encuentre la probabilidad de que la ensalada contenga más de cinco vegetales

- (a) en un día dado;
 (b) en tres de los siguientes cuatro días;
 (c) por primera vez en abril el día 5.

14. La probabilidad de que una persona muera de cierta infección respiratoria es 0.002. Encuentre la probabilidad de que mueran menos de cinco de los siguientes 2000 infectados de esta forma.

15. Suponga que, en promedio, una persona en 1000 comete un error numérico al preparar su declaración de impuestos. Si se seleccionan 10,000 formas al azar y se examinan, encuentre la probabilidad de que 6, 7 u 8 de las formas contengan un error.

16. Se sabe que la probabilidad de que un estudiante de una preparatoria local presente escoliosis (curvatura de la espina dorsal) es 0.004. De los siguientes 1875 estudiantes que se revisen en búsqueda de escoliosis, encuentre la probabilidad de que

- (a) menos de cinco presenten el problema;
 (b) ocho, nueve o 10 presenten el problema.

17. (a) Encuentre la media y la varianza de la variable aleatoria X , que representa el número de personas entre 2000 que mueren de la infección respiratoria en el ejercicio 14.

(b) De acuerdo con el teorema de Chebyshev, hay una probabilidad de al menos $3/4$ de que el número de personas que morirán entre las 2000 infectadas caiga dentro de ¿cuál intervalo?

18. (a) Encuentre la media y la varianza de la variable aleatoria X , que representa el número de personas entre 10,000 que cometen un error al preparar su declaración de impuestos en el ejercicio 15.

(b) De acuerdo con el teorema de Chebyshev, hay una probabilidad de al menos $8/9$ de que el número de personas que cometerán errores al preparar sus declaraciones de impuestos entre 10,000 esté dentro de ¿cuál intervalo?

19. Un fabricante de automóviles se preocupa por una falla en el mecanismo de freno de un modelo particular. La falla puede causar en raras ocasiones una catástrofe a alta velocidad. Suponga que la distribución del número de autos por año que experimentará la falla es una variable aleatoria de Poisson con $\lambda = 5$.

- (a) ¿Cuál es la probabilidad de que, a lo más, tres autos por año sufran una catástrofe?
 (b) ¿Cuál es la probabilidad de que más de un auto por año experimente una catástrofe?

20. Los cambios en los procedimientos de los aeropuertos requieren una planeación considerable. Los índices de llegadas de los aviones es un factor importante que se debe tomar en cuenta. Suponga que los aviones pequeños llegan a cierto aeropuerto, de acuerdo con un proceso de Poisson, con un índice de seis por hora. De esta manera, el parámetro de Poisson para las llegadas en un periodo de t horas es $\lambda = 6t$.

- (a) ¿Cuál es la probabilidad de que exactamente cuatro aeronaves pequeñas lleguen durante un periodo de una hora?
 (b) ¿Cuál es la probabilidad de que al menos cuatro lleguen durante un periodo de una hora?
 (c) Si definimos un día laboral como 12 horas, ¿cuál es la probabilidad de que al menos 75 pequeñas aeronaves lleguen durante un día?

21. El número de clientes que llegan por hora a ciertas instalaciones de servicio automotriz se supone que sigue una distribución de Poisson con media $\lambda = 7$.

- (a) Calcule la probabilidad de que más de 10 clientes lleguen en un periodo de dos horas.
 (b) ¿Cuál es el número medio de llegadas durante un periodo de dos horas?

22. Considere el ejercicio 16 de la página 140. ¿Cuál es el número medio de estudiantes que presentan el problema?

23. La probabilidad de que una persona muera cuando contrae una infección respiratoria es 0.002. De los siguientes 2000 infectados con este tipo de enfermedad, ¿cuál es el número medio que morirá?

24. Una compañía compra lotes grandes de cierta clase de dispositivo electrónico. Se utiliza un método que rechaza un lote si se encuentran dos o más unidades defectuosas en una muestra aleatoria de 100 unidades.

- (a) ¿Cuál es el número medio de unidades defectuosas que se encuentran en una muestra de 100 unidades si el lote tiene 1% de defectuosas?
 (b) ¿Cuál es la varianza?

Ejercicios de repaso

1. Durante un proceso de producción se seleccionan al azar 15 unidades cada día de la línea de ensamble para verificar el porcentaje de defectuosos en el proceso. Se sabe a partir de información histórica que la probabilidad de tener una unidad defectuosa es 0.05. En cualquier momento en que dos o más unidades defectuosas se encuentran en la muestra de 15, el proceso se detiene. Este procedimiento se utiliza para proporcionar una señal en caso de que la probabilidad de unidades defectuosas aumente.

- (a) ¿Cuál es la probabilidad de que en un día dado el proceso de producción se detenga? (Suponga 5% de unidades defectuosas.)
 (b) Suponga que la probabilidad de una unidad defectuosa aumenta a 0.07. ¿Cuál es la probabilidad de que en algún día dado el proceso de producción no se detenga?

2. Se considera para la producción una máquina automática de soldar. Se considerará para su compra si es exitosa en 99% de sus soldaduras. De otra manera, no se considerará eficiente. Se lleva a cabo la prueba de un prototipo que realizará 100 soldaduras. La máquina se aceptará para la producción si no falla en más de tres soldaduras.

- (a) ¿Cuál es la probabilidad de que se rechace una buena máquina?
 (b) ¿Cuál es la probabilidad de que se acepte una máquina ineficiente con 95% de soldaduras exitosas?

3. Una agencia de renta de coches en un aeropuerto local tiene disponibles cinco Ford, siete Chevrolet, cuatro Dodge, tres Honda y cuatro Toyota. Si la agencia selecciona al azar nueve de estos autos para transportar delegados del aeropuerto al centro de convenciones del centro de la ciudad, encuentre la probabilidad de que se utilicen dos Ford, tres Chevrolet, un Dodge, un Honda y dos Toyota.

4. Las llamadas de servicio llegan a un centro de mantenimiento de acuerdo con un proceso de Poisson con un promedio de 2.7 llamadas por minuto. Encuentre la probabilidad de que

- (a) lleguen no más de cuatro llamadas en cualquier minuto;
 (b) lleguen menos de dos llamadas en cualquier minuto;
 (c) lleguen más de 10 llamadas en un periodo de cinco minutos.

5. Una empresa de electrónica afirma que la proporción de unidades defectuosas de cierto proceso es 5%. Un comprador tiene un procedimiento estándar para inspeccionar 15 unidades que selecciona al azar de un lote grande. En una ocasión particular, el comprador encuentra cinco artículos defectuosos.

- (a) ¿Cuál es la probabilidad de esta ocurrencia, dado que la afirmación de 5% de defectuosos es correcta?
 (b) ¿Cuál sería su reacción si fuera el comprador?

6. Un dispositivo electrónico de conmutación ocasionalmente falla y puede ser necesario su reemplazo. Se sabe que el dispositivo es satisfactorio si, en promedio, no comete más de 0.20 errores por hora. Se elige un periodo particular de cinco horas como "prueba" del dispositivo. Si no ocurre más de un error, el dispositivo se considera satisfactorio.

- (a) ¿Cuál es la probabilidad de que un dispositivo satisfactorio se considere que no lo es sobre la base de la prueba? Suponga que existe un proceso de Poisson.
 (b) ¿Cuál es la probabilidad de que un dispositivo se acepte como satisfactorio cuando, de hecho, el número medio de errores es 0.25? De nuevo, suponga que existe un proceso de Poisson.

7. Una compañía por lo general compra lotes grandes de cierta clase de dispositivo electrónico. Se utiliza un método que rechaza un lote si se encuentran dos o más unidades defectuosas en una muestra aleatoria de 100 unidades.

- (a) ¿Cuál es la probabilidad de rechazar un lote que tiene 1% de unidades defectuosas?
 (b) ¿Cuál es la probabilidad de aceptar un lote que tiene 5% de unidades defectuosas?

8. El propietario de una farmacia local sabe que, en promedio, llegan a su farmacia 100 personas por hora.

- (a) Encuentre la probabilidad de que en un periodo dado de tres minutos nadie entre a la farmacia.
 (b) Encuentre la probabilidad de que en un periodo dado de tres minutos entren más de cinco personas a la farmacia.

9. (a) Suponga que lanza cuatro dados. Encuentre la probabilidad de que obtenga al menos un 1.

- (b) Suponga que lanza 24 veces dos dados. Encuentre la probabilidad de que obtenga al menos uno (1, 1), es decir, lanza "ojos de serpiente"

[NOTA: La probabilidad de la parte (a) es mayor que la de la parte (b).]

10. Suponga que se venden 500 billetes de lotería. Entre ellos, 200 billetes pagan *al menos* el costo del billete. Suponga ahora que compra cinco billetes. En-

cuentre la probabilidad de que gane al menos el costo de tres billetes.

11. Las imperfecciones en las tarjetas de circuitos y los chips para computadora se prestan por sí mismos a tratamiento estadístico. Para un tipo particular de tarjeta la probabilidad de falla de un diodo es 0.03. Suponga que una tarjeta de circuitos contiene 200 diodos.

- (a) ¿Cuál es el número medio de fallas entre los diodos?
 (b) ¿Cuál es la varianza?
 (c) La tarjeta funcionará si no hay diodos defectuosos. ¿Cuál es la probabilidad de que una tarjeta funcione?

12. El comprador potencial de un motor particular requiere (entre otras cosas) que el motor encienda exitosamente 10 veces consecutivas. Suponga que la probabilidad de un encendido exitoso es 0.990. Supongamos que los resultados de intentos de encendido son independientes.

- (a) ¿Cuál es la probabilidad de que el motor sea aceptado después de sólo 10 encendidos?
 (b) ¿Cuál es la probabilidad de que se realicen 12 intentos de encendido durante el proceso de aceptación?

13. El esquema de aceptación para comprar lotes que contienen un número grande de baterías es probar no más de 75 baterías seleccionadas al azar y rechazar un lote si falla una sola batería. Suponga que la probabilidad de una falla es 0.001.

- (a) ¿Cuál es la probabilidad de que se acepte un lote?
 (b) ¿Cuál es la probabilidad de que se rechace un lote en la 20ª prueba?
 (c) ¿Cuál es la probabilidad de que se rechace en 10 o menos pruebas?

14. Una compañía perforadora de pozos petroleros se arriesga en varios sitios, y su éxito o fracaso es independiente de un sitio a otro. Suponga que la probabilidad de éxito en cualquier sitio específico es 0.25.

- (a) ¿Cuál es la probabilidad de que un perforador barrene 10 sitios y tenga un éxito?
 (b) El perforador siente que irá a la quiebra si perfora 10 veces antes de que ocurra el primer éxito. ¿Cuáles son las perspectivas del perforador para la ruina?

15. Considere la información del ejercicio 14. El perforador siente que "dará en el clavo" si el segundo éxito ocurre en o antes del sexto intento. ¿Cuál es la probabilidad de que el perforador "dé en el clavo"?

Algunas distribuciones continuas de probabilidad

6.1 Distribución uniforme continua

Una de las distribuciones continuas más simples en la estadística es la **distribución uniforme continua**. Esta distribución se caracteriza por una función de densidad que es "plana", y por ello la probabilidad es uniforme en un intervalo cerrado, digamos $[A, B]$. Aunque las aplicaciones de la distribución uniforme continua no son tan abundantes como lo son para otras distribuciones que se presentan en este capítulo, es apropiado para el principiante comenzar esta introducción a las distribuciones continuas con la distribución uniforme.

Distribución uniforme

La función de densidad de la variable aleatoria uniforme continua X en el intervalo $[A, B]$ es

$$f(x; A, B) = \begin{cases} \frac{1}{B - A}, & A \leq x \leq B \\ = 0 & \text{en cualquier otro caso.} \end{cases}$$

Se debe recalcar al lector que la función de densidad forma un rectángulo con base $B - A$ y **altura constante** $\frac{1}{B - A}$. Como resultado, la distribución uniforme a menudo se llama **distribución rectangular**.

En la figura 6.1 se muestra la función de densidad para una variable aleatoria uniforme en el intervalo $[1, 3]$.

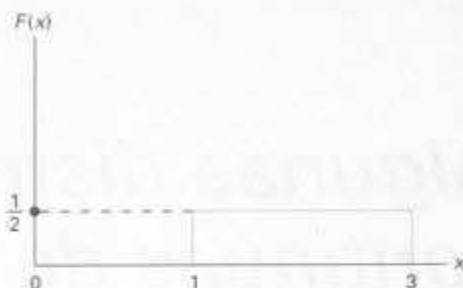


Figura 6.1 Función de densidad para una variable aleatoria en el intervalo $[1, 3]$.

Es sencillo calcular las probabilidades para la distribución uniforme debido a la naturaleza simple de la función de densidad. Sin embargo, note que la aplicación de esta distribución se basa en la suposición de que la probabilidad de caer en un intervalo de longitud fija dentro de $[A, B]$ es constante.

Ejemplo 6.1 Suponga que se puede reservar una sala de conferencias grande para cierta compañía por no más de cuatro horas. Sin embargo, el uso de la sala de conferencias es tal que muy a menudo tienen lugar conferencias largas y cortas. De hecho, se puede suponer que la duración X de una conferencia tiene una distribución uniforme en el intervalo $[0, 4]$.

- (a) ¿Cuál es la función de densidad de la probabilidad?
 (b) ¿Cuál es la probabilidad de que cualquier conferencia dada dure al menos tres horas?

SOLUCIÓN

- (a) La función de densidad apropiada para la variable aleatoria distribuida uniformemente X en esta situación es

$$f(x) = \begin{cases} \frac{1}{4}, & 0 \leq x \leq 4 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(b) $\Pr[X \geq 3] = \int_3^4 \left(\frac{1}{4}\right) dx = \frac{1}{4}$. ■

Teorema 6.1

La media y la varianza de la distribución uniforme son

$$\mu = \frac{A + B}{2} \quad \text{y} \quad \sigma^2 = \frac{(B - A)^2}{12}.$$

Las demostraciones de los teoremas se dejan al lector. Véase el ejercicio 20 de la página 159.

6.2 Distribución normal

La distribución continua de probabilidad más importante en todo el campo de la estadística es la **distribución normal**. Su gráfica, que se denomina **curva normal**, es la curva con forma de campana de la figura 6.2, la cual describe aproximadamente muchos fenómenos que ocurren en la naturaleza, la industria y la investigación. Las mediciones físicas en áreas como los experimentos meteorológicos, estudios de lluvia y mediciones de partes fabricadas a menudo se explican más que adecuadamente con una distribución normal. Además, los errores en las mediciones científicas se aproximan extremadamente bien mediante una distribución normal. En 1733, Abraham DeMoivre desarrolló la ecuación matemática de la curva normal. Proporciona una base sobre la que se fundamenta gran parte de la teoría de la estadística inductiva. La distribución normal a menudo se denomina **distribución gaussiana**, en honor de Karl Friedrich Gauss (1777-1855), quien también derivó su ecuación a partir de un estudio de errores en mediciones repetidas de la misma cantidad.

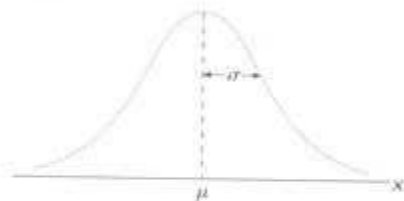


Figura 6.2 Curva normal.

Una variable aleatoria continua X que tiene la distribución en forma de campana de la figura 6.2 se llama **variable aleatoria normal**. La ecuación matemática para la distribución de probabilidad de la variable normal depende de los dos parámetros μ y σ , su media y desviación estándar. De aquí, denotamos los valores de la densidad de X con $n(x; \mu, \sigma)$.

Distribución normal

La función de densidad de la variable aleatoria normal X , con media μ y varianza σ^2 , es

$$n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2)(x-\mu)^2/\sigma^2}, \quad -\infty < x < \infty,$$

donde $\pi = 3.14159 \dots$ y $e = 2.71828 \dots$

Una vez que se especifican μ y σ , la curva normal queda determinada por completo. Por ejemplo, si $\mu = 50$ y $\sigma = 5$, entonces se pueden calcular las ordenadas $n(x; 50, 5)$ para varios valores de x y dibujar la curva. En la figura 6.3 dibujamos dos curvas normales que tienen la misma desviación estándar pero diferentes medias. Las dos curvas son idénticas en forma pero están centradas en diferentes posiciones a lo largo del eje horizontal.

En la figura 6.4 trazamos dos curvas normales con la misma media pero con diferentes desviaciones estándar. Esta vez vemos que las dos curvas están centradas exactamente en la misma posición sobre el eje horizontal, pero la curva con la ma-

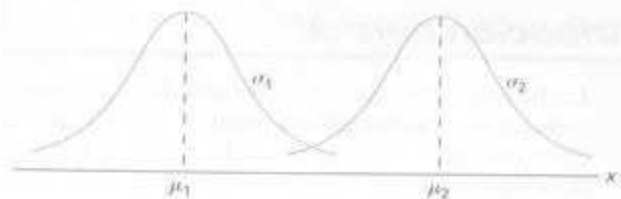


Figura 6.3 Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 = \sigma_2$.

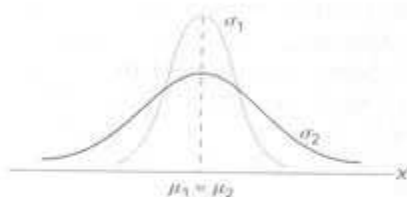


Figura 6.4 Curvas normales con $\mu_1 = \mu_2$ y $\sigma_1 < \sigma_2$.

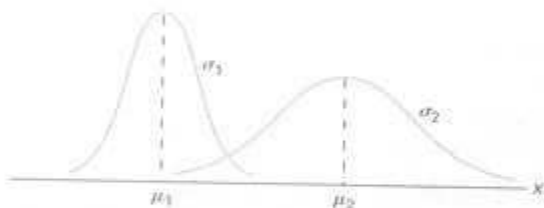


Figura 6.5 Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 < \sigma_2$.

yor desviación estándar es más baja y se extiende más lejos. Recuerde que el área bajo una curva de probabilidad debe ser igual a 1, y entre más variable sea el conjunto de observaciones más baja y ancha será la curva correspondiente.

La figura 6.5 muestra el resultados de trazar dos curvas normales que tienen diferentes medias y diferentes desviaciones estándar. Claramente, están centradas en posiciones diferentes sobre el eje horizontal y sus formas reflejan los dos diferentes valores de σ .

De una inspección de las figuras 6.2 a 6.5 y al examinar la primera y segunda derivadas de $n(x; \mu, \sigma)$, listamos las siguientes propiedades de la curva normal:

1. La moda, que es el punto sobre el eje horizontal donde la curva es un máximo, ocurre en $x = \mu$.
2. La curva es simétrica alrededor de un eje vertical a través de la media μ .
3. La curva tiene sus puntos de inflexión en $x = \mu \pm \sigma$, es cóncava hacia abajo si $\mu - \sigma < X < \mu + \sigma$, y es cóncava hacia arriba en cualquier otro punto.

- La curva normal se aproxima al eje horizontal de manera asintótica conforme nos alejamos de la media en cualquier dirección.
- El área total bajo la curva y sobre el eje horizontal es igual a 1.

Mostraremos ahora que los parámetros μ y σ^2 son realmente la media y la varianza de la distribución normal. Para evaluar la media, escribimos

$$E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-(1/2)((x-\mu)/\sigma)^2} dx.$$

Al hacer $z = (x - \mu)/\sigma$ y $dx = \sigma dz$, obtenemos

$$\begin{aligned} E(X) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma z) e^{-z^2/2} dz \\ &= \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2} dz. \end{aligned}$$

El primer término de la derecha es μ veces el área bajo una curva normal con media cero y varianza 1, y por ello igual a μ . Por integración directa, el segundo término es igual a 0. De aquí que

$$E(X) = \mu.$$

La varianza de la distribución normal está dada por

$$E[(X - \mu)^2] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-(1/2)((x-\mu)/\sigma)^2} dx.$$

Nuevamente al hacer $z = (x - \mu)/\sigma$ y $dx = \sigma dz$, obtenemos

$$E[(X - \mu)^2] = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz.$$

Al integrar por partes con $u = z$ y $dv = z e^{-z^2/2} dz$ de modo que $du = dz$ y $v = -e^{-z^2/2}$, encontramos que

$$E[(X - \mu)^2] = \frac{\sigma^2}{\sqrt{2\pi}} \left(-z e^{-z^2/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-z^2/2} dz \right) = \sigma^2(0 + 1) = \sigma^2.$$

Muchas variables aleatorias tienen distribuciones de probabilidad que se pueden describir de manera adecuada mediante la curva normal una vez que se especifiquen μ y σ^2 . En este capítulo supondremos que se conocen estos dos parámetros, quizá de investigaciones previas. Más tarde haremos inferencias estadísticas cuando se desconozcan μ y σ^2 y se estimen a partir de los datos experimentales disponibles.

Señalamos en un principio el papel que juega la distribución normal como una aproximación razonable de variables científicas en experimentos de la vida real. Hay otras aplicaciones de la distribución normal que el lector apreciará conforme avance en el libro. La distribución normal encuentra una gran aplicación como *distribución limitante*. Bajo ciertas condiciones la distribución normal proporciona una

buena aproximación continua a las distribuciones binomial e hipergeométrica. El caso de la aproximación a la binomial se cubre en la sección 6.5. En el capítulo 8 el lector aprenderá acerca de las **distribuciones muestrales**. Resulta que la distribución limitante de promedios muestrales es normal. Esto proporciona una base amplia para la inferencia estadística que es muy valiosa para el analista de datos interesado en la estimación y prueba de hipótesis. Las importantes áreas del análisis de varianza (capítulos 13, 14 y 15) y del control de calidad (capítulo 17) tienen su teoría basada en suposiciones que hacen uso de la distribución normal.

6.3 Áreas bajo la curva normal

La curva de cualquier distribución continua de probabilidad o función de densidad se construye de modo que el área bajo la curva limitada por las dos ordenadas $x = x_1$ y $x = x_2$ es igual a la probabilidad de que la variable aleatoria X tome un valor entre $x = x_1$ and $x = x_2$. Así, para la curva normal de la figura 6.6,

$$\begin{aligned} P(x_1 < X < x_2) &= \int_{x_1}^{x_2} n(x; \mu, \sigma) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-(1/2)((x-\mu)/\sigma)^2} dx \end{aligned}$$

está representada por el área de la región sombreada.

En las figuras 6.3, 6.4 y 6.5 vemos cómo la curva normal depende de la media y de la desviación estándar de la distribución bajo investigación. El área bajo la curva entre cualesquiera dos ordenadas también debe depender de los valores μ y σ . Esto es evidente en la figura 6.7, donde sombrearemos las regiones que corresponden a $P(x_1 < X < x_2)$ para dos curvas con diferentes medias y varianzas. La $P(x_1 < X < x_2)$, donde X es la variable aleatoria que describe la distribución I, se indica por el área sombreada más oscura. Si X es la variable aleatoria que describe la región II, entonces $P(x_1 < X < x_2)$ está dada por toda la región sombreada. Obviamente, las dos regiones sombreadas tienen tamaños diferentes; por tanto, la probabilidad asociada con cada distribución será diferente para los dos valores dados de X .

La dificultad que se encuentra al resolver las integrales de funciones de densidad normal necesita de la tabulación de las áreas de la curva normal para una refe-

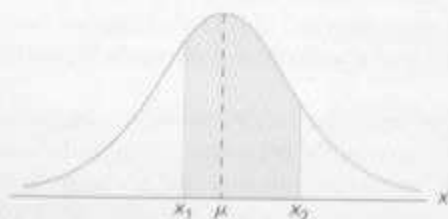


Figura 6.6 $P(x_1 < X < x_2) = \text{área de la región sombreada}$.

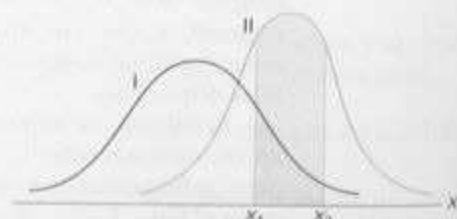


Figura 6.7 $P(x_1 < X < x_2)$ para diferentes curvas normales.

rencia rápida. Sin embargo, sería una tarea sin fin intentar establecer tablas separadas para cada valor concebible de μ y σ . Afortunadamente, somos capaces de transformar todas las observaciones de cualquier variable aleatoria normal X a un nuevo conjunto de observaciones de una variable aleatoria normal Z con media cero y varianza 1. Esto se puede realizar por medio de la transformación

$$Z = \frac{X - \mu}{\sigma}.$$

Siempre que X tome un valor x , el valor correspondiente de Z está dado por $z = (x - \mu)/\sigma$. Por tanto, si X cae entre los valores $x = x_1$ y $x = x_2$, la variable aleatoria Z caerá entre los valores correspondientes $z_1 = (x_1 - \mu)/\sigma$ y $z_2 = (x_2 - \mu)/\sigma$. En consecuencia, podemos escribir

$$\begin{aligned} P(x_1 < X < x_2) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{x_1}^{x_2} e^{-(1/2)(x-\mu)^2/\sigma^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-z^2/2} dz \\ &= \int_{z_1}^{z_2} n(z; 0, 1) dz = P(z_1 < Z < z_2), \end{aligned}$$

donde Z se ve como una variable aleatoria normal con media cero y varianza 1.

Definición 6.1

La distribución de una variable aleatoria normal con media cero y varianza 1 se llama **distribución normal estándar**.

Las distribuciones original y transformada se ilustran en la figura 6.8. Como todos los valores de X caen entre x_1 y x_2 tienen valores z correspondientes entre z_1 y z_2 , el área bajo la curva X entre las ordenadas $x = x_1$ y $x = x_2$ de la figura 6.8 es igual al área bajo la curva Z entre las ordenadas transformadas $z = z_1$ y $z = z_2$.

Hemos reducido ahora el número que se requiere de tablas de áreas de curva normal a una, la de la distribución normal estándar. La tabla A.3 indica el área bajo la curva normal estándar que corresponde a $P(Z < z)$ para valores de z que van de -3.49 a 3.49 . Para ilustrar el uso de esta tabla, encontremos la probabilidad de que Z sea menor que 1.74 . Primero, localizamos un valor de z igual a 1.7 en la co-

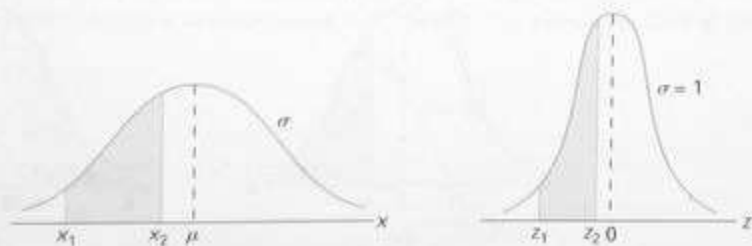


Figura 6.8 Distribuciones normales original y transformada.

lunna izquierda; después nos movemos a lo largo del renglón a la columna bajo 0.04; donde leemos 0.9591. Por tanto, $P(Z < 1.74) = 0.9591$. Para encontrar un valor z que corresponda a una probabilidad dada, el proceso se invierte. Por ejemplo, el valor z que deja un área de 0.2148 bajo la curva a la izquierda de z se ve que es -0.79 .

Ejemplo 6.2 Dada una distribución normal estándar, encuentre el área bajo la curva que yace (a) a la derecha de $z = 1.84$ y (b) entre $z = -1.97$ y $z = 0.86$

SOLUCIÓN

- (a) El área en la figura 6.9(a) a la derecha de $z = 1.84$ es igual a 1 menos el área en la tabla A.3 a la izquierda de $z = 1.84$ a saber, $1 - 0.9671 = 0.0329$.
- (b) El área en la figura 6.9(b) entre $z = -1.97$ y $z = 0.86$ es igual al área a la izquierda de $z = 0.86$ menos el área a la izquierda de $z = -1.97$. De la tabla A.3 encontramos que el área que se desea es $0.8051 - 0.0244 = 0.7807$. ■

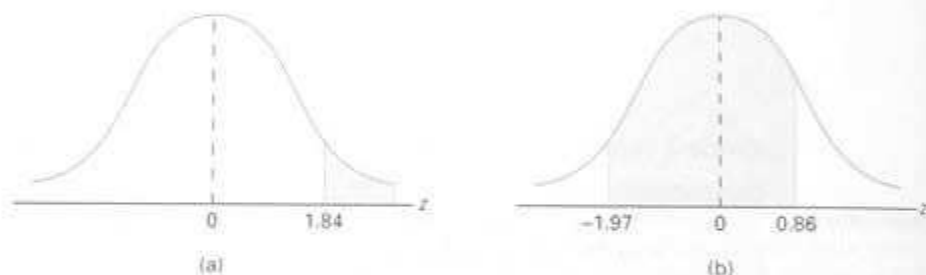


Figura 6.9 Áreas para el ejemplo 6.2.

Ejemplo 6.3 Dada una distribución normal estándar, encuentre el valor de k tal que (a) $P(Z > k) = 0.3015$, y (b) $P(k < Z < -0.18) = 0.4197$.

SOLUCIÓN

- (a) En la figura 6.10(a) vemos que el valor k que deja un área de 0.3015 a la derecha debe dejar entonces un área de 0.6985 a la izquierda. De la tabla A.3 se sigue que $k = 0.52$.

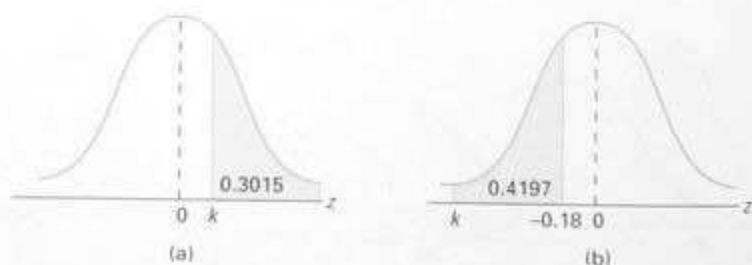


Figura 6.10 Áreas para el ejemplo 6.3

- (b) De la tabla A.3 notamos que el área total a la izquierda de -0.18 es igual a 0.4286 . En la figura 6.10(b) vemos que el área entre k y -0.18 es 0.4197 por lo que el área a la izquierda de k debe ser $0.4286 - 0.4197 = 0.0089$. Por tanto, de la tabla A.3, tenemos $k = -2.37$. ■

Ejemplo 6.4 Dada una distribución normal con $\mu = 50$ y $\sigma = 10$, encuentre la probabilidad de que X tome un valor entre 45 y 62.

SOLUCIÓN

Los valores z que corresponden a $x_1 = 45$ y $x_2 = 62$ son

$$z_1 = \frac{45 - 50}{10} = -0.5 \quad \text{y} \quad z_2 = \frac{62 - 50}{10} = 1.2.$$

Por tanto,

$$P(45 < X < 62) = P(-0.5 < Z < 1.2).$$

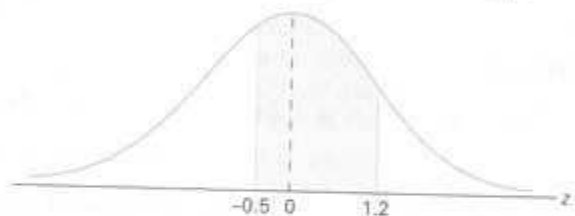


Figura 6.11 Área para el ejemplo 6.4.

La $P(-0.5 < Z < 1.2)$ se muestra por el área de la región sombreada de la figura 6.11. Esta área se puede encontrar al restar el área a la izquierda de la ordenada $z = -0.5$ de toda el área a la izquierda de $z = 1.2$. Con el uso de la tabla A.3, tenemos

$$\begin{aligned} P(45 < X < 62) &= P(-0.5 < Z < 1.2) \\ &= P(Z < 1.2) - P(Z < -0.5) = 0.8849 - 0.3085 = 0.5764. \quad \blacksquare \end{aligned}$$

Ejemplo 6.5 Dada una distribución normal con $\mu = 300$ y $\sigma = 50$, encuentre la probabilidad de que X tome un valor mayor que 362.

SOLUCIÓN

La distribución de probabilidad normal que muestra el área que se desea se representa en la figura 6.12. Para encontrar la $P(X > 362)$, necesitamos evaluar el área bajo la curva normal a la derecha de $x = 362$. Esto se puede realizar al transformar

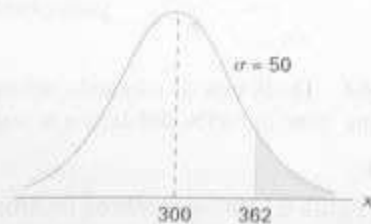


Figura 6.12 Área para el ejemplo 6.5.

$x = 362$ al valor z correspondiente, al obtener el área a la izquierda de z de la tabla A.3 y después restar esta área de 1. Encontramos que

$$z = \frac{362 - 300}{50} = 1.24.$$

De aquí

$$P(X > 362) = P(Z > 1.24) = 1 - P(Z < 1.24) = 1 - 0.8925 = 0.1075. \blacksquare$$

De acuerdo con el teorema de Chebyshev, la probabilidad de que una variable aleatoria tome un valor dentro de dos desviaciones estándar de la media es al menos $3/4$. Si la variable aleatoria tiene una distribución normal, los valores z que corresponden a $x_1 = \mu - 2\sigma$ y $x_2 = \mu + 2\sigma$ se calculan fácilmente y son

$$z_1 = \frac{(\mu - 2\sigma) - \mu}{\sigma} = -2 \quad \text{y} \quad z_2 = \frac{(\mu + 2\sigma) - \mu}{\sigma} = 2.$$

De aquí

$$\begin{aligned} P(\mu - 2\sigma < X < \mu + 2\sigma) &= P(-2 < Z < 2) = P(Z < 2) - P(Z < -2) \\ &= 0.9772 - 0.0228 = 0.9544, \end{aligned}$$

que es una afirmación mucho más fuerte que la que se establece mediante el teorema de Chebyshev.

En ocasiones se nos pide encontrar el valor de z que corresponde a una probabilidad específica que cae entre los valores que se listan en la tabla A.3 (véase el ejemplo 6.6). Por conveniencia, siempre elegiremos el valor z que corresponde a la probabilidad tabular que está más cerca de la probabilidad que se especifica. Sin embargo, si la probabilidad dada cae a la mitad entre dos probabilidades tabulares, debemos elegir para z el valor que cae a la mitad entre los valores correspondientes de z . Por ejemplo, para encontrar el valor z que corresponde a una probabilidad de 0.7975, que cae entre 0.7967 y 0.7995 en la tabla A.3, elegimos $z = 0.83$, pues 0.7975 es más cercano a 0.7967. Por otro lado, para una probabilidad de 0.7981, que cae a la mitad entre 0.7967 y 0.7995, tomamos $z = 0.835$.

Los anteriores dos ejemplos se resolvieron al ir primero de un valor de x a un valor z y después calcular el área que se desea. En el ejemplo 6.6 invertimos el proceso y comenzamos con un área o probabilidad conocida, encontramos el valor z y después determinamos x reacomodando la fórmula

$$z = \frac{x - \mu}{\sigma} \quad \text{para obtener} \quad x = \sigma z + \mu.$$

Ejemplo 6.6 Dada una distribución normal con $\mu = 40$ y $\sigma = 6$, encuentre el valor de x que tiene (a) 45% del área a la izquierda y (b) 14% del área a la derecha.

SOLUCIÓN

(a) En la figura 6.13(a) se sombrea un área de 0.45 a la izquierda del valor x deseado. Requerimos un valor z que deja un área de 0.45 a la izquierda. De la tabla

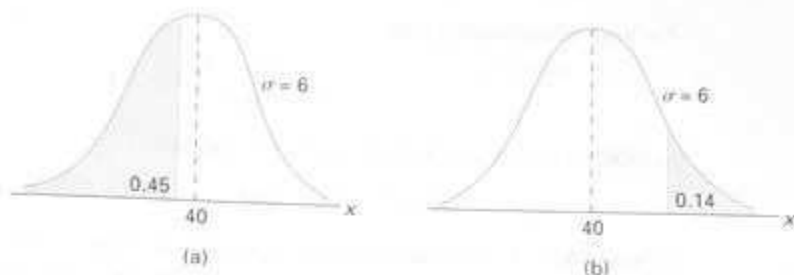


Figura 6.13 Áreas para el ejemplo 6.6.

A.3 encontramos $P(Z < -0.13) = 0.45$ por lo que el valor z deseado es -0.13 . De aquí

$$x = (6)(-0.13) + 40 = 39.22.$$

- (b) En la figura 6.13(b) sombreamos un área igual a 0.14 a la derecha del valor x que se desea. Esta vez requerimos un valor z que deje 0.14 del área a la derecha y por ello un área de 0.86 a la izquierda. De nuevo, de la tabla A.3, encontramos $P(Z < 1.08) = 0.86$ por lo que el valor z que se desea es 1.08 y

$$x = (6)(1.08) + 40 = 46.48. \quad \blacksquare$$

6.4 Aplicaciones de la distribución normal

Algunos de los muchos problemas para los que es aplicable la distribución normal se tratan en los siguientes ejemplos. El uso de la curva normal para aproximar probabilidades binomiales se considera en la sección 6.5.

Ejemplo 6.7 Cierta tipo de batería de almacenamiento dura, en promedio, 3.0 años, con una desviación estándar de 0.5 años. Suponga que las duraciones de la batería se distribuyen normalmente, encuentre la probabilidad de que una batería dada dure menos de 2.3 años.

SOLUCIÓN

Primero construya un diagrama como el de la figura 6.14, que muestra la distribución dada de duraciones de las baterías y el área que se desea. Para encontrar la $P(X < 2.3)$, necesitamos evaluar el área bajo la curva normal a la izquierda de 2.3. Esto se logra al encontrar el área a la izquierda del valor x correspondiente.

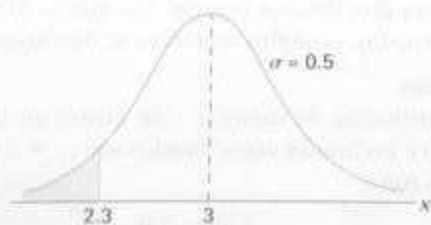


Figura 6.14 Área para el ejemplo 6.7.

De aquí encontramos que

$$z = \frac{2.3 - 3}{0.5} = -1.4$$

y entonces con el uso de la tabla A.3 tenemos

$$P(X < 2.3) = P(Z < -1.4) = 0.0808. \blacksquare$$

Ejemplo 6.8 Una empresa eléctrica fabrica focos que tienen una duración, antes de fundirse, que se distribuye normalmente con media igual a 800 horas y una desviación estándar de 40 horas. Encuentre la probabilidad de que un foco se funda entre 778 y 834 horas.

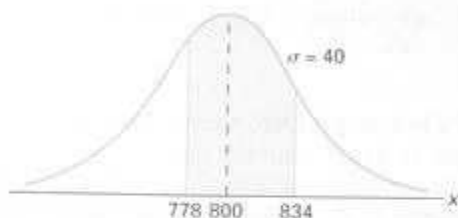


Figura 6.15 Área para el ejemplo 6.8.

SOLUCIÓN

La distribución de los focos se ilustra en la figura 6.15. Los valores z que corresponden a $x_1 = 778$ y $x_2 = 834$ son

$$z_1 = \frac{778 - 800}{40} = -0.55 \quad \text{y} \quad z_2 = \frac{834 - 800}{40} = 0.85.$$

De aquí

$$\begin{aligned} P(778 < X < 834) &= P(-0.55 < Z < 0.85) \\ &= P(Z < 0.85) - P(Z < -0.55) \\ &= 0.8023 - 0.2912 = 0.5111. \blacksquare \end{aligned}$$

Ejemplo 6.9 En un proceso industrial el diámetro de un cojinete es una parte importante. El comprador establece que las especificaciones en el diámetro sean 3.0 ± 0.01 cm. La implicación es que ninguna parte que caiga fuera de estas especificaciones se aceptará. Se sabe que en el proceso el diámetro de un cojinete tiene una distribución normal con media 3.0 y una desviación estándar $\sigma = 0.005$. En promedio, ¿cuántos cojinetes se descartarán?

SOLUCIÓN

La distribución de diámetros se ilustra en la figura 6.16. Los valores que corresponden a los límites especificados son $x_1 = 2.99$ y $x_2 = 3.01$, los valores z correspondientes son

$$z_1 = \frac{2.99 - 3.0}{0.005} = -2.0, \quad z_2 = \frac{3.01 - 3.00}{0.005} = +2.0.$$

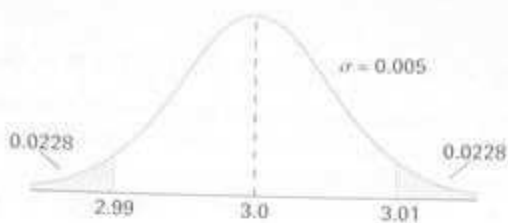


Figura 6.16 Área para el ejemplo 6.9.

De aquí

$$P(2.99 < X < 3.01) = P(-2.0 < Z < 2.0).$$

De la tabla A.3, $P(Z < -2.0) = 0.0228$. Debido a la simetría de la distribución normal, encontramos que

$$P(-2.0 < Z < 2.0) = 1 - 2(0.0228) = 0.9544.$$

Como resultado se anticipa que, en promedio, 4.56% de los cojinetes fabricados se descartarán. ■

Ejemplo 6.10 Se utilizan medidores para rechazar todos los componentes donde cierta dimensión no está dentro de la especificación $1.50 \pm d$. Se sabe que esta medición se distribuye de forma normal con media 1.50 y desviación estándar 0.2. Determine el valor d tal que las especificaciones "cubran" 95% de las mediciones.

SOLUCIÓN

De la tabla A.3 sabemos que

$$P(-1.96 < Z < 1.96) = 0.95.$$

Por tanto,

$$1.96 = \frac{(1.50 + d) - 1.50}{0.2}$$

de la que obtenemos

$$d = (0.2)(1.96) = 0.392.$$

Una ilustración de las especificaciones se muestra en la figura 6.17. ■

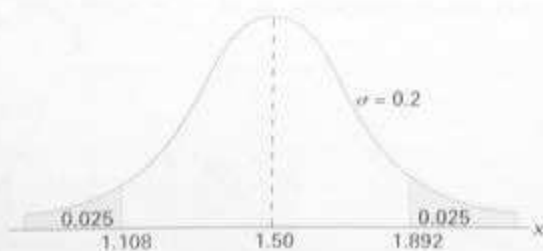


Figura 6.17 Especificaciones para el ejemplo 6.10.

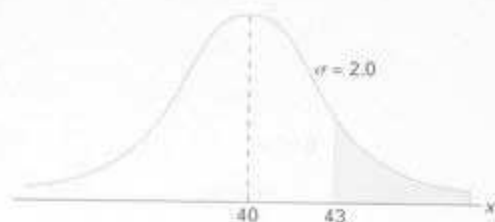


Figura 6.18 Área para el ejemplo 6.11.

Ejemplo 6.11 Cierta máquina fabrica resistores eléctricos que tienen una resistencia media de 40 ohmios y una desviación estándar de dos ohmios. Suponga que la resistencia sigue una distribución normal y se puede medir con cualquier grado de precisión. ¿qué porcentaje de resistores tendrán una resistencia que exceda 43 ohmios?

SOLUCIÓN

Se encuentra un porcentaje al multiplicar la frecuencia relativa por 100%. Como la frecuencia relativa para un intervalo es igual a la probabilidad de caer en el intervalo, debemos encontrar el área a la derecha de $x = 43$ en la figura 6.18. Esto se puede hacer al transformar $x = 43$ al valor z correspondiente, con lo que se obtiene el área a la izquierda de z de la tabla A.3, y después se resta esta área de 1. Encontramos que

$$z = \frac{43 - 40}{2} = 1.5.$$

De aquí

$$P(X > 43) = P(Z > 1.5) = 1 - P(Z < 1.5) = 1 - 0.9332 = 0.0668.$$

Por tanto, 6.68% de los resistores tendrán una resistencia que excede 43 ohmios. ■

Ejemplo 6.12 Encuentre el porcentaje de resistores que exceden 43 ohmios para el ejemplo 6.11 si la resistencia se mide al ohm más cercano.

SOLUCIÓN

Este problema difiere del ejemplo 6.11 pues ahora asignamos una medida de 43 ohmios a todos los resistores cuyas resistencias sean mayores que 42.5 y menores que 43.5. Realmente aproximamos una distribución discreta por medio de una distribución continua normal. El área que se requiere es la región sombreada a la derecha de 43.5 en la figura 6.19. Encontramos ahora que

$$z = \frac{43.5 - 40}{2} = 1.75.$$

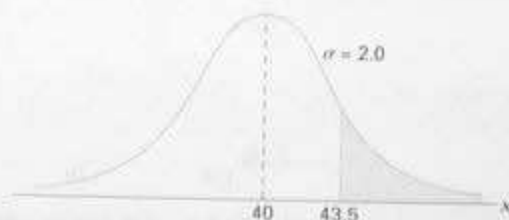


Figura 6.19 Área para el ejemplo 6.12.

De aquí

$$P(X > 43.5) = P(Z > 1.75) = 1 - P(Z < 1.75) = 1 - 0.9599 = 0.0401.$$

Por tanto, 4.01% de los resistores exceden 43 ohmios cuando se miden al ohmio más cercano. La diferencia $6.68\% - 4.01\% = 2.67\%$ entre esta respuesta y la del ejemplo 6.11 representa todos los resistores que tienen una resistencia mayor que 43 y menor que 43.5 que ahora se registran como de 43 ohmios. ■

Ejemplo 6.13 La calificación promedio para un examen es 74 y la desviación estándar es 7. Si 12% de la clase obtiene A y las calificaciones siguen una curva que tiene una distribución normal, ¿cuál es la A más baja posible y la B más alta posible?

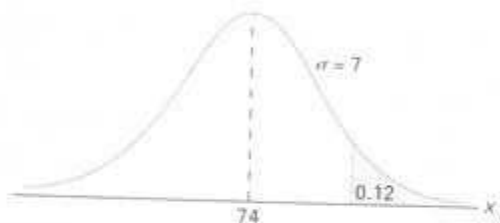


Figura 6.20 Área para el ejemplo 6.13.

SOLUCIÓN

En este ejemplo comenzamos con una área de probabilidad conocida, encontramos el valor z y después determinamos x de la fórmula $x = \sigma z + \mu$. Un área de 0.12, que corresponde a la fracción de estudiantes que reciben A, se sombrea en la figura 6.20. Requerimos un valor z que deje 0.12 del área a la derecha y, por ello, un área de 0.88 a la izquierda. De la tabla A.3, $P(Z < 1.175) = 0.88$, por lo que el valor z que se desea es 1.175. De aquí

$$x = (7)(1.175) + 74 = 82.225.$$

Por tanto, la A más baja es 83 y la B más alta es 82. ■

Ejemplo 6.14 Refiérase al ejemplo 6.13 y encuentre el sexto decil.

SOLUCIÓN

El sexto decil, escrito como D_6 , es el valor x que deja 60% del área a la izquierda como se muestra en la figura 6.21. De la tabla A.3 encontramos $P(Z < 0.25) = 0.6$, por lo que el valor z que se desea es 0.25. Ahora bien, $x = (7)(0.25) + 74 = 75.75$. De aquí $D_6 = 75.75$. Es decir, 60% de las calificaciones son 75 o menos. ■

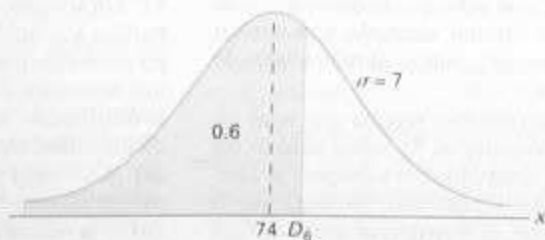


Figura 6.21 Área para el ejemplo 6.14.

Ejercicios

1. Dada una distribución normal estándar, encuentre el área bajo la curva que está

- (a) a la izquierda de $z = 1.43$;
- (b) a la derecha de $z = -0.89$;
- (c) entre $z = -2.16$ y $z = -0.65$;
- (d) a la izquierda de $z = -1.39$;
- (e) a la derecha de $z = 1.96$;
- (f) entre $z = -0.48$ y $z = 1.74$.

2. Encuentre el valor de z si el área bajo una curva normal estándar

- (a) a la derecha de z es 0.3622;
- (b) a la izquierda de z es 0.1131;
- (c) entre 0 y z , con $z > 0$, es 0.4838;
- (d) entre $-z$ y z , con $z > 0$, es 0.9500.

3. Dada una distribución normal estándar, encuentre el valor de k tal que

- (a) $P(Z < k) = 0.0427$;
- (b) $P(Z > k) = 0.2946$;
- (c) $P(-0.93 < Z < k) = 0.7235$.

4. Dada una distribución normal con $\mu = 30$ y $\sigma = 6$, encuentre

- (a) el área de la curva normal a la derecha de $x = 17$;
- (b) el área de la curva normal a la izquierda de $x = 22$;
- (c) el área de la curva normal entre $x = 32$ y $x = 41$;
- (d) el valor de x que tiene 80% del área de la curva normal a la izquierda;
- (e) los dos valores de x que contienen 75% central del área de la curva normal.

5. Dada la variable X normalmente distribuida con media 18 y desviación estándar 2.5, encuentre

- (a) $P(X < 15)$;
- (b) el valor de k tal que $P(X < k) = 0.2236$;
- (c) el valor de k tal que $P(X > k) = 0.1814$;
- (d) $P(17 < X < 21)$.

6. De acuerdo con el teorema de Chebyshev, la probabilidad de que cualquier variable aleatoria tome un valor dentro de tres desviaciones estándar de la media es al menos 8/9. Si se sabe que la distribución de probabilidad de una variable aleatoria X es normal con media μ y varianza σ^2 , ¿cuál es el valor exacto de $P(\mu - 3\sigma < X < \mu + 3\sigma)$?

7. Un investigador científico reporta que unos ratones vivirán un promedio de 40 meses cuando sus dietas se restringen drásticamente y después se enriquecen con vitaminas y proteínas. Suponga que la vida de tales ratones se distribuyen normalmente con una desviación estándar de 6.3 meses, encuentre la probabilidad de que un ratón dado viva

- (a) más de 32 meses;
- (b) menos de 28 meses;
- (c) entre 37 y 49 meses.

8. La barras de pan de centeno que cierta panadería distribuye a las tiendas locales tienen una longitud promedio de 30 centímetros y una desviación estándar de dos centímetros. Suponga que las longitudes están distribuidas normalmente, ¿qué porcentaje de las barras son

- (a) más largas que 31.7 centímetros?
- (b) entre 29.3 y 33.5 centímetros de longitud?
- (c) más cortas que 25.5 centímetros?

9. Se regula una máquina despachadora de refresco para que sirva un promedio de 200 mililitros por vaso. Si la cantidad de bebida se distribuye normalmente con una desviación estándar igual a 15 mililitros,

- (a) ¿qué fracción de los vasos contendrán más de 224 mililitros?
- (b) ¿cuál es la probabilidad de que un vaso contenga entre 191 y 209 mililitros?
- (c) ¿cuántos vasos probablemente se derramarán si se utilizan vasos de 230 mililitros para las siguientes 1000 bebidas?
- (d) ¿por debajo de qué valor obtendremos 25% de las bebidas más pequeñas?

10. El diámetro interior del anillo de un pistón ya terminado se distribuye normalmente con una media de 10 centímetros y una desviación estándar de 0.03 centímetros.

- (a) ¿Qué proporción de anillos tendrán diámetros interiores que excedan 10.075 centímetros?
- (b) ¿Cuál es la probabilidad de que el anillo de un pistón tenga un diámetro interior entre 9.97 y 10.03 centímetros?
- (c) ¿Por debajo de qué valor del diámetro interior caerá el 15% de los anillos de pistón?

11. Un abogado va todos los días de su casa en los suburbios a su oficina en el centro de la ciudad. El tiempo promedio para un viaje de ida es 24 minutos, con una desviación estándar de 3.8 minutos. Suponga que la distribución de los tiempos de viaje está distribuida normalmente.

- (a) ¿Cuál es la probabilidad de que un viaje tome al menos 1/2 hora?
- (b) Si la oficina abre a las 9:00 A.M. y él sale diario de su casa a las 8:45 A.M., ¿qué porcentaje de las veces llega tarde al trabajo?

(c) Si sale de su casa a las 8:35 A.M. y el café se sirve en la oficina de 8:50 A.M. a 9:00 A.M., cuál es la probabilidad de que pierda el café?

(d) Encuentre la longitud de tiempo por arriba de la cual encontramos el 15% de los viajes más lentos.

(e) Encuentre la probabilidad de que dos de los siguientes tres viajes tomen al menos 1/2 hora.

12. En el ejemplar de noviembre de 1990 del *Chemical Engineering Progress* un estudio discute sobre el porcentaje de pureza del oxígeno de cierto proveedor. Suponga que la media fue 99.61 con una desviación estándar de 0.08. Suponga que la distribución del porcentaje de pureza fue aproximadamente normal.

(a) ¿Qué porcentaje de los valores de pureza esperaría que estuvieran entre 99.5 y 99.7?

(b) ¿Qué valor de pureza esperaría que excediera exactamente 5% de la población?

13. La vida promedio de cierto tipo de motor pequeño es 10 años con una desviación estándar de dos años. El fabricante reemplaza gratis todos los motores que fallen dentro del tiempo de garantía. Si está dispuesto a reemplazar sólo 3% de los motores que fallan, ¿de qué duración debe ser la garantía que ofrezca? Suponga que la duración de un motor sigue una distribución normal.

14. Las alturas de 1000 estudiantes se distribuyen normalmente con una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros. Suponga que las alturas se registran al medio centímetro más cercano, ¿cuántos de estos estudiantes esperaría que tuvieran alturas

(a) menores que 160.0 centímetros?

(b) entre 171.5 y 182.0 centímetros inclusive?

(c) igual a 175.0 centímetros?

(d) mayor que o igual a 188.0 centímetros?

15. Una compañía paga a sus empleados un salario promedio de \$15.90 por hora con una desviación estándar de \$1.50. Si los salarios se distribuyen aproximadamente de forma normal y se pagan al centavo más próximo,

(a) ¿qué porcentaje de los trabajadores reciben salarios entre \$13.75 y \$16.22 inclusive por hora?

(b) ¿el 5% más alto de los salarios por hora de los empleados es mayor a qué cantidad?

16. Los pesos de un número grande de perros poodle miniatura se distribuyen aproximadamente de forma normal con una media de ocho kilogramos y una desviación estándar de 0.9 kilogramos. Si las mediciones se registran al décimo de kilogramo más cercano, encuentre la fracción de estos poodle con pesos

(a) por arriba de 9.5 kilogramos;

(b) a lo más 8.6 kilogramos;

(c) entre 7.3 y 9.1 kilogramos inclusive.

17. La resistencia a la tracción de cierto componente de metal se distribuye normalmente con una media de 10,000 kilogramos por centímetro cuadrado y una desviación estándar de 100 kilogramos por centímetro cuadrado. Las mediciones se registran a los 50 kilogramos por centímetro cuadrado más cercanos.

(a) ¿Qué proporción de estos componentes excede 10,150 kilogramos por centímetro cuadrado de resistencia a la tracción?

(b) Si las especificaciones requieren que todos los componentes tengan resistencia a la tracción entre 9800 y 10,200 kilogramos por centímetro cuadrado inclusive, ¿qué proporción de piezas esperaría que se descartará?

18. Si un conjunto de observaciones se distribuye de manera normal, ¿qué porcentaje de estas difieren de la media en

(a) más de 1.3σ ?

(b) menos de 0.52σ ?

19. Los CI de 600 aspirantes de cierta universidad se distribuyen aproximadamente de forma normal con una media de 115 y una desviación estándar de 12. Si la universidad requiere un CI de al menos 95, ¿cuántos de estos estudiantes serán rechazados sobre esta base sin importar sus otras calificaciones?

20. Dada una distribución continua uniforme, demuestre que

$$(a) \mu = \frac{A + B}{2}$$

$$(b) \sigma^2 = \frac{(B - A)^2}{12}$$

21. La cantidad diaria de café, en litros, que sirve una máquina que se localiza en el vestíbulo de un aeropuerto es una variable aleatoria X que tiene una distribución continua uniforme con $A = 7$ y $B = 10$. Encuentre la probabilidad de que en un día dado la cantidad de café que sirve esta máquina sea

(a) a lo más 8.8 litros;

(b) más de 7.4 litros pero menos de 9.5 litros;

(c) al menos 8.5 litros.

22. Un autobús llega cada 10 minutos a una parada. Se supone que el tiempo de espera para un individuo particular es una variable aleatoria con distribución uniforme.

(a) ¿Cuál es la probabilidad de que el individuo espere más de siete minutos?

6.5 Aproximación normal a la binomial

Las probabilidades asociadas con experimentos binomiales se obtienen fácilmente a partir de la fórmula $b(x; n, p)$ de la distribución binomial o de la tabla A.1 cuando n es pequeña. Además, las probabilidades binomiales están fácilmente disponibles en muchos paquetes de software. Sin embargo, es instructivo aprender la relación entre la distribución binomial y la normal. En la sección 5.6 ilustramos cómo la distribución de Poisson se puede utilizar para aproximar probabilidades binomiales cuando n es bastante grande y p es muy cercana a 0 o a 1. Las distribuciones binomial y de Poisson son discretas ambas. La primera aplicación de una distribución continua de probabilidad para aproximar probabilidades sobre un espacio muestral discreto se demuestra en el ejemplo 6.12, donde se utiliza la curva normal. La distribución normal a menudo es una buena aproximación a una distribución discreta cuando la última adquiere una forma de campana simétrica. Desde un punto de vista teórico, algunas distribuciones convergen a la normal conforme sus parámetros se aproximan a ciertos límites. La distribución normal es una distribución de aproximación conveniente pues la función de distribución acumulada se tabula muy fácil. La distribución binomial se aproxima bien por la normal en problemas prácticos cuando se trabaja con la función de distribución acumulada. Establezcemos ahora un teorema que nos permitirá utilizar áreas bajo la curva normal para aproximar propiedades binomiales cuando n es suficientemente grande.

Teorema 6.2

Si X es una variable aleatoria binomial con media $\mu = np$ y varianza $\sigma^2 = npq$, entonces la forma limitante de la distribución de

$$Z = \frac{X - np}{\sqrt{npq}}$$

conforme $n \rightarrow \infty$, es la distribución normal estándar $n(z; 0, 1)$.

Resulta que la distribución normal con $\mu = np$ y $\sigma^2 = np(1 - p)$ no sólo proporciona una aproximación muy precisa a la distribución binomial cuando n es grande y p no está extremadamente cercana a 0 o 1, sino que también proporciona una aproximación bastante buena aun cuando n es pequeña y p está razonablemente cercana a 1/2.

Para ilustrar la aproximación normal a la distribución binomial, primero dibujamos el histograma para $b(x; 15, 0.4)$ y después superponemos la curva normal particular que tenga la misma media y varianza que la variable binomial X . De aquí dibujamos una curva normal con

$$\mu = np = (15)(0.4) = 6 \quad \text{y} \quad \sigma^2 = npq = (15)(0.4)(0.6) = 3.6$$

El histograma de $b(x; 15, 0.4)$ y la curva normal superpuesta correspondiente, que está determinada por completo por su media y su varianza, se ilustran en la figura 6.22.

La probabilidad exacta de que la variable aleatoria binomial X tome un valor dado x es igual al área cuya base se centra en x . Por ejemplo, la probabilidad exacta de que X tome el valor 4 es igual al área del rectángulo con base centrada en $x = 4$. Con el uso de la tabla A.1, encontramos que esta área es

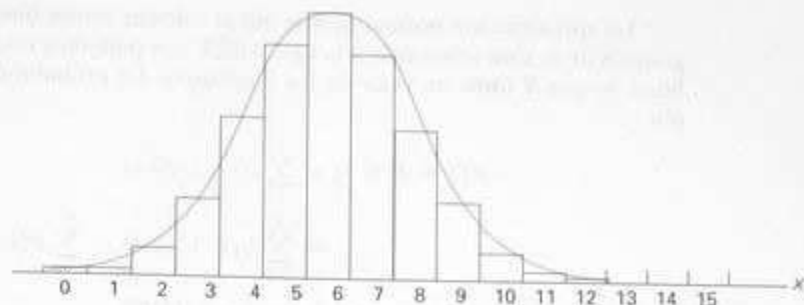


Figura 6.22 Aproximación normal de $b(x; 15, 0.4)$.

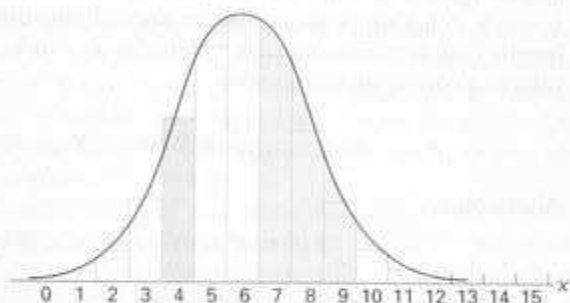


Figura 6.23 Aproximación normal de $b(4; 15, 0.4)$ y $\sum_{x=2}^x b(x; 15, 0.4)$.

$$P(X = 4) = b(4; 15, 0.4) = 0.1268,$$

que es aproximadamente igual al área de la región sombreada bajo la curva normal entre las dos ordenadas $x_1 = 3.5$ y $x_2 = 4.5$ en la figura 6.23. Al convertir a valores z , tenemos

$$z_1 = \frac{3.5 - 6}{1.897} = -1.32 \quad \text{y} \quad z_2 = \frac{4.5 - 6}{1.897} = -0.79.$$

Si X es una variable aleatoria binomial y Z una variable normal estándar, entonces

$$\begin{aligned} P(X = 4) &= b(4; 15, 0.4) \\ &\approx P(-1.32 < Z < -0.79) \\ &= P(Z < -0.79) - P(Z < -1.32) \\ &= 0.2148 - 0.0934 \\ &= 0.1214. \end{aligned}$$

Esto está bastante de acuerdo con el valor exacto de 0.1268.

La aproximación normal es más útil al calcular sumas binomiales para valores grandes de n . Con referencia a la figura 6.23, nos podemos interesar en la probabilidad de que X tome un valor de 7 a 9 inclusive. La probabilidad exacta está dada por

$$\begin{aligned} P(7 \leq X \leq 9) &= \sum_{x=7}^9 b(x; 15, 0.4) \\ &= \sum_{x=7}^9 b(x; 15, 0.4) - \sum_{x=6}^6 b(x; 15, 0.4) \\ &= 0.9662 - 0.6098 \\ &= 0.3564. \end{aligned}$$

que es igual a la suma de las áreas de los rectángulos con bases centradas en $x = 7, 8,$ y 9 . Para la aproximación normal encontramos el área de la región sombreada bajo la curva entre las ordenadas $x_1 = 6.5$ y $x_2 = 9.5$ en la figura 6.23. Los valores z correspondientes son

$$z_1 = \frac{6.5 - 6}{1.897} = 0.26, \quad \text{y} \quad z_2 = \frac{9.5 - 6}{1.897} = 1.85.$$

Ahora bien,

$$\begin{aligned} P(7 \leq X \leq 9) &= P(0.26 < Z < 1.85) \\ &= P(Z < 1.85) - P(Z < 0.26) \\ &= 0.9678 - 0.6026 \\ &= 0.3652. \end{aligned}$$

Una vez más, la aproximación de la curva normal proporciona un valor que está muy de acuerdo con el valor exacto de 0.3564. El grado de precisión, que depende de qué tan bien se ajusta la curva al histograma, aumentará conforme n aumente. Esto es particularmente cierto cuando p no es muy cercana $1/2$ y el histograma ya no es simétrico. Las figuras 6.24 y 6.25 muestran los histogramas para $b(x; 6, 0.2)$ y $b(x; 15, 0.2)$, respectivamente. Es evidente que una curva normal se ajustará al histograma cuando $n = 15$ considerablemente mejor que cuando $n = 6$.

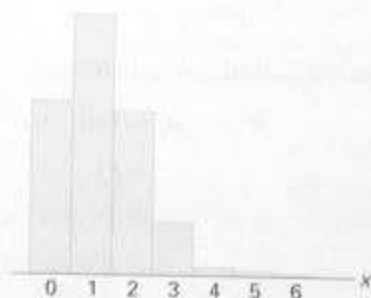


Figura 6.24 Histograma para $b(x; 6, 0.2)$.

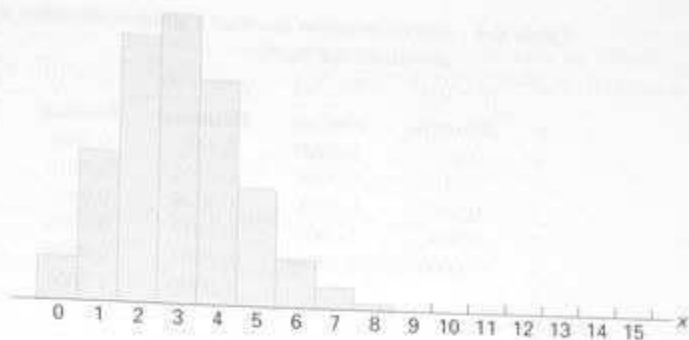


Figura 6.25 Histograma para $b(x; 15, 0.2)$.

En resumen, utilizamos la aproximación normal para evaluar probabilidades binomiales siempre que p no sea cercana a 0 o a 1. La aproximación es excelente cuando n es grande y muy buena para valores pequeños de n si p es razonablemente cercana a $1/2$. Al calcular np y nq se proporciona una posible guía para determinar cuándo se puede utilizar la aproximación normal. Si np y nq son mayores que o iguales a 5, la aproximación será buena.

Como indicamos antes, la calidad de la aproximación es bastante buena para n grande. Si p es cercana a $1/2$, un tamaño de la muestra moderado o pequeño será suficiente para una aproximación razonable. Ofrecemos la tabla 6.1 en la página 164 como una indicación de la calidad de la aproximación. Se dan tanto la aproximación normal como las probabilidades binomiales acumuladas reales. Nótese que en $p = 0.05$ y $p = 0.10$, la aproximación es bastante gruesa para $n = 10$. Sin embargo, aun para $n = 10$, nótese la mejoría para $p = 0.50$. Por otro lado, cuando p es fija en $p = 0.05$, nótese la mejoría de la aproximación conforme vamos de $n = 20$ a $n = 100$.

Ejemplo 6.15 La probabilidad de que un paciente se recupere de una rara enfermedad de la sangre es 0.4. Si se sabe que 100 personas contraen esta enfermedad, ¿cuál es la probabilidad de que menos de 30 sobrevivan?

SOLUCIÓN

Represéntese con la variable binomial X el número de pacientes que sobreviven. Como $n = 100$, debemos obtener resultados bastante precisos con el uso de la aproximación de la curva normal con

$$\mu = np = (100)(0.4) = 40$$

y

$$\sigma = \sqrt{npq} = \sqrt{(100)(0.4)(0.6)} = 4.899.$$

Para obtener la probabilidad deseada, tenemos que encontrar el área a la izquierda de $x = 29.5$. El valor z que corresponde a 29.5 es

$$z = \frac{29.5 - 40}{4.899} = -2.14,$$

Tabla 6.1 Aproximación normal y probabilidades binomiales acumuladas reales

r	$p = 0.05, n = 10$		$p = 0.10, n = 10$		$p = 0.50, n = 10$	
	Binomial	Normal	Binomial	Normal	Binomial	Normal
0	0.5987	0.5000	0.3487	0.2981	0.0010	0.0022
1	0.9139	0.9265	0.7361	0.7019	0.0107	0.0136
2	0.9885	0.9981	0.9298	0.9429	0.0547	0.0571
3	0.9990	1.0000	0.9872	0.9959	0.1719	0.1711
4	1.0000	1.0000	0.9984	0.9999	0.3770	0.3745
5			1.0000	1.0000	0.6230	0.6255
6					0.8281	0.8289
7					0.9453	0.9429
8					0.9893	0.9864
9					0.9990	0.9978
10					1.0000	0.9997

r	$p = 0.05$					
	$n = 20$		$n = 50$		$n = 100$	
	Binomial	Normal	Binomial	Normal	Binomial	Normal
0	0.3585	0.3015	0.0769	0.0968	0.0059	0.0197
1	0.7358	0.6985	0.2794	0.2578	0.0371	0.0537
2	0.9245	0.9382	0.5405	0.5000	0.1183	0.1251
3	0.9841	0.9948	0.7604	0.7422	0.2578	0.2451
4	0.9974	0.9998	0.8964	0.9032	0.4360	0.4090
5	0.9997	1.0000	0.9622	0.9744	0.6160	0.5910
6	1.0000	1.0000	0.9882	0.9953	0.7660	0.7549
7			0.9968	0.9994	0.8720	0.8749
8			0.9992	0.9999	0.9369	0.9463
9			0.9998	1.0000	0.9718	0.9803
10			1.0000	1.0000	0.9885	0.9941

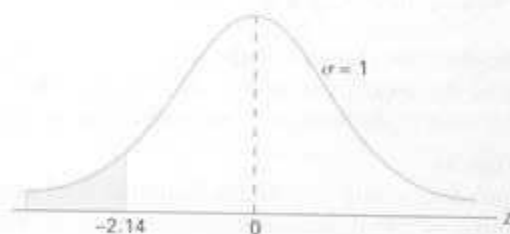


Figura 6.26 Área para el ejemplo 6.15.

y la probabilidad de que menos de 30 de los 100 pacientes sobrevivan está dada por la región sombreada en la figura 6.26. De aquí,

$$P(X < 30) = P(Z < -2.14) = 0.0162. \blacksquare$$

Ejemplo 6.16 Una prueba de opción múltiple tiene 200 preguntas cada una con cuatro respuestas posibles de las que sólo una es la correcta. ¿Cuál es la probabilidad de que con puras conjeturas se obtengan de 25 a 30 respuestas correctas para 80 de los 200 problemas acerca de los que el estudiante no tiene conocimientos?

SOLUCIÓN

La probabilidad de una respuesta correcta para cada una de las 80 preguntas es $p = 1/4$. Si X representa el número de respuestas correctas debidas a conjeturas, entonces

$$P(25 \leq X \leq 30) = \sum_{x=25}^{30} b(x; 80, \frac{1}{4}).$$

Al usar la aproximación de la curva normal con

$$\mu = np = (80)(\frac{1}{4}) = 20$$

y

$$\sigma = \sqrt{npq} = \sqrt{(80)(\frac{1}{4})(\frac{3}{4})} = 3.873,$$

necesitamos el área entre $x_1 = 24.5$ y $x_2 = 30.5$. Los valores z correspondientes son

$$z_1 = \frac{24.5 - 20}{3.873} = 1.16 \quad \text{y} \quad z_2 = \frac{30.5 - 20}{3.873} = 2.71.$$

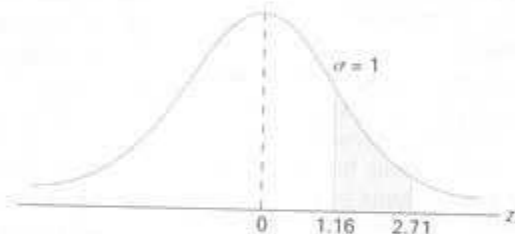


Figura 6.27 Áreas para el ejemplo 6.16.

La probabilidad de adivinar correctamente de 25 a 30 preguntas está dada por la región sombreada de la figura 6.27. De la tabla A.3 encontramos que

$$\begin{aligned} P(25 \leq X \leq 30) &= \sum_{x=25}^{30} b(x; 80, \frac{1}{4}) \\ &= P(1.16 < Z < 2.71) \\ &= P(Z < 2.71) - P(Z < 1.16) \\ &= 0.9966 - 0.8770 \\ &= 0.1196. \quad \blacksquare \end{aligned}$$

Ejercicios

1. Evalúe $P(1 \leq X \leq 4)$ para una variable binomial con $n = 15$ y $p = 0.2$ con el uso de
 - (a) la tabla A.1 del apéndice;
 - (b) la aproximación de la curva normal.
2. Se lanza 400 veces una moneda. Utilice la aproximación de la curva normal para encontrar la probabilidad de obtener
 - (a) entre 185 y 210 caras inclusive;
 - (b) exactamente 205 caras;
 - (c) menos de 176 o más de 227 caras.
3. Un proceso para fabricar un componente electrónico tiene 1% de defectuosos. Un plan de control de calidad es seleccionar 100 artículos del proceso, y si

ninguno está defectuoso el proceso continúa. Use la aproximación normal a la binomial para encontrar

- (a) la probabilidad de que el proceso continúe con el plan de muestreo que se describe;
 (b) la probabilidad de que el proceso continúe aun si éste está mal (es decir, si la frecuencia de componentes defectuosos cambia a 5.0% de defectuosos).

4. Un proceso produce 10% de artículos defectuosos. Si se seleccionan al azar 100 artículos del proceso, ¿cuál es la probabilidad de que el número de defectuosos

- (a) exceda de 13?
 (b) sea menor que 8?

5. La probabilidad de que un paciente se recupere de una delicada operación de corazón es 0.9. De los siguientes 100 pacientes que tienen esta operación, ¿cuál es la probabilidad de que

- (a) sobrevivan entre 84 y 95 inclusive?
 (b) sobrevivan menos de 86?

6. Investigadores de la Universidad George Washington y del Instituto Nacional de Salud reportan que aproximadamente 75% de las personas creen que "los tranquilizantes funcionan muy bien para hacer que una persona esté más tranquila y relajada". De las siguientes 80 personas entrevistadas, ¿cuál es la probabilidad de que

- (a) al menos 50 sean de esta opinión?
 (b) a lo más 56 tengan esta opinión?

7. Si 20% de los residentes de una ciudad de Estados Unidos prefiere un teléfono blanco sobre cualquier otro color disponible, ¿cuál es la probabilidad de que entre los siguientes 1000 teléfonos que se instalan en esta ciudad

- (a) entre 170 y 185 inclusive sean blancos?
 (b) al menos 210 pero no más de 225 sean blancos?

8. Un fabricante de medicamentos sostiene que cierto medicamento cura una enfermedad de la sangre en promedio en 80% de los casos. Para verificar la aseveración, inspectores gubernamentales utilizan el medicamento en una muestra de 100 individuos y deciden aceptar la afirmación si 75 o más se curan.

(a) ¿Cuál es la probabilidad de que la afirmación se rechace cuando la probabilidad de curación es, de hecho, 0.8?

(b) ¿Cuál es la probabilidad de que el gobierno acepte la afirmación cuando la probabilidad de curación sea tan baja como 0.7?

9. Un sexto de los estudiantes hombres de primer ingreso que entran a una escuela estatal grande son de otros estados. Si los estudiantes se asignan al azar a los dormitorios, 180 en un edificio, ¿cuál es la probabilidad de que en un dormitorio dado al menos un quinto de los estudiantes sea de otro estado?

10. Una compañía farmacéutica sabe que aproximadamente 5% de sus píldoras anticonceptivas tienen un ingrediente que está por debajo de la dosis mínima, lo que vuelve ineficaz a la píldora. ¿Cuál es la probabilidad de que menos de 10 en una muestra de 200 píldoras sean ineficaces?

11. Estadísticas publicadas por la Administración Nacional de Seguridad de Tránsito en Carreteras y el Consejo de Seguridad Nacional muestran que en una noche promedio de fin de semana, uno de cada 10 conductores está ebrio. Si se verifican 400 conductores al azar la siguiente noche de sábado, ¿cuál es la probabilidad de que el número de conductores ebrios sea

- (a) menor que 32?
 (b) más de 49?
 (c) al menos 35 pero menos de 47?

12. Se lanza 180 veces un par de dados. ¿Cuál es la probabilidad de que ocurra un total de siete

- (a) al menos 25 veces?
 (b) entre 33 y 41 veces inclusive?
 (c) exactamente 30 veces?

13. Una compañía produce componentes para un motor. Las especificaciones de las partes sugieren que 95% de los artículos cumplen con las especificaciones. Las partes se embarcan en lotes de 100 a los clientes.

- (a) ¿Cuál es la probabilidad de que más de dos artículos estén defectuosos en un lote dado?
 (b) ¿Cuál es la probabilidad de que más de 10 artículos estén defectuosos en un lote?

6.6 Distribuciones gamma y exponencial

Aunque la distribución normal se puede utilizar para resolver muchos problemas en ingeniería y en la ciencia, hay aún numerosas situaciones que requieren diferentes tipos de funciones de densidad. Dos de estas funciones de densidad, las **distribuciones gamma y exponencial**, se estudian en esta sección.

Resulta que la distribución exponencial es un caso especial de la distribución gamma. Ambas encuentran un gran número de aplicaciones. Las distribuciones exponencial y gamma juegan un papel importante en teoría de colas y problemas de confiabilidad. Los tiempos entre llegadas en instalaciones de servicio, y tiempo de falla de partes componentes y sistemas eléctricos, a menudo quedan bien modelados mediante la distribución exponencial. La relación entre la gamma y la exponencial permite que la gamma se involucre en tipos de problemas similares. En la sección 6.7 se proporcionarán más detalles e ilustraciones.

La distribución gamma deriva su nombre de la bien conocida **función gamma**, que se estudia en muchas áreas de las matemáticas. Antes de que procedamos con la distribución gamma, revisemos esta función y algunas de sus propiedades importantes.

Definición 6.2

La **función gamma** se define como

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

para $\alpha > 0$.

Al integrar por partes con $u = x^{\alpha-1}$ y $dv = e^{-x} dx$, obtenemos

$$\Gamma(\alpha) = -e^{-x}x^{\alpha-1} \Big|_0^{\infty} + \int_0^{\infty} e^{-x}(\alpha-1)x^{\alpha-2} dx = (\alpha-1) \int_0^{\infty} x^{\alpha-2} e^{-x} dx$$

para $\alpha > 1$, que produce la fórmula recursiva

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1).$$

La aplicación repetida de la fórmula recursiva da

$$\Gamma(\alpha) = (\alpha-1)(\alpha-2)\Gamma(\alpha-2) = (\alpha-1)(\alpha-2)(\alpha-3)\Gamma(\alpha-3),$$

y así sucesivamente. Nótese que cuando $\alpha = n$, donde n es un entero positivo,

$$\Gamma(n) = (n-1)(n-2)\dots\Gamma(1).$$

Sin embargo, por la definición 6.2,

$$\Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$$

y de aquí

$$\Gamma(n) = (n-1)!.$$

Una propiedad importante de la función gamma, que se deja al lector para su verificación (véase el ejercicio 3 de la página 177), es que $\Gamma(1/2) = \sqrt{\pi}$.

Incluiremos ahora la función gamma en nuestra definición de la distribución gamma.

Distribución Gamma

La variable aleatoria continua X tiene una **distribución gamma**, con parámetros α y β , si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

cuando $\alpha > 0$ y $\beta > 0$.

En la figura 6.28 se muestran gráficas de varias distribuciones gamma para ciertos valores específicos de los parámetros α y β . La distribución gamma especial para la que $\alpha = 1$ se llama **distribución exponencial**.

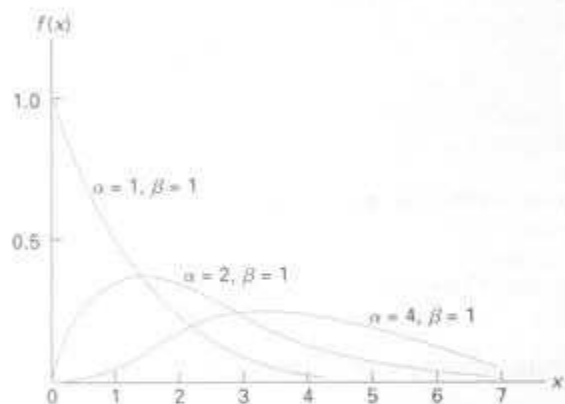


Figura 6.28 Distribuciones gamma.

Distribución exponencial

La variable aleatoria continua X tiene una **distribución exponencial**, con parámetro β , si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

donde $\beta > 0$.

El siguiente teorema y corolario dan la media y la varianza de las distribuciones gamma y exponencial.

Teorema 6.3

La media y la varianza de la distribución gamma son

$$\mu = \alpha\beta \quad \text{y} \quad \sigma^2 = \alpha\beta^2.$$

PRUEBA Para encontrar la media de la distribución gamma, escribimos

$$\mu = E(X) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty x^\alpha e^{-x/\beta} dx.$$

Ahora bien, sea $y = x/\beta$, que da

$$\mu = \frac{\beta}{\Gamma(\alpha)} \int_0^\infty y^\alpha e^{-y} dy = \frac{\beta \Gamma(\alpha + 1)}{\Gamma(\alpha)} = \alpha\beta.$$

Para encontrar la varianza de la distribución gamma, procedemos como antes para obtener:

$$E(X^2) = \frac{\beta^2 \Gamma(\alpha + 2)}{\Gamma(\alpha)} = (\alpha + 1)\alpha\beta^2,$$

y entonces

$$\sigma^2 = E(X^2) - \mu^2 = (\alpha + 1)\alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2.$$

Corolario 1

La media y la varianza de la distribución exponencial son

$$\mu = \beta \quad \text{y} \quad \sigma^2 = \beta^2.$$

Relación con el proceso de Poisson

Proseguiremos con las aplicaciones de la distribución exponencial y después regresaremos a la distribución gamma. Las aplicaciones más importantes de la distribución exponencial son situaciones donde se aplica el proceso de Poisson (véase el capítulo 5). El lector debe recordar que el proceso de Poisson permite el uso de la distribución discreta llamada distribución de Poisson. Recuérdese que la distribución de Poisson se utiliza para calcular la probabilidad de números específicos de "eventos" durante un *periodo o espacio* particular. En muchas aplicaciones, el tiempo o la cantidad de espacio es la variable aleatoria. Por ejemplo, un ingeniero industrial se puede interesar en modelar el tiempo T entre llegadas a una intersección congestionada durante una hora pico en una ciudad grande. Una llegada representa el evento de Poisson.

La relación entre la distribución exponencial (a menudo denominada exponencial negativa) y el proceso de Poisson es bastante simple. En el capítulo 5 se desarrolló la distribución de Poisson como una distribución de un solo parámetro con parámetro λ , donde λ se puede interpretar como el número medio de eventos *por unidad de "tiempo"*. Considérese ahora la *variable aleatoria* descrita por el tiempo que se requiere para que ocurra el primer evento. Con el uso de la distribución de Poisson, encontramos que la probabilidad de que no ocurra algún evento, en el periodo hasta el tiempo t está dada por

$$p(0; \lambda t) = \frac{e^{-\lambda t} (\lambda t)^0}{0!} = e^{-\lambda t}.$$

Podemos ahora utilizar lo anterior y hacer que X sea el tiempo para el primer evento de Poisson. La probabilidad de que la duración del tiempo hasta el primer evento exceda x es la misma que la probabilidad de que no ocurra algún evento de Poisson en x . Esto último, por supuesto, está dado por $e^{-\lambda x}$. Como resultado,

$$P(X \geq x) = e^{-\lambda x},$$

Así la función de distribución acumulada para X está dada por

$$P(0 \leq X \leq x) = 1 - e^{-\lambda x}.$$

Ahora bien, a fin de que reconozcamos la presencia de la distribución exponencial, podemos diferenciar la función de distribución acumulada anterior para obtener la función de densidad

$$f(x) = \lambda e^{-\lambda x},$$

que es la función de densidad de la distribución exponencial con $\lambda = 1/\beta$.

6.7 Aplicaciones de las distribuciones exponencial y gamma

En lo anteriormente mencionado proporcionamos las bases para la aplicación de la distribución exponencial en el “tiempo de llegada” o tiempo para problemas de eventos de Poisson. Ilustraremos aquí y después procederemos a discutir el papel de la distribución gamma en estas aplicaciones de modelado. Nótese que la media de la distribución exponencial es el parámetro β , el recíproco del parámetro en la distribución de Poisson. El lector debe recordar que con frecuencia se dice que la distribución de Poisson no tiene memoria, lo que implica que las ocurrencias en periodos sucesivos son independientes. El parámetro β importante es el tiempo medio entre eventos. En teoría de confiabilidad, donde la falla de equipo a menudo se ajusta a este proceso de Poisson, β se llama **tiempo medio entre fallas**. Muchas decomposturas de equipo siguen el proceso de Poisson, y por ello se aplica la distribución exponencial. Otras aplicaciones incluyen tiempos de sobrevivencia en experimentos biomédicos y tiempo de respuesta de computadoras.

En el siguiente ejemplo mostramos una aplicación simple de la distribución exponencial a un problema de confiabilidad. La distribución binomial también juega un papel en la solución.

Ejemplo 6.17 Suponga que un sistema contiene cierto tipo de componente cuyos tiempo de falla en años está dado por T . La variable aleatoria T se modela bien mediante la distribución exponencial con tiempo medio para la falla $\beta = 5$. Si se instalan cinco de estos componentes en diferentes sistemas, ¿cuál es la probabilidad de que al menos dos aún funcionen al final de ocho años?

SOLUCIÓN

La probabilidad de que un componente dado aún funcione después de ocho años está dada por

$$P(T > 8) = \frac{1}{5} \int_8^{\infty} e^{-t/5} dt = e^{-8/5} = 0.2.$$

Represéntese con X el número de componentes que funcionan después de ocho años. Entonces con el uso de la distribución binomial

$$P(X \geq 2) = \sum_{x=2}^5 b(x; 5, 0.2) = 1 - \sum_{x=0}^1 b(x; 5, 0.2) = 1 - 0.7373 = 0.2627.$$

Los ejercicios 1, 7, 8 y el ejercicio de repaso 2 proporcionan ejemplos adicionales del uso de la distribución exponencial en problemas de tiempo de espera y de confiabilidad.

La importancia de la distribución gamma está en el hecho de que define una familia de la que otras distribuciones son casos especiales. Pero la gamma misma tiene aplicaciones importantes en tiempo de espera y teoría de confiabilidad. Mientras que la distribución exponencial describe el tiempo hasta la ocurrencia de un evento de Poisson (o el tiempo entre eventos de Poisson), el tiempo (o espacio) que transcurre hasta que *ocurre un número específico de eventos de Poisson* es una variable aleatoria cuya función de densidad está descrita por la de la distribución gamma. Este número específico de eventos es el parámetro α en la función de densidad gamma. Así se vuelve fácil comprender que cuando $\alpha = 1$, ocurre el caso especial de la distribución exponencial. La densidad gamma se puede desarrollar de su relación con el proceso de Poisson de la misma manera en que lo hicimos con la densidad exponencial. Los detalles se dejan al lector. El siguiente es un ejemplo numérico del uso de la distribución gamma en una aplicación de tiempo de espera. ■

Ejemplo 6.18 Suponga que las llamadas telefónicas que llegan a un conmutador particular siguen un proceso de Poisson con un promedio de cinco llamadas que lleguen por minuto. ¿Cuál es la probabilidad de que pase más de un minuto hasta que lleguen dos llamadas al conmutador?

SOLUCIÓN

El proceso de Poisson se aplica al tiempo que pasa hasta la ocurrencia de dos eventos de Poisson que siguen una distribución gamma con $\beta = 1/5$ y $\alpha = 2$. Sea la variable aleatoria X el tiempo en minutos que transcurre antes de que lleguen dos llamadas. La probabilidad que se requiere está dada por

$$P(X \leq x) = \int_0^x \frac{1}{\beta^2} x e^{-x/\beta} dx$$

$$P(X \leq 1) = 25 \int_0^1 x e^{-5x} dx = [1 - e^{-5(1)}(1 + 5)] = 0.96.$$

Mientras el origen de la distribución gamma trata con el tiempo (o espacio) hasta la ocurrencia de α eventos de Poisson, hay muchos ejemplos donde una distribución gamma trabaja muy bien aunque no exista una estructura de Poisson clara. Esto es particularmente cierto para problemas de **tiempo de sobrevivencia** en aplicaciones de ingeniería y biomédicas. ■

Ejemplo 6.19 En un estudio biomédico con ratas se usa una investigación de respuesta a la dosis para determinar el efecto de la dosis de un tóxico en su tiempo de sobrevivencia. El tóxico es uno que se descarga con frecuencia en la atmósfera desde el combustible de los aviones. Para cierta dosis del tóxico el estudio determina que el tiempo de sobrevivencia, en semanas, tiene una distribución gamma con $\alpha = 5$ y $\beta = 10$. ¿Cuál es la probabilidad de que una rata no sobreviva más de 60 semanas?

SOLUCIÓN

Sea la variable aleatoria X el tiempo de sobrevivencia (tiempo para morir). La probabilidad que se requiere es

$$P(X \leq x) = \int_0^x \frac{1}{\beta^\alpha} \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)} dx,$$

$$P(X \leq 60) = \frac{1}{\beta^5} \int_0^{60} \frac{x^{4} e^{-x/\beta}}{\Gamma(5)} dx.$$

La integral anterior se puede resolver a través del uso de la **función gamma incompleta**, que resulta ser la función de distribución acumulada para la distribución gamma. Esta función se escribe como

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy.$$

Si hacemos $y = X/\beta$, $X = \beta y$ tenemos

$$P(X \leq 60) = \int_0^{60/\beta} \frac{y^4 e^{-y}}{\Gamma(5)} dy,$$

que se denota como $F(6; 5)$ en la tabla de la función gamma incompleta del apéndice A.24. Nótese que esto permite un cálculo rápido de las probabilidades para la distribución gamma. En realidad, para este problema la probabilidad de que la rata no sobreviva más de 60 días está dada por

$$P(X \leq 60) = F(6; 5) = 0.715. \quad \blacksquare$$

6.8 Distribución ji cuadrada

Otro caso especial muy importante de la distribución gamma se obtiene al hacer $\alpha = v/2$ y $\beta = 2$, donde v es un entero positivo. Este resultado se llama **distribución ji cuadrada**. La distribución tiene un solo parámetro, v , llamado **grados de libertad**.

Distribución ji cuadrada

La variable aleatoria continua X tiene una distribución ji cuadrada, con v grados de libertad, si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{2^{v/2} \Gamma(v/2)} x^{v/2-1} e^{-x/2}, & x > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

donde v es un entero positivo.

La distribución ji cuadrada juega un papel vital en la inferencia estadística. Tiene una aplicación considerable en la metodología y en la teoría. Aunque no discutimos aplicaciones con detalle en este capítulo, es importante saber que los capítulos 8, 9 y 16 contienen aplicaciones importantes. La distribución ji cuadrada es un componente importante de la prueba de hipótesis y de la estimación estadísticas.

Los temas que tratan con distribuciones de muestreo, análisis de varianza y estadística no paramétrica implican el uso extenso de la distribución ji cuadrada.

Corolario 2

La media y la varianza de la distribución ji cuadrada son

$$\mu = v \quad \text{y} \quad \sigma^2 = 2v.$$

6.9 Distribución logarítmica normal

La distribución logarítmica normal se utiliza en una amplia variedad de aplicaciones. La distribución se aplica en casos donde una transformación de logaritmo natural tiene como resultado una distribución normal.

Distribución logarítmica normal

La variable aleatoria continua X tiene una distribución logarítmica normal si la variable aleatoria $Y = \ln(X)$ tiene una distribución normal con media μ y desviación estándar σ . La función de densidad de X que resulta es

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x) - \mu]^2 / (2\sigma^2)}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Corolario 3

La media y varianza de la distribución logarítmica normal son

$$E(X) = e^{\mu + \sigma^2/2}, \quad \text{Var}(X) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$$

La función de distribución acumulada es bastante simple debido a su relación con la distribución normal. El uso de la función de distribución se ilustra mediante el siguiente ejemplo.

Ejemplo 6.20 Se sabe que históricamente la concentración de contaminantes producidos por plantas químicas exhiben un comportamiento que parece una distribución logarítmica normal. Esto es importante cuando se consideran problemas con respecto a la obediencia de las regulaciones gubernamentales. Suponga que se asume que la concentración de cierto contaminante, en partes por millón, tiene una distribución logarítmica normal con parámetros $\mu = 3.2$ y $\sigma = 1$. ¿Cuál es la probabilidad de que la concentración exceda ocho partes por millón?

SOLUCIÓN

Sea la variable aleatoria X la concentración de contaminantes

$$P[X > 8] = 1 - P[X \leq 8].$$

Como $\ln(X)$ tiene una distribución normal con media $\mu = 3.2$ y desviación estándar $\sigma = 1$,

$$P\{X \leq 8\} = \Phi\left[\frac{\ln(8) - 3,2}{1}\right] = \Phi(-1,12) = 0,1314.$$

Aquí, utilizamos la notación Φ para denotar la función de distribución acumulada de la distribución normal estándar. Como resultado, la probabilidad de que la concentración del contaminante exceda ocho partes por millón es 0.1314. ■

6.10 Distribución de Weibull

La tecnología moderna nos capacita para diseñar muchos sistemas complicados cuya operación, o quizá seguridad, depende de la confiabilidad de los diversos componentes que forman los sistemas. Por ejemplo, un fusible se puede quemar, una columna de acero se puede torcer o un dispositivo sensor de calor puede fallar. Componentes idénticos sujetos a idénticas condiciones ambientales fallarán en momentos diferentes e impredecibles. Vimos el papel que las distribuciones gamma y exponencial juegan en estos tipos de problemas. Otra distribución que se utiliza con amplitud en años recientes para tratar con tales problemas es la **distribución de Weibull**, que introdujo el físico sueco Waloddi Weibull en 1939.

Distribución de Weibull

La variable aleatoria continua X tiene una **distribución de Weibull**, con parámetros α y β si su función de densidad está dada por

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, & x > 0 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

donde $\alpha > 0$ y $\beta > 0$.

En la figura 6.29 se ilustran las gráficas de la distribución de Weibull para $\alpha = 1$ y diversos valores del parámetro β . Vemos que las curvas cambian de forma de manera considerable para diferentes valores de los parámetros, en particular el pará-

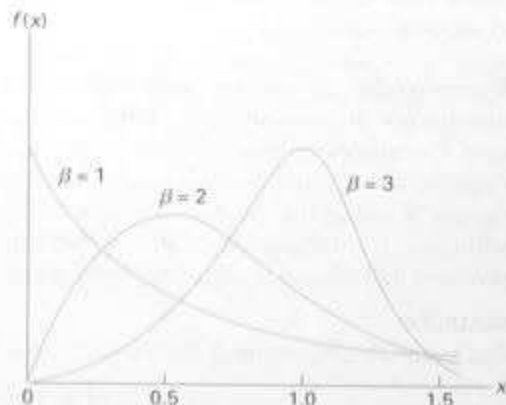


Figura 6.29 Distribuciones de Weibull ($\alpha = 1$).

metro β . Si hacemos $\beta = 1$, la distribución de Weibull se reduce a la distribución exponencial. Para valores de $\beta > 1$, las curvas se vuelven un poco en forma de campana y recuerdan las curvas normales, pero muestran asimetría.

La media y varianza de la distribución de Weibull se establecen en el siguiente teorema. Se pide al lector que proporcione la demostración en el ejercicio 10 de la página 177.

Teorema 6.4

La media y la varianza de la distribución de Weibull son

$$\mu = \alpha^{-1/\beta} \Gamma\left(1 + \frac{1}{\beta}\right), \quad \sigma^2 = \alpha^{-2/\beta} \left\{ \Gamma\left(1 + \frac{2}{\beta}\right) - \left[\Gamma\left(1 + \frac{1}{\beta}\right) \right]^2 \right\}.$$

Como la distribución gamma y la exponencial, la distribución de Weibull también se aplica a problemas de confiabilidad y de prueba de vida como los de **tiempo de falla o duración de la vida** de un componente, medido de algún tiempo específico hasta que falla. Representemos este tiempo de falla mediante la variable aleatoria continua T , con función de densidad de probabilidad $f(t)$, donde $f(t)$ es la distribución de Weibull.

Para aplicar la distribución de Weibull a la teoría de confiabilidad, definamos primero la confiabilidad de un componente o producto como la *probabilidad de que funcione apropiadamente por al menos un tiempo específico bajo condiciones experimentales específicas*. Por tanto, si $R(t)$ se define como la confiabilidad del componente dado en el tiempo t , podemos escribir

$$R(t) = P(T > t) = \int_t^{\infty} f(t) dt = 1 - F(t),$$

donde $F(t)$ es la distribución acumulada de T . La probabilidad condicional de que un componente caiga en el intervalo de $T = t$ a $T = t + \Delta t$, dado que sobrevive al tiempo t , es

$$\frac{F(t + \Delta t) - F(t)}{R(t)}.$$

Al dividir esta proporción entre Δt y tomar el límite cuando $\Delta t \rightarrow 0$, obtenemos la **tasa de falla**, denotada por $Z(t)$. De aquí

$$Z(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} \frac{1}{R(t)} = \frac{F'(t)}{R(t)} = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)},$$

que expresa la tasa de falla en términos de la distribución del tiempo de falla.

Del hecho de que $R(t) = 1 - F(t)$ y entonces $R'(t) = -F'(t)$, podemos escribir la ecuación diferencial

$$Z(t) = \frac{-R'(t)}{R(t)} = \frac{-d[\ln R(t)]}{dt}$$

y entonces resolver,

$$\ln R(t) = - \int Z(t) dt + \ln c \quad \text{o} \quad R(t) = ce^{-\int Z(t) dt}.$$

donde c satisface la suposición inicial de que $R(0) = 1$ o $F(0) = 1 - R(0) = 0$. Veamos así que un conocimiento de la función de densidad $f(t)$ o la tasa de falla $Z(t)$ determina unívocamente a la otra.

Ejemplo 6.21 Muestre que la función de tasa de falla está dada por

$$Z(t) = \alpha\beta t^{\beta-1}, \quad t > 0,$$

si y sólo si la distribución del tiempo para la falla es la distribución de Weibull

$$f(t) = \alpha\beta t^{\beta-1}e^{-at^\beta}, \quad t > 0.$$

SOLUCIÓN

Suponga que $Z(t) = \alpha\beta t^{\beta-1}$, $t > 0$. Entonces podemos escribir

$$f(t) = Z(t)R(t),$$

donde

$$R(t) = ce^{-\int Z(t)dt} = ce^{-\int \alpha\beta t^{\beta-1}dt} = ce^{-at^\beta}.$$

De la condición de que $R(0) = 1$, encontramos que $c = 1$. De aquí

$$R(t) = e^{-at^\beta}$$

y

$$f(t) = \alpha\beta t^{\beta-1}e^{-at^\beta}, \quad t > 0.$$

Ahora, si suponemos que

$$f(t) = \alpha\beta t^{\beta-1}e^{-at^\beta}, \quad t > 0,$$

entonces $Z(t)$ está determinada al escribir

$$Z(t) = \frac{f(t)}{R(t)},$$

donde

$$R(t) = 1 - F(t) = 1 - \int_0^t \alpha\beta x^{\beta-1}e^{-ax^\beta} dx = 1 + \int_0^t de^{-ax^\beta} = e^{-at^\beta}.$$

Entonces

$$Z(t) = \frac{\alpha\beta t^{\beta-1}e^{-at^\beta}}{e^{-at^\beta}} = \alpha\beta t^{\beta-1}, \quad t > 0.$$

En el ejemplo 6.21 se ve que la tasa de falla disminuye con el tiempo si $\beta < 1$, aumenta con el tiempo si $\beta > 1$, y es constante si $\beta = 1.0$. Es fácil ver, entonces, que la forma más simple, a saber, la distribución exponencial, se aplica para una tasa de falla constante (vía el proceso de Poisson). ■

Ejercicios

1. Si una variable aleatoria X tiene la distribución gamma con $\alpha = 2$ y $\beta = 1$, encuentre $P(1.8 < X < 2.4)$.

2. En cierta ciudad, el consumo diario de agua (en millones de litros) sigue aproximadamente una distribución gamma con $\alpha = 2$ y $\beta = 3$. Si la capacidad diaria de dicha ciudad es nueve millones de litros de agua, ¿cuál es la probabilidad de que en cualquier día dado el suministro de agua sea inadecuado?

3. Utilice la función gamma con $y = \sqrt{2x}$ para mostrar que $\Gamma(1/2) = \sqrt{\pi}$.

4. Suponga que el tiempo, en horas, que toma reparar una bomba de calor es una variable aleatoria X que tiene una distribución gamma con parámetros $\alpha = 2$ y $\beta = 1/2$. ¿Cuál es la probabilidad de que la siguiente llamada de servicio requiera

- (a) a lo más una hora para reparar la bomba de calor?
 (b) al menos dos horas para reparar la bomba de calor?

5. (a) Encuentre la media y la varianza del consumo diario de agua del ejercicio 2.

(b) De acuerdo con el teorema de Chebyshev, hay una probabilidad de al menos $3/4$ de que el consumo de agua en cualquier día dado caiga dentro de qué intervalo?

6. En cierta ciudad, el consumo diario de energía eléctrica, en millones de kilowatts hora, es una variable aleatoria X que tiene una distribución gamma con media $\mu = 6$ y varianza $\sigma^2 = 12$.

- (a) Encuentre los valores de α y β .
 (b) Encuentre la probabilidad de que en cualquier día dado el consumo de energía diario exceda 12 millones de kilowatts hora.

7. La longitud de tiempo para que un individuo sea atendido en una cafetería es una variable aleatoria que tiene una distribución exponencial con una media de cuatro minutos. ¿Cuál es la probabilidad de que una persona sea atendida en menos de tres minutos en al menos cuatro de los siguientes seis días?

8. La vida en años de cierto interruptor eléctrico tiene una distribución exponencial con una vida promedio de $\beta = 2$. de estos interruptores se instalan en diferentes sistemas, ¿cuál es la probabilidad de que a lo más 30 fallen durante el primer año?

9. Suponga que la vida de servicio, en años, de la batería de un aparato para sordos es una variable

aleatoria que tiene una distribución de Weibull con $\alpha = 1/2$ y $\beta = 2$.

- (a) ¿Cuánto tiempo se puede esperar que dure tal batería?
 (b) ¿Cuál es la probabilidad de que tal batería esté en operación después de dos años?

10. Derive la media y varianza de la distribución de Weibull.

11. Las vidas de ciertos sellos de automóvil tienen la distribución de Weibull con tasa de falla $Z(t) = 1/\sqrt{t}$. Encuentre la probabilidad de que tal sello aún esté en uso después de cuatro años.

12. La variable aleatoria continua X tiene la distribución beta con parámetros α y β si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

donde $\alpha > 0$ y $\beta > 0$. Si la proporción de una marca de televisores que requiere servicio durante el primer año de operación es una variable aleatoria que tiene una distribución beta con $\alpha = 3$ y $\beta = 2$, ¿cuál es la probabilidad de que al menos 80% de los nuevos modelos de esta marca que se vendieron este año requerirán servicio durante su primer año de operación?

13. En una actividad de investigación biomédica se determinó que el tiempo de sobrevivencia en semanas de un animal cuando se le somete a cierta exposición de radiación gamma tiene una distribución gamma con $\alpha = 5$ y $\beta = 10$.

- (a) ¿Cuál es el tiempo medio de sobrevivencia de un animal seleccionado al azar del tipo que se utilizó en el experimento?
 (b) ¿Cuál es la desviación estándar del tiempo de sobrevivencia?
 (c) ¿Cuál es la probabilidad de que un animal sobreviva más de 30 semanas?

14. Se sabe que la duración de cierto tipo de transistor sigue una distribución gamma con media de 10 semanas y desviación estándar de $\sqrt{50}$ semanas.

- (a) ¿Cuál es la probabilidad de que el transistor dure a lo más 50 semanas?
 (b) ¿Cuál es la probabilidad de que el transistor no sobreviva las primeras 10 semanas?

15. El tiempo de respuesta de una computadora es una aplicación importante de las distribuciones gamma y exponencial. Suponga que un estudio de cierto sistema de computadoras revela que el tiempo de respuesta en segundos tiene una distribución exponencial con una media de tres segundos.

(a) ¿Cuál es la probabilidad de que el tiempo de respuesta exceda cinco segundos?

(b) ¿Cuál es la probabilidad de que el tiempo de respuesta exceda 10 segundos?

16. Los datos de porcentaje a menudo siguen una distribución logarítmica normal. Se estudia el uso promedio de potencia (dB por hora) para una compañía particular y se sabe que tiene una distribución logarítmica normal con parámetros $\mu = 4$ y $\sigma = 2$. ¿Cuál es la probabilidad de que la compañía utilice más de 270 dB durante cualquier hora particular?

17. Para el ejercicio 16 ¿cuál es el uso de potencia media (dB promedio por hora)? ¿Cuál es la variancia?

18. El número de automóviles que llegan a cierta intersección por minuto tiene una distribución de Poisson con una media de 10. El interés se centra alrededor del tiempo que transcurre antes de que 15 automóviles aparezcan en la intersección.

(a) ¿Cuál es la probabilidad de que más de 15 automóviles aparezcan en la intersección durante cualquier minuto dado?

(b) ¿Cuál es la probabilidad de que se requieran más de dos minutos antes de que lleguen 15 autos?

19. Considere la información del ejercicio 18.

(a) ¿Cuál es la probabilidad de que transcurra más de un minuto entre llegadas?

(b) ¿Cuál es el número medio de minutos que transcurren entre llegadas?

Ejercicios de repaso

1. De acuerdo con un estudio publicado por un grupo de sociólogos de la Universidad de Massachusetts, aproximadamente 49% de los consumidores de Valium en el estado de Massachusetts son empleados de oficina. ¿Cuál es la probabilidad de que entre 482 y 510, incluso, de los siguientes 100 consumidores de Valium seleccionados al azar de este Estado sean empleados de oficina?

2. La distribución exponencial se aplica con frecuencia a los tiempos de espera entre éxitos en un proceso de Poisson. Si el número de llamadas que se reciben por hora en un servicio de contestación telefónica es una variable aleatoria de Poisson con parámetro $\lambda = 6$, sabemos que el tiempo, en horas, entre llamadas sucesivas tiene una distribución exponencial con parámetro $\beta = 1/6$. ¿Cuál es la probabilidad de esperar más de 15 minutos entre cualesquiera dos llamadas sucesivas?

3. Cuando α es un entero positivo n , la distribución gamma también se conoce como **distribución de Erlang**. Al hacer $\alpha = n$ en la distribución gamma de la página 168, la distribución de Erlang es

$$f(x) = \begin{cases} \frac{x^{n-1} e^{-x/\beta}}{\beta^n (n-1)!}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Se puede mostrar que si los tiempos entre eventos sucesivos son independientes, y cada uno tiene una distribución exponencial con parámetro b , entonces el tiempo total de espera X transcurrido hasta que ocurren n eventos tiene la distribución de Erlang. Con referencia al ejercicio 2, ¿cuál es la probabilidad de que las siguientes tres llamadas se reciban dentro de los siguientes 30 minutos?

4. Un fabricante de cierto tipo de máquina grande desea comprar remaches de uno de dos fabricantes. Es importante que la resistencia de ruptura de cada remache exceda 10,000 psi. Dos fabricantes (A y B) ofrecen este tipo de remache. Ambos tienen remaches cuya resistencia a la ruptura está distribuida de forma normal. Las resistencias de ruptura medias para los fabricantes A y B son 14,000 psi y 13,000 psi, respectivamente. Las desviaciones estándar son 2000 psi y 1000 psi, respectivamente. ¿Cuál fabricante producirá, en promedio, el menor número de remaches defectuosos?

5. De acuerdo con un censo reciente, casi 65% de todos los hogares en Estados Unidos se componen de una o dos personas. Suponga que este porcentaje aún es válido hoy en día, ¿cuál es la probabilidad de que entre 490 y 515 incluso de los siguientes 1000 hogares seleccionados al azar en Estados Unidos consistan en una o dos personas?

6. La vida de cierto tipo de dispositivo tiene una tasa de falla anunciada de 0.01 por hora. La tasa de falla es constante y se aplica la distribución exponencial.
- ¿Cuál es el tiempo medio de falla?
 - ¿Cuál es la probabilidad de que pasen 200 horas antes de que se observe una falla?
7. En una planta de proceso químico es importante que el rendimiento de cierto tipo de producto en lote permanezca por arriba de 80%. Si permanece por debajo de 80% por un tiempo prolongado, la compañía pierde dinero. Los lotes producidos ocasionalmente con defectos son de poco interés. Pero si varios lotes por día resultan defectuosos, la planta se detiene y se hacen ajustes. Se sabe que el rendimiento se distribuye normalmente con desviación estándar de 4%.
- ¿Cuál es la probabilidad de una "falsa alarma" (rendimiento por debajo del 80%) cuando el rendimiento medio es de 85%?
 - ¿Cuál es la probabilidad de que un lote producido tenga un rendimiento que exceda 80% cuando de hecho el rendimiento medio es de 79%?
8. Considere una tasa de falla de un componente eléctrico de una vez cada cinco horas. Es importante considerar el tiempo que transcurre para que fallen dos componentes.
- Suponga que se aplica la distribución gamma, ¿cuál es el tiempo medio que transcurre para la falla de dos componentes?
 - ¿Cuál es la probabilidad de que transcurran 12 horas antes de que fallen dos componentes?
9. Se establece que la elongación de una barra de acero bajo una carga particular se distribuye normalmente con una media de 0.05 pulgadas y $\sigma = 0.01$ pulgadas. Encuentre la probabilidad de que la elongación esté
- por arriba de 0.1 pulgadas;
 - por abajo de 0.04 pulgadas;
 - entre 0.025 y 0.065 pulgadas.
10. Se sabe que un satélite controlado tiene un error (distancia del objetivo) que se distribuye normalmente con media cero y desviación estándar de cuatro pies. El fabricante del satélite define un "éxito" como un disparo en el que el satélite llega a 10 pies del objetivo. Calcule la probabilidad de que falle el satélite.
11. Un técnico planea probar cierto tipo de resina desarrollada en el laboratorio para determinar la naturaleza del tiempo que transcurre antes de que la unión se lleve a cabo. Se sabe que el tiempo medio para la unión es tres horas y la desviación estándar es 0.5 horas. Se considerará un producto indeseable si el tiempo de unión es menor que una hora o más de cuatro horas. Comente sobre la utilidad de la resina. ¿Con qué frecuencia su desempeño se considera indeseable? Suponga que el tiempo para la unión se distribuye normalmente.
12. Considere la información del ejercicio de repaso 6. ¿Cuál es la probabilidad de que transcurran menos de 200 horas antes de que ocurran dos fallas?
13. Para el ejercicio 12, ¿cuál es la media y la varianza del tiempo que transcurre antes de que ocurran dos fallas?
14. Se sabe que la tasa promedio de uso de agua (miles de galones por hora) en cierta comunidad implica la distribución logarítmica normal con parámetros $\mu = 5$ y $\sigma = 2$. Es importante para propósitos de planeación obtener una apreciación de los periodos de alta utilización. ¿Cuál es la probabilidad de que, para cualquier hora dada, se usen 50,000 galones de agua?
15. Para el ejercicio 14, ¿cuál es la probabilidad media del uso de agua por hora promedio en miles de galones?
16. En el ejercicio 14, sección 6.10, se supone que la duración de un transistor tiene una distribución gamma con media de 10 semanas y desviación estándar de $\sqrt{50}$ semanas. Suppose that the gamma distribution assumption is incorrect. Suponga que la suposición de distribución gamma es incorrecta. Suponga que la distribución es normal.
- ¿Cuál es la probabilidad de que el transistor dure a lo más 50 semanas?
 - ¿Cuál es la probabilidad de que el transistor no sobreviva las primeras 10 semanas?
 - Comente la diferencia entre sus resultados aquí y los que se encuentran en el ejercicio 14, sección 6.10.

Funciones de variables aleatorias

7.1 Introducción

Este capítulo contiene un amplio espectro de material. Los capítulos 5 y 6 tratan con tipos específicos de distribuciones, tanto discretas como continuas. Estas son distribuciones que encuentran uso en muchas aplicaciones objeto de la materia, que incluyen confiabilidad, control de calidad y muestreo de aceptación. En este capítulo comenzamos con un tema más general, el de las distribuciones de funciones de variables aleatorias. Se introducen las técnicas generales y se ilustran con ejemplos. Estas son seguidas por un concepto relacionado, *funciones generadoras de momentos*, que puede ser útil en el estudio de distribuciones de funciones lineales de variables aleatorias.

En los métodos estadísticos estándar, el resultado de la prueba de hipótesis estadísticas, estimación o incluso las gráficas estadísticas no involucra a una sola variable aleatoria sino, más bien, a *funciones de una o más variables aleatorias*. Como resultado, la inferencia estadística requiere la distribución de estas funciones. Por ejemplo, es común el uso de **promedios de variables aleatorias**. Además, las sumas y combinaciones lineales más generales son importantes. A menudo nos interesa la distribución de sumas de los cuadrados de variables aleatorias, particularmente en el uso de las técnicas del análisis de varianza que se discuten en los capítulos 12-14.

7.2 Transformaciones de variables

Con frecuencia, en estadística, se encuentra la necesidad de derivar la distribución de probabilidad de una función de una o más variables aleatorias. Por ejemplo, su-

ponga que X es una variable aleatoria discreta con distribución de probabilidad $f(x)$ y suponga además que $Y = u(X)$ define una transformación uno a uno entre los valores de X y Y . Deseamos encontrar la distribución de probabilidad de Y . Es importante notar que la transformación uno a uno implica que cada valor x está relacionado con un, y sólo un, valor $y = u(x)$ que cada valor y está relacionado con un, y sólo un, valor $x = w(y)$, donde $w(y)$ se obtiene al resolver $y = u(x)$ para x en términos de y .

De nuestra discusión de las distribuciones de probabilidad discreta en el capítulo 3 es claro que la variable aleatoria Y toma el valor y cuando X toma el valor $w(y)$. En consecuencia, la distribución de probabilidad de Y está dada por

$$g(y) = P(Y = y) = P[X = w(y)] = f[w(y)].$$

Teorema 7.1

Suponga que X es una variable aleatoria **discreta** con distribución de probabilidad $f(x)$. Definamos con $Y = u(X)$ una transformación uno a uno entre los valores de X y Y de modo que la ecuación $y = u(x)$ se pueda resolver unívocamente para x en términos de y , digamos $x = w(y)$. Entonces la distribución de probabilidad de Y es

$$g(y) = f[w(y)].$$

Ejemplo 7.1 Sea X una variable aleatoria geométrica con distribución de probabilidad $f(x) = \frac{1}{4}(\frac{1}{4})^{x-1}$, $x = 1, 2, 3, \dots$. Encuentre la distribución de probabilidad de la variable aleatoria $Y = X^2$.

SOLUCIÓN

Como los valores de X son todos positivos, la transformación define una correspondencia uno a uno entre los valores x y y , $y = x^2$ y $x = \sqrt{y}$. De aquí

$$g(y) = \begin{cases} f(\sqrt{y}) = \frac{1}{4}(\frac{1}{4})^{y-1}, & y = 1, 4, 9, \dots \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Considere un problema donde X_1 y X_2 son dos variables aleatorias discretas con distribución de probabilidad conjunta $f(x_1, x_2)$ y deseamos encontrar la distribución de probabilidad conjunta $g(y_1, y_2)$ de las dos variables aleatorias nuevas

$$Y_1 = u_1(X_1, X_2) \quad \text{y} \quad Y_2 = u_2(X_1, X_2),$$

que definen una transformación uno a uno entre el conjunto de puntos (x_1, x_2) y (y_1, y_2) . Al resolver las ecuaciones $y_1 = u_1(x_1, x_2)$ y $y_2 = u_2(x_1, x_2)$ de forma simultánea, obtenemos la solución inversa única

$$x_1 = w_1(y_1, y_2) \quad \text{y} \quad x_2 = w_2(y_1, y_2).$$

De aquí las variables aleatorias Y_1 y Y_2 toman los valores y_1 y y_2 , respectivamente, cuando X_1 toma el valor $w_1(y_1, y_2)$ y X_2 toma el valor $w_2(y_1, y_2)$. La distribución de probabilidad conjunta de Y_1 y Y_2 es entonces

$$\begin{aligned} g(y_1, y_2) &= P(Y_1 = y_1, Y_2 = y_2) \\ &= P[(X_1 = w_1(y_1, y_2), X_2 = w_2(y_1, y_2))] \\ &= f[w_1(y_1, y_2), w_2(y_1, y_2)]. \quad \blacksquare \end{aligned}$$

Teorema 7.2

Suponga que X_1 y X_2 son variables aleatorias **discretas** con distribución de probabilidad conjunta $f(x_1, x_2)$. Definamos con $Y_1 = u_1(X_1, X_2)$ y $Y_2 = u_2(X_1, X_2)$ una transformación uno a uno entre los puntos (x_1, x_2) y (y_1, y_2) de modo que las ecuaciones

$$y_1 = u_1(x_1, x_2) \quad \text{y} \quad y_2 = u_2(x_1, x_2)$$

se pueden resolver unívocamente para x_1 y x_2 en términos de y_1 y y_2 , digamos $x_1 = w_1(y_1, y_2)$ y $x_2 = w_2(y_1, y_2)$. Entonces la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2(y_1, y_2)].$$

El teorema 7.2 es en extremo útil para encontrar la distribución de alguna variable aleatoria $Y_1 = u_1(X_1, X_2)$, donde X_1 y X_2 son variables aleatorias discretas con distribución de probabilidad conjunta $f(x_1, x_2)$. Definimos simplemente una segunda función, digamos $Y_2 = u_2(X_1, X_2)$, y mantenemos una correspondencia uno a uno entre los puntos (x_1, x_2) y (y_1, y_2) , obtenemos la distribución de probabilidad conjunta $g(y_1, y_2)$. La distribución de Y_1 es precisamente la distribución marginal de $g(y_1, y_2)$, que se encuentra al sumar los valores y_2 . Al denotar la distribución de Y_1 por $h(y_1)$, podemos escribir

$$h(y_1) = \sum_{y_2} g(y_1, y_2)$$

Ejemplo 7.2 Sean X_1 y X_2 dos variables aleatorias independientes que tienen distribuciones de Poisson con parámetros μ_1 y μ_2 , respectivamente. Encuentre la distribución de la variable aleatoria $Y_1 = X_1 + X_2$.

SOLUCIÓN

Como X_1 y X_2 son independientes, podemos escribir

$$f(x_1, x_2) = f(x_1)f(x_2) = \frac{e^{-\mu_1}\mu_1^{x_1}}{x_1!} \frac{e^{-\mu_2}\mu_2^{x_2}}{x_2!} = \frac{e^{-(\mu_1+\mu_2)}\mu_1^{x_1}\mu_2^{x_2}}{x_1!x_2!},$$

donde $x_1 = 0, 1, 2, \dots$ y $x_2 = 0, 1, 2, \dots$. Definamos ahora una segunda variable aleatoria, digamos $Y_2 = X_2$. Las funciones inversas están dadas por $x_1 = y_1 - y_2$ y $x_2 = y_2$. Con el uso del teorema 7.2, encontramos que la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$g(y_1, y_2) = \frac{e^{-(\mu_1+\mu_2)}\mu_1^{y_1-y_2}\mu_2^{y_2}}{(y_1 - y_2)!y_2!},$$

donde $y_1 = 0, 1, 2, \dots$ y $y_2 = 0, 1, 2, \dots, y_1$. Nótese que como $x_1 \geq 0$, la transformación $x_1 = y_1 - y_2$ implica que x_2 y por tanto y_2 siempre deben ser menores que o iguales a y_1 . En consecuencia, la distribución marginal de probabilidad de Y_1 es

$$\begin{aligned} h(y_1) &= \sum_{y_2=0}^{y_1} g(y_1, y_2) = e^{-(\mu_1+\mu_2)} \sum_{y_2=0}^{y_1} \frac{\mu_1^{y_1-y_2}\mu_2^{y_2}}{(y_1 - y_2)!y_2!} \\ &= \frac{e^{-(\mu_1+\mu_2)}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!}{y_2!(y_1 - y_2)!} \mu_1^{y_1-y_2}\mu_2^{y_2} \end{aligned}$$

$$= \frac{e^{-(\mu_1 + \mu_2)}}{y_1!} \sum_{r_2=0}^{y_1} \binom{y_1}{r_2} \mu_1^{r_1 - r_2} \mu_2^{r_2}.$$

Al reconocer esta suma como la expansión binomial de $(\mu_1 + \mu_2)^{y_1}$, obtenemos

$$h(y_1) = \frac{e^{-(\mu_1 + \mu_2)} (\mu_1 + \mu_2)^{y_1}}{y_1!}, \quad y_1 = 0, 1, 2, \dots,$$

de lo cual concluimos que la suma de las dos variables aleatorias independientes que tienen distribuciones de Poisson, con parámetros μ_1 y μ_2 , tiene una distribución de Poisson con parámetro $\mu_1 + \mu_2$. ■

Para encontrar la distribución de probabilidad de la variable aleatoria $Y = u(X)$ cuando X es una variable aleatoria continua y la transformación es uno a uno, necesitaremos el teorema 7.3.

Teorema 7.3

Suponga que X es una variable aleatoria **continua** con distribución de probabilidad $f(x)$. Definamos con $Y = u(X)$ una correspondencia uno a uno entre los valores de X y Y de modo que la ecuación $y = u(x)$ se pueda resolver unívocamente para x en términos de y , digamos $x = w(y)$. Entonces la distribución de probabilidad de Y es

$$g(y) = f[w(y)]|J|,$$

donde $J = w'(y)$ y se llama **jacobiano** de la transformación.

PRUEBA Suponga que $y = u(x)$ es una función creciente como la de la figura 7.1. En-

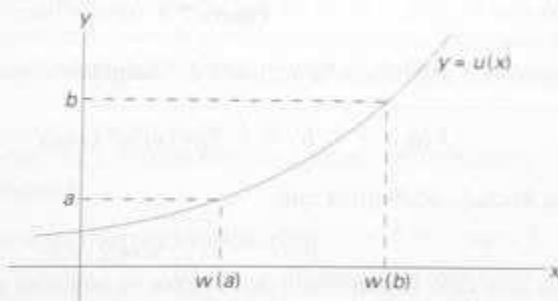


Figura 7.1 Función creciente.

tonces, vemos que siempre que Y cae entre a y b , la variable aleatoria X debe caer entre $w(a)$ y $w(b)$. De aquí

$$P(a < Y < b) = P[w(a) < X < w(b)] = \int_{w(a)}^{w(b)} f(x) dx.$$

Al cambiar la variable de integración de x a y mediante la relación $x = w(y)$, obtenemos $dx = w'(y) dy$, y de aquí

$$P(a < Y < b) = \int_a^b f[w(y)]w'(y) dy.$$

Como la integral da la probabilidad que se desea para toda $a < b$ dentro del conjunto permisible de valores y , entonces la distribución de probabilidad de Y es

$$g(y) = f[w(y)]w'(y) = f[w(y)]J.$$

Si reconocemos $J = w'(y)$ como el recíproco de la pendiente de la línea tangente a la curva de la función creciente $y = u(x)$, entonces es obvio que $J = |J|$. De aquí

$$g(y) = f[w(y)]|J|.$$

Suponga que $y = u(x)$ es una función decreciente como la de la figura 7.2. Entonces escribimos

$$P(a < Y < b) = P[w(b) < X < w(a)] = \int_{w(b)}^{w(a)} f(x) dx.$$

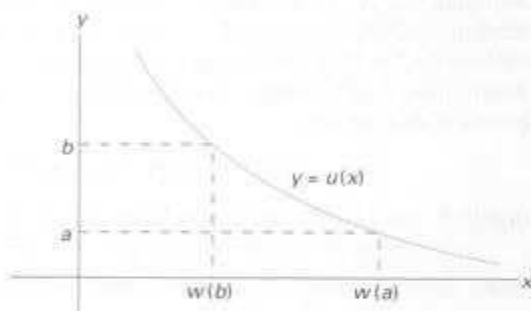


Figura 7.2 Función decreciente.

De nuevo al cambiar la variable de integración por y , obtenemos

$$P(a < Y < b) = \int_b^a f[w(y)]w'(y) dy = - \int_a^b f[w(y)]w'(y) dy,$$

de lo cual concluimos que

$$g(y) = -f[w(y)]w'(y) = -f[w(y)]J.$$

En este caso la pendiente de la curva es negativa y $J = -|J|$. Entonces

$$g(y) = f[w(y)]|J|,$$

como antes.

Ejemplo 7.3 Sea X una variable aleatoria continua con distribución de probabilidad

$$f(x) = \begin{cases} \frac{x}{12}, & 1 < x < 5 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = 2X - 3$.

SOLUCIÓN

La solución inversa de $y = 2x - 3$ da $x = (y + 3)/2$, de la que obtenemos $J = w'(y) = dx/dy = 1/2$. Por tanto, con el uso del teorema 7.3, encontramos que la función de densidad de Y es

$$g(y) = \begin{cases} \frac{(y+3)/2}{12} \left(\frac{1}{2}\right) = \frac{y+3}{48}, & -1 < y < 7 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Para encontrar la distribución de probabilidad conjunta de las variables aleatorias $Y_1 = u_1(X_1, X_2)$ y $Y_2 = u_2(X_1, X_2)$ cuando X_1 y X_2 son continuas y la transformación es uno a uno, necesitamos un teorema adicional, análogo al teorema 7.2, que establecemos sin demostración. ■

Teorema 7.4

Suponga que X_1 y X_2 son variables aleatorias **continuas** con distribución de probabilidad conjunta $f(x_1, x_2)$. Definamos con $Y_1 = u_1(X_1, X_2)$ y $Y_2 = u_2(X_1, X_2)$ una transformación uno a uno entre los puntos (x_1, x_2) y (y_1, y_2) de modo que las ecuaciones $y_1 = u_1(x_1, x_2)$ y $y_2 = u_2(x_1, x_2)$ se pueden resolver unívocamente para x_1 y x_2 en términos de y_1 y y_2 , digamos $x_1 = w_1(y_1, y_2)$ y $x_2 = w_2(y_1, y_2)$. Entonces la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2(y_1, y_2)] |J|,$$

donde el jacobiano es el determinante 2×2

$$J = \begin{vmatrix} \partial x_1 / \partial y_1 & \partial x_1 / \partial y_2 \\ \partial x_2 / \partial y_1 & \partial x_2 / \partial y_2 \end{vmatrix}$$

y $\partial x_1 / \partial y_1$ es simplemente la derivada de $x_1 = w_1(y_1, y_2)$ con respecto a y_1 , y y_2 permanece constante, que en cálculo se denomina derivada parcial de x_1 con respecto a y_1 . Las otras derivadas parciales se definen de manera similar.

Ejemplo 7.4 Sean X_1 y X_2 dos variables aleatorias continuas con distribución de probabilidad conjunta

$$f(x_1, x_2) = \begin{cases} 4x_1x_2, & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad conjunta de $Y_1 = X_1^2$ y $Y_2 = X_1X_2$.

SOLUCIÓN

Las soluciones inversas de $y_1 = x_1^2$ y $y_2 = x_1x_2$ son $x_1 = \sqrt{y_1}$ y $x_2 = y_2/\sqrt{y_1}$, de las que obtenemos

$$J = \begin{vmatrix} 1/(2\sqrt{y_1}) & 0 \\ -y_2/2y_1^{3/2} & 1/\sqrt{y_1} \end{vmatrix} = \frac{1}{2y_1}.$$

Para determinar el conjunto B de puntos en el plano y_1, y_2 en el que se traza el conjunto A de puntos en el plano x_1, x_2 , escribimos

$$x_1 = \sqrt{y_1} \quad y \quad x_2 = y_2/\sqrt{y_1}$$

y después al hacer $x_1 = 0$, $x_2 = 0$, $x_1 = 1$, y $x_2 = 1$, las fronteras del conjunto A se transforman a $y_1 = 0$, $y_2 = 0$, $y_1 = 1$, y $y_2 = \sqrt{y_1}$ o $y_2^2 = y_1$. Las dos regiones se ilustran en la figura 7.3. Claramente, la transformación es uno a uno, al trazar el con-

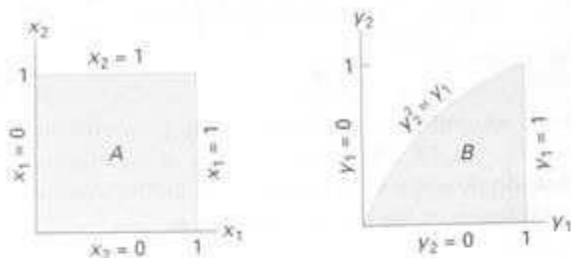


Figura 7.3 Trazo del conjunto A en el conjunto B.

junto $A = \{(x_1, x_2) | 0 < x_1 < 1, 0 < x_2 < 1\}$ en el conjunto $B = \{(y_1, y_2) | y_2^2 < y_1 < 1, 0 < y_2 < 1\}$. Del teorema 7.4, la distribución de probabilidad conjunta de Y_1 y Y_2 es

$$g(y_1, y_2) = 4(\sqrt{y_1}) \frac{y_2}{\sqrt{y_1}} \frac{1}{2y_1} = \begin{cases} \frac{2y_2}{y_1}, & y_2^2 < y_1 < 1, 0 < y_2 < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

A menudo surgen problemas cuando deseamos encontrar la distribución de probabilidad de la variable aleatoria $Y = u(X)$ cuando X es una variable aleatoria continua y la transformación no es uno a uno. Es decir, para cada valor x corresponde exactamente un valor y , pero a cada valor y corresponde más de un valor x . Por ejemplo, suponga que $f(x)$ es positiva en el intervalo $-1 < x < 2$ y cero en cualquier otro caso.

Considere la transformación $y = x^2$. En este caso $x = \pm\sqrt{y}$ para $0 < y < 1$ y $x = \sqrt{y}$ para $1 < y < 4$. Para el intervalo $1 < y < 4$, la distribución de probabilidad de Y se encuentra como antes, con el uso del teorema 7.3. Es decir,

$$g(y) = f[w(y)]|J| = \frac{f(\sqrt{y})}{2\sqrt{y}}, \quad 1 < y < 4.$$

Sin embargo, cuando $0 < y < 1$, podemos dividir el intervalo $-1 < x < 1$ para obtener las dos funciones inversas

$$x = -\sqrt{y} \quad -1 < x < 0 \quad y \quad x = \sqrt{y}, \quad 0 < x < 1.$$

Entonces para todo valor y corresponde un solo valor x para cada partición. De la figura 7.4 vemos que

$$\begin{aligned} P(a < Y < b) &= P(-\sqrt{b} < X < -\sqrt{a}) + P(\sqrt{a} < X < \sqrt{b}) \\ &= \int_{-\sqrt{b}}^{-\sqrt{a}} f(x) dx + \int_{\sqrt{a}}^{\sqrt{b}} f(x) dx. \end{aligned}$$

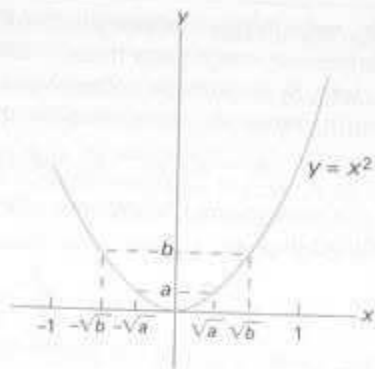


Figura 7.4 Función decreciente y creciente.

Al cambiar la variable de integración de x a y , obtenemos

$$\begin{aligned} P(a < Y < b) &= \int_0^a f(-\sqrt{y})J_1 dy + \int_a^b f(\sqrt{y})J_2 dy \\ &= -\int_a^b f(-\sqrt{y})J_1 dy + \int_a^b f(\sqrt{y})J_2 dy, \end{aligned}$$

donde

$$J_1 = \frac{d(-\sqrt{y})}{dy} = \frac{-1}{2\sqrt{y}} = -|J_1| \quad \text{y} \quad J_2 = \frac{d(\sqrt{y})}{dy} = \frac{1}{2\sqrt{y}} = |J_2|.$$

De aquí, podemos escribir

$$P(a < Y < b) = \int_a^b [f(-\sqrt{y})|J_1| + f(\sqrt{y})|J_2|] dy,$$

y entonces

$$g(y) = f(-\sqrt{y})|J_1| + f(\sqrt{y})|J_2| = \frac{[f(-\sqrt{y}) + f(\sqrt{y})]}{2\sqrt{y}}, \quad 0 < y < 1.$$

La distribución de probabilidad de Y para $0 < y < 4$ se puede escribir ahora como

$$g(y) = \begin{cases} \frac{[f(-\sqrt{y}) + f(\sqrt{y})]}{2\sqrt{y}}, & 0 < y < 1 \\ \frac{f(\sqrt{y})}{2\sqrt{y}}, & 1 < y < 4 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Este procedimiento para encontrar $g(y)$ cuando $0 < y < 1$ se generaliza en el teorema 7.5 para k funciones inversas. Para transformaciones que no son uno a uno de funciones de diversas variables, se remite al lector a *Introduction to Mathematical Statistics* de Hogg y Craig (véase la bibliografía). ■

Teorema 7.5

Suponga que X es una variable aleatoria **continua** con distribución de probabilidad $f(x)$. Sea $Y = u(X)$ una transformación entre los valores de X y Y que no es uno a uno. Si el intervalo sobre el que se define X se puede dividir en k conjuntos mutuamente disjuntos de modo que cada uno de las funciones inversas

$$x_1 = w_1(y), \quad x_2 = w_2(y), \quad \dots, \quad x_k = w_k(y)$$

de $y = u(x)$ defina una correspondencia uno a uno, entonces la distribución de probabilidad de Y es

$$g(y) = \sum_{i=1}^k f[w_i(y)] |J_i|,$$

donde $J_i = w_i'(y)$, $i = 1, 2, \dots, k$.

Ejemplo 7.5 Muestre que $Y = (X - \mu)^2 / \sigma^2$ tiene una distribución ji cuadrada con 1 grado de libertad cuando X tiene una distribución normal con media μ y varianza σ^2 .

SOLUCIÓN

Sea $Z = (X - \mu) / \sigma$, donde la variable aleatoria Z tiene la distribución normal estándar

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Encontraremos ahora la distribución de la variable aleatoria $Y = Z^2$. Las soluciones inversas de $y = z^2$ son $z = \pm\sqrt{y}$. Si designamos $z_1 = -\sqrt{y}$ y $z_2 = \sqrt{y}$, entonces $J_1 = -1/2\sqrt{y}$ y $J_2 = 1/2\sqrt{y}$. De aquí, por el teorema 7.5, tenemos

$$\begin{aligned} g(y) &= \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| \frac{-1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-y/2} \left| \frac{1}{2\sqrt{y}} \right| \\ &= \frac{1}{2^{1/2}\sqrt{\pi}} y^{1/2-1} e^{-y/2}, \quad y > 0, \end{aligned}$$

Como $g(y)$ es una función de densidad, se sigue que

$$\begin{aligned} 1 &= \frac{1}{2^{1/2}\sqrt{\pi}} \int_0^{\infty} y^{1/2-1} e^{-y/2} dy = \frac{\Gamma(1/2)}{\sqrt{\pi}} \int_0^{\infty} \frac{1}{2^{1/2}\Gamma(1/2)} y^{1/2-1} e^{-y/2} dy \\ &= \frac{\Gamma(1/2)}{\sqrt{\pi}}, \end{aligned}$$

la integral es el área bajo una curva de probabilidad gamma con parámetros $\alpha = 1/2$ y $\beta = 2$. Por tanto, $\sqrt{\pi} = \Gamma(1/2)$ y la distribución de probabilidad de Y está dada por

$$g(y) = \begin{cases} \frac{1}{2^{1/2}\Gamma(1/2)} y^{1/2-1} e^{-y/2}, & y > 0, \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

que es una distribución ji cuadrada con 1 grado de libertad. ■

7.3 Momentos y funciones generadoras de momentos

En esta sección nos concentramos en aplicaciones de las funciones generadoras de momentos. El propósito obvio de la función generadora de momentos es la determinación de los momentos de las distribuciones. Sin embargo, la contribución más importante es establecer distribuciones de funciones de variables aleatorias.

Si $g(X) = X^r$ for $r = 0, 1, 2, 3, \dots$, la definición 7.1 da un valor esperado que se denomina r -ésimo **momento alrededor del origen** de la variable aleatoria X , que denotamos como μ'_r .

Definición 7.1

El r -ésimo **momento alrededor del origen** de la variable aleatoria X está dado por

$$\mu'_r = E(X^r) = \begin{cases} \sum_x x^r f(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} x^r f(x) dx & \text{si } X \text{ es continua.} \end{cases}$$

Como el primer y segundo momentos alrededor del origen están dados por $\mu'_1 = E(X)$ y $\mu'_2 = E(X^2)$, podemos escribir la media y la varianza de una variable aleatoria como

$$\mu = \mu'_1 \quad \text{y} \quad \sigma^2 = \mu'_2 - \mu^2.$$

Aunque los momentos de una variable aleatoria se pueden determinar directamente de la definición 7.1, existe un procedimiento alternativo. Este procedimiento requiere la utilización de una **función generadora de momentos**.

Definición 7.2

La **función generadora de momentos** de la variable aleatoria X está dada por $E(e^{tx})$ y se denota con $M_X(t)$. De aquí

$$M_X(t) = E(e^{tx}) = \begin{cases} \sum_x e^{tx} f(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{si } X \text{ es continua.} \end{cases}$$

Las funciones generadoras de momentos existirán sólo si la suma o integral de la definición 7.2 converge. Si existe una función generadora de momentos de una variable aleatoria X , se puede utilizar para generar todos los momentos de dicha variable. El método se describe en el teorema 7.6.

Teorema 7.6

Sea X una variable aleatoria con función generadora de momentos $M_X(t)$. Entonces

$$\left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0} = \mu'_r.$$

PRUEBA Suponga que podemos diferenciar dentro de la sumatoria y de la integral, obtenemos

$$\frac{dM_X(t)}{dt} = \begin{cases} \sum_x x' e^{tx} f(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} x' e^{tx} f(x) dx & \text{si } X \text{ es continua.} \end{cases}$$

Al hacer $t = 0$, vemos que ambos casos se reducen a $E(X') = \mu'$.

Ejemplo 7.6 Encuentre la función generadora de momentos de la variable aleatoria binomial X y después utilicela para verificar que $\mu = np$ y $\sigma^2 = npq$.

SOLUCIÓN

De la definición 7.2 tenemos

$$M_X(t) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^t)^x q^{n-x}.$$

Si reconocemos a esta última suma como la expansión binomial de $(pe^t + q)^n$, obtenemos

$$M_X(t) = (pe^t + q)^n.$$

Ahora bien,

$$\frac{dM_X(t)}{dt} = n(pe^t + q)^{n-1} pe^t$$

y

$$\frac{d^2M_X(t)}{dt^2} = np[e^t(n-1)(pe^t + q)^{n-2} pe^t + (pe^t + q)^{n-1} e^t].$$

Al hacer $t = 0$, obtenemos

$$\mu'_1 = np \quad \text{y} \quad \mu'_2 = np[(n-1)p + 1].$$

Por tanto,

$$\mu = \mu'_1 = np \quad \text{y} \quad \sigma^2 = \mu'_2 - \mu^2 = np(1-p) = npq,$$

que está de acuerdo con los resultados que se obtuvieron en el capítulo 5. ■

Ejemplo 7.7 Muestre que la función generadora de momentos de la variable aleatoria X que tiene una distribución de probabilidad normal con media μ y varianza σ^2 está dada por

$$M_X(t) = \exp\left(\frac{\mu t + \sigma^2 t^2}{2}\right).$$

SOLUCIÓN

De la definición 7.2 la función generadora de momentos de la variable aleatoria normal X es

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left\{\frac{(x-\mu)}{\sigma}\right\}^2\right] dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{[x^2 - 2(\mu + t\sigma^2)x + \mu^2]}{2\sigma^2}\right] dx. \end{aligned}$$

Al completar los cuadrados en el exponente, podemos escribir

$$x^2 - 2(\mu + t\sigma^2)x + \mu^2 = [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4$$

y entonces

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{[x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4}{2\sigma^2}\right) dx \\ &= \exp\left(\frac{\mu t + \sigma^2 t^2}{2}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(1/2)[x - (\mu + t\sigma^2)]^2}{\sigma^2}\right) dx. \end{aligned}$$

Sea $w = [x - (\mu + t\sigma^2)]/\sigma$; entonces $dx = \sigma dw$ y

$$\begin{aligned} M_X(t) &= \exp\left(\frac{\mu t + \sigma^2 t^2}{2}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \\ &= \exp\left(\frac{\mu t + \sigma^2 t^2}{2}\right), \end{aligned}$$

pues la última integral representa el área bajo una curva de densidad normal estándar y por ello es igual a 1. ■

Ejemplo 7.8 Muestre que la función generadora de momentos de la variable aleatoria X que tiene una distribución ji cuadrada con v grados de libertad es $M_X(t) = (1 - 2t)^{-v/2}$.

SOLUCIÓN

La distribución ji cuadrada se obtuvo como un caso especial de la distribución gamma al hacer $\alpha = v/2$ y $\beta = 2$. Al sustituir por $f(x)$ en la definición 7.2, obtenemos

$$\begin{aligned} M_X(t) &= \int_0^{\infty} e^{tx} \frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1} e^{-x/2} dx \\ &= \frac{1}{2^{v/2}\Gamma(v/2)} \int_0^{\infty} x^{v/2-1} e^{-x(1-2t)/2} dx. \end{aligned}$$

Al escribir $y = x(1 - 2t)/2$ y $dx = [2/(1 - 2t)]dy$, obtenemos

$$M_X(t) = \frac{1}{2^{v/2}\Gamma(v/2)} \int_0^\infty \left(\frac{2y}{1-2t}\right)^{v/2-1} e^{-y} \frac{2}{1-2t} dy$$

$$= \frac{1}{\Gamma(v/2)(1-2t)^{v/2}} \int_0^\infty y^{v/2-1} e^{-y} dy = (1-2t)^{-v/2},$$

debido a que la última integral es igual a $\Gamma(v/2)$.

Aunque el método de transformación de variables proporciona una forma efectiva de encontrar la distribución de una función de diversas variables, hay un procedimiento alternativo y que a menudo se prefiere cuando la función en cuestión es una combinación lineal de variables aleatorias independientes. Este procedimiento utiliza las propiedades de las funciones generadoras de momentos que se presentan en los siguientes cuatro teoremas. Para conservar el alcance matemático de este libro, establecemos el teorema 7.7 sin demostración. ■

Teorema 7.7

(teorema de unicidad) Sean X y Y dos variables aleatorias con funciones generadoras de momentos $M_X(t)$ y $M_Y(t)$, respectivamente. Si $M_X(t) = M_Y(t)$ para todos los valores de t , entonces X y Y tienen la misma distribución de probabilidad.

Teorema 7.8

$$M_{X+a}(t) = e^{at}M_X(t).$$

PRUEBA

$$M_{X+a}(t) = E[e^{t(X+a)}] = e^{at}E[e^{tX}] = e^{at}M_X(t).$$

Teorema 7.9

$$M_{aX}(t) = M_X(at).$$

PRUEBA

$$M_{aX}(t) = E[e^{t(aX)}] = E[e^{(at)X}] = M_X(at).$$

Teorema 7.10

Si X_1, X_2, \dots, X_n son variables aleatorias independientes con funciones generadoras de momentos $M_{X_1}(t), M_{X_2}(t), \dots, M_{X_n}(t)$, respectivamente, y $Y = X_1 + X_2 + \dots + X_n$, entonces

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t) \cdots M_{X_n}(t).$$

PRUEBA. Para el caso continuo

$$M_Y(t) = E(e^{tY}) = E[e^{t(X_1+X_2+\dots+X_n)}]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t(X_1+X_2+\dots+X_n)} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n.$$

Como las variables son independientes, tenemos

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \cdots f_n(x_n)$$

y entonces

$$\begin{aligned} M_Y(t) &= \int_{-\infty}^{\infty} e^{t_1 x_1} f_1(x_1) dx_1 \int_{-\infty}^{\infty} e^{t_2 x_2} f_2(x_2) dx_2 \cdots \int_{-\infty}^{\infty} e^{t_n x_n} f_n(x_n) dx_n \\ &= M_{X_1}(t) M_{X_2}(t) \cdots M_{X_n}(t). \end{aligned}$$

La demostración para el caso discreto se obtiene de manera similar al reemplazar las integrales con sumatorias.

Los teoremas 7.7 al 7.10 son vitales para comprender las funciones generadoras de momentos. A continuación se presenta un ejemplo como ilustración. Hay muchas situaciones en las que necesitamos conocer la distribución de la suma de las variables aleatorias. Podemos utilizar los teoremas 7.7 y 7.10 y el resultado del ejercicio 19 que sigue de esta sección para encontrar la distribución de una suma de dos variables aleatorias independientes de Poisson con funciones generadoras de momento dadas por

$$M_{X_1}(t) = e^{\mu_1(e^t - 1)} \quad \text{y} \quad M_{X_2}(t) = e^{\mu_2(e^t - 1)},$$

respectivamente. De acuerdo con el teorema 7.10, la función generadora de momento de la variable aleatoria $Y_1 = X_1 + X_2$ es

$$M_{Y_1}(t) = M_{X_1}(t)M_{X_2}(t) = e^{\mu_1(e^t - 1)}e^{\mu_2(e^t - 1)} = e^{(\mu_1 + \mu_2)(e^t - 1)}$$

que de inmediato identificamos como la función generadora de momentos de una variable aleatoria que tiene una distribución de Poisson con el parámetro $\mu_1 + \mu_2$. Por ello, de acuerdo con el teorema 7.7, de nuevo concluimos que la suma de dos variables aleatorias independientes que tienen distribuciones de Poisson, con parámetros μ_1 y μ_2 , tiene una distribución de Poisson con parámetro $\mu_1 + \mu_2$.

Combinaciones lineales de variables aleatorias

En estadística aplicada a menudo se necesita conocer la distribución de probabilidad de una combinación lineal de variables aleatorias normales independientes. Obtenemos la distribución de la variable aleatoria $Y = a_1 X_1 + a_2 X_2$ cuando X_1 es una variable normal con media μ_1 y varianza σ_1^2 y X_2 también es una variable normal pero independiente de X_1 , con media μ_2 y varianza σ_2^2 . Primero, por el teorema 7.10, encontramos

$$M_Y(t) = M_{a_1 X_1}(t) M_{a_2 X_2}(t),$$

y después, con el uso del teorema 7.9,

$$M_Y(t) = M_{X_1}(a_1 t) M_{X_2}(a_2 t).$$

Al sustituir $a_i t$ por t en una función generadora de momentos de la distribución normal que se deriva en el ejemplo 7.7 y después $a_i t$ por t , tenemos

$$\begin{aligned} M_Y(t) &= \exp(a_1 \mu_1 t + a_1^2 \sigma_1^2 t^2 / 2 + a_2 \mu_2 t + a_2^2 \sigma_2^2 t^2 / 2) \\ &= \exp[(a_1 \mu_1 + a_2 \mu_2) t + (a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2) t^2 / 2]. \end{aligned}$$

que reconocemos como la función generadora de momentos de una distribución que es normal con $a_1\mu_1 + a_2\mu_2$ y varianza $a_1^2\sigma_1^2 + a_2^2\sigma_2^2$.

Para generalizar al caso de n variables normales independientes, establecemos el siguiente resultado.

Teorema 7.11

Si X_1, X_2, \dots, X_n son variables aleatorias independientes que tienen distribuciones normales con medias $\mu_1, \mu_2, \dots, \mu_n$ y varianzas $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ respectivamente, entonces la variable aleatoria

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$$

tiene una distribución normal con media

$$\mu_Y = a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n$$

y varianza

$$\sigma_Y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2.$$

Ahora queda claro que la distribución de Poisson y la distribución normal poseen una propiedad reproductiva en el sentido de que las sumas de variables aleatorias independientes que tienen cualquiera de estas distribuciones es una variable aleatoria que también tiene el mismo tipo de distribución. La distribución ji cuadrada también posee esta propiedad reproductiva.

Teorema 7.12

Si X_1, X_2, \dots, X_n son variables aleatorias mutuamente independientes que tienen, respectivamente, distribuciones ji cuadrada con v_1, v_2, \dots, v_n grados de libertad, entonces la variable aleatoria

$$Y = X_1 + X_2 + \dots + X_n$$

tiene una distribución ji cuadrada con $v = v_1 + v_2 + \dots + v_n$ grados de libertad.

PRUEBA. Por el teorema 7.10,

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t).$$

Del ejemplo 7.8,

$$M_{X_i}(t) = (1 - 2t)^{-v_i/2}, \quad i = 1, 2, \dots, n.$$

Por tanto,

$$\begin{aligned} M_Y(t) &= (1 - 2t)^{-v_1/2}(1 - 2t)^{-v_2/2}\cdots(1 - 2t)^{-v_n/2} \\ &= (1 - 2t)^{-(v_1+v_2+\cdots+v_n)/2} \end{aligned}$$

que reconocemos como la función generadora de momentos de una distribución ji cuadrada con $v = v_1 + v_2 + \dots + v_n$ grados de libertad.

Corolario

Si X_1, X_2, \dots, X_n son variables aleatorias independientes que tienen distribuciones normales idénticas con media μ y varianza σ^2 , entonces la variable aleatoria

$$Y = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2$$

tiene una distribución ji cuadrada con $\nu = n$ grados de libertad.

Este corolario es una consecuencia inmediata del ejemplo 7.5, que establece que cada una de las n variables aleatorias independientes $[(X_i - \mu)/\sigma]^2, i = 1, 2, \dots, n$, tiene una distribución ji cuadrada con 1 grado de libertad. Este corolario es en extremo importante. Establece una relación entre la muy importante distribución ji cuadrada y la distribución normal. También debe proporcionar al lector una idea clara de lo que queremos decir con el parámetro que denominamos grados de libertad. A medida que avancemos en los capítulos siguientes, la noción de grados de libertad jugará un papel de importancia creciente. Vemos del corolario que si Z_1, Z_2, \dots, Z_n son variables aleatorias normales estándar independientes, entonces $\sum_{i=1}^n Z_i^2$ tiene una distribución ji cuadrada y el *parámetro único*, ν , los grados de libertad, es n , el *número* de variables normales estándar.

Ejercicios

1. Sea X una variable aleatoria con distribución de probabilidad

$$f(x) = \begin{cases} \frac{1}{2}, & x = 1, 2, 3 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = 2X - 1$.

2. Sea X una variable aleatoria binomial con distribución de probabilidad

$$f(x) = \begin{cases} \binom{3}{x} \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{3-x}, & x = 0, 1, 2, 3 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = X^2$.

3. Sean X_1 y X_2 variables aleatorias discretas con la distribución multinomial

$$f(x_1, x_2) = \binom{2}{x_1, x_2, 2-x_1-x_2} \left(\frac{1}{4}\right)^{x_1} \left(\frac{1}{3}\right)^{x_2} \left(\frac{5}{12}\right)^{2-x_1-x_2}$$

para $x_1 = 0, 1, 2; x_2 = 0, 1, 2; x_1 + x_2 \leq 2$; y cero en cualquier otro caso. Encuentre la distribución de probabilidad conjunta de $Y_1 = X_1 + X_2$ y $Y_2 = X_1 - X_2$.

4. Sean X_1 y X_2 variables aleatorias discretas con la distribución de probabilidad conjunta

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{18}, & x_1 = 1, 2; x_2 = 1, 2, 3 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = X_1 X_2$.

5. Si X tiene la distribución de probabilidad

$$f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Muestre que la variable aleatoria $Y = -2 \ln X$ tiene una distribución ji cuadrada con 2 grados de libertad.

6. Dada la variable aleatoria X con distribución de probabilidad

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

encuentre la distribución de probabilidad de Y , donde $Y = 8X^3$.

7. La velocidad de una molécula en un gas uniforme en equilibrio es una variable aleatoria V cuya distribución de probabilidad está dada por

$$f(v) = \begin{cases} kv^2 e^{-bv}, & v > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

donde k es una constante apropiada y b depende de la temperatura absoluta y de la masa de la molécula. Encuentre la distribución de probabilidad de la energía cinética de la molécula W , donde $W = mV^2/2$.

8. La utilidad de un distribuidor, en unidades de \$5000, sobre un automóvil nuevo está dada por $Y = X^2$, donde X es una variable aleatoria que tiene la función de densidad

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre la función de densidad de probabilidad de la variable aleatoria Y .

(b) Con el uso de la función de densidad de Y , encuentre la probabilidad de que la utilidad sea menor que \$500 sobre el siguiente automóvil nuevo que venda este distribuidor.

9. El periodo hospitalario, en días, para pacientes que siguen un tratamiento para cierto tipo de enfermedad del riñón es una variable aleatoria $Y = X + 4$, donde X tiene la función de densidad

$$f(x) = \begin{cases} \frac{32}{(x+4)^3}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre la función de densidad de probabilidad de la variable aleatoria Y .

(b) Con el uso de la función de densidad de Y , encuentre la probabilidad de que el periodo hospitalario para un paciente que sigue este tratamiento exceda de ocho días.

10. Las variables aleatorias X y Y , que representan los pesos de cremas y chiclosos en cajas de un kilogramo de chocolates que contienen una combinación de cremas, chiclosos y envinados, tienen la función de densidad conjunta

$$f(x, y) = \begin{cases} 24xy, & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ & x + y \leq 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

(a) Encuentre la función de densidad de probabilidad de la variable aleatoria $Z = X + Y$.

(b) Con el uso de la función de densidad de Z , encuentre la probabilidad de que en una caja dada la suma de las cremas y los chiclosos sea al menos $1/2$ pero menos que $3/4$ del peso total.

11. La cantidad de queroseno, en miles de litros, en un tanque al inicio de cualquier día es una cantidad aleatoria Y de la que una cantidad aleatoria X se vende durante ese día. Suponga que la función de densidad conjunta de estas variables está dada por

$$f(x, y) = \begin{cases} 2, & 0 < x < y, 0 < y < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la función de densidad de probabilidad para la cantidad de queroseno que queda en el tanque al final del día.

12. Sean X_1 y X_2 variables aleatorias independientes que tiene cada una la distribución de probabilidad

$$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Muestre que las variables aleatorias Y_1 y Y_2 son independientes cuando $Y_1 = X_1 + X_2$ y $Y_2 = X_1/(X_1 + X_2)$.

13. Una corriente de I amperes que fluye a través de una resistencia de R ohmios varía de acuerdo con la distribución de probabilidad

$$f(i) = \begin{cases} 6i(1-i), & 0 \leq i < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Si la resistencia varía independientemente de la corriente de acuerdo con la distribución de probabilidad

$$g(r) = \begin{cases} 2r, & 0 < r < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

encuentre la distribución de probabilidad para la potencia $W = I^2 R$ watts.

14. Sea X una variable aleatoria con distribución de probabilidad

$$f(x) = \begin{cases} \frac{1+x}{2}, & -1 < x < 1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = X^2$.

15. Si X tiene la distribución de probabilidad

$$f(x) = \begin{cases} \frac{2(x+1)}{9}, & -1 < x < 2 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Encuentre la distribución de probabilidad de la variable aleatoria $Y = X^2$.

16. Muestre que el r -ésimo momento alrededor del origen de la distribución gamma es

$$\mu'_r = \frac{\beta \Gamma(\alpha + r)}{\Gamma(\alpha)}$$

SUGERENCIA: sustituya $y = x/\beta$ en la integral que define μ'_r y después utilice la función gamma para evaluar la integral.

17. Una variable aleatoria X tiene la distribución uniforme discreta

$$f(x; k) = \begin{cases} \frac{1}{k}, & x = 1, 2, 3, \dots, k \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Muestre que la función generadora de momentos de X es

$$M_X(t) = \frac{e^t(1 - e^{kt})}{k(1 - e^t)}$$

18. Una variable aleatoria X tiene la distribución geométrica $g(x; p) = pq^{x-1}$ para $x = 1, 2, 3, \dots$. Muestre que la función generadora de momentos de X es

$$M_X(t) = \frac{pe^t}{1 - qe^t}$$

y después use $M_X(t)$ para encontrar la media y la varianza de la distribución geométrica.

19. Una variable aleatoria X tiene la distribución de Poisson $p(x; \mu) = e^{-\mu} \mu^x / x!$ para $x = 1, 2, 3, \dots$. Muestre que la función generadora de momentos de X es

$$M_X(t) = e^{\mu(e^t - 1)}$$

Con el uso $M_X(t)$, encuentre la media y la varianza de la distribución de Poisson.

20. La función generadora de momentos de cierta variable aleatoria de Poisson X está dada por

$$M_X(t) = e^{\mu(e^t - 1)}$$

Encuentre $P(\mu - 2\sigma < X < \mu + 2\sigma)$.

21. Con el uso de la función generadora de momentos del ejemplo 7.8, muestre que la media y la varianza de la distribución ji cuadrada con v grados de libertad son, respectivamente, v y $2v$.

22. Mediante la expansión de x^x en una serie de Maclaurin y la integración término por término, muestre que

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x) dx \\ &= 1 + \mu t + u'_2 \frac{t^2}{2!} + \dots + u'_r \frac{t^r}{r!} + \dots \end{aligned}$$

Distribuciones fundamentales de muestreo y descripciones de datos

8.1 Muestreo aleatorio

El resultado de un experimento estadístico se puede registrar como un valor numérico o como una representación descriptiva. Cuando se lanza un par de dados y el total es el resultado de interés, registramos un valor numérico. Sin embargo, si a los estudiantes de cierta escuela se les hacen pruebas de sangre y el tipo sanguíneo es de interés, entonces una representación descriptiva puede ser la más útil. La sangre de una persona se puede clasificar de ocho maneras. Puede ser AB, A, B u O, con un signo más o uno menos, que depende de la presencia o ausencia del antígeno Rh.

El estadístico se interesa en primer lugar en el análisis de los datos numéricos. Para la clasificación de los tipos sanguíneos, puede convenir utilizar números del 1 al 8 para representar los tipos de sangre y entonces registrar el número apropiado para cada estudiante. En cualquier estudio, el número de observaciones posibles puede ser pequeño, grande pero finito o infinito. Por ejemplo, en la clasificación de tipos sanguíneos sólo podemos tener tantas observaciones como estudiantes haya en la escuela. El proyecto, por tanto, tiene como resultado un número finito de observaciones. Por otro lado, si pudiéramos lanzar un par de dados indefinidamente y registrar los totales que ocurriesen, obtendríamos un conjunto infinito de valores, donde cada valor representaría el resultado de un solo lanzamiento del par de dados.

En este capítulo nos enfocamos en el muestreo de distribuciones o poblaciones y estudiamos cantidades tan importantes como la *media de la muestra* y la *varianza de la muestra*, que son de importancia vital para los capítulos siguientes. Además, intentamos dar al lector una introducción al papel que jugarán la media y la varian-

za de la muestra en los próximos capítulos sobre inferencia estadística. Se dan introducciones a las *representaciones de datos*, a través de las que se pueden realizar evaluaciones gráficas de lo que se puede aprender de los datos experimentales.

El uso de las computadoras modernas de alta velocidad permite al científico o al ingeniero aumentar enormemente su uso de la inferencia estadística formal con técnicas gráficas. La mayor parte de las veces la inferencia formal parece bastante árida y quizá incluso abstracta para el practicante o el administrador que desea que el análisis estadístico sea una guía para la toma de decisiones.

Comenzamos esta sección con la presentación de las nociones de *poblaciones* y *muestras*. Ambas se mencionan de manera amplia en la introducción al capítulo 1. La totalidad de observaciones que nos interesan, finitas o infinitas, constituye lo que llamamos **población**. En el pasado la palabra *población* se refería a observaciones que se obtenían de estudios estadísticos con personas. En la actualidad, el estadístico utiliza el término para referirse a observaciones relevantes de cualquier cosa de interés, ya sea de grupos de personas, animales o todos los resultados posibles de algún sistema biológico o de ingeniería complicado.

Definición 8.1

Una **población** consiste en la totalidad de las observaciones en las que estamos interesados.

El número de observaciones en la población se define como el **tamaño** de la población. Si en la escuela hay 600 estudiantes que clasificamos de acuerdo con su tipo sanguíneo, decimos que tenemos una población de tamaño 600. Los números en las cartas de una baraja, las alturas de los residentes de cierta ciudad y las longitudes de los peces en un lago particular son ejemplos de poblaciones de tamaño finito. En cada caso el número total de observaciones es un número finito. El experimento de lanzamiento de un dado genera una población cuyo tamaño es infinito. De manera similar, las observaciones que se obtienen al medir la presión atmosférica cada día del pasado al futuro, o todas las mediciones de la profundidad de un lago desde cualquier posición concebible, son ejemplos de poblaciones cuyos tamaños son infinitos. Algunas poblaciones finitas son tan grandes que en teoría las suponemos infinitas. Esto es cierto si se considera la población de la duración de cierto tipo de batería de almacenamiento que se fabrica para su distribución masiva en todo el país.

Cada observación en una población es un valor de una variable aleatoria X que tiene alguna distribución de probabilidad $f(x)$. Si se inspeccionan artículos que salen de una línea de montaje en busca de defectos, entonces cada observación en la población podría ser un valor 0 o 1 de la variable aleatoria binomial X con distribución de probabilidad

$$b(x; 1, p) = p^x q^{1-x}, \quad x = 0, 1,$$

donde 0 indica un artículo no defectuoso y 1 indica uno defectuoso. Por supuesto, se supone que p , la probabilidad de que cualquier artículo esté defectuoso, permanece constante de una prueba a otra. En el experimento de tipo sanguíneo la variable aleatoria X representa el tipo de sangre al tomar un valor del 1 al 8. A cada estudiante se le asigna uno de los valores de la variable aleatoria discreta. Las

duraciones de las baterías de almacenamiento son valores que toma una variable aleatoria continua que quizá tiene una distribución normal. De aquí en adelante cuando hagamos referencia a una "población binomial", una "población normal" o, en general, la "población $f(x)$ ", nos referiremos a una población cuyas observaciones son valores de una variable aleatoria que tiene una distribución binomial, una distribución normal o la distribución de probabilidad $f(x)$. Por ello a la media y la varianza de una variable aleatoria o distribución de probabilidad también se les denomina la media y la varianza de la población correspondiente.

En el campo de la inferencia estadística el estadístico se interesa en llegar a conclusiones con respecto a la población cuando es imposible o poco práctico observar todo el conjunto de observaciones que constituyen la población. Por ejemplo, al intentar determinar la longitud promedio de la vida de cierta marca de foco, sería imposible probar todos los focos si tenemos que venderlos. Los costos exorbitantes también pueden ser un factor prohibitivo para estudiar toda la población. Por tanto, debemos depender de un subconjunto de observaciones de la población para ayudarnos a hacer inferencias con respecto a la misma población. Esto nos lleva a considerar la noción de muestreo.

Definición 8.2

Una **muestra** es un subconjunto de una población.

Si nuestras inferencias a partir de la muestra para la población han de ser válidas, debemos obtener muestras que sean representativas de la población. Con mucha frecuencia estamos tentados a elegir una muestra mediante la selección de los miembros más convenientes de la población. Tal procedimiento puede conducir a inferencias erróneas con respecto a la población. Cualquier procedimiento de muestreo que produzca inferencias que sobreestimen o subestimen de forma consistente alguna característica de la población se dice que está **sesgado**. Para eliminar cualquier posibilidad de sesgo en el procedimiento de muestreo, es deseable elegir una **muestra aleatoria** en el sentido de que las observaciones se realizan de forma independiente y al azar.

Para seleccionar una muestra aleatoria de tamaño n de una población $f(x)$, definamos la variable aleatoria X_i , $i = 1, 2, \dots, n$, que represente la i -ésima medición o valor de la muestra que observemos. Las variables aleatorias X_1, X_2, \dots, X_n constituirán entonces una muestra aleatoria de la población $f(x)$ con valores numéricos x_1, x_2, \dots, x_n si las mediciones se obtienen al repetir el experimento n veces independientes bajo esencialmente las mismas condiciones. Debido a las condiciones idénticas bajo las que se seleccionan los elementos de la muestra, es razonable suponer que las n variables aleatorias X_1, X_2, \dots, X_n son independientes y que cada una tiene la misma distribución de probabilidad $f(x)$. Es decir, las distribuciones de probabilidad de X_1, X_2, \dots, X_n son, respectivamente, $f(x_1), f(x_2), \dots, f(x_n)$ y su distribución de probabilidad conjunta es

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \cdots f(x_n).$$

El concepto de muestra aleatoria se describe de manera formal mediante la siguiente definición.

Definición 8.3

Sean X_1, X_2, \dots, X_n variables aleatorias independientes, cada una con la misma distribución de probabilidad $f(x)$. Definimos entonces a X_1, X_2, \dots, X_n como una **muestra aleatoria** de tamaño n de la población $f(x)$ y escribimos su distribución de probabilidad conjunta como

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2)\cdots f(x_n).$$

Si se realiza una selección aleatoria de $n = 8$ baterías de almacenamiento de un proceso de fabricación, que mantiene las mismas especificaciones, y registramos la duración de cada batería con la primera medición x_1 como un valor de X_1 , la segunda medición x_2 como un valor de X_2 , etcétera, entonces x_1, x_2, \dots, x_n son los valores de la muestra aleatoria X_1, X_2, \dots, X_n . Si suponemos que la población de duraciones de las baterías es normal, los valores posibles de cualquier $X_i, i = 1, 2, \dots, 8$, serán precisamente los mismos que los de la población original, y por ello X_i tiene la misma distribución normal idéntica que X .

8.2 Algunas estadísticas importantes

Nuestro propósito principal al seleccionar muestras aleatorias es obtener información acerca de los parámetros desconocidos de la población. Suponga, por ejemplo, que deseamos llegar a una conclusión con respecto a la proporción de personas bebedoras de café en Estados Unidos que prefieren cierta marca de café. Sería imposible preguntar a cada bebedor de café estadounidense a fin de calcular el valor del parámetro p que representa la proporción de la población. En su lugar, se selecciona una muestra aleatoria grande y se calcula la proporción \hat{p} de personas en esta muestra que prefieren la marca de café en cuestión. El valor \hat{p} se utiliza ahora para hacer una inferencia con respecto a la proporción p verdadera.

Ahora bien, \hat{p} es una función de los valores observados en la muestra aleatoria; como son posibles muchas muestras aleatorias a partir de la misma población, esperaríamos que \hat{p} variara algo de una muestra a otra. Es decir, \hat{p} es un valor de una variable aleatoria que representamos con \hat{P} . Tal variable aleatoria se llama **estadística**.

Definición 8.4

Cualquier función de las variables aleatorias que forman una muestra aleatoria se llama **estadística**.

Tendencia central en la muestra

En el capítulo 4 introducimos los dos parámetros μ y σ^2 , que miden el centro de localización y la variabilidad de una distribución de probabilidad. Estos son parámetros poblacionales constantes y de ninguna manera resultan afectados o influidos por las observaciones de una muestra aleatoria. Definiremos, sin embargo, algunas estadísticas importantes que describen las medidas correspondientes de una muestra aleatoria. Las estadísticas que por lo general se utilizan más para medir el centro de un conjunto de datos, acomodados en orden de magnitud, son la **media**, **mediana** y

moda. La media y la mediana se exponen en el capítulo 1, pero se definen de nuevo aquí. La más importante de éstas y que consideraremos primero es la media.

Definición 8.5

Si X_1, X_2, \dots, X_n representan una muestra aleatoria de tamaño n , entonces la **media de la muestra** se define mediante la estadística

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Nótese que la estadística \bar{X} toma el valor $\bar{x} = \sum_{i=1}^n x_i/n$ cuando X_1 toma el valor x_1 , X_2 toma el valor x_2 y así sucesivamente. En la práctica al valor de una estadística por lo general se le da el mismo nombre de la estadística. Por ejemplo, el término *media de la muestra* se aplica tanto a la estadística \bar{X} como a su valor calculado \bar{x} .

Hay una referencia anterior a la media de la muestra en el capítulo 1. En ese contexto discutimos los valores calculados con base en conjuntos específicos de datos.

Ejemplo 8.1 Un inspector de alimentos examina una muestra aleatoria de siete latas de cierta marca de atún para determinar el porcentaje de impurezas externas. Se registran los siguientes datos: 1.8, 2.1, 1.7, 1.6, 0.9, 2.7, y 1.8. Calcule la media de la muestra.

SOLUCIÓN

El valor observado \bar{x} de la estadística \bar{X} es

$$\bar{x} = \frac{1.8 + 2.1 + 1.7 + 1.6 + 0.9 + 2.7 + 1.8}{7} = 1.8\%$$

La segunda estadística más útil para medir el centro de un conjunto de datos es la mediana. Designaremos a la mediana con el símbolo \tilde{X} . ■

Definición 8.6

Si X_1, X_2, \dots, X_n representa una muestra aleatoria de tamaño n , acomodada en orden creciente de magnitud, entonces la **mediana de la muestra** se define mediante la estadística

$$\tilde{X} = \begin{cases} X_{(n+1)/2} & \text{si } n \text{ es impar} \\ \frac{X_{n/2} + X_{(n/2)+1}}{2} & \text{si } n \text{ es par} \end{cases}$$

Ejemplo 8.2 El número de barcos extranjeros que llegan a un puerto de la costa este en siete días seleccionados al azar son 8, 3, 9, 5, 6, 8, y 5. Encuentre la mediana de la muestra.

SOLUCIÓN

Al acomodar las observaciones en orden creciente de magnitud, obtenemos

$$3 \quad 5 \quad 5 \quad 6 \quad 8 \quad 8 \quad 9$$

y de aquí $\tilde{x} = 6$. ■

Ejemplo 8.3 Se encuentra que el contenido de nicotina para una muestra aleatoria de seis cigarrillos de cierta marca son 2.3, 2.7, 2.5, 2.9, 3.1, y 1.9 miligramos. Encuentre la mediana.

SOLUCIÓN

Si acomodamos estos contenidos de nicotina en orden creciente de magnitud, obtenemos

$$1.9 \quad 2.3 \quad 2.5 \quad 2.7 \quad 2.9 \quad 3.1$$

y la mediana es entonces la media de 2.5 y 2.7. Por tanto,

$$\bar{x} = \frac{2.5 + 2.7}{2} = 2.6 \text{ miligramos.}$$

La tercera y última estadística para medir el centro de una muestra aleatoria que presentaremos es la moda, que se designa con la estadística M . ■

Definición 8.7

Si X_1, X_2, \dots, X_n , no necesariamente diferentes, representan una muestra aleatoria de tamaño n , entonces la **moda** M es aquel valor de la muestra que ocurre más a menudo o con la mayor frecuencia. La moda puede no existir, y cuando existe no necesariamente es única.

La moda no siempre existe. Esto es realmente cierto cuando todas las observaciones ocurren con la misma frecuencia. Para algunos conjuntos de datos habrá varios valores que ocurran con la mayor frecuencia, en cuyo caso tenemos más de una moda.

Ejemplo 8.4 Si las donaciones de una muestra aleatoria de residentes de Fairway Forest para la Asociación del Pulmón de Virginia se registran como 9, 10, 5, 9, 9, 7, 8, 6, 10, y 11 dólares, entonces la moda es $m = \$9$, el valor que ocurre con la mayor frecuencia. ■

Ejemplo 8.5 El número de películas vistas el mes pasado por una muestra aleatoria de 12 estudiantes de preparatoria se registraron como sigue: 2, 0, 3, 1, 2, 4, 2, 5, 4, 0, 1, y 4. En este caso hay dos modas, 2 y 4, pues ambas ocurren con la mayor frecuencia. Se dice que la distribución es **bimodal**. ■

Ejemplo 8.6 No existe moda para el contenido de nicotina del ejemplo 8.3, puesto que cada medición ocurre sólo una vez. ■

En resumen, consideremos los méritos relativos de la media, mediana y moda. La media es la medida de localización central más comúnmente utilizada en estadística. Emplea toda la información disponible. Las distribuciones de medias que se obtienen en muestreos repetidos de una población son bien conocidos, y en consecuencia los métodos que se utilizan en la inferencia estadística para estimar μ se basan en la media de la muestra. Esto será evidente en la sección 8.3. La única desventaja real de la media es que puede resultar afectada de manera adversa por valores extremos. En el ejemplo 8.4 la contribución media a la Asociación del Pul-

món de Virginia es \$8.40, que se acerca bastante a la moda o la mediana, que son \$9 ambas. Sin embargo si una de las contribuciones hubiera sido mucho mayor, digamos \$90 en lugar de \$11, entonces la contribución media sería \$16.30, un valor considerablemente más alto que la mayoría de los donativos.

La mediana tiene la ventaja de ser fácil de calcular si el número de observaciones es relativamente pequeño. No resulta influida por valores extremos y en consecuencia en el ejemplo 8.4 da un mejor *centro de los datos*, a saber \$9, si la contribución más alta es \$90 en lugar de \$11. Al tratar con muestras que se seleccionan de poblaciones, las medias de las muestras por lo general no variarán tanto de una muestra a otra como las medianas. Por tanto, si intentamos estimar el centro de una población con base en un valor de la muestra, la media es más estable que la mediana. Por tanto una media de la muestra probablemente esté más cercana que la mediana de la muestra a la media de la población.

La moda es la medición que se utiliza menos de las tres. Para conjuntos pequeños de datos su valor casi no tiene utilidad si es que, de hecho, existe. Sólo en el caso de una gran cantidad de datos tiene un sentido significativo. Sus dos ventajas principales son que 1) no requiere cálculo y 2) se puede usar para datos cualitativos así como para cuantitativos. Entonces, si correr es la forma de ejercicio que la mayoría de las personas expresa como preferida, decimos que correr es la **selección modal**.

Variabilidad en la muestra

Las tres medidas de localización central que se definieron antes no dan por sí mismas una descripción adecuada de nuestros datos. Necesitamos conocer cómo se dispersan las observaciones del promedio. Se remite al lector al capítulo 1 para una discusión más amplia. Es muy posible tener dos conjuntos de observaciones con la misma media o mediana que difieran de manera considerable en la variabilidad de sus mediciones alrededor del promedio.

Considere las siguientes mediciones, en litros, para dos muestras de jugo de naranja embotellado por las compañías A y B:

Muestra A	0.97	1.00	0.94	1.03	1.11
Muestra B	1.06	1.01	0.88	0.91	1.14

Ambas muestras tienen la misma media, 1.00 litros. Es bastante obvio que la compañía A embotella el jugo de naranja con un contenido más uniforme que la B. Decimos que la variabilidad o la dispersión de las observaciones del promedio es menor para la muestra A que para la muestra B. Por tanto, al comprar jugo de naranja, tendremos más confianza de que la botella seleccionada esté más cerca del promedio anunciado si la compramos de la compañía A.

Las estadísticas más importantes para medir la variabilidad de una muestra aleatoria son el **rango** y la **varianza**. La más simple de calcular es el rango:

Definición 8.8

El **rango** de una muestra aleatoria X_1, X_2, \dots, X_n se define con la estadística $X_{(n)} - X_{(1)}$, donde $X_{(n)}$ y $X_{(1)}$ son, respectivamente, las observaciones más grande y más pequeña en la muestra.

En el caso de las compañías que embotellan jugo de naranja, el rango para la compañía A es 0.17 litros comparado con un rango de 0.26 litros para la B, lo que indica una mayor dispersión de los valores para la compañía B.

Ejemplo 8.7 Los CI de una muestra aleatoria de cinco miembros de una asociación estudiantil femenina son 108, 112, 127, 118 y 113. Encuentre el rango.

SOLUCIÓN

El rango de los cinco CI es $127 - 108 = 19$.

El rango puede ser una medida pobre de la variabilidad, en particular si el tamaño de la muestra o población es grande. Considera sólo los valores extremos y no nos dice nada acerca de la distribución de los valores intermedios. Considere, por ejemplo, los dos conjuntos de datos siguientes, ambos con un rango de 12:

	Datos observados								
Conjunto A	3	4	5	6	8	9	10	12	15
Conjunto B	3	7	7	7	8	8	8	9	15

En el primer conjunto la media y la mediana son ocho, pero los números varían a lo largo de todo el intervalo de 3 a 15. En el segundo conjunto la media y la mediana también son ocho, pero la mayor parte de los valores están más cerca del centro de los datos. Aunque el rango falla al medir esta variabilidad entre la observación superior y la inferior, tiene algunas aplicaciones útiles. En la industria el rango se puede predeterminar al especificar por adelantado que una medición particular de los artículos que salen de una línea de producción deba caer dentro de cierto intervalo. En tanto que todas las mediciones caigan dentro del intervalo especificado, se dice que el proceso está bajo control. El uso del rango de la muestra para las *tablas de control* se discute con amplitud en el capítulo 17.

Para compensar la desventaja del rango, consideraremos una medida de variabilidad, a saber, la **varianza de la muestra**, que considera la posición de cada observación en relación con la media de la muestra. ■

Definición 8.9

Si X_1, X_2, \dots, X_n representa una muestra aleatoria de tamaño n , entonces la **varianza de la muestra** se define con la estadística

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

El valor calculado de s^2 para una muestra dada se denota con s^2 . Nótese que s^2 se define esencialmente como el promedio de los cuadrados de las desviaciones

de las observaciones de su media. La razón para utilizar $n - 1$ como divisor en lugar de la elección más obvia n quedará más claro en el capítulo 9.

Ejemplo 8.8 Una comparación de precios de café en cuatro tiendas de abarrotes seleccionadas al azar en San Diego muestra aumentos en comparación con el mes anterior de 12, 15, 17, y 20 centavos para una bolsa de una libra. Encuentre la varianza de esta muestra aleatoria de aumentos de precio.

SOLUCIÓN

Al calcular la media de la muestra, obtenemos

$$\bar{x} = \frac{12 + 15 + 17 + 20}{4} = 16 \text{ centavos.}$$

Por tanto,

$$\begin{aligned} s^2 &= \frac{\sum_{i=1}^4 (x_i - 16)^2}{3} \\ &= \frac{(12 - 16)^2 + (15 - 16)^2 + (17 - 16)^2 + (20 - 16)^2}{3} \\ &= \frac{(-4)^2 + (-1)^2 + (1)^2 + (4)^2}{3} = \frac{34}{3}. \end{aligned}$$

Si \bar{x} es un número decimal que se redondea, acumulamos un error grande cuando utilizamos la fórmula de la varianza de la muestra de la definición 8.9. Para evitar esto, derivemos una fórmula muy útil en el siguiente teorema. ■

Teorema 8.1

Si S^2 es la varianza de una muestra aleatoria de tamaño n , podemos escribir

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}{n(n-1)},$$

PRUEBA. Por definición,

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n (X_i^2 - 2\bar{X}X_i + \bar{X}^2)}{n-1} \\ &= \frac{\sum_{i=1}^n X_i^2 - 2\bar{X} \sum_{i=1}^n X_i + n\bar{X}^2}{n-1} \end{aligned}$$

Al reemplazar \bar{X} por $\sum_{i=1}^n X_i/n$ y multiplicar el numerador y el denominador por n , obtenemos la fórmula de cálculo más útil

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}{n(n-1)}$$

Definición 8.10

La **desviación estándar de la muestra**, que se denota con S , es la raíz cuadrada positiva de la varianza de la muestra.

Ejemplo 8.9 Encuentre la varianza de los datos 3, 4, 5, 6, 6 y 7, que representan el número de truchas atrapadas por una muestra aleatoria de seis pescadores el 19 de junio de 1996, en el lago Muskoka.

SOLUCIÓN

Encontramos que $\sum_{i=1}^6 x_i^2 = 171$, $\sum_{i=1}^6 x_i = 31$, $n = 6$. De aquí

$$s^2 = \frac{(6)(171) - (31)^2}{(6)(5)} = \frac{13}{6} \quad \blacksquare$$

Ejercicios

1. Defina las poblaciones adecuadas de las que se seleccionaron las siguientes muestras:

(a) Se llamó por teléfono a personas de 200 casas en la ciudad de Richmond y se les pidió nombrar al candidato por el que votarían en la elección para la mesa directiva de la escuela.

(b) Se lanzó 100 veces una moneda y se registraron 34 cruces.

(c) Se probaron 200 pares de un nuevo tipo de tenis en un torneo profesional y, en promedio, duraron cuatro meses.

(d) En cinco ocasiones diferentes tomó a una abogada 21, 26, 24, 22 y 21 minutos manejar de su casa en los suburbios a su oficina en el centro de la ciudad.

2. El número de infracciones emitidas por violaciones de tránsito por ocho oficiales estatales durante el fin de semana del Día de Conmemoración de los Caídos son 5, 4, 7, 7, 6, 3, 8 y 6.

(a) Si estos valores representan el número de infracciones levantadas por una muestra aleatoria de ocho oficiales estatales del condado de Montgomery en Virginia, defina una población adecuada.

(b) Si los valores representan el número de infracciones levantadas por una muestra aleatoria de ocho oficiales estatales de Carolina del Sur, defina una población adecuada.

3. El número de respuestas incorrectas en un examen de competencia de verdadero o falso para una muestra

aleatoria de 15 estudiantes se registraron como sigue: 2, 1, 3, 0, 1, 3, 6, 0, 3, 3, 5, 2, 1, 4 y 2. Encuentre

(a) la media;

(b) la mediana;

(c) la moda.

4. Las longitudes de tiempo, en minutos, que 10 pacientes esperan en el consultorio de un doctor antes de recibir tratamiento se registraron como sigue: 5, 11, 9, 5, 10, 15, 6, 10, 5 y 10. Trate los datos como una muestra aleatoria y encuentre

(a) la media;

(b) la mediana;

(c) la moda.

5. Los tiempos de reacción para una muestra aleatoria de nueve sujetos para un estimulante se registraron como 2.5, 3.6, 3.1, 4.3, 2.9, 2.3, 2.6, 4.1 y 3.4 segundos. Calcule

(a) la media;

(b) la mediana.

6. De acuerdo con la escritora de ecología Jacqueline Killeen, los fosfatos contenidos en los detergentes de uso casero pasan directo a través de nuestros sistemas de desagüe, ocasionando que los lagos se conviertan en pantanos que eventualmente se secan y se convierten en desiertos. Los siguientes datos muestran la cantidad de fosfatos por carga de lavado, en gramos, para una muestra aleatoria de diversos tipos

de detergentes que se usan de acuerdo con las instrucciones prescritas:

Detergente de lavandería	Fosfato por carga (gramos)
A & P Vela azul	48
Dash	47
Concentrado	42
Agua fría	42
Briza	41
Oxydol	34
Ajax	31
Sears	30
Fab	29
Poder frío	29
Bold	29
Rinso	26

Para los datos de fosfato dados, encuentre

- (a) la media;
 (b) la mediana;
 (c) la moda.

7. Una muestra aleatoria de empleados de una planta de manufacturas local prometieron los siguientes donativos, en dólares, al United Fund: 100, 40, 75, 15, 20, 100, 75, 50, 30, 10, 55, 75, 25, 50, 90, 80, 15, 25, 45 y 100. Calcule

- (a) la media;
 (b) la moda.

8. Encuentre la media, la mediana y la moda para la muestra cuyas observaciones, 15, 7, 8, 95, 19, 12, 8, 22 y 14, representan el número de días con incapacidad médica reportados en nueve declaraciones de impuestos federales sobre ingresos. ¿Qué valor parece ser la mejor medición del centro de nuestros datos? Expresé las razones de su preferencia.

9. Con referencia a los tiempos que esperan 10 pacientes en un consultorio médico antes de recibir tratamiento en el ejercicio 4, encuentre

- (a) el rango;
 (b) la desviación estándar.

10. Con referencia a la muestra de tiempos de reacción para los nueve sujetos que reciben el estimulante en el ejercicio 5, calcule

- (a) el rango;
 (b) la varianza con el uso de la fórmula de la definición 8.9.

11. Con referencia a la muestra aleatoria de respuestas incorrectas en un examen de competencia de verdadero y falso para los 15 estudiantes en el ejercicio 3, calcule la varianza con el uso de la fórmula

- (a) de la definición 8.9;
 (b) del teorema 8.1.

12. El contenido de alquitrán de ocho marcas de cigarrillos que se seleccionan al azar de la lista más reciente publicada por la Comisión de Comercio Federal es el que sigue: 7.3, 8.6, 10.4, 16.1, 12.2, 15.1, 14.5 y 9.3 miligramos. Calcule

- (a) la media;
 (b) la varianza.

13. Los promedios de los puntos por grado de 20 estudiantes universitarios del último año seleccionados al azar de una clase que se va a graduar son los siguientes:

3.2	1.9	2.7	2.4
2.8	2.9	3.8	3.0
2.5	3.3	1.8	2.5
3.7	2.8	2.0	3.2
2.3	2.1	2.5	1.9

Calcule la desviación estándar.

14. (a) Muestre que la varianza de la muestra permanece sin cambio si se suma o se resta una constante c de cada valor de la muestra.

(b) Muestre que la varianza de la muestra se hace c^2 veces su valor original si cada observación en la muestra se multiplica por c .

15. Verifique que la varianza de la muestra 4, 9, 3, 6, 4 y 7 es 5.1, y con el uso de este hecho junto con los resultados del ejercicio 14, encuentre

- (a) la varianza de la muestra 12, 27, 9, 18, 12 y 21;
 (b) la varianza de la muestra 9, 14, 8, 11, 9 y 12.

8.3 Presentaciones de datos y métodos gráficos

En la sección 3.4 presentamos al lector las distribuciones empíricas. La motivación es el uso de presentaciones creativas para extraer información acerca de las propiedades de un conjunto de datos. Por ejemplo, los diagramas de tronco y hojas proporcionan a quien los ve una imagen de la simetría y de otras propiedades de los datos. En este capítulo tratamos con muestras que, por supuesto, son colecciones de datos

experimentales de los que extraemos conclusiones sobre las poblaciones. A menudo la apariencia de la muestra proporciona información acerca de la distribución de la que se toman los datos. Por ejemplo, en el capítulo 1 ilustramos la naturaleza general de pares de muestras con gráficas de puntos que presentan una comparación relativa entre la tendencia central y la variabilidad entre las dos muestras.

En los capítulos siguientes, con frecuencia hacemos la suposición de que la distribución es normal. La información gráfica con respecto a la validez de esta suposición se puede obtener de presentaciones como los diagramas de tronco y hojas y los histogramas de frecuencia. Además, presentaremos en esta sección la noción de *gráficas de probabilidad normal* y *gráficas de cuantiles*. Estas gráficas se utilizan en estudios que tienen grados de complejidad que varían, con el objetivo principal de que las gráficas proporcionen una verificación diagnóstica de la suposición de que los datos vienen de una distribución normal.

Podemos caracterizar el análisis estadístico como el proceso de extraer conclusiones acerca de los sistemas en presencia de variabilidad del sistema. La intención de un ingeniero de aprender acerca de un proceso químico a menudo se ve empañado por la *variabilidad del proceso*. Un estudio que involucra el número de artículos defectuosos en un proceso de producción con frecuencia se hace más difícil por la variabilidad en el método de fabricación de los artículos. En todo lo anterior, aprendimos acerca de las muestras y las estadísticas que expresan el centro de localización y la variabilidad en la muestra. Estas estadísticas proporcionan medidas simples, mientras que una presentación gráfica agrega información adicional en términos de una imagen.

Gráfico de caja y extensión o gráfica de caja

Otra presentación que es útil para reflejar propiedades de una muestra es la **gráfica de caja y extensión**. Esta gráfica encierra el *rango intercuartil* de los datos en una caja que tiene la mediana representada dentro. El rango intercuartil tiene como extremos el percentil 75 (cuartil superior) y el percentil 25 (cuartil inferior). Además de la caja se prolongan extensiones, que muestran las *observaciones extremas* en la muestra. Para muestras razonablemente grandes la presentación muestra el centro de la localización, la variabilidad y el grado de asimetría.

Además, una variación que se llama **gráfica de caja** puede proporcionar a quien la ve información con respecto a cuáles observaciones son **datos apartados**. Los datos apartados son observaciones que se consideran inusualmente alejadas de la masa de datos. Hay muchas pruebas estadísticas diseñadas para detectar datos apartados. Técnicamente, se puede considerar a un dato apartado como una observación que representa un "evento raro"; es decir, existe una probabilidad pequeña de obtener un valor tan alejado de la masa de datos. El concepto de datos apartados resurge en el capítulo 12 en el contexto del análisis de regresión.

La información visual en la caja y extensión y en la caja no intenta ser una prueba formal de datos alejados. Más bien, se ve como una herramienta de diagnóstico. Mientras que la determinación de cuáles observaciones son datos apartados varía con el tipo de software que se use, un procedimiento común es utilizar un **múltiplo del rango intercuartil**. Por ejemplo, si la distancia desde la caja excede 1.5 veces el rango intercuartil (en cualquier dirección) la observación se puede etiquetar como

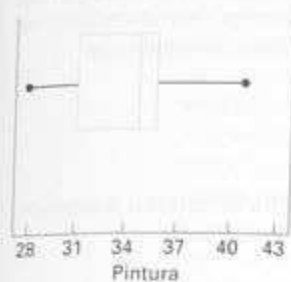


Figura 8.2 Gráfica de caja y extensión para el espesor de las "asas" de las latas de pintura.

Muestra	Mediciones				Muestra	Mediciones					
1	29	36	39	34	34	16	35	30	35	29	37
2	29	29	28	32	31	17	40	31	38	35	31
3	34	34	39	38	37	18	35	36	30	33	32
4	35	37	33	38	41	19	35	34	35	30	36
5	30	29	31	38	29	20	35	35	31	38	36
6	34	31	37	39	36	21	32	36	36	32	36
7	30	35	33	40	36	22	36	37	32	34	34
8	28	28	31	34	30	23	29	34	33	37	35
9	32	36	38	38	35	24	36	36	35	37	37
10	35	30	37	35	31	25	36	30	35	33	31
11	35	30	35	38	35	26	35	30	29	38	35
12	38	34	35	35	31	27	35	36	30	34	36
13	34	35	33	30	34	28	35	30	36	29	35
14	40	35	34	33	35	29	38	36	35	31	31
15	34	35	38	35	30	30	30	34	40	28	30

ra este conjunto asimétrico de datos. Nótese que el bloque izquierdo es considerablemente más grande que el bloque de la derecha. La mediana es 35. El cuartil inferior es 31, mientras que el superior es 36. Nótese también que la observación extrema en la derecha está más alejada de la caja que la observación extrema de la izquierda. No hay datos apartados en este conjunto de datos.

Hay formas adicionales en las que las gráficas de caja y extensión y otras presentaciones gráficas pueden ayudar al analista. Las muestras múltiples se pueden comparar de forma gráfica. Las gráficas de datos pueden sugerir relaciones entre variables. Las gráficas pueden ayudar en la detección de anomalías o de observaciones de datos apartados en las muestras.

Otro tipo de gráfica que en particular puede ser útil para caracterizar la naturaleza de un conjunto de datos es la *gráfica de cuantiles*. Como en el caso de la gráfica de caja y extensión, se puede hacer uso de las ideas básicas de la gráfica de cuantiles para *comparar muestras de datos*, donde el objetivo del analista es encontrar diferencias. Ilustraciones adicionales de este tipo de utilización se darán en los capítulos siguientes donde se presenta la inferencia formal de la estadística asociada con la comparación de las muestras. En ese momento se demostrarán estudios de caso en los que se expone al lector la inferencia formal y las gráficas de diagnóstico para el mismo conjunto de datos. ■

Gráfica de cuantiles

El propósito de las gráficas de cuantiles es describir, en forma de muestra, la función de distribución acumulada que se presentó en el capítulo 3.

Definición 8.11

Un **cuantil** de una muestra, $q(f)$, es un valor para el que una fracción específica f de los valores de los datos es menor que o igual a $q(f)$.

Obviamente, un cuantil representa una estimación de una característica de una población, o más bien, la distribución teórica. La mediana de la muestra es $q(0.5)$. El percentil 75 (cuartil superior) es $q(0.75)$ y el cuartil inferior es $q(0.25)$.

Una **gráfica de cuantiles** simplemente grafica los valores de los datos en el eje vertical contra una evaluación empírica de la fracción de observaciones excedidas por los valores de los datos. Para propósitos teóricos esta fracción se calcula con

$$f_i = \frac{i - \frac{3}{4}}{n + \frac{1}{4}},$$

donde i es el orden de las observaciones cuando se clasifican de inferior a superior. En otras palabras, si denotamos las observaciones clasificadas como

$$y_{(1)} \leq y_{(2)} \leq y_{(3)} \cdots \leq y_{(n-1)} \leq y_{(n)},$$

entonces la gráfica de cuantiles describe una gráfica de $y_{(i)}$ contra f_i . En la figura 8.3 se presenta la gráfica de cuantiles para las "asas" de las latas de pintura analizadas con anterioridad.

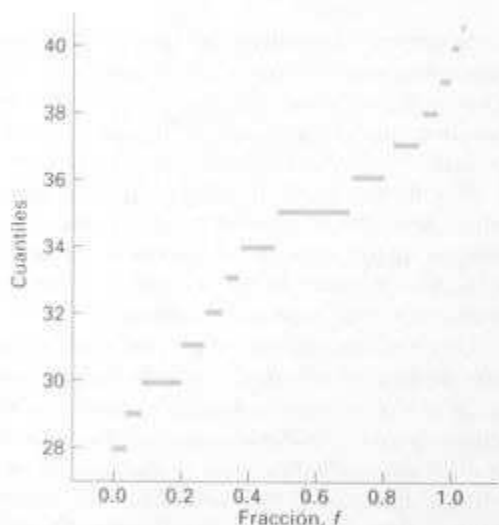


Figura 8.3 Gráfica de cuantiles para los datos de latas de pintura de la página 211.

A diferencia de la gráfica de caja y extensión, la gráfica de cuantiles realmente muestra todas las observaciones. Todos los cuantiles, incluida la mediana y los cuantiles superior e inferior, se pueden aproximar de forma visual. Por ejemplo, fácilmente observamos una mediana de 35 y un cuartil superior de alrededor de 36. Las indicaciones de agrupaciones relativamente grandes alrededor de valores específicos se indican por pendientes cercanas a cero, mientras que los datos dispersos en ciertas áreas producen pendientes más abruptas. La figura 8.3 muestra dispersión de datos de los valores 28 a 30 pero una densidad relativamente alta de 36 a 38. En los capítulos 9 y 10 proseguimos con las gráficas de cuantiles mediante la ilustración de formas útiles de comparación de distintas muestras.

Detección de desviaciones de la normalidad

Debe ser algo evidente para el lector que la detección de si un conjunto de datos viene o no de una distribución normal puede ser una herramienta importante para el analista de datos. Como indicamos al principio de esta sección, a menudo hacemos la suposición de que la totalidad o subconjuntos de las observaciones en un conjunto de datos son realizaciones de variables aleatorias normales independientes idénticamente distribuidas. Una vez más, la gráfica de diagnóstico a menudo puede agregar (con propósitos de presentación) una *prueba de la bondad del ajuste* formal de los datos. Las pruebas de bondad de ajuste se presentan en el capítulo 10. Para el lector de un artículo o reporte científico, la información de diagnóstico es mucho más clara, menos árida y quizá no aburrida. En los capítulos siguientes (capítulos 9 a 13) nos enfocamos de nuevo en los métodos de detección de desviaciones de la normalidad como un agregado de la inferencia estadística formal. La siguiente subsección proporciona una presentación e ilustración de una gráfica de diagnóstico que se llama *gráfica de cuantiles-cuantiles normales*.

Gráfica de cuantiles-Cuantiles normales

La gráfica de cuantiles-cuantiles normales toma ventaja de lo que se conoce acerca de los cuantiles de la distribución normal. La metodología incluye una gráfica de los cuantiles empíricos recién presentados contra el cuantil correspondiente de la distribución normal. Ahora bien, la expresión para un cuantil de una variable aleatoria $N(\mu, \sigma)$ es muy complicada. Sin embargo, una buena aproximación está dada por

$$q_{\mu, \sigma}(f) = \mu + \sigma \{4.91[f^{0.14} - (1-f)^{0.14}]\}.$$

La expresión entre llaves (el múltiplo de σ) es la aproximación para el cuantil correspondiente para la variable aleatoria $N(0, 1)$ es decir,

$$q_{0,1}(f) = 4.91[f^{0.14} - (1-f)^{0.14}].$$

Definición 8.12

La **gráfica de cuantiles-cuantiles normales** es una gráfica de $y_{(i)}$ (observaciones ordenadas) contra $q_{0,1}(f_i)$, donde $f_i = \frac{i - \frac{3}{8}}{n + \frac{1}{4}}$.

Una relación cercana a una línea recta sugiere que los datos provienen de una distribución normal. La intersección en el eje vertical es una estimación de la media de la población y la pendiente es una estimación de la desviación estándar σ . La figura 8.4 en la página 214 muestra una gráfica de cuantiles-cuantiles normales para los datos de las latas de pintura.

Graficación de la probabilidad normal

Nótese cómo la desviación de la normalidad se hace clara a partir de la apariencia de la gráfica. La asimetría que exhiben los datos tiene como resultado cambios en la pendiente.

Las ideas de graficación de la probabilidad se manifiestan en gráficas diferentes a la gráfica de cuantiles-cuantiles normales que se presentó aquí. Por ejemplo, se da mucha atención a la llamada **gráfica de probabilidad normal**, en la que el eje

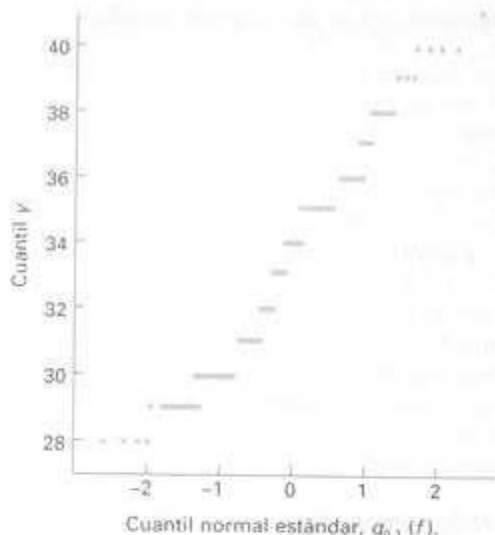


Figura 8.4 Gráfica de cuantiles-cuantiles normales para los datos de las latas de pintura.

vertical contiene f graficada en un papel especial y la escala utiliza los resultados en una línea recta cuando se grafican contra los valores de los datos ordenados. Además, una gráfica alternativa hace uso de los valores esperados de las observaciones clasificadas para la distribución normal y grafica las observaciones clasificadas contra su valor esperado, con la suposición de datos de $N(\mu, \sigma)$. Una vez más, la línea recta es el criterio gráfico que se utiliza. Continuamos para sugerir que tener como base los métodos analíticos gráficos que se desarrollan en esta sección ayuda en la ilustración de los métodos formales de distinción entre muestras distintas de datos.

Ejemplo 8.12 Considere los datos del ejercicio 23 en la página 328 en el capítulo 10. En un estudio, *Retención de nutrientes y respuesta de comunidades de macroinvertebrados a la presión de aguas residuales en un ecosistema de río*, que se llevó a cabo en el departamento de zoología del Instituto Politécnico y Universidad Estatal de Virginia, se reunieron datos sobre mediciones de densidad (número de organismos por metro cuadrado) en dos diferentes estaciones colectoras. En el capítulo 10 se dan detalles con respecto a los métodos analíticos de comparación de muestras para determinar si ambas son de la misma distribución $N(\mu, \sigma)$. Los datos son los siguientes:

Número de organismos por metro cuadrado			
Estación 1		Estación 2	
5,030	4,980	2,800	2,810
13,700	11,910	4,670	1,330
10,730	8,130	6,890	3,320
11,400	26,850	7,720	1,230
860	17,660	7,030	2,130
2,200	22,800	7,330	2,190
4,250	1,130		
15,040	1,690		

Construya una gráfica de cuantiles-cuantiles normales y extraiga conclusiones con respecto a si es razonable o no suponer que las dos muestras son de la misma distribución $N(\mu, \sigma)$.

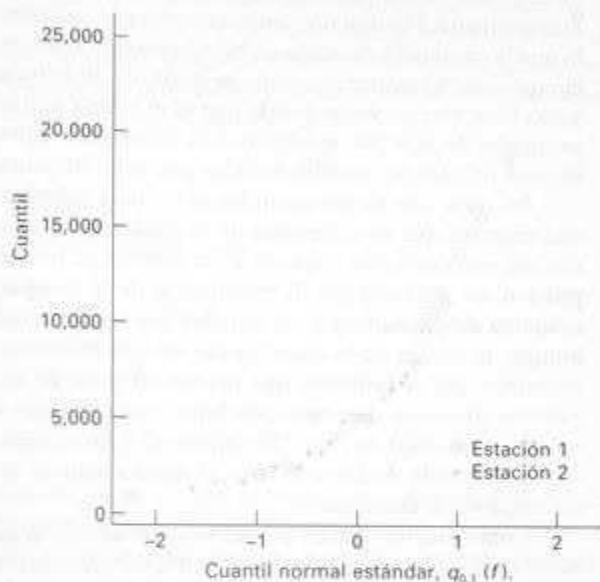


Figura 8.5 Gráfica de cuantiles-cuantiles normales para los datos de densidad del ejemplo 8.12.

SOLUCIÓN

La figura 8.5 muestra la gráfica de cuantiles-cuantiles normales para las mediciones de densidad. La gráfica muestra una apariencia que está lejos de una sola línea recta. De hecho, los datos de la estación 1 reflejan pocos valores en la cola inferior de la distribución y varios en la cola superior. El “agrupamiento” de observaciones hace parecer improbable que las dos muestras vengan de una distribución común $N(\mu, \sigma)$. ■

Aunque concentramos nuestro desarrollo e ilustración en la graficación de la probabilidad para distribuciones normales, podemos enfocarnos en cualquier distribución. Solamente necesitaríamos calcular cantidades de forma analítica para la distribución teórica en cuestión.

8.4 Distribuciones muestrales

El campo de la inferencia estadística trata básicamente con generalizaciones y predicciones. Por ejemplo, podemos afirmar, con base en las opiniones de varias personas entrevistadas en la calle, que en una próxima elección 60% de los votantes en la ciudad de Detroit favorecen a cierto candidato. En este caso tratamos con una muestra aleatoria de opiniones de una población finita muy grande. Como segunda

ilustración podemos afirmar que el costo promedio para construir una residencia en Charleston, Carolina del Sur, está entre \$130,000 y 135,000, con base en las estimaciones de tres contratistas seleccionados al azar de 30 que construyen actualmente en esta ciudad. La población que será muestreada aquí nuevamente es finita pero muy pequeña. Finalmente, consideremos una máquina despachadora de refresco en la que la cantidad promedio de bebida servida se mantiene en 240 mililitros. Un funcionario de la compañía calcula la media de 40 bebidas y obtiene $\bar{x} = 236$ mililitros y con base en este valor decide que la máquina aún sirve bebidas con un contenido promedio de $\mu = 240$ mililitros. Las 40 bebidas representan una muestra de la población infinita de posibles bebidas que esta máquina servirá.

En cada uno de los ejemplos anteriores calculamos una estadística a partir de una muestra que se selecciona de la población, y de estas estadísticas hacemos varias afirmaciones con respecto a los valores de los parámetros de la población que pueden ser ciertos o no. El funcionario de la compañía toma la decisión de que la máquina despachadora sirve bebidas con un contenido promedio de 240 mililitros, aunque la media de la muestra fue de 236 mililitros, porque sabe de la teoría de muestreo que es probable que ocurra tal valor de la muestra. De hecho, si realiza pruebas similares, digamos cada hora, esperaría que los valores de \bar{x} fluctuaran por arriba y por abajo de $\mu = 240$ mililitros. Sólo cuando el valor de \bar{x} es considerablemente diferente de 240 mililitros el funcionario de la compañía iniciará una acción para ajustar la máquina.

Como una estadística es una variable aleatoria que depende sólo de la muestra observada, debe tener una distribución de probabilidad.

Definición 8.13

La distribución de probabilidad de una estadística se llama **distribución muestral**.

La distribución de probabilidad de \bar{X} se llama **distribución muestral de la media**.

La distribución muestral de una estadística depende del tamaño de la población, el tamaño de las muestras y el método de elección de las muestras. En el resto de este capítulo estudiaremos varias de las más importantes distribuciones muestrales de las estadísticas que se utilizan con más frecuencia. Las aplicaciones de estas distribuciones muestrales a problemas de inferencia estadística se consideran en la mayoría de los capítulos restantes.

Se deben ver las distribuciones muestrales de \bar{X} y S^2 como el mecanismo a partir del cual haremos finalmente inferencias de los parámetros μ y σ^2 . La distribución muestral de \bar{X} con tamaño muestral n es la distribución que resulta cuando un experimento se lleva a cabo una y otra vez (siempre con tamaño muestral n) y resultan los diversos valores de \bar{X} . Esta distribución muestral, entonces, describe la variabilidad de los promedios muestrales alrededor de la media de la población μ . En el caso de la máquina de refresco, el conocimiento de la distribución muestral de \bar{X} da al analista el conocimiento de una discrepancia "típica" entre un valor \bar{x} observado y μ . Se aplica el mismo principio en el caso de la distribución de S^2 . La distribución muestral produce información acerca de la variabilidad de los valores de s^2 alrededor de σ^2 en experimentos que se repiten.

8.5 Distribuciones muestrales de medias

La primera distribución muestral importante a considerar es la de la media \bar{X} . Suponga que una muestra aleatoria de n observaciones se toma de una población normal con media μ y varianza σ^2 . Cada observación $X_i, i = 1, 2, \dots, n$, de la muestra aleatoria tendrá entonces la misma distribución normal que la población que se muestrea. De aquí, por la propiedad reproductiva de la distribución normal que se establece en el teorema 7.11, concluimos que

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

tiene distribución normal con media

$$\mu_{\bar{X}} = \frac{\mu + \mu + \dots + \mu}{n} = \mu$$

y varianza

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2 + \sigma^2 + \dots + \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Si tomamos muestras de una población con distribución desconocida, finita o infinita, la distribución muestral de \bar{X} aún será aproximadamente normal con media μ y varianza σ^2/n siempre que el tamaño de la muestra sea grande. Este asombroso resultado es una consecuencia inmediata del siguiente teorema, que se llama teorema del límite central. La demostración se describe en el ejercicio 11 de la página 224.

Teorema 8.2

Teorema del límite central Si \bar{X} es la media de una muestra aleatoria de tamaño n tomada de una población con media μ y varianza finita σ^2 , entonces la forma límite de la distribución de

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

conforme $n \rightarrow \infty$, es la distribución normal estándar $n(z; 0, 1)$.

La aproximación normal para \bar{X} por lo general será buena si $n \geq 30$ sin importar la forma de la población. Si $n < 30$, la aproximación es buena sólo si la población no es muy diferente de una distribución normal y, como se estableció antes, si se sabe que la población es normal, la distribución muestral de \bar{X} seguirá una distribución normal exacta, no importa qué tan pequeño sea el tamaño de las muestras.

Ejemplo 8.13 Una empresa eléctrica fabrica focos que tienen una duración que se distribuye aproximadamente en forma normal, con media de 800 horas y desviación estándar de 40 horas. Encuentre la probabilidad de que una muestra aleatoria de 16 focos tenga una vida promedio de menos de 775 horas.

SOLUCIÓN

La distribución muestral de \bar{X} será aproximadamente normal, con $\mu_{\bar{X}} = 800$ y $\sigma_{\bar{X}} = 40/\sqrt{16} = 10$. La probabilidad que se desea está dada por el área de la región sombreada de la figura 8.6.

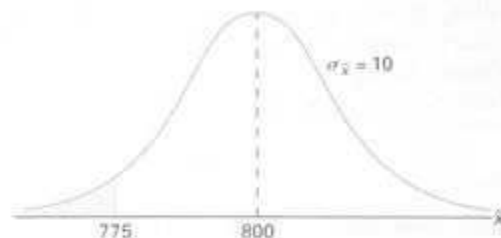


Figura 8.6 Área para el ejemplo 8.13.

En correspondencia con $\bar{x} = 775$, encontramos que

$$z = \frac{775 - 800}{10} = -2.5,$$

y por tanto

$$P(\bar{X} < 775) = P(Z < -2.5) = 0.0062. \quad \blacksquare$$

Inferencias sobre la media de la población

Una aplicación muy importante del teorema del límite central es la determinación de valores razonables de la media de la población μ . Temas como la prueba de hipótesis, estimación, control de calidad y otros hacen uso del teorema del límite central. El siguiente ejemplo ilustra el uso del teorema del límite central a este respecto, aunque la aplicación formal de los temas anteriores se deja para capítulos futuros.

Ejemplo 8.14 Un importante proceso de fabricación produce partes de componentes cilíndricos para la industria automotriz. Es importante que el proceso produzca partes que tengan una media de 5 milímetros. El ingeniero involucrado hace la conjetura de que la media de la población es de 5.0 milímetros. Se lleva a cabo un experimento en el que 100 partes elaboradas por el proceso se seleccionan al azar y se mide el diámetro de cada una de ellas. Se sabe que la desviación estándar de la población es $\sigma = 0.1$. El experimento indica un diámetro promedio de la muestra $\bar{x} = 5.027$ milímetros. ¿Esta información de la muestra parece apoyar o refutar la conjetura del ingeniero?

SOLUCIÓN

Este ejemplo refleja la clase de problema que se plantea a menudo y que se resuelve con la maquinaria de prueba de hipótesis que se introduce en los capítulos posteriores. No utilizaremos aquí el formalismo asociado con la prueba de hipótesis pero ilustraremos los principios y la lógica que se utilizan.

Si los datos apoyan o rechazan la conjetura depende de la probabilidad de que

datos similares a los que se obtuvieron en este experimento ($\bar{x} = 5.027$) pueden ocurrir con facilidad cuando de hecho $\mu = 5.0$ (figura 8.7). En otras palabras, ¿qué tan probable es que se pueda obtener $\bar{x} \geq 5.027$ con $n = 100$ si la media de la población

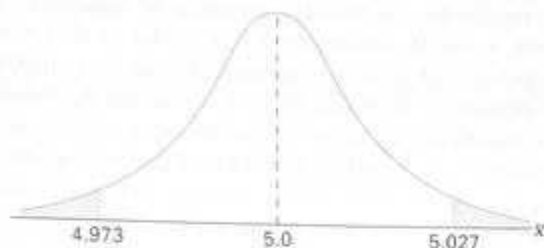


Figura 8.7 Área para el ejemplo 8.14.

es $\mu = 5.0$? Si esta probabilidad sugiere que $\bar{x} = 5.027$ no es poco razonable, la conjetura no se rechaza. Si la probabilidad es bastante baja, se puede argumentar con certidumbre que los datos no apoyan la conjetura de que $\mu = 5.0$. La probabilidad que elijamos calcular está dada por

$$\Pr[|(\bar{X} - 5)| \geq 0.027].$$

En otras palabras, si la media μ es 5, ¿cuál es la posibilidad de que \bar{X} se desvíe a lo más en 0.027 milímetros?

$$\begin{aligned} P[|(\bar{X} - 5)| \geq 0.027] &= P[(\bar{X} - 5) \geq 0.027] + P[(\bar{X} - 5) \leq -0.027] \\ &= 2P\left(\frac{\bar{X} - 5}{0.1/\sqrt{100}} \geq 2.7\right) \approx 2.7. \end{aligned}$$

Aquí simplemente estandarizamos \bar{X} de acuerdo con el teorema del límite central.

Si la conjetura $\mu = 5.0$ es cierta, $\frac{\bar{X} - 5.0}{0.1/\sqrt{100}}$ es $N(0, 1)$. Así

$$2P\left[\frac{\bar{X} - 5.0}{0.1/\sqrt{100}} \geq 2.7\right] = 2P[Z \geq 2.7] = 2(0.0035) = 0.007.$$

De esta manera se experimentaría por casualidad una \bar{x} que está a 0.027 milímetros de la media en sólo siete de 1000 experimentos. Como resultado, este experimento con $\bar{x} = 5.027$ ciertamente no proporciona evidencia que apoye la conjetura de que $\mu = 5.0$. ■

Distribución muestral de la diferencia entre dos promedios

La ilustración del ejemplo 8.14 trata con nociones de inferencia estadística sobre una sola media μ . El ingeniero estaba interesado en apoyar una conjetura con respecto a una sola media de la población. Una aplicación mucho más importante incluye dos poblaciones. Un científico o ingeniero se interesa en un experimento

comparativo el que se cotejan dos métodos de producción, 1 y 2. La base para esta comparación es $\mu_1 - \mu_2$, la diferencia de las medias de las poblaciones.

Suponga que tenemos dos poblaciones, la primera con media μ_1 y varianza σ_1^2 , y la segunda con media μ_2 y varianza σ_2^2 . Representemos con la estadística \bar{X}_1 la media de una muestra aleatoria de tamaño n_1 seleccionada de la primera población, y con la estadística \bar{X}_2 la media de una muestra aleatoria seleccionada de la segunda población, independiente de la muestra de la primera población. ¿Qué podemos decir acerca de la distribución de muestreo de la diferencia $\bar{X}_1 - \bar{X}_2$ para muestras repetidas de tamaño n_1 y n_2 ? De acuerdo con el teorema 8.2, las variables \bar{X}_1 y \bar{X}_2 están distribuidas aproximadamente de forma normal con medias μ_1 y μ_2 y varianzas σ_1^2/n_1 y σ_2^2/n_2 , respectivamente. Esta aproximación mejora conforme n_1 y n_2 aumentan. Al elegir muestras independientes de las dos poblaciones las variables \bar{X}_1 y \bar{X}_2 serán independientes, y entonces con el uso del teorema 7.11, con $a_1 = 1$ y $a_2 = -1$, podemos concluir que $\bar{X}_1 - \bar{X}_2$ está distribuida aproximadamente de forma normal con media

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

y varianza

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Teorema 8.3

Si se extraen al azar muestras independientes de tamaño n_1 y n_2 de dos poblaciones, discretas o continuas, con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, entonces la distribución muestral de las diferencias de las medias, $\bar{X}_1 - \bar{X}_2$, está distribuida aproximadamente de forma normal con media y varianza dadas por

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 \quad \text{y} \quad \sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

De aquí

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$$

es aproximadamente una variable normal estándar.

Si n_1 y n_2 son mayores que o iguales a 30, la aproximación normal para la distribución de $\bar{X}_1 - \bar{X}_2$ es muy buena sin importar las formas de las dos poblaciones. Sin embargo, aun cuando n_1 y n_2 sean menores que 30, la aproximación normal es razonablemente buena excepto cuando las poblaciones no son definitivamente normales. Por supuesto, si ambas poblaciones son normales, entonces $\bar{X}_1 - \bar{X}_2$ tiene una distribución normal sin importar cuáles son los tamaños de n_1 y n_2 .

Ejemplo 8.15 Se llevan a cabo dos experimentos independientes en los que se comparan dos tipos diferentes de pintura. Se pintan 18 especímenes con el tipo A y

en cada uno se registra el tiempo de secado en horas. Lo mismo se hace con el tipo B . Se sabe que las desviaciones estándar de la población son ambas 1.0.

Suponga que el tiempo medio de secado es igual para los dos tipos de pintura, encuentre $\Pr(\bar{X}_A - \bar{X}_B > 1.0)$, donde \bar{X}_A y \bar{X}_B son los tiempos promedio de secado para muestras de tamaño $n_A = n_B = 18$.

SOLUCIÓN

De la distribución de muestreo de $\bar{X}_A - \bar{X}_B$, sabemos que la distribución es normal con media

$$\mu_{\bar{X}_A - \bar{X}_B} = \mu_A - \mu_B = 0$$

y varianza

$$\sigma_{\bar{X}_A - \bar{X}_B}^2 = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} = \frac{1}{18} + \frac{1}{18} = \frac{1}{9}$$

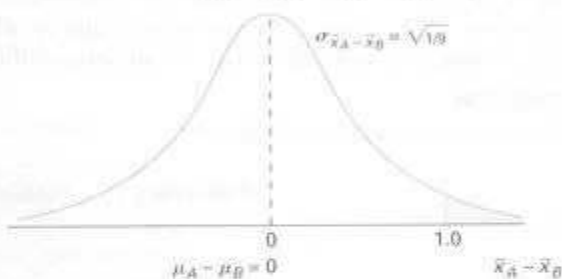


Figura 8.8 Área para el ejemplo 8.15.

La probabilidad que se desea está dada por la región sombreada en la figura 8.8. En correspondencia con el valor $\bar{X}_A - \bar{X}_B = 1.0$, tenemos

$$Z = \frac{1 - (\mu_A - \mu_B)}{\sqrt{1/9}} = \frac{1 - 0}{\sqrt{1/9}} = 3.0;$$

por lo que

$$\Pr(z > 3.0) = 1 - \Pr(z < 3.0) = 1 - 0.9987 = 0.0013.$$

¿Qué podemos aprender del resultado anterior? La maquinaria de cálculo se basa en la suposición de que $\mu_A = \mu_B$. Suponga, sin embargo, que el experimento realmente se lleva a cabo con el propósito de realizar una inferencia con respecto a la igualdad de μ_A y μ_B , los tiempos medios de secado de las dos poblaciones. Si los dos promedios difieren por una hora (o más), esto claramente es una evidencia que nos conducirá a concluir que el tiempo medio de secado de la población no es igual para los dos tipos de pintura. Por otro lado, suponga que la diferencia en los dos promedios muestrales es más pequeña que, digamos, 15 minutos. Si $\mu_A = \mu_B$,

$$\begin{aligned} \Pr\{(\bar{X}_A - \bar{X}_B) > 0.25 \text{ hora}\} &= \Pr\left(\frac{\bar{X}_A - \bar{X}_B - 0}{\sqrt{1/9}} > \frac{3}{4}\right) \\ &= \Pr\left(Z > \frac{3}{4}\right) = 1 - \Pr(Z < 0.75) \\ &= 1 - 0.7734 = 0.2266. \end{aligned}$$

Como esta probabilidad no es excesivamente baja, concluimos que una diferencia de 15 minutos en la media muestral puede ocurrir por casualidad (es decir, sucede con frecuencia aunque $\mu_A = \mu_B$). Como resultado, este tipo de diferencia en el tiempo promedio de secado ciertamente *no es una señal clara* de que $\mu_A \neq \mu_B$.

Como indicamos al principio, en los capítulos siguientes se proporcionará más formalismo con respecto a éste y a otros tipos de inferencia estadística (por ejemplo, la prueba de hipótesis). El teorema del límite central y las distribuciones de muestreo que se presentan en las siguientes tres secciones también jugarán un papel vital. ■

Ejemplo 8.16 Los cinescopios para televisión del fabricante A tienen una duración media de 6.5 años y una desviación estándar de 0.9 años, mientras que los del fabricante B tienen una duración media de 6.0 años y una desviación estándar de 0.8 años. ¿Cuál es la probabilidad de que una muestra aleatoria de 36 cinescopios del fabricante A tengan una duración media que sea al menos de un año más que la duración media de una muestra de 49 cinescopios del fabricante B?

SOLUCIÓN

Se nos da la siguiente información:

Población 1	Población 2
$\mu_1 = 6.5$	$\mu_2 = 6.0$
$\sigma_1 = 0.9$	$\sigma_2 = 0.8$
$n_1 = 36$	$n_2 = 49$

Si utilizamos el teorema 8.3, la distribución muestral de $\bar{X}_1 - \bar{X}_2$ será aproximadamente normal y tendrá una media y una desviación estándar de

$$\mu_{\bar{X}_1 - \bar{X}_2} = 6.5 - 6.0 = 0.5 \quad \text{y} \quad \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{0.81}{36} + \frac{0.64}{49}} = 0.189.$$

La probabilidad de que la media de 36 cinescopios del fabricante A sea al menos un año mayor que la media de 49 cinescopios del fabricante B está dada por el área de la región sombreada de la figura 8.9. Con respecto al valor $\bar{x}_1 - \bar{x}_2 = 1.0$, encontramos que

$$z = \frac{1.0 - 0.5}{0.189} = 2.65,$$

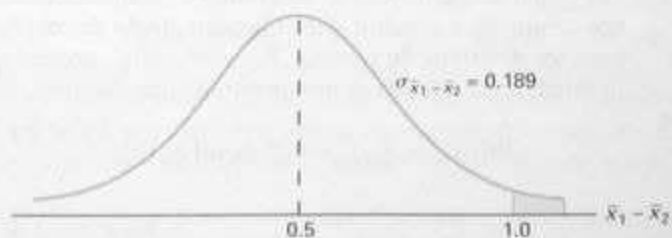


Figura 8.9 Área para el ejemplo 8.16.

y de aquí

$$P(\bar{X}_1 - \bar{X}_2 \geq 1.0) = P(Z > 2.65) = 1 - P(Z < 2.65) = 1 - 0.9960 \\ = 0.0040. \blacksquare$$

Ejercicios

1. Si se extraen todas las muestras posibles de tamaño 16 de una población normal con media igual a 50 y desviación estándar igual a 5, ¿cuál es la probabilidad de que una media muestral \bar{X} caiga en el intervalo que va de $\mu_X - 1.9\sigma_{\bar{X}}$ a $\mu_X + 0.4\sigma_{\bar{X}}$? Suponga que las medias muestrales se pueden medir con cualquier grado de precisión.

2. Dada la población uniforme discreta

$$f(x) = \begin{cases} \frac{1}{3}, & x = 2, 4, 6 \\ 0, & \text{en cualquier otro caso,} \end{cases}$$

encuentre la probabilidad de que una muestra aleatoria de tamaño 54, seleccionada con reemplazo, dé una media muestral mayor que 4.1 pero menor que 4.4. Suponga que las medias se miden al décimo más cercano.

3. Se fabrica cierto tipo de hilo con una resistencia a la tracción media de 78.3 kilogramos y una desviación estándar de 5.6 kilogramos. ¿Cómo cambia la varianza de la media muestral cuando el tamaño de la muestra

(a) aumenta de 64 a 196?
 (b) disminuye de 784 a 49?

4. Si la desviación estándar de la media para la distribución muestral de muestras aleatorias de tamaño 36 de una población grande o infinita es 2, ¿qué tan grande debe ser el tamaño de la muestra si la desviación estándar se reduce a 1.2?

5. Una máquina de refrescos se ajusta para que la cantidad de bebida que sirve promedie 240 mililitros con una desviación estándar de 15 mililitros. La máquina se verifica periódicamente tomando una muestra de 40 bebidas y se calcula el contenido promedio. Si la media de las 40 bebidas es un valor dentro del intervalo $\mu_{\bar{X}} \pm 2\sigma_{\bar{X}}$, se piensa que la máquina opera satisfactoriamente; de otra forma, se ajusta. En la sección 8.4, el funcionario de la compañía encuentra que la media de 40 bebidas es $\bar{x} = 236$ mililitros y concluye que la máquina no necesita un ajuste. ¿Esta fue una decisión razonable?

6. Las estaturas de 1000 estudiantes están distribuidas aproximadamente de forma normal con una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros. Si se extraen 200 muestras aleatorias

de tamaño 25 de esta población y las medias se registran al décimo de centímetro más cercano, determine (a) la media y la desviación estándar de la distribución muestral de \bar{X} ;

(b) el número de las medias muestrales que caen entre 172.5 y 175.8 centímetros inclusive;
 (c) el número de medias muestrales que caen por debajo de 172.0 centímetros.

7. La variable aleatoria X , que representa el número de cerezas en una empanada, tiene la siguiente distribución de probabilidad:

x	4	5	6	7
$P(X = x)$	0.2	0.4	0.3	0.1

(a) Encuentre la media μ y la varianza σ^2 de X .
 (b) Encuentre la media $\mu_{\bar{X}}$ y la varianza $\sigma_{\bar{X}}^2$ de la media \bar{X} para muestras aleatorias de 36 empanadas de cereza.
 (c) Encuentre la probabilidad de que el número promedio de cerezas en 36 empanadas sea menor que 5.5.

8. Si cierta máquina fabrica resistores eléctricos que tienen una resistencia media de 40 ohmios y una desviación estándar de dos ohmios, ¿cuál es la probabilidad de que una muestra aleatoria de 36 de estos resistores tenga una resistencia combinada de más de 1458 ohmios?

9. La vida media de una máquina para hacer pasta es de siete años, con una desviación estándar de un año. Suponga que las vidas de estas máquinas siguen aproximadamente una distribución normal, encuentre

(a) la probabilidad de que la vida media de una muestra aleatoria de nueve de estas máquinas caiga entre 6.4 y 7.2 años;

(b) el valor de \bar{x} a la derecha del cual caería el 15% de las medias calculadas de muestras aleatorias de tamaño nueve.

10. El tiempo que el cajero de un banco con servicio en el automóvil atiende a un cliente es una variable aleatoria con media $\mu = 3.2$ minutos y una desviación estándar $\sigma = 1.6$ minutos. Si se observa una

- muestra aleatoria de 64 clientes, encuentre la probabilidad de que su tiempo medio con el cajero sea
- a lo más 2.7 minutos;
 - más de 3.5 minutos;
 - al menos 3.2 minutos pero menos de 3.4 minutos.

11. Teorema del límite central

- (a) Con el uso de los teoremas 7.8, 7.9 y 7.10, muestre que

$$M_{(\bar{X}-\mu)/(\sigma/\sqrt{n})}(t) = e^{-t^2/2} \left[M_X\left(\frac{t}{\sigma/\sqrt{n}}\right) \right]^n,$$

donde \bar{X} es la media de una muestra aleatoria de tamaño n de una población $f(x)$ con media μ y varianza σ^2 , y de aquí

$$\ln M_{(\bar{X}-\mu)/(\sigma/\sqrt{n})}(t) = -\frac{t^2}{2} + n \ln M_X\left(\frac{t}{\sigma/\sqrt{n}}\right).$$

- (b) Use el resultado del ejercicio 22 que está en la página 197 para expandir $M_X(t/\sigma/\sqrt{n})$ como una serie infinita en potencias de t . Podemos escribir entonces $M_X(t/\sigma/\sqrt{n}) = 1 + v$, donde v es una serie infinita.
- (c) Suponga que n es suficientemente grande, expanda $\ln(1+v)$ en una serie de Maclaurin y después muestre que

$$\lim_{n \rightarrow \infty} \ln M_{(\bar{X}-\mu)/(\sigma/\sqrt{n})}(t) = -\frac{t^2}{2}$$

y de aquí

$$\lim_{n \rightarrow \infty} M_{(\bar{X}-\mu)/(\sigma/\sqrt{n})}(t) = e^{-t^2/2}.$$

12. Se toma una muestra aleatoria de tamaño 25 de una población normal que tiene una media de 80 y una desviación estándar de 5. Una segunda muestra aleatoria de tamaño 36 se toma de una población normal diferente que tiene una media de 75 y una desviación estándar de 3. Encuentre la probabilidad de que la media muestral calculada de las 25 mediciones exceda la media muestral calculada de las 36 mediciones por al menos 3.4 pero en menos de 5.9. Suponga que las medias se miden al décimo más cercano.

13. La distribución de alturas de cierta raza de perros terrier tiene una altura media de 72 centímetros y una

desviación estándar de 10 centímetros, mientras que la distribución de alturas de cierta raza de poodles tiene una altura media de 28 centímetros con una desviación estándar de 5 centímetros. Suponga que las medias muestrales se pueden medir con cualquier grado de precisión, encuentre la probabilidad de que la media muestral para una muestra aleatoria de alturas de 64 terriers exceda la media muestral para una muestra aleatoria de alturas de 100 poodles por cuando mucho 44.2 centímetros.

14. La calificación media de estudiantes de primer año en un examen de aptitud en cierta universidad es 540, con una desviación estándar de 50. ¿Cuál es la probabilidad de que dos grupos de estudiantes seleccionados al azar, que consisten en 32 y 50 estudiantes, respectivamente, difieran en sus calificaciones medias por
- más de 20 puntos?
 - una cantidad entre 5 y 10 puntos?
- Suponga que las medias se miden con cualquier grado de precisión.

15. Construya una gráfica de cuantiles para estos datos. Las duraciones en horas de cincuenta lámparas incandescentes de 40 watts a 110 voltios enfriadas internamente tomadas de pruebas en condiciones adversas:

919	1196	785	1126	936	918	1156	920	948
1067	1092	1162	1170	929	950	905	972	1035
1045	855	1195	1195	1340	1122	938	970	1257
956	1102	1157	978	832	1009	1157	1151	1009
765	958	902	1022	1333	811	1217	1085	896
958	1311	1037	702	923				

16. Construya una gráfica de cuantiles-cuantiles normales de estos datos. Diámetros de 36 cabezas de remache en 1/100 de pulgada:

6.72	6.77	6.82	6.70	6.78	6.70	6.62	6.75	6.66
6.66	6.64	6.76	6.73	6.80	6.72	6.76	6.76	6.68
6.66	6.62	6.72	6.76	6.70	6.78	6.76	6.67	6.70
6.72	6.74	6.81	6.79	6.78	6.66	6.76	6.76	6.72

8.6 Distribución muestral de S^2

En la sección anterior aprendimos acerca de la distribución de muestreo de \bar{X} . El teorema del límite central nos permitió hacer uso del hecho de que

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

tiende a $N(0, 1)$ conforme crece el tamaño de la muestra. Los ejemplos 8.15 y 8.16 ilustran la aplicación del teorema del límite central, que nos permite extraer conclusiones acerca de la media de la población o la diferencia en dos medias de población. Las *distribuciones muestrales de estadísticas importantes* nos permiten conocer información sobre los parámetros. Por lo general, los parámetros son la contraparte de la estadística en cuestión. Si un ingeniero se interesa en la resistencia media de la población de cierto tipo de resistor, la distribución muestral de \bar{X} se explotará una vez que se reúna la información de la muestra. Por otro lado, si se estudia la variabilidad en el resistor, claramente la distribución muestral de S^2 se utilizará para conocer la contraparte paramétrica, la varianza de la población σ^2 .

Si se extrae una muestra aleatoria de tamaño n de una población normal con media μ y varianza σ^2 , y se calcula la varianza muestral S^2 obtenemos un valor de la estadística S^2 . Procederemos a considerar la distribución de la estadística $(n-1)S^2/\sigma^2$.

Mediante la suma y resta de la media muestral \bar{X} , es fácil ver que

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n [(X_i - \bar{X}) + (\bar{X} - \mu)]^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + 2(\bar{X} - \mu) \sum_{i=1}^n (X_i - \bar{X}) \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2. \end{aligned}$$

Al dividir cada término de la igualdad entre σ^2 y sustituir $(n-1)S^2$ por $\sum_{i=1}^n (X_i - \bar{X})^2$, obtenemos

$$\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} + \frac{(\bar{X} - \mu)^2}{\sigma^2/n}.$$

Ahora, de acuerdo con el corolario del teorema 7.12 sabemos que

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}$$

es una variable aleatoria ji cuadrada con n grados de libertad. El segundo término de la derecha de la igualdad es el cuadrado de una variable normal estándar, pues \bar{X} es una variable aleatoria normal con media $\mu_{\bar{X}} = \mu$ y varianza $\sigma_{\bar{X}}^2 = \sigma^2/n$. Por tanto, podemos concluir del ejemplo 8.1 que $(\bar{X} - \mu)^2/(\sigma^2/n)$ es una variable aleatoria ji cuadrada con 1 grado de libertad. Con el uso de técnicas avanzadas que están fuera del alcance de este libro, se puede mostrar también que las dos variables ji cuadradas

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \quad \text{y} \quad \frac{(\bar{X} - \mu)^2}{(\sigma^2/n)}$$

son independientes. Debido a la propiedad reproductiva de las variables aleatorias ji cuadrada, que se establece en el teorema 7.12, parece razonable suponer que $(n-1)S^2/\sigma^2$ también es una variable aleatoria ji cuadrada con $v = n - 1$ grados de libertad. Establezcamos este resultado en el siguiente teorema.

Teorema 8.4

Si S^2 es la varianza de una muestra aleatoria de tamaño n que se toma de una población normal que tiene la varianza σ^2 , entonces la estadística

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}$$

tiene una distribución ji cuadrada con $v = n - 1$ grados de libertad.

Los valores de la variable aleatoria χ^2 se calculan de cada muestra mediante la fórmula

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

La probabilidad de que una muestra aleatoria produzca un valor χ^2 mayor que algún valor específico es igual al área bajo la curva a la derecha de este valor. Se acostumbra representar con χ_α^2 el valor χ^2 por arriba del que encontramos un área de α . Esto se ilustra mediante la región sombreada en la figura 8.10.

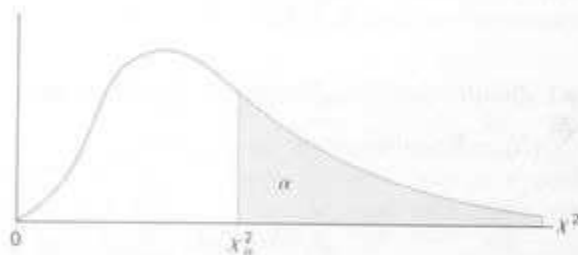


Figura 8.10 Valores tabulados de la distribución ji cuadrada.

La tabla A.5 da los valores de χ_α^2 para diversos valores de α y v . Las áreas, α , son los encabezados de las columnas; los grados de libertad, v , se dan en la columna izquierda, y las entradas de la tabla son los valores χ^2 . De aquí, el valor χ^2 con 7 grados de libertad, que deja un área de 0.05 a la derecha, es $\chi_{0.05}^2 = 14.067$. Debido a la falta de simetría, debemos usar también las tablas para encontrar $\chi_{0.95}^2 = 2.167$ para $v = 7$.

Exactamente 95% de una distribución ji cuadrada yace entre $\chi_{0.975}^2$ y $\chi_{0.025}^2$. Un valor χ^2 que cae a la derecha de $\chi_{0.025}^2$ no es probable que ocurra a menos que nuestro valor supuesto de σ^2 sea demasiado pequeño. De manera similar, un valor χ^2 que cae a la izquierda de $\chi_{0.975}^2$ es improbable a menos que nuestro valor supuesto de σ^2 sea demasiado grande. En otras palabras, es posible tener un valor χ^2 a la izquierda de $\chi_{0.975}^2$ o a la derecha de $\chi_{0.025}^2$ cuando σ^2 es correcta, pero si esto debe ocurrir, es más probable que el valor supuesto de σ^2 sea un error.

Ejemplo 8.17 Un fabricante de baterías para auto garantiza que sus baterías durarán, en promedio, tres años con una desviación estándar de un año. Si cinco de estas baterías tienen duraciones de 1.9, 2.4, 3.0, 3.5 y 4.2 años, ¿el fabricante aún está convencido de que sus baterías tienen una desviación estándar de un año? Suponga que la duración de la batería sigue una distribución normal.

SOLUCIÓN

Encontramos primero la varianza de la muestra:

$$s^2 = \frac{(5)(48.26) - (15)^2}{(5)(4)} = 0.815.$$

Entonces

$$\chi^2 = \frac{(4)(0.815)}{1} = 3.26$$

es un valor de una distribución ji cuadrada con 4 grados de libertad. Como 95% de los valores χ^2 con 4 grados de libertad caen entre 0.484 y 11.143, el valor calculado con $\sigma^2 = 1$ es razonable y por tanto el fabricante no tiene razón para sospechar que la desviación estándar es diferente a un año. ■

Grados de libertad como medición de la información muestral

El lector puede obtener algunos conocimientos al considerar el teorema 8.4 y el corolario que sigue del teorema 7.12 en la sección 7.3. Sabemos que con las condiciones del teorema 7.12, a saber, una muestra aleatoria que se toma de una distribución normal, que la variable aleatoria

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2}$$

tiene una distribución χ^2 con n grados de libertad. Nótese ahora el teorema 8.4, que indica que con las mismas condiciones del teorema 7.12, la variable aleatoria

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2}$$

tiene una distribución χ^2 con $n - 1$ grados de libertad. El lector puede recordar que el término *grados de libertad*, que se utiliza en este contexto idéntico, se presenta en el capítulo 1.

Como indicamos al principio, no se dará la demostración del teorema 8.4. Sin embargo, el lector puede ver el teorema 8.4 como una indicación de que cuando no se conoce μ y se considera la distribución de

$$\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2},$$

hay 1 grado de libertad menos, o se pierde un grado de libertad en la estimación de μ (es decir, cuando μ se reemplaza por \bar{x}). En otras palabras, hay n grados de libertad o piezas de información independientes en la muestra aleatoria de la distribu-

ción normal. Cuando los datos (los valores en la muestra) se utilizan para calcular la media, hay 1 grado de libertad menos en la información que se utiliza para estimar σ^2 .

8.7 Distribución t

En la sección 8.6 se presentó la utilidad del teorema del límite central. Sus aplicaciones giran alrededor de las inferencias sobre una media de la población o la diferencia entre dos medias de población. El uso del teorema del límite central y la distribución normal es ciertamente útil en este contexto. Sin embargo, se supuso que la desviación estándar de la población *se conoce*. Esta suposición puede no ser irrazonable en situaciones donde el ingeniero está bastante familiarizado con el sistema o proceso. Sin embargo, en muchos escenarios experimentales el conocimiento de σ ciertamente no es más razonable que el conocimiento de la media de la población μ . A menudo, de hecho, una estimación de σ la debe proporcionar la misma información muestral que produce el promedio muestral \bar{x} . Como resultado, una estadística natural a considerar para tratar con las inferencias sobre μ es:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

puesto que S es el análogo de la muestra para σ . Si el tamaño de la muestra es pequeño, los valores de S^2 fluctúan de forma considerable de una muestra a otra (véase el ejercicio 7 de la página 236) y la distribución de T se desvía de forma apreciable de la de una distribución normal estándar.

Si el tamaño de la muestra es suficientemente grande, digamos $n \geq 30$, la distribución de T no difiere de manera considerable de la normal estándar. Sin embargo, para $n < 30$, es útil tratar con la distribución exacta de T . Para desarrollar la distribución muestral de T supondremos que nuestra muestra aleatoria se seleccionó de una población normal. Podemos escribir entonces

$$T = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{S^2/\sigma^2}} = \frac{Z}{\sqrt{V/(n-1)}}$$

donde

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

tiene la distribución normal estándar, y

$$V = \frac{(n-1)S^2}{\sigma^2}$$

tiene una distribución ji cuadrada con $v = n - 1$ grados de libertad. Al muestrear de poblaciones normales, se puede mostrar que \bar{X} y S^2 son independientes, y en consecuencia lo son Z y V . El siguiente teorema da la función de densidad de la variable aleatoria T .

Teorema 8.5

Sea Z una variable aleatoria normal estándar y V una variable aleatoria ji cuadrada con v grados de libertad. Si Z y V son independientes, entonces la distribución de la variable aleatoria T , donde

$$T = \frac{Z}{\sqrt{V/v}},$$

está dada por

$$h(t) = \frac{\Gamma[(v+1)/2]}{\Gamma(v/2)\sqrt{\pi v}} \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2}, \quad -\infty < t < \infty.$$

Esta se conoce como la **distribución t** con v grados de libertad.

Corolario

Sean X_1, X_2, \dots, X_n variables aleatorias independientes que son todas normales con media μ y desviación estándar σ . Sea

$$\bar{X} = \sum_{i=1}^n \frac{X_i}{n} \quad \text{y} \quad s^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}.$$

Entonces la variable aleatoria $T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$ tiene una distribución t con $v = n - 1$ grados de libertad.

La distribución de probabilidad de T se publicó por primera vez en 1908 en un artículo de W. S. Gosset. En esa época, Gosset era empleado de una cervecería irlandesa que desaprobaba la publicación de investigaciones de sus empleados. Para evadir esta prohibición, publicó su trabajo en secreto bajo el nombre "Student". En consecuencia, la distribución de T normalmente se llama distribución t de Student, o simplemente distribución t . Para derivar la ecuación de esta distribución, Gosset supone que las muestras se seleccionan de una población normal. Aunque esto parecería una suposición muy restrictiva, se puede mostrar que las poblaciones no normales que poseen distribuciones en forma casi de campana aún proporcionan valores de T que se aproximan muy de cerca a la distribución t .

La distribución de T es similar a la distribución de Z , pues ambas son simétricas alrededor de una media de cero. Ambas distribuciones tienen forma de campana, pero la distribución t es más variable, debido al hecho de que los valores T dependen de las fluctuaciones de dos cantidades, \bar{X} y S^2 , mientras que los valores Z dependen sólo de los cambios de \bar{X} de una muestra a otra. La distribución de T difiere de la de Z en que la varianza de T depende del tamaño de la muestra n y siempre es mayor que 1. Únicamente cuando el tamaño de la muestra $n \rightarrow \infty$ las dos distribuciones serán la misma. En la figura 8.11 mostramos la relación entre una distribución normal estándar ($v = \infty$) y las distribuciones t con 2 y 5 grados de libertad.

Se acostumbra representar con t_α el valor t por arriba del cual encontramos un área igual a α . De aquí, el valor t con 10 grados de libertad que deja un área de 0.025 a la derecha es $t = 2.228$. Como la distribución t es simétrica alrededor de una me-

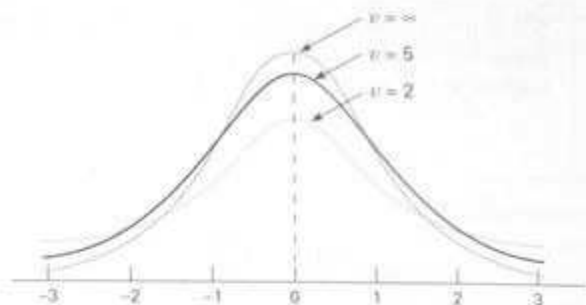


Figura 8.11 Curvas de la distribución t para $v = 2, 5$ y ∞ .

dia de cero, tenemos $t_{1-\alpha} = -t_{\alpha}$; es decir, el valor t que deja un área de $1 - \alpha$ a la derecha y por tanto un área de α a la izquierda, es igual al valor t negativo que deja un área de α en la cola derecha de la distribución (véase la figura 8.12). Esto es, $t_{0.95} = -t_{0.05}$, $t_{0.99} = -t_{0.01}$, etcétera.

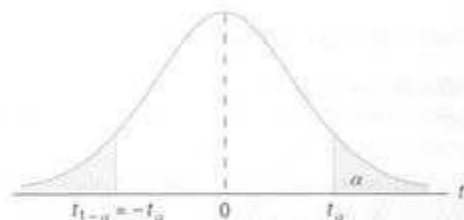


Figura 8.12 Propiedad de simetría de la distribución t .

Ejemplo 8.18 El valor t con $v = 14$ grados de libertad que deja un área de 0.025 a la izquierda, y por tanto un área de 0.975 a la derecha, es

$$t_{0.975} = -t_{0.025} = -2.145. \quad \blacksquare$$

Ejemplo 8.19 Encuentre $P(-t_{0.025} < T < t_{0.05})$.

SOLUCIÓN

Como $t_{0.05}$ deja un área de 0.05 a la derecha, y $-t_{0.025}$ deja un área de 0.025 a la izquierda, encontramos un área total de

$$1 - 0.05 - 0.025 = 0.925$$

entre $-t_{0.025}$ y $t_{0.05}$. De aquí

$$P(-t_{0.025} < T < t_{0.05}) = 0.925. \quad \blacksquare$$

Ejemplo 8.20 Encuentre k tal que $P(k < T < -1.761) = 0.045$, para una muestra aleatoria de tamaño 15 que se selecciona de una distribución normal.

SOLUCIÓN

De la tabla A.4 notamos que 1.761 corresponde a $t_{0.05}$ cuando $v = 14$. Por tanto, $-t_{0.05} = -1.761$. Como k en el enunciado de probabilidad original está a la izquierda de $-t_{0.05} = -1.761$, sea $k = -t_{\alpha}$. Entonces, de la figura 8.13, tenemos

$$0.045 = 0.05 - \alpha \quad \text{o} \quad \alpha = 0.005.$$

Por ello, de la tabla A.4 con $v = 14$,

$$k = -t_{0.005} = -2.977 \quad \text{y} \quad P(-2.977 < T < -1.761) = 0.045.$$

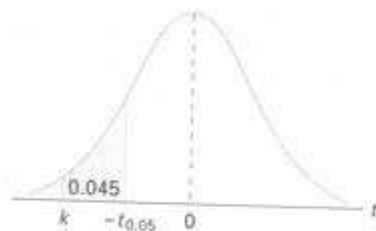


Figura 8.13 Valores t para el ejemplo 8.20.

Exactamente 95% de los valores de una distribución t con $v = n - 1$ grados de libertad caen entre $-t_{0.025}$ y $t_{0.025}$. Por supuesto, hay otros valores t que contienen 95% de la distribución, como $-t_{0.02}$ y $t_{0.03}$, pero estos valores no aparecen en la tabla A.4, y además el intervalo más corto posible se obtiene al elegir valores t que dejen exactamente la misma área en las dos colas de nuestra distribución. Un valor t que cae por debajo de $-t_{0.025}$ o por arriba de $t_{0.025}$ tiende a hacernos creer que ha ocurrido un evento muy raro o quizá que nuestra suposición acerca de μ es un error. Si esto ocurre, tomaremos la última decisión y afirmaremos que nuestro valor supuesto de μ es erróneo. De hecho, un valor t que cae por debajo de $-t_{0.01}$ o por arriba de $t_{0.01}$ proporcionaría incluso fuerte evidencia de que nuestro valor supuesto de μ es bastante improbable. En el capítulo 10 se tratarán procedimientos generales para probar afirmaciones con respecto al valor del parámetro μ . El siguiente ejemplo ilustra un aspecto preliminar de las bases de estos procedimientos. ■

Ejemplo 8.21 Un ingeniero químico afirma que el rendimiento medio de la población de cierto proceso en lotes es 500 gramos por milímetro de materia prima. Para verificar esta afirmación muestrea 25 lotes cada mes. Si el valor t calculado cae entre $-t_{0.05}$ y $t_{0.05}$, queda satisfecho con su afirmación. ¿Qué conclusión extraería de una muestra que tiene una media $\bar{x} = 518$ gramos por milímetro y una desviación estándar muestral $s = 40$ gramos? Suponga que la distribución de rendimientos es aproximadamente normal.

SOLUCIÓN

De la tabla A.4 encontramos que $t_{0.05} = 1.711$ para 24 grados de libertad. Por tanto, el fabricante queda satisfecho con esta afirmación si una muestra de 25 lotes rinde un valor t entre -1.711 y 1.711 . Si $\mu = 500$, entonces

$$t = \frac{518 - 500}{40/\sqrt{25}} = 2.25,$$

un valor muy por arriba de 1.711. La probabilidad de obtener un valor t , con $v = 24$, igual o mayor que 2.25 es aproximadamente 0.02. Si $\mu > 500$, el valor de t calculado de la muestra sería más razonable. De aquí que es probable que el fabricante concluya que el proceso produce un mejor producto del que piensa. ■

¿Para qué se utiliza la distribución T ?

La distribución t se usa de manera extensa en problemas que tienen que ver con inferencia acerca de la media de la población (como se ilustra en el ejemplo 8.21) o en problemas que implican muestras comparativas (es decir, en casos donde se trata de determinar si las medias de dos muestras son significativamente diferentes). El lector debe notar que el uso de la distribución t para la estadística

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

requiere que X_1, X_2, \dots, X_n sea normal. El uso de la distribución t y la consideración del tamaño de la muestra no se relacionan con el teorema del límite central. El uso de la distribución normal estándar en lugar de T para $n \geq 30$ solamente implica que S es un estimador suficientemente bueno de σ en este caso. En los capítulos que siguen la distribución t encuentra un uso extenso.

8.8 Distribución F

Motivamos la distribución t en parte sobre la base de la aplicación a problemas en los que hay muestreo comparativo (es decir, comparación entre dos medias muestrales). Algunos de nuestros ejemplos en los capítulos siguientes proporcionarán el formalismo. Un ingeniero químico reúne datos de dos catalizadores. Un biólogo colecta datos sobre dos medios de crecimiento. Un químico reúne datos sobre dos métodos de material de recubrimiento para suprimir la corrosión. Aunque es de interés que la información muestral arroje luz sobre dos medias de poblaciones, es frecuente el caso en el que una comparación, en algún sentido, de la variabilidad sea igualmente importante, si no es que más. La distribución F encuentra enorme aplicación en la comparación de varianzas muestrales. Las aplicaciones de la distribución F se encuentran en problemas que involucran dos o más muestras.

La estadística F se define como la razón de dos variables aleatorias j cuadrada independientes, dividida cada una entre su número de grados de libertad. De aquí, podemos escribir

$$F = \frac{U/v_1}{V/v_2},$$

donde U y V son variables aleatorias independientes que tienen distribuciones j cuadradas con v_1 y v_2 grados de libertad, respectivamente. Estableceremos ahora la distribución muestral de F .

Teorema 8.6

Sean U y V dos variables aleatorias independientes que tienen distribuciones χ^2 cuadradas con v_1 y v_2 grados de libertad, respectivamente. Entonces la distribución de la variable aleatoria $F = \frac{U/v_1}{V/v_2}$ está dada por

$$h(f) = \begin{cases} \frac{\Gamma[(v_1 + v_2)/2](v_1/v_2)^{v_1/2}}{\Gamma(v_1/2)\Gamma(v_2/2)} \frac{f^{v_1/2-1}}{(1 + v_1f/v_2)^{(v_1+v_2)/2}} & 0 < f < \infty \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Ésta se conoce como la **distribución F** con v_1 y v_2 grados de libertad.

La curva de la distribución F depende no sólo de los dos parámetros v_1 y v_2 sino también del orden en el que se establecen. Una vez que se dan estos dos valores, podemos identificar la curva. En la figura 8.14 se muestran curvas F típicas.

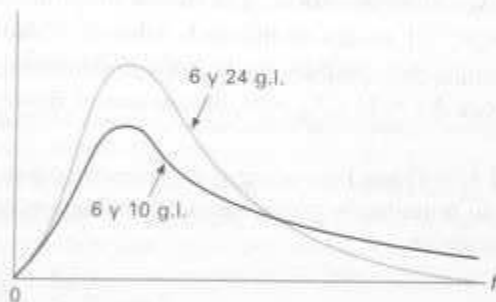


Figura 8.14 Distribuciones F típicas.

Sea f_α por arriba del cual encontramos un área igual a α . Esto se ilustra mediante la región sombreada de la figura 8.15. La tabla A.6 da valores de f_α sólo para $\alpha = 0.05$ y $\alpha = 0.01$ para varias combinaciones de los grados de libertad v_1 y v_2 . De aquí, el valor f con 6 y 10 grados de libertad, que deja un área de 0.05 a la derecha, es $f_{0.05} = 3.22$. Por medio del siguiente teorema, la tabla A.6 también se puede utilizar para encontrar valores de $f_{0.95}$ y $f_{0.99}$. La demostración se deja al lector.

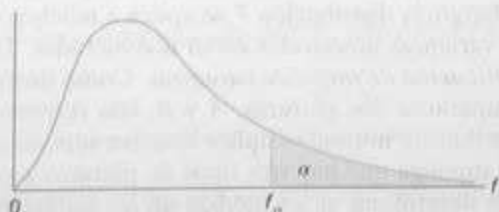


Figura 8.15 Valores tabulados de la distribución F.

Teorema 8.7

Al escribir $f_{\alpha}(v_1, v_2)$ para f_{α} con v_1 y v_2 grados de libertad, obtenemos

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_{\alpha}(v_2, v_1)}$$

Así, el valor f con 6 y 10 grados de libertad, que deja un área de 0.95 a la derecha, es

$$f_{0.95}(6, 10) = \frac{1}{f_{0.05}(10, 6)} = \frac{1}{4.06} = 0.246.$$

Suponga que las muestras aleatorias de tamaño n_1 y n_2 se seleccionan de dos poblaciones normales con varianzas σ_1^2 y σ_2^2 , respectivamente. Del teorema 8.4, sabemos que

$$X_1^2 = \frac{(n_1 - 1)S_1^2}{\sigma_1^2} \quad \text{y} \quad X_2^2 = \frac{(n_2 - 1)S_2^2}{\sigma_2^2}$$

son variables aleatorias que tienen distribuciones ji cuadradas con $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$ grados de libertad. Además, como las muestras se seleccionan al azar, tratamos con variables aleatorias independientes, y entonces con el uso del teorema 8.6 con $X_1^2 = U$ y $X_2^2 = V$, obtenemos el siguiente resultado.

Teorema 8.8

Si S_1^2 y S_2^2 son las varianzas de muestras aleatorias independientes de tamaño n_1 y n_2 tomadas de poblaciones normales con varianzas σ_1^2 y σ_2^2 , respectivamente, entonces

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

tiene una distribución F con $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$ grados de libertad.

¿Para qué se utiliza la distribución F ?

Contestamos esta pregunta, en parte, al inicio de esta sección. La distribución F se usa en situaciones de dos muestras para extraer inferencias acerca de las varianzas de población. Esto implica la aplicación del resultado del teorema 8.8. Sin embargo, la distribución F se aplica a muchos otros tipos de problemas en los que las varianzas muestrales están involucradas. De hecho, la distribución F se llama *distribución de razón de varianzas*. Como ilustración, considere el ejemplo 8.15. Se compararon dos pinturas, A y B , con respecto a su tiempo medio de secado. La distribución normal se aplica bien (se supone que se conocen σ_A y σ_B). Sin embargo, suponga que hay tres tipos de pinturas a comparar, digamos A , B y C . Deseamos determinar si las medias de las poblaciones son equivalentes. Suponga, de hecho, que importante información resumida del experimento es la siguiente:

Pintura	Media muestral	Varianza muestral	Tamaño muestral
A	$\bar{X}_A = 4.5$	$s_A^2 = 0.2$	10
B	$\bar{X}_B = 5.5$	$s_B^2 = 0.14$	10
C	$\bar{X}_C = 6.5$	$s_C^2 = 0.11$	10

El problema se centra alrededor de si los promedios muestrales ($\bar{x}_A, \bar{x}_B, \bar{x}_C$) están suficientemente alejados o no. La implicación de "suficientemente alejados" es muy importante. Parecería razonable que si la variabilidad entre los promedios muestrales es mayor que lo que se esperaría por casualidad, los datos no apoyan la conclusión de que $\mu_A = \mu_B = \mu_C$. Que estos promedios muestrales pudieran ocurrir por casualidad depende de la *variabilidad dentro de las muestras*, como cuantifican s_A^2 , s_B^2 y s_C^2 . La noción de los componentes importantes de la variabilidad se ve mejor por medio de algunas gráficas simples. Considere la gráfica de los datos originales de las muestras A, B y C, que se muestran en la figura 8.16. Estos datos podrían generar con facilidad la información resumen anterior.

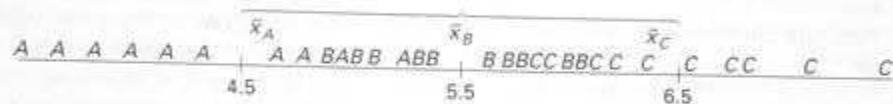


Figura 8.16 Datos de tres muestras diferentes.

Parece evidente que los datos vienen de distribuciones con diferentes medias de población, aunque hay alguna superposición entre las muestras. Un análisis que incluya todos los datos intentaría determinar si la variabilidad entre los promedios muestrales y la variabilidad dentro de las muestras podría haber ocurrido conjuntamente *si de hecho las poblaciones tienen una media común*. Nótese que la clave para este análisis se centra alrededor de las dos siguientes fuentes de variabilidad.

1. Variabilidad dentro de las muestras (entre observaciones en distintas muestras).
2. Variabilidad entre muestras (entre promedios muestrales).

Claramente, si la variabilidad en (1) es considerablemente mayor que la de (2) habrá una considerable superposición en los datos de la muestra y una señal de que los datos podrían provenir de una distribución común. Se encuentra un ejemplo en el conjunto de datos que contienen tres muestras, y que se puede ver en la figura 8.17. Por otro lado, es muy improbable que los datos de una distribución con una media común puedan tener variabilidad entre promedios muestrales que sea considerablemente mayor que la variabilidad dentro de las muestras.



Figura 8.17 Datos que con facilidad podrían venir de la misma población.

Las fuentes de variabilidad en (1) y (2) anteriores generan importantes razones de *varianzas muestrales* y las razones se utilizan junto con la distribución F. El procedimiento general involucrado se llama **análisis de varianza**. Es interesante que en

el ejemplo de la pintura que se describe aquí tratamos con inferencias acerca de tres medias de población; pero se utilizan dos fuentes de variabilidad. No proporcionaremos detalles aquí pero en los capítulos 13, 14 y 15 haremos un uso extenso del análisis de varianza y, por supuesto, la distribución F juega un papel importante.

Ejercicios

- Para una distribución ji cuadrada encuentre
 - $\chi_{0.025}^2$ cuando $\nu = 15$;
 - $\chi_{0.01}^2$ cuando $\nu = 7$;
 - $\chi_{0.05}^2$ cuando $\nu = 24$.
- Para una distribución ji cuadrada encuentre lo siguiente:
 - $\chi_{0.005}^2$ cuando $\nu = 5$;
 - $\chi_{0.05}^2$ cuando $\nu = 19$;
 - $\chi_{0.01}^2$ cuando $\nu = 12$.
- Para una distribución ji cuadrada encuentre χ_{α}^2 tal que
 - $P(X^2 > \chi_{\alpha}^2) = 0.99$ cuando $\nu = 4$;
 - $P(X^2 > \chi_{\alpha}^2) = 0.025$ cuando $\nu = 19$;
 - $P(37.652 < X^2 < \chi_{\alpha}^2) = 0.045$ cuando $\nu = 25$.
- Para una distribución ji cuadrada encuentre χ_{α}^2 tal que
 - $P(X^2 > \chi_{\alpha}^2) = 0.01$ cuando $\nu = 21$;
 - $P(X^2 < \chi_{\alpha}^2) = 0.95$ cuando $\nu = 6$;
 - $P(\chi_{\alpha}^2 < X^2 < 23.209) = 0.015$ cuando $\nu = 10$.
- Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones, de una población normal con varianza $\sigma^2 = 6$, tenga una varianza s^2
 - mayor que 9.1;
 - entre 3.462 y 10.745.
 Suponga que las varianzas muestrales son mediciones continuas.
- Las calificaciones de un examen de colocación que se aplicó a estudiantes de primer año de licenciatura durante los últimos cinco años están aproximadamente distribuidas de forma normal con una media $\mu = 74$ y una varianza $\sigma^2 = 8$. ¿Consideraría aun que $\sigma^2 = 8$ es un valor válido de la varianza si una muestra aleatoria de 20 estudiantes que realizan este examen de colocación este año obtienen un valor de $s^2 = 20$?
- Muestre que la varianza de S^2 para muestras aleatorias de tamaño n de una población normal disminuye conforme n se hace grande. SUGERENCIA: primero encuentre la varianza de $(n-1)S^2/\sigma^2$.
 - Encuentre $t_{0.025}$ cuando $\nu = 14$.
 - Encuentre $-t_{0.10}$ cuando $\nu = 10$.
 - Encuentre $t_{0.995}$ cuando $\nu = 7$.
 - Encuentre $P(T < 2.365)$ cuando $\nu = 7$.
 - Encuentre $P(T > 1.318)$ cuando $\nu = 24$.
 - Encuentre $P(-1.356 < T < 2.179)$ cuando $\nu = 12$.
 - Encuentre $P(T > -2.567)$ cuando $\nu = 17$.
 - Encuentre $P(-t_{0.005} < T < t_{0.01})$.
 - Encuentre $P(T > -t_{0.025})$.
- Dada una muestra aleatoria de tamaño 24 de una distribución normal, encuentre k tal que
 - $P(-2.069 < T < k) = 0.965$;
 - $P(k < T < 2.807) = 0.095$;
 - $P(-k < T < k) = 0.90$.
- Una empresa manufacturera afirma que las baterías que utiliza en sus juegos electrónicos duran un promedio de 30 horas. Para mantener este promedio, se prueban 16 baterías cada mes. Si el valor t que se calcula cae entre $-t_{0.025}$ y $t_{0.025}$, la empresa queda satisfecha con su afirmación. ¿Qué conclusiones extraería la empresa de una muestra que tiene una media de $\bar{x} = 27.5$ horas y una desviación estándar de $s = 5$ horas? Suponga que la distribución de las duraciones de las baterías es aproximadamente normal.
- Una población normal con varianza desconocida tiene una media de 20. ¿Se tiene posibilidad de obtener una muestra aleatoria de tamaño 9 de esta población con una media de 24 y una desviación estándar de 4.1? Si no, ¿qué conclusión sacaría?
- Un fabricante de cierta marca de barras de cereal bajo en grasa afirma que su contenido promedio de grasa saturada es 0.5 gramos. En una muestra aleatoria de 8 barras de cereal de esta marca el contenido de grasa saturada fue 0.6, 0.7, 0.7, 0.3, 0.4, 0.5, 0.4 y 0.2. ¿Estaría de acuerdo con la afirmación?
- Para una distribución F encuentre:
 - $f_{0.05}$ con $\nu_1 = 7$ y $\nu_2 = 15$;
 - $f_{0.05}$ con $\nu_1 = 15$ y $\nu_2 = 7$;
 - $f_{0.01}$ con $\nu_1 = 24$ y $\nu_2 = 19$;
 - $f_{0.95}$ con $\nu_1 = 19$ y $\nu_2 = 24$;
 - $f_{0.99}$ con $\nu_1 = 28$ y $\nu_2 = 12$.
- Pruebas de resistencia a la tracción sobre 10 cables conductores soldados para un dispositivo semi-

conductor dan los siguientes resultados en libras fuerza requeridas para romper la unión:

19.8	12.7	13.2	16.9	10.6
18.8	11.1	14.3	17.0	12.5

Otro conjunto de ocho cables conductores se probó después del encapsulado para determinar si la resistencia a la tracción había aumentado debido al encapsulado del dispositivo, con los siguientes resultados:

24.9	22.8	23.6	22.1	20.4	21.6	21.8	22.5
------	------	------	------	------	------	------	------

Haga comentarios sobre la evidencia disponible con respecto a la igualdad de las dos varianzas de población.

17. Considere las siguientes mediciones de la capacidad de producción de calor del carbón producido por dos minas (en millones de calorías por tonelada):

Mina 1:	8260	8130	8350	8070	8340	
Mina 2:	7950	7890	7900	8140	7920	7840

¿Se puede concluir que las dos varianzas de población son iguales?

Ejercicios de repaso

1. Considere los datos que se muestran en el ejercicio 3 a continuación de la sección 3.4. Construya una gráfica de caja y extensión y comente la naturaleza de la muestra. Calcule la media muestral y la desviación estándar de la muestra.

2. Si X_1, X_2, \dots, X_n son variables aleatorias independientes que tienen distribuciones exponenciales idénticas con parámetro q , muestre que la función de densidad de la variable aleatoria $Y = X_1 + X_2 + \dots + X_n$ es la de una distribución gamma con parámetros $\alpha = n$ y $\beta = \theta$.

3. Al probar el monóxido de carbono en cierta marca de cigarrillos, los datos, en miligramos por cigarrillo, se codificaron al restar 12 de cada observación. Utilice los resultados del ejercicio 14 de la página 208 para encontrar la desviación estándar del contenido de monóxido de carbono de una muestra aleatoria de 15 cigarrillos de esta marca si las mediciones codificadas son 3.8, -0.9, 5.4, 4.5, 5.2, 5.6, 2.7, -0.1, -0.3, -1.7, 5.7, 3.3, 4.4, -0.5, y 1.9.

4. Si S_1^2 y S_2^2 representan las varianzas de muestras aleatorias independientes de tamaño $n_1 = 8$ y $n_2 = 12$, tomadas de poblaciones normales con varianzas iguales, encuentre $P(S_1^2/S_2^2 < 4.89)$.

5. Una muestra aleatoria de cinco presidentes de bancos indican salarios anuales de \$163,000, \$148,000, \$152,000, \$135,000 y \$141,000. Encuentre la varianza de este conjunto.

6. Si el número de huracanes que azotan cierta área del este de Estados Unidos por año es una variable aleatoria que tiene una distribución de Poisson con $\mu = 6$, encuentre la probabilidad de que esta área sea azotada por

- (a) exactamente 15 huracanes en dos años;
 (b) a lo más 9 huracanes en dos años.

7. Una compañía de taxis prueba una muestra aleatoria de 10 neumáticos radiales con cinturones de acero de cierta marca y registra las siguientes duraciones: 48,000, 53,000, 45,000, 61,000, 59,000, 56,000, 63,000, 49,000, 53,000, y 54,000 kilómetros. Utilice los resultados del ejercicio 14 en la página 208 para encontrar la desviación estándar de este conjunto de datos al dividir primero cada observación entre 1000 y después restar 55.

8. Considere los datos del ejercicio 2 que sigue de la sección 3.4. Construya una gráfica de caja y extensión. Haga comentarios. Calcule la media muestral y la desviación estándar muestral.

9. Si S_1^2 y S_2^2 representan las varianzas de muestras aleatorias independientes de tamaño $n_1 = 25$ y $n_2 = 31$, tomadas de poblaciones normales con varianzas $\sigma_1^2 = 10$ y $\sigma_2^2 = 15$, respectivamente, encuentre $P(S_1^2/S_2^2 > 1.26)$.

10. Considere el ejercicio 9. Comente cualquier dato apartado.

11. Considere la revisión del ejercicio 1. Comente cualquier dato apartado en el conjunto.

Problemas de estimación de una y dos muestras

9.1 Introducción

En los capítulos anteriores enfatizamos las propiedades del muestreo de la media muestral y de la varianza muestral. También enfatizamos las representaciones de datos en varias formas. El propósito de estas presentaciones es la construcción de las bases que permitan a los estadísticos extraer conclusiones acerca de los parámetros de la población a partir de los datos experimentales. Por ejemplo, el teorema del límite central proporciona información sobre la distribución de la media muestral \bar{X} . La distribución involucra la media de la población μ . Así, cualesquiera conclusiones que se extraigan con respecto a μ a partir de un promedio muestral observado deben depender del conocimiento de su distribución muestral. Comentarios similares se podrían aplicar a S^2 y σ^2 . Resulta claro que cualesquiera conclusiones que obtengamos acerca de la varianza de una distribución normal probablemente involucraría la distribución muestral de S^2 .

En este capítulo comenzamos por señalar de manera formal el propósito de la inferencia estadística. Seguimos con la presentación del problema de la **estimación de los parámetros de la población**. Restringiremos nuestros desarrollos formales de los procedimientos específicos de estimación a problemas que involucren una y dos muestras.

9.2 Inferencia estadística

En el capítulo 1 presentamos la filosofía general de la inferencia estadística formal. La teoría de la **inferencia estadística** consiste en aquellos métodos por los que se rea-

lizan inferencias o generalizaciones acerca de una población. La tendencia actual es la distinción entre el **método clásico** de estimación de un parámetro de la población, por medio del cual las inferencias se basan de manera estricta en información que se obtiene de una muestra aleatoria seleccionada de la población, y el **método bayesiano**, que utiliza el conocimiento subjetivo previo sobre la distribución de probabilidad de los parámetros desconocidos junto con la información que proporcionan los datos de la muestra. A lo largo de la mayor parte de este capítulo utilizaremos los métodos clásicos para estimar los parámetros de la población desconocidos como la media, la proporción y la varianza mediante el cálculo de estadísticas de muestras aleatorias y la aplicación de la teoría de las distribuciones muestrales, mucho de lo cual se cubre en el capítulo 8. La estimación bayesiana se discute en la sección 9.13.

La inferencia estadística se puede dividir en dos áreas principales: **estimación y pruebas de hipótesis**. Trataremos estas dos áreas por separado, en este capítulo se tratarán la teoría y las aplicaciones de la estimación; y la prueba de hipótesis, en el capítulo 10. Para distinguir claramente entre las dos áreas, considere los siguientes ejemplos. Un candidato a un puesto público puede desear estimar la verdadera proporción de votantes que lo favorecerán mediante la obtención de las opiniones de una muestra aleatoria de 100 votantes. La fracción de votantes en la muestra que favorecerán al candidato se podría utilizar como una estimación de la verdadera proporción en la población de votantes. El conocimiento de la distribución muestral de una proporción nos permite establecer el grado de precisión de nuestra estimación. Este problema cae en el área de la estimación.

Considere ahora el caso en el que se está interesado en encontrar si la marca A de cera para piso es más resistente al desgaste que la marca B. Se puede plantear la hipótesis de que la marca A es mejor que la marca B y, después de la prueba apropiada, aceptar o rechazar esta hipótesis. En este ejemplo no intentamos estimar un parámetro sino que en realidad tratamos de llegar a una decisión correcta acerca de una hipótesis preestablecida. Una vez más dependemos de la teoría del muestreo y del uso de datos que nos proporcionen alguna medición de la precisión de nuestra decisión.

9.3 Métodos clásicos de estimación

Una **estimación puntual** de algún parámetro de la población θ es un solo valor $\hat{\theta}$ de una estadística $\hat{\theta}$. Por ejemplo, el valor \bar{x} de la estadística \bar{X} , que se calcula a partir de la muestra de tamaño n , es una estimación puntual del parámetro poblacional μ . De manera similar, $\hat{p} = x/n$ es una estimación puntual de la verdadera proporción p para un experimento binomial.

No se espera que un estimador realice la estimación del parámetro poblacional sin error. No esperamos que \bar{X} estime μ exactamente, sino que en realidad esperamos que no esté muy alejado. Para una muestra particular es posible obtener un estimado más cercano de μ mediante el uso de la mediana de la muestra \bar{X} como un estimador. Considere, por ejemplo, una muestra que consista en los valores 2, 5 y 11 de una población cuya media es 4 pero supuestamente se le desconoce. Estimáramos a μ como $\bar{x} = 6$, con el uso de la media muestral como nuestra estimación, o $\bar{x} = 5$, con el uso de la mediana muestral como nuestra estimación. En este caso el estimador \bar{X} produce una estimación más cercana al verdadero parámetro que la del estimador \bar{X} .

Por otro lado, si nuestra muestra aleatoria contiene los valores 2, 6 y 7, entonces $\bar{x} = 5$ y $\bar{y} = 6$, por lo que \bar{X} ahora es el mejor estimador. Al no conocer el valor real de μ , debemos decidir de antemano si se utiliza \bar{X} o \bar{Y} como nuestro estimador.

Estimador insesgado

¿Cuáles son las propiedades deseables de una "buena" función de decisión que influirán sobre nosotros para elegir un estimador en lugar de otro? Sea $\hat{\theta}$ un estimador cuyo valor $\hat{\theta}$ es una estimación puntual de algún parámetro poblacional desconocido θ . Ciertamente, desearíamos que la distribución muestral de $\hat{\theta}$ tuviera una media igual al parámetro estimado. Se dice que un estimador que posee esta propiedad es **insesgado**.

Definición 9.1

Se dice que una estadística $\hat{\theta}$ es un estimador **insesgado** del parámetro θ si

$$\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta.$$

Ejemplo 9.1 Muestre que S^2 es un estimador insesgado del parámetro σ^2 .

SOLUCIÓN

Escribamos

$$\begin{aligned} \sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n [(X_i - \mu) - (\bar{X} - \mu)]^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - 2(\bar{X} - \mu) \sum_{i=1}^n (X_i - \mu) + n(\bar{X} - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2. \end{aligned}$$

Ahora bien

$$\begin{aligned} E(S^2) &= E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right] = \frac{1}{n-1} \left[\sum_{i=1}^n E(X_i - \mu)^2 - nE(\bar{X} - \mu)^2 \right] \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n \sigma_{X_i}^2 - n\sigma_{\bar{X}}^2 \right). \end{aligned}$$

Sin embargo,

$$\sigma_{X_i}^2 = \sigma^2 \quad \text{para } i = 1, 2, \dots, n \quad \text{y} \quad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}.$$

Por tanto

$$E(S^2) = \frac{1}{n-1} \left(n\sigma^2 - n\frac{\sigma^2}{n} \right) = \sigma^2.$$

Aunque S^2 es un estimador insesgado de σ^2 , S es, por otro lado, un estimador sesgado de σ siendo insignificante el sesgo en muestras grandes. Este ejemplo ilustra **por qué dividimos entre $n - 1$** en lugar de n cuando se estima la varianza. ■

Varianza de un estimador puntual

Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores insesgados del mismo parámetro poblacional θ , elegiríamos el estimador cuya distribución muestral tuviera la menor varianza. De aquí, si $\sigma_{\hat{\theta}_1}^2 < \sigma_{\hat{\theta}_2}^2$, decimos que $\hat{\theta}_1$ es un **estimador más eficiente** de θ que $\hat{\theta}_2$.

Definición 9.2

Si consideramos todos los posibles estimadores insesgados de algún parámetro θ , el de menor varianza se llama **estimador más eficiente** de θ .

En la figura 9.1 ilustramos las distribuciones muestrales de tres diferentes estimadores $\hat{\theta}_1$, $\hat{\theta}_2$, y $\hat{\theta}_3$, todos para θ . Es claro que sólo $\hat{\theta}_1$ y $\hat{\theta}_2$ son insesgados, pues sus distribuciones están centradas en θ . El estimador $\hat{\theta}_1$ tiene una varianza menor que $\hat{\theta}_2$ y por tanto es más eficiente. De aquí que nuestra elección de un estimador de θ , entre los tres que se consideran, sería $\hat{\theta}_1$.

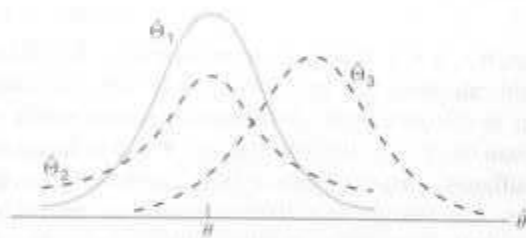


Figura 9.1 Distribuciones muestrales de diferentes estimadores de θ .

Para poblaciones normales se puede mostrar que \bar{X} y \tilde{X} son estimadores insesgados de la media de la población μ , pero la varianza de \bar{X} es más pequeña que la varianza de \tilde{X} . De esta manera las estimaciones \bar{x} y \tilde{x} serán, en promedio, iguales a la media poblacional μ , pero \bar{x} es probable que esté más cerca de μ para una muestra dada, y por ello \bar{X} es más eficiente que \tilde{X} .

Es improbable que incluso el estimador insesgado más eficiente estime el parámetro poblacional con exactitud. Es cierto que nuestra precisión aumenta con muestras grandes, pero no hay razón por la cual debemos esperar que una estimación **puntual** de una muestra dada sea exactamente igual al parámetro poblacional que se supone estima. Hay muchas situaciones en las que es preferible determinar un intervalo dentro del cual esperaríamos encontrar el valor del parámetro. Tal intervalo se llama **intervalo de estimación**.

Estimación por intervalo

Una estimación por intervalo de un parámetro poblacional θ es un intervalo de la forma $\hat{\theta}_L < \theta < \hat{\theta}_U$, donde $\hat{\theta}_L$ y $\hat{\theta}_U$ dependen del valor de la estadística $\hat{\Theta}$ para una muestra particular y también de la distribución de muestreo de $\hat{\Theta}$. Así, una muestra aleatoria de calificaciones verbales SAT para estudiantes universitarios de una clase de primer año podría producir un intervalo de 530 a 550 dentro del cual esperamos encontrar el promedio real de todas las calificaciones verbales del SAT para este grupo. Los valores de los puntos extremos, 530 y 550, dependerán de la media muestral calculada \bar{x} y de la distribución de muestreo de \bar{X} . A medida que el tamaño de la muestra aumenta, sabemos que $\sigma_{\bar{X}}^2 = \sigma^2/n$ disminuye, y en consecuencia es probable que nuestra estimación esté cercana al parámetro μ , lo que tiene como resultado un intervalo más pequeño. De esta manera el intervalo estimado indica, por su longitud, la precisión de la estimación puntual. Un ingeniero obtendrá una idea de la proporción de la población de artículos defectuosos al tomar una muestra y calcular la *proporción de defectuosos de la muestra*. Pero una estimación por intervalo podría ser más informativa.

Como muestras distintas por lo general darán valores diferentes de $\hat{\Theta}$ y, por tanto, valores diferentes de $\hat{\theta}_L$ y $\hat{\theta}_U$, estos puntos extremos del intervalo son valores de las variables aleatorias correspondientes $\hat{\Theta}_L$ y $\hat{\Theta}_U$. De la distribución muestral de $\hat{\Theta}$ seremos capaces de determinar $\hat{\theta}_L$ y $\hat{\theta}_U$ de modo que $P(\hat{\Theta}_L < \theta < \hat{\Theta}_U)$ sea igual a algún valor fraccional positivo que queramos especificar. Si, por ejemplo, encontramos $\hat{\theta}_L$ y $\hat{\theta}_U$ tales que

$$P(\hat{\Theta}_L < \theta < \hat{\Theta}_U) = 1 - \alpha,$$

Para $0 < \alpha < 1$, tenemos entonces una probabilidad de $1 - \alpha$ de seleccionar una variable aleatoria que produzca un intervalo que contenga θ . El intervalo $\hat{\theta}_L < \theta < \hat{\theta}_U$, que se calcula a partir de la muestra seleccionada, se llama entonces **intervalo de confianza** de $(1 - \alpha)100\%$ la fracción $1 - \alpha$ se llama **coeficiente de confianza** o **grado de confianza**, y los extremos, $\hat{\theta}_L$ y $\hat{\theta}_U$, se denominan **límites de confianza** inferior y superior. Así, cuando $\alpha = 0.05$, tenemos un intervalo de confianza de 95%, y cuando $\alpha = 0.01$ obtenemos un intervalo de confianza más amplio de 99%. Entre más amplio sea el intervalo de confianza podemos tener más confianza de que el intervalo dado contenga el parámetro desconocido. Por supuesto, es mejor tener una confianza de 95% de que la vida promedio de cierto transistor de televisor está entre seis y siete años que tener una confianza de 99% de que está entre tres y 10 años. Idealmente, preferimos un intervalo corto con un alto grado de confianza. Algunas veces las restricciones en el tamaño de nuestra muestra nos impiden tener intervalos cortos sin sacrificar algo de nuestro grado de confianza.

En las secciones que siguen proseguimos con las nociones de estimación puntual y por intervalo, donde cada sección representa un caso especial diferente. El lector debe notar que mientras la estimación puntual y por intervalo representan diferentes aproximaciones para obtener información con respecto a un parámetro, también se relacionan en el sentido de que los estimadores del intervalo de confianza se basan en estimadores puntuales. En la siguiente sección, por ejemplo, veremos que el estimador \bar{X} es un estimador puntual de μ muy razonable. Como resultado, el importante estimador del intervalo de confianza de μ depende del conocimiento de la distribución muestral de \bar{X} .

9.4 Una sola muestra: estimación de la media

La distribución muestral de \bar{X} está centrada en μ y en la mayoría de las aplicaciones la varianza es más pequeña que la de cualesquiera otros estimadores de μ . Así, la media muestral \bar{x} se utilizará como una estimación puntual para la media de la población μ . Recuerde que $\sigma_{\bar{X}}^2 = \sigma^2/n$, por lo que una muestra grande dará un valor de \bar{X} que proviene de una distribución de muestreo con varianza pequeña. De aquí que \bar{x} es probablemente una estimación muy precisa de μ cuando n es grande.

Consideremos ahora la estimación por intervalo de μ . Si nuestra muestra se selecciona a partir de una población normal o, a falta de ésta, si n es suficientemente grande, podemos establecer un intervalo de confianza para μ al considerar la distribución muestral de \bar{X} . De acuerdo con el teorema del límite central, podemos esperar que la distribución muestral de \bar{X} esté distribuida de forma aproximadamente normal con media $\mu_{\bar{X}} = \mu$ y desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Al escribir $z_{\alpha/2}$ para el valor z por arriba del cual encontramos un área de $\alpha/2$, podemos ver de la figura 9.2 que

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha,$$

donde

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

Por ello

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1 - \alpha.$$

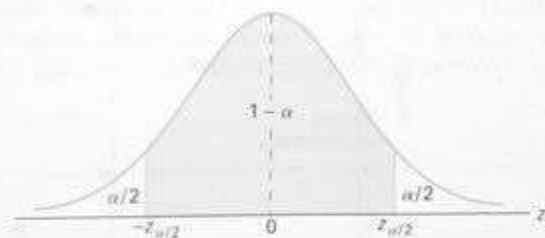


Figura 9.2 $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$.

Al multiplicar cada término en la desigualdad por σ/\sqrt{n} , y después restar \bar{X} de cada término y multiplicar por -1 (para invertir el sentido de las desigualdades), obtenemos

$$P\left(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Se selecciona una muestra aleatoria de tamaño n de una población cuya varianza σ^2 se conoce y se calcula la media \bar{x} para obtener el siguiente intervalo de confianza de $(1 - \alpha)100\%$. Es importante enfatizar que recurrimos al teorema del límite central. Como resultado es importante anotar las condiciones para las aplicaciones que siguen.

Intervalo de confianza de μ ; con σ conocida

Si \bar{x} es la media de una muestra aleatoria de tamaño n de una población con varianza σ^2 , conocida, un intervalo de confianza de $(1 - \alpha)$ 100% para μ está dado por

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}},$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Para muestras pequeñas que se seleccionan de poblaciones no normales, no podemos esperar que nuestro grado de confianza sea preciso. Sin embargo, para muestras de tamaño $n \geq 30$, sin importar la forma de la mayor parte de las poblaciones, la teoría de muestreo garantiza buenos resultados.

Claramente, los valores de las variables aleatorias $\hat{\theta}_L$ y $\hat{\theta}_U$, que se definen en la sección 9.3, son los límites de confianza

$$\hat{\theta}_L = \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{y} \quad \hat{\theta}_U = \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Muestras diferentes darán valores diferentes de \bar{x} y por tanto producirán diferentes intervalos de estimación del parámetro μ como se muestra en la figura 9.3. Los pun-

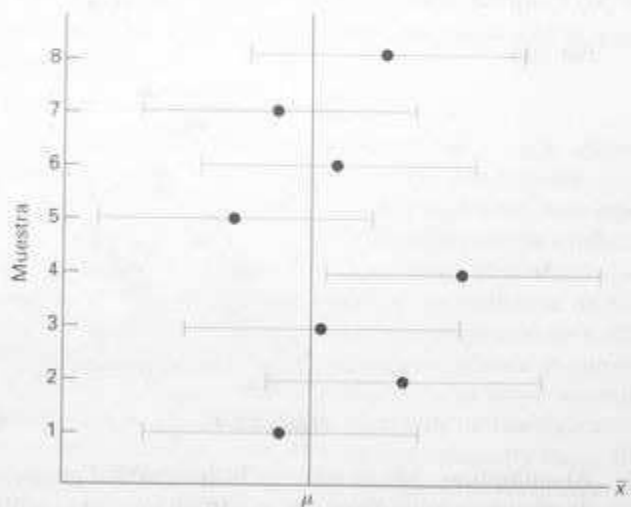


Figura 9.3 Estimaciones por intervalo de μ para muestras diferentes.

tos circulares al centro de cada intervalo indican la posición de la estimación puntual \bar{x} para cada muestra aleatoria. Se ve que la mayor parte de los intervalos contienen a μ , pero no todos los casos. Nótese que todos los intervalos son del mismo ancho, pues esto depende sólo de la elección de $z_{\alpha/2}$ una vez que se determina σ . Entre más grande elegimos el valor de $z_{\alpha/2}$, hacemos más anchos todos los intervalos y podemos tener más confianza en que la muestra particular que se seleccione producirá un intervalo que contenga el parámetro desconocido μ .

Ejemplo 9.2 Se encuentra que la concentración promedio de zinc que se saca del agua a partir de una muestra de mediciones de zinc en 36 sitios diferentes es 2.6 gramos por mililitro. Encuentre los intervalos de confianza de 95% y 99% para la concentración media de zinc en el río. Suponga que la desviación estándar de la población es 0.3.

SOLUCIÓN

La estimación puntual de μ es $\bar{x} = 2.6$. El valor z , que deja un área de 0.025 a la derecha y por tanto un área de 0.975 a la izquierda, es $z_{0.025} = 1.96$ (Tabla A.3). De aquí que el intervalo de confianza de 95% es

$$2.6 - (1.96)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (1.96)\left(\frac{0.3}{\sqrt{36}}\right),$$

que se reduce a $2.50 < \mu < 2.70$. Para encontrar un intervalo de confianza de 99%, encontramos el valor z que deja un área de 0.005 a la derecha y de 0.995 a la izquierda. Por tanto, con el uso de la tabla A.3 de nuevo, $z_{0.005} = 2.575$, y el intervalo de confianza de 99% es

$$2.6 - (2.575)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (2.575)\left(\frac{0.3}{\sqrt{36}}\right).$$

o simplemente

$$2.47 < \mu < 2.73.$$

Vemos ahora que se requiere un intervalo más grande para estimar μ con un grado más alto de precisión.

El intervalo de confianza de $(1 - \alpha)100\%$ proporciona una estimación de la precisión de nuestra estimación puntual. Si μ es realmente el valor central del intervalo, entonces \bar{x} estima μ sin error. La mayor parte de las veces, sin embargo, \bar{x} no será exactamente igual a μ y la estimación puntual es errónea. La magnitud de

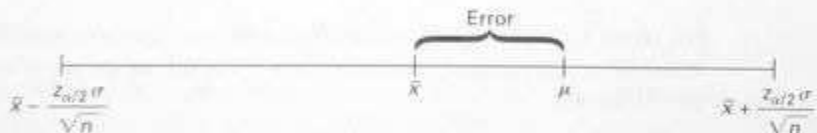


Figura 9.4 Error en la estimación de μ mediante \bar{x} .

este error será el valor absoluto de la diferencia entre μ y \bar{x} , y podemos tener $(1 - \alpha)100\%$ de confianza de que esta diferencia no excederá $z_{\alpha/2}\sigma/\sqrt{n}$. Esto se puede ver con facilidad si dibujamos un diagrama de un intervalo de confianza hipotético como el de la figura 9.4. ■

Teorema 9.1

Si se utiliza \bar{x} como una estimación de μ , podemos tener una confianza de $(1 - \alpha)100\%$ de que el error no excederá de $z_{\alpha/2}\sigma/\sqrt{n}$.

En el ejemplo 9.2 tenemos el 95% de confianza de que la media muestral $\bar{x} = 2.6$ difiere de la media real μ por una cantidad menor que 0.1 y 99% de confianza de que la diferencia es menor que 0.13.

Con frecuencia, deseamos saber qué tan grande necesita ser una muestra para asegurar que el error al estimar μ será menor que una cantidad específica e . Por el teorema 9.1 esto significa que debemos elegir n de modo que $z_{\alpha/2}\sigma/\sqrt{n} = e$. Al resolver esta ecuación se obtiene la siguiente fórmula para n .

Teorema 9.2

Si \bar{x} se usa como estimación de μ , podemos tener $(1 - \alpha)$ 100% de confianza de que el error no excederá una cantidad específica e cuando el tamaño de la muestra es

$$n = \left(\frac{z_{\alpha/2}\sigma}{e} \right)^2.$$

Cuando se resuelve para el tamaño de la muestra, n , todos los valores fraccionales se redondean al siguiente número entero. Si se sigue este principio, podemos estar seguros que nuestro grado de confianza nunca cae por debajo de $(1 - \alpha)$ 100%.

Hablando estrictamente, la fórmula del teorema 9.2 se aplica sólo si conocemos la varianza de la población de la que seleccionamos nuestra muestra. A falta de esta información, podemos tomar una muestra preliminar de tamaño $n \geq 30$ que proporcione una estimación de σ . Después, con el uso de s como aproximación para σ en el teorema 9.2 podemos determinar aproximadamente cuántas observaciones se necesitan para proporcionar el grado de precisión que se desea.

Ejemplo 9.3 ¿Qué tan grande se requiere una muestra en el ejemplo 9.2 si queremos tener 95% de confianza de que nuestra estimación de μ difiera por menos de 0.05?

SOLUCIÓN

La desviación estándar de la población es $\sigma = 0.3$. Entonces, por el teorema 9.2,

$$n = \left[\frac{(1.96)(0.3)}{0.05} \right]^2 = 138.3.$$

Por tanto, podemos tener una confianza 95% de que una muestra aleatoria de tamaño 139 proporcionará una estimación \bar{x} que difiere de μ por una cantidad menor que 0.05. ■

Con frecuencia intentamos estimar la media de una población cuando se desconoce la varianza. El lector debe recordar que en el capítulo 8 aprendimos que si tenemos una muestra aleatoria a partir de una *distribución normal*, entonces la variable aleatoria

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

tiene una distribución *t* de Student con $n - 1$ grados de libertad. Aquí S es la desviación estándar de la muestra. En esta situación en que se desconoce σ se puede utilizar T para construir un intervalo de confianza de μ . El procedimiento es el mismo que cuando se conoce σ excepto que σ se reemplaza por S y la distribución normal estándar se reemplaza por la distribución *t*. Con referencia a la figura 9.5, podemos asegurar que

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha,$$

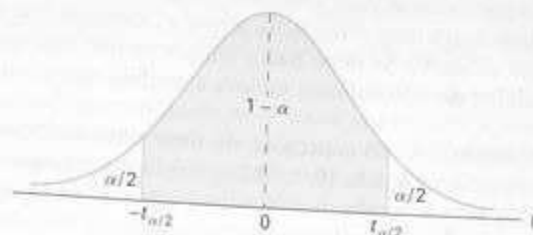


Figura 9.5 $P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha$.

Donde $t_{\alpha/2}$ es el valor t con $n - 1$ grados de libertad, arriba del cual encontramos un área de $\alpha/2$. Debido a la simetría, un área igual de $\alpha/2$ caerá a la izquierda de $-t_{\alpha/2}$. Al sustituir por T , escribimos

$$P\left(-t_{\alpha/2} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2}\right) = 1 - \alpha.$$

Al multiplicar cada término en la desigualdad por S/\sqrt{n} , y después restar \bar{X} de cada término y multiplicar por -1 , obtenemos

$$P\left(\bar{X} - t_{\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2} \frac{S}{\sqrt{n}}\right) = 1 - \alpha.$$

Para nuestra muestra aleatoria particular de tamaño n , se calculan la media \bar{x} y la desviación estándar s y se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% para μ .

Intervalo de confianza para μ ; con σ desconocida

Si \bar{x} y s son la media y la desviación estándar de una muestra aleatoria de una población normal con varianza σ^2 , desconocida, un intervalo de confianza de $(1 - \alpha)$ 100% para μ es

$$\bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n}},$$

donde $t_{\alpha/2}$ es el valor t con $v = n - 1$ grados de libertad, que deja un área de $\alpha/2$ a la derecha.

Hacemos una distinción entre los casos de σ conocida y σ desconocida al calcular las estimaciones del intervalo de confianza. Debemos enfatizar que para el caso de σ conocida se utiliza el teorema del límite central, mientras que para σ desconocida hacemos uso de la distribución muestral de la variable aleatoria T . Sin embargo, el uso de la distribución t se basa en la premisa de que el muestreo se realiza de una distribución normal. En tanto que la distribución tenga forma aproximada de campana, los intervalos de confianza se pueden calcular cuando σ^2 se desconoce mediante el uso de la distribución t y se pueden esperar muy buenos resultados.

Con mucha frecuencia los estadísticos recomiendan que aun cuando la normalidad no se pueda suponer, con σ desconocida y $n \geq 30$, s puede reemplazar a σ y se puede utilizar el intervalo de confianza

$$\bar{x} \pm z_{\alpha/2} s / \sqrt{n}$$

Por lo general éste se denomina como un *intervalo de confianza de muestra grande*. La justificación yace sólo en la presunción de que con una muestra tan grande como 30, \bar{y} estará muy cerca de la σ real y de esta manera el teorema del límite central sigue valiendo. Se debe hacer énfasis en que esto es sólo una aproximación y que la calidad de este enfoque mejora a medida que el tamaño de la muestra crece más.

Ejemplo 9.4 El contenido de siete contenedores similares de ácido sulfúrico son 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, y 9.6 litros. Encuentre un intervalo de confianza del 95% para la media de todos los contenedores si se supone una distribución aproximadamente normal.

SOLUCIÓN

La media muestral y la desviación estándar para los datos dados son

$$\bar{x} = 10.0 \quad y \quad s = 0.283$$

Con el uso de la tabla A.4, encontramos $t_{0.025} = 2.447$ para $v = 6$ grados de libertad. De aquí, el intervalo de confianza de 95% para μ es

$$10.0 - (2.477) \left(\frac{0.283}{\sqrt{7}} \right) < \mu < 10.0 + (2.447) \left(\frac{0.283}{\sqrt{7}} \right),$$

que se reduce a $9.74 < \mu < 10.26$. ■

9.5 Error estándar de una estimación puntual

Hacemos una distinción bastante clara entre los objetivos de las estimaciones puntuales y las estimaciones del intervalo de confianza. Los primeros proporcionan un solo número que se extrae de un conjunto de datos experimentales y los últimos proporcionan un intervalo *dados los datos experimentales* que es razonable para el parámetro; es decir, $100(1 - \alpha)\%$ de tales intervalos que se calculan “cubren” el parámetro.

Estas dos aproximaciones a la estimación se relacionan entre sí. El “hilo común” es la distribución muestral del estimador puntual. Considere, por ejemplo, el estimador \bar{X} de μ con σ conocida. Indicamos antes que una medida de la calidad de un estimador insesgado es su varianza. La varianza de \bar{X} es

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}.$$

De esta forma la desviación estándar de \bar{X} o *error estándar de \bar{X}* es σ/\sqrt{n} . De manera simple, el error estándar de un estimador es su desviación estándar. Para el caso de \bar{X} , el límite de confianza que se calcula

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \text{se escribe como} \quad \bar{x} \pm z_{\alpha/2} \text{s.e.}(\bar{x}),$$

donde s.e. es el error estándar. El punto importante a considerar es que el ancho del intervalo de confianza de μ depende de la calidad del estimador puntual a través de su error estándar. En el caso donde σ se desconoce y el muestreo es sobre una distribución normal, s reemplaza a σ y se incluye el *error estándar estimado* s/\sqrt{n} . De esta forma los límites de confianza de μ son

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = \bar{x} \pm t_{\alpha/2} \hat{s.e.}(\bar{x}).$$

De nuevo, el intervalo de confianza *no es mejor* (en términos de anchura) que la *calidad de la estimación puntual*, en este caso a través de su error estándar estimado. A menudo la paquetería para computadora se refiere a los errores estándar estimados simplemente como “errores estándar”.

A medida que nos movemos hacia intervalos de confianza más complejos, prevalece el concepto de que el ancho de los intervalos de confianza se hace más corto conforme mejora la calidad de la estimación puntual correspondiente; aunque no siempre es tan simple como aquí se ilustra. Se puede argumentar que un intervalo de confianza es simplemente una ampliación de la estimación puntual para tomar en cuenta la precisión de ésta.

9.6 Límites de tolerancia

La estimación del intervalo de confianza puede ser muy útil en un intento de determinar con precisión el valor de un parámetro, digamos una media poblacional. Sin embargo, la interpretación de un intervalo de confianza sobre una media a menudo es confuso para el ingeniero o el científico. Con frecuencia es el caso que el analista es atraído por los *límites del error* dados por un intervalo de confianza para una media cuando, de hecho, el problema requiere algo diferente por completo.

En la sección 9.3 definimos un intervalo de confianza para un parámetro θ que depende de la distribución muestral de $\bar{\theta}$. Es decir, calculamos los límites de confianza de modo que una fracción específica de los intervalos que se calculan a partir de todas las muestras posibles del mismo tamaño contenga el parámetro poblacional θ . Por ejemplo, cuando se seleccionan todas las muestras posibles del mismo tamaño n a partir de una distribución normal, 95% de los intervalos determinados por los límites de confianza $\bar{x} \pm 1.96\sigma/\sqrt{n}$ contendrán el parámetro μ . Este hecho se demuestra en la figura 9.3. Por ello tenemos una confianza de 95% de que el intervalo $\bar{x} \pm 1.96\sigma/\sqrt{n}$, que se calcula a partir de una muestra particular, contendrá el parámetro μ .

Límites de observaciones individuales futuras

Muy a menudo el científico o el ingeniero está menos interesado en la estimación de parámetros y más preocupado por obtener una noción acerca de dónde podrían caer las *observaciones individuales* o mediciones. Si un ingeniero fabrica partes componentes y tiene especificaciones sobre cierta dimensión para la fabricación de la parte, hay poca preocupación sobre la media de la dimensión. En su lugar, se debe intentar determinar los límites que en sentido probabilístico “cubren” los valores individuales de la población (es decir, los valores de la dimensión que se miden).

Un método para establecer un límite sobre valores individuales de la población es determinar un intervalo de confianza sobre una *proporción fija* de las mediciones. Esto se motiva mejor al visualizar una situación en la que hacemos un muestreo aleatorio de una distribución normal con media μ conocida y varianza σ^2 . Claramente, un límite que cubre el 95% central de la población de observaciones es

$$\mu \pm 1.96\sigma.$$

Esto se llama **intervalo de tolerancia**, y en realidad es exacta la cobertura de 95% de las observaciones medidas. Sin embargo, en la práctica μ y σ rara vez se conocen; por ello el usuario debe aplicar

$$\bar{x} \pm ks,$$

y ahora, por supuesto, el intervalo es una variable aleatoria y de aquí que la *cobertura* de una proporción de la población disfruta por el intervalo no es exacta. Como resultado se aplica un intervalo de confianza de $100(1 - \gamma)\%$ al planteamiento pues no se puede esperar que todo el tiempo $\bar{x} \pm ks$ cubra cualquier proporción específica. Como resultado tenemos la siguiente definición.

Límites de tolerancia

Para una distribución normal de mediciones con media μ y desviación estándar σ , desconocidas, los **límites de tolerancia** están dados por $\bar{x} \pm ks$, donde k se determina de modo que se pueda asegurar con una confianza de $100(1 - \gamma)\%$ que los límites dados contienen al menos la proporción $1 - \alpha$ de las mediciones.

La tabla A.7 da valores de k para $1 - \alpha = 0.90, 0.95, 0.99$; $\gamma = 0.05, 0.01$; y para valores seleccionados de n de 2 a 1000.

Ejemplo 9.5 Una máquina produce piezas de metal de forma cilíndrica. Se toma una muestra de estas piezas y se encuentra que los diámetros son 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01, y 1.03 centímetros. Encuentre los límites de tolerancia de 99% que contendrán 95% de las piezas de metal que produce esta máquina, suponga una distribución aproximadamente normal.

SOLUCIÓN

La media muestral y la desviación estándar para los datos dados son

$$\bar{x} = 1.0056 \quad y \quad s = 0.0245.$$

De la tabla A.7 para $n = 9$, $1 - \gamma = 0.99$, y $1 - \alpha = 0.95$, encontramos $k = 4.550$ para los límites de los dos lados. Por aquí los límites de tolerancia de 99% son

$$1.0056 \pm (4.550)(0.0245).$$

Aquí tenemos 99% de confianza de que el intervalo de tolerancia de 0.894 a 1.117 contendrá 95% de las piezas de metal que produce esta máquina. Es interesante notar que el correspondiente intervalo de confianza de 99% para μ (véase el ejercicio 13 de la página 252) tiene un límite inferior de 0.978 y un límite superior de 1.033, lo que verifica nuestro planteamiento anterior de que un intervalo de tolerancia debe necesariamente ser mayor que un intervalo de confianza con el mismo grado de confianza. ■

Distinción entre intervalos de confianza e intervalos de tolerancia

Es importante enfatizar nuevamente la diferencia entre límites de confianza y límites de tolerancia y explicar dónde se utiliza cada uno. El intervalo de confianza sobre la media no es útil a menos que el analista de datos esté interesado en la media poblacional. La media de la población se necesita estimar y el intervalo de confianza produce los límites apropiados. El intervalo de tolerancia está mucho más atento a *dónde caen las observaciones individuales*. ¿Dónde estará la mayor parte de los valores? Considere el ejercicio 15 de la página 252. Hay un proceso de ingeniería que produce esas agujas de tejer. La capacidad del proceso es bastante importante para el ingeniero, y la variabilidad en la dureza de Rockwell puede ser muy molesta. En realidad, el comprador probablemente fije una especificación como límite inferior por debajo del cual no aceptará ninguna aguja. Aquí, por supuesto, los límites de tolerancia son considerablemente más relevantes y lo que el proceso produce en promedio debe tomar un segundo plano. Suponga que un proceso produce piezas cilíndricas para automóvil (como en el ejercicio 13, página 252) y el objetivo es 1.00 centímetros de diámetro. Además, suponga que los límites de especificación se fijan en 1.00 ± 0.05 . Después se toma una muestra del proceso, los límites del 99% de tolerancia calculados ($\gamma = 0.01$) que contienen, digamos, 99% de los diámetros individuales están dados por el intervalo $[0.96, 1.02]$. Los límites de tolerancia son más estrictos que los requisitos del proceso. El ingeniero puede entonces estar confiado por completo de que sería bastante raro tener una parte que cayera fuera de los requisitos que fija el consumidor.

El lector debe comprender que el intervalo de tolerancia tiene características que no son diferentes de las del intervalo de confianza. De hecho, el intervalo de tolerancia del 100 $(1 - \alpha)\%$ sobre, digamos, la proporción 0.95 se puede ver como un intervalo de confianza sobre el 95% central de la distribución normal correspondiente. Los límites de tolerancia de un solo lado también son relevantes. Por ejemplo, en la ilustración del coeficiente de dureza de Rockwell se desearía que el límite de tolerancia reflejara lo siguiente: "Tenemos, digamos, 99% de confianza de que al menos 99% de los valores de la dureza de Rockwell excederá el valor calculado"; la forma del intervalo es entonces

$$\bar{x} - ks.$$

Aquí la tolerancia de un solo lado es un límite de confianza de un solo lado sobre el 1% más bajo de la distribución normal.

Ejercicios

1. Definamos $S^{*2} = \sum_{i=1}^n (X_i - \bar{X})^2/n$. Muestre que

$$E(S^{*2}) = [(n-1)/n]\sigma^2,$$

y de aquí que S^{*2} es un estimador sesgado para σ^2 .

2. Si X es una variable aleatoria binomial, muestre que

(a) $\hat{p} = X/n$ es un estimador insesgado de p ;

(b) $p^* = \frac{X + \sqrt{n}/2}{n + \sqrt{n}}$ es un estimador sesgado de p .

3. Muestre que el estimador P^* del ejercicio 2(b) se vuelve insesgado conforme $n \rightarrow \infty$.

4. Una empresa eléctrica fabrica focos que tienen una duración aproximadamente distribuida de forma normal con una desviación estándar de 40 horas. Si una muestra de 30 focos tiene una duración promedio de 780 horas, encuentre un intervalo de confianza de 96% para la media de la población de todos los focos que produce esta empresa.

5. A muchos pacientes con problemas cardíacos se les implantó un marcapasos para controlar su ritmo cardíaco. Se monta un módulo conector de plástico sobre la parte superior del marcapasos. Suponga una desviación estándar de 0.0015 y una distribución aproximadamente normal, encuentre un intervalo de confianza de 95% para la media de todos los módulos conectores que fabrica cierta compañía. Una muestra aleatoria de 75 módulos tiene un promedio de 0.310 pulgadas.

6. Las estaturas de una muestra aleatoria de 50 estudiantes universitarios muestra una media de 174.5 centímetros y una desviación estándar de 6.9 centímetros.

(a) Construya un intervalo de confianza de 98% para la estatura media de todos los estudiantes de la universidad.

(b) ¿Qué podemos afirmar con 98% de confianza sobre el tamaño posible de nuestro error si estimamos que la estatura media de todos los estudiantes de la universidad es 174.5 centímetros?

7. Una muestra aleatoria de 100 propietarios de automóviles muestra que, en el estado de Virginia, un automóvil se maneja, en promedio, 23,500 kilómetros por año con una desviación estándar de 3900 kilómetros.

(a) Construya un intervalo de confianza de 99% para el número promedio de kilómetros que se maneja un automóvil anualmente en Virginia.

(b) ¿Qué puede afirmar con 99% de confianza sobre el tamaño posible de nuestro error si estimamos que el número promedio de kilómetros manejados por los propietarios de automóviles en Virginia es 23,500 kilómetros por año?

8. ¿De qué tamaño se necesita una muestra en el ejercicio 4 si deseamos tener 96% de confianza que nuestra media muestral esté dentro de 10 horas de la media real?

9. ¿De qué tamaño se necesita una muestra en el ejercicio 5 si deseamos tener 95% de confianza que nuestra media muestral esté dentro de 0.005 pulgadas de la media real?

10. Un experto en eficiencia desea determinar el tiempo promedio que toma perforar tres hoyos en cierta placa metálica. ¿De qué tamaño se necesitará una muestra para tener 95% de confianza de que esta media muestral esté dentro de 15 segundos de la media real? Suponga que se sabe de estudios previos que $\sigma = 40$ segundos.

11. Un investigador de la UCLA afirma que la vida de los ratones se puede extender en 25% cuando se reducen las calorías en su alimento en aproximadamente 40% desde el momento en que se les desteta. Las dietas restringidas se enriquecen a niveles normales con

vitaminas y proteínas. Suponga que se sabe de estudios previos que $\sigma = 5.8$ meses, ¿cuántos ratones se deben incluir en nuestra muestra si deseamos tener 99% de confianza de que la vida media de la muestra esté dentro de dos meses de la media de la población para todos los ratones sujetos a esta dieta reducida?

12. El consumo regular de cereales preendulzados contribuye a la caída de los dientes, enfermedades del corazón y otras enfermedades degenerativas de acuerdo con estudios llevados a cabo por el doctor W. H. Bowen del Instituto Nacional de Salud y el doctor J. Yudben, profesor de nutrición y dietética de la Universidad de Londres. En una muestra aleatoria de 20 porciones sencillas similares del cereal Alpha-Bits, el contenido promedio de azúcar fue 11.3 gramos con una desviación estándar de 2.45 gramos. Suponga que el contenido de azúcar está distribuido normalmente, construya un intervalo de confianza de 95% para el contenido medio de azúcar para porciones sencillas del cereal Alpha-Bits.

13. Una máquina produce piezas metálicas de forma cilíndrica. Se toma una muestra de las piezas y los diámetros son 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01 y 1.03 centímetros. Encuentre un intervalo de confianza de 99% para el diámetro medio de las piezas de esta máquina, suponga una distribución aproximadamente normal.

14. Una muestra aleatoria de 10 barras de chocolate energético de cierta marca tiene, en promedio, 230 calorías con una desviación estándar de 15 calorías. Construya un intervalo de confianza de 99% para el contenido medio real de calorías de esta marca de barras de chocolate energético. Suponga que la distribución de las calorías es aproximadamente normal.

15. Se toma una muestra aleatoria de 12 agujas de tejer en un estudio de prueba de dureza por el método de Rockwell de la cabeza de las agujas. Se realizaron mediciones de la dureza de Rockwell para cada una de las 12, lo que dio un valor promedio de 48.50 con una desviación estándar de 1.5. Suponga que las mediciones se distribuyen de forma normal, construya un intervalo de confianza de 90% para la dureza de Rockwell media.

16. Una muestra aleatoria de 12 graduadas de cierta escuela secretarial escriben a máquina un promedio de 79.3 palabras por minuto con una desviación estándar de 7.8 palabras por minuto. Suponga una distribución normal para el número de palabras que se escribe por minuto, encuentre un intervalo de confianza de 95% para el número promedio de palabras escritas por todas las graduadas de esta escuela.

17. Una muestra aleatoria de 25 botellas de aspirinas contiene, en promedio, 325.05 mg de aspirina con una

desviación estándar de 0.5. Encuentre los límites de tolerancia del 95% que contendrán 90% del contenido de aspirina para esta marca. Suponga que el contenido de aspirina se distribuye normalmente.

18. Las siguientes mediciones se registraron para el tiempo de secado, en horas, de cierta marca de pintura de látex:

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

Suponga que las mediciones representan una muestra aleatoria de una población normal, encuentre los límites de tolerancia del 99% que contendrán 95% de los tiempos de secado.

19. Refiérase al ejercicio 7, construya un intervalo de tolerancia del 99% que contenga 99% de las millas que recorren los automóviles anualmente en Virginia. Suponga que la distribución de las mediciones es aproximadamente normal.

20. Refiérase al ejercicio 15, construya un intervalo del 95% que contenga 90% de las mediciones.

9.7 Dos muestras: estimación de la diferencia entre dos medias

Si tenemos dos poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, un estimador puntual de la diferencia entre μ_1 y μ_2 está dado por la estadística $\bar{X}_1 - \bar{X}_2$. Por tanto, para obtener una estimación puntual de $\mu_1 - \mu_2$, seleccionaremos dos muestras aleatorias independientes, una de cada población, de tamaño n_1 y n_2 , y calculamos la diferencia $\bar{x}_1 - \bar{x}_2$, de las medias muestrales. Claramente, debemos considerar las distribuciones muestrales de $\bar{X}_1 - \bar{X}_2$.

De acuerdo con el teorema 8.3, podemos esperar que la distribución muestral de $\bar{X}_1 - \bar{X}_2$ esté distribuida de forma aproximadamente normal con media $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$ y desviación estándar $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}$. Por tanto, podemos asegurar con una probabilidad de $1 - \alpha$ que la variable normal estándar

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$$

caerá entre $-z_{\alpha/2}$ y $z_{\alpha/2}$. Con referencia, una vez más, a la figura 9.2, escribimos

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha.$$

Al sustituir para Z , establecemos de manera equivalente que

$$P\left[-z_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} < z_{\alpha/2}\right] = 1 - \alpha,$$

que conduce al siguiente intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_1 - \mu_2$.

De nuevo, se invoca el teorema del límite central y de esta forma el intervalo de confianza que sigue es aproximado a menos que las dos poblaciones sean normales.

Si \bar{x}_1 y \bar{x}_2 son las medias de muestras aleatorias independientes de tamaño n_1 y n_2 de poblaciones con varianzas conocidas σ_1^2 y σ_2^2 , respectivamente, un intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_1 - \mu_2$ está dado por

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Intervalo de confianza para $\mu_1 - \mu_2$ con σ_1^2 y σ_2^2 conocida

El grado de confianza es exacto cuando las muestras se seleccionan de poblaciones normales. Para poblaciones no normales el teorema del límite central permite una buena aproximación para muestras de tamaños razonables.

Para el caso de la estimación de un intervalo de confianza sobre la diferencia entre dos medias, necesitamos considerar las condiciones experimentales del proceso de recolección de datos. Se supone que tenemos dos muestras aleatorias independientes de distribuciones con medias μ_1 y μ_2 , respectivamente. Es importante que las condiciones experimentales simulen este "ideal" descrito por las suposiciones tan cerca como sea posible. Muy a menudo el experimentador debe planear la estrategia del experimento en consecuencia. Para casi cualquier estudio de este tipo, hay una así llamada *unidad experimental*, que es la parte del experimento que produce el error experimental y es responsable de la varianza de la población que denominamos σ^2 . En un estudio de medicina, la unidad experimental es el paciente o sujeto. En un experimento de agricultura, puede ser una parcela de tierra. En un experimento químico, puede ser una cantidad de materias primas. Es importante que las diferencias entre estas unidades tengan un impacto mínimo sobre los resultados. El experimentador tendrá un grado de seguridad de que las unidades experimentales no sesgarán los resultados si las condiciones que definen a las dos poblaciones se *asignan al azar* a las unidades experimentales. De nuevo nos concentraremos en la aleatoriedad en los capítulos futuros que tratan de la prueba de hipótesis.

Ejemplo 9.6 Se lleva a cabo un experimento en que se comparan dos tipos de motores, A y B . Se mide el rendimiento en millas por galón de gasolina. Se realizan 50 experimentos con el motor tipo A y 75 con el motor tipo B . La gasolina que se utiliza y las demás condiciones se mantienen constantes. El rendimiento promedio de gasolina para el motor A es de 36 millas por galón y el promedio para el motor B es 42 millas por galón. Encuentre un intervalo de confianza de 96% sobre $\mu_B - \mu_A$, donde μ_B y μ_A son el rendimiento de gasolina medio poblacional para los motores B y A , respectivamente. Suponga que las desviaciones estándar poblacionales son seis y ocho para los motores A y B , respectivamente.

SOLUCIÓN

La estimación puntual de $\mu_B - \mu_A$ es $\bar{x}_B - \bar{x}_A = 42 - 36 = 6$. Con el uso de $\alpha = 0.04$, encontramos $z_{0.02} = 2.05$ de la tabla A.3. De aquí con la sustitución en la fórmula anterior, el intervalo de confianza de 96% es

$$6 - 2.05\sqrt{\frac{64}{75} + \frac{36}{50}} < \mu_1 - \mu_2 < 6 + 2.05\sqrt{\frac{64}{75} + \frac{36}{50}}$$

o simplemente $3.43 < \mu_B - \mu_A < 8.57$. ■

Este procedimiento para estimar la diferencia entre dos medias se aplica si se conocen σ_1^2 y σ_2^2 . Si las varianzas no se conocen y las dos distribuciones implicadas son aproximadamente normales, la distribución t resulta implicada como en el caso de una sola muestra. Si no se está dispuesto a suponer normalidad, muestras grandes (digamos mayores que 30) permitirán el uso de s_1 y s_2 en lugar de σ_1 y σ_2 , res-

pectivamente, con la explicación de que $s_1 = \sigma_1$ y $s_2 = \sigma_2$. De nuevo, por supuesto, el intervalo de confianza es aproximado.

Varianzas desconocidas

Considere el caso donde se desconocen σ_1^2 y σ_2^2 . Si $\sigma_1^2 = \sigma_2^2 = \sigma^2$, obtenemos una variable normal estándar de la forma

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2[(1/n_1) + (1/n_2)]}}$$

De acuerdo con el teorema 8.4, las dos variables aleatorias

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \quad \text{y} \quad \frac{(n_2 - 1)S_2^2}{\sigma^2}$$

tienen distribuciones ji cuadrada con $n_1 - 1$ y $n_2 - 1$ grados de libertad, respectivamente. Además, son variables ji cuadrada independientes, puesto que las muestras aleatorias se seleccionaron de forma independiente. En consecuencia, su suma

$$V = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2}$$

tiene una distribución ji cuadrada con $v_1 = n_1 + n_2 - 2$ grados de libertad.

Como se puede mostrar que las expresiones anteriores para Z y V son independientes, se sigue del teorema 8.5 que la estadística

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2[(1/n_1) + (1/n_2)]}} \bigg/ \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2(n_1 + n_2 - 2)}}$$

tiene la distribución t con $v = n_1 + n_2 - 2$ grados de libertad.

Se puede obtener una estimación puntual de la varianza común desconocida σ^2 al unir las varianzas muestrales. Denotemos al estimador de unión por S_p^2 , escribamos entonces

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Al sustituir S_p^2 en la estadística T , obtenemos la forma menos incómoda

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{(1/n_1) + (1/n_2)}}$$

Con el uso de la estadística T , tenemos

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha,$$

donde $t_{\alpha/2}$ es el valor t con $n_1 + n_2 - 2$ grados de libertad, por arriba del cual encontramos un área de $\alpha/2$. Al sustituir para T en la desigualdad, escribamos

$$P\left[-t_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{(1/n_1) + (1/n_2)}} < t_{\alpha/2}\right] = 1 - \alpha.$$

Después de llevar a cabo las manipulaciones matemáticas de costumbre, se calculan la diferencia de las medias muestrales $\bar{x}_1 - \bar{x}_2$ y la varianza unida

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

y entonces se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_1 - \mu_2$.

Se ve con facilidad que el valor de s_p^2 anterior es un promedio ponderado de las dos varianzas muestrales s_1^2 y s_2^2 , donde los pesos son los grados de libertad.

Intervalo de confianza para

$$\mu_1 - \mu_2;$$

$$\sigma_1^2 = \sigma_2^2$$

pero desconocidas

Si \bar{x}_1 y \bar{x}_2 son las medias de muestras aleatorias independientes con tamaños n_1 y n_2 , respectivamente, de poblaciones aproximadamente normales con varianzas iguales pero desconocidas, un intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_1 - \mu_2$ está dado por

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

donde s_p es la estimación de unión de la desviación estándar poblacional y $t_{\alpha/2}$ es el valor t con $v = n_1 + n_2 - 2$ grados de libertad, que deja un área de $\alpha/2$ a la derecha.

Ejemplo 9.7 En el artículo "Macroinvertebrate Community Structure as an Indicator of Acid Mine Pollution" publicado en la *Revista de Contaminación Ambiental* se da un reporte sobre una investigación realizada en Cane Creek, Alabama, para determinar la relación entre parámetros fisicoquímicos seleccionados y diversas mediciones de la estructura de la comunidad de macroinvertebrados. Una faceta de la investigación fue una evaluación de la efectividad de un índice numérico de la diversidad de especies para indicar la degradación del agua debida al desagüe ácido de una mina. Conceptualmente, un índice alto de la diversidad de especies macroinvertebradas debe indicar un sistema acuático no contaminado, mientras que un índice de diversidad baja indicaría un sistema acuático contaminado.

Se eligieron dos estaciones de muestreo independientes para este estudio, una que se localiza corriente abajo del punto de descarga de la mina de ácido y la otra localizada corriente arriba. Para 12 muestras mensuales reunidas en la estación corriente abajo el índice de diversidad de especies tuvo un valor medio $\bar{x}_1 = 3.11$ y una desviación estándar $s_1 = 0.771$, mientras que 10 muestras mensuales reunidas en la estación corriente arriba tuvieron un valor medio del índice $\bar{x}_2 = 2.04$ y una desviación estándar $s_2 = 0.448$. Encuentre un intervalo de confianza de 90% para la diferencia entre las medias poblacionales para los dos sitios, suponga que las poblaciones están distribuidas de forma aproximadamente normal con varianzas iguales.

SOLUCIÓN

Representemos con μ_1 y μ_2 las medias poblacionales, respectivamente, para los índices de diversidad de especies en las estaciones corriente abajo y corriente arriba. Deseamos encontrar un intervalo de confianza de 90% para $\mu_1 - \mu_2$. Nuestra estimación puntual de $\mu_1 - \mu_2$ es

$$\bar{x}_1 - \bar{x}_2 = 3.11 - 2.04 = 1.07.$$

La estimación de la unión, s_p^2 , de la varianza común, σ^2 , es

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(11)(0.771^2) + (9)(0.448^2)}{12 + 10 - 2} = 0.417$$

Al tomar la raíz cuadrada, obtenemos $s_p = 0.646$. Con el uso de $\alpha = 0.1$, encontramos en la tabla A.4 que $t_{0.05} = 1.725$ para $v = n_1 + n_2 - 2 = 20$ grados de libertad. Por tanto, el intervalo de confianza de 90% para $\mu_1 - \mu_2$ es

$$1.07 - (1.725)(0.646)\sqrt{\frac{1}{12} + \frac{1}{10}} < \mu_1 - \mu_2 < 1.07 + (1.725)(0.646)\sqrt{\frac{1}{12} + \frac{1}{10}},$$

que se simplifica a $0.593 < \mu_1 - \mu_2 < 1.547$. ■

Interpretación del intervalo de confianza

Para el caso de un solo parámetro, el intervalo de confianza simplemente produce límites de error sobre el parámetro. Los valores contenidos en el intervalo se deben ver como valores razonables dados los datos experimentales. En el caso de una diferencia entre dos medias, la interpretación se puede extender a una de comparación de las dos medias. Por ejemplo, si tenemos gran confianza de que una diferencia $\mu_1 - \mu_2$ es positiva, realmente inferiremos que $\mu_1 > \mu_2$ con poco riesgo de caer en un error. De esta forma, en el ejemplo 9.7, tenemos una confianza de 90% de que el intervalo de 0.593 a 1.547 contiene la diferencia de las medias poblacionales para valores del índice de diversidad de especies en las dos estaciones. El hecho de que ambos límites de confianza sean positivos indica que, en promedio, el índice para la estación que se localiza corriente abajo del punto de descarga es mayor que el índice para la estación que se localiza corriente arriba.

Tamaños iguales de muestras

El procedimiento para construir intervalos de confianza para $\mu_1 - \mu_2$ con $\sigma_1 = \sigma_2 = \sigma$ desconocidas requiere la suposición de que las poblaciones son normales. Desviaciones ligeras de la suposición de varianzas iguales o de normalidad no alteran seriamente el grado de confianza de nuestro intervalo. (En el capítulo 10 se presenta un procedimiento para probar la igualdad de dos varianzas poblacionales desconocidas basado en la información que proporcionan las varianzas muestrales.) Si las varianzas poblacionales son considerablemente diferentes, aún obtenemos resultados razonables cuando las poblaciones son normales, dado que $n_1 = n_2$. Por tanto, en un experimento planeado, se debe hacer un esfuerzo para igualar el tamaño de las muestras.

Varianzas distintas

Consideremos ahora el problema de encontrar una estimación por intervalos de $\mu_1 - \mu_2$ cuando no es probable que las varianzas poblacionales desconocidas sean iguales. La estadística que se utiliza con más frecuencia en este caso es

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$$

que tiene aproximadamente una distribución t con v grados de libertad, donde

$$v = \frac{(s_1^2/n + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

Como v rara vez es un entero, lo redondeamos al número entero más cercano.

Con el uso de la estadística T , escribimos

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha,$$

donde $t_{\alpha/2}$ es el valor de la distribución t con v grados de libertad, arriba del cual encontramos un área de $\alpha/2$. Al sustituir para T en la desigualdad, y seguir los pasos exactos como antes, establecemos el resultado final.

Intervalo de confianza para $\mu_1 - \mu_2$ desconocidas

Si \bar{x}_1 y s_1^2 , y \bar{x}_2 y s_2^2 , son las medias y varianzas de muestras pequeñas independientes de tamaño n_1 y n_2 , respectivamente, de distribuciones aproximadamente normales con varianzas desconocidas y diferentes, un intervalo de confianza aproximado del $(1 - \alpha)$ 100% para $\mu_1 - \mu_2$ es

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

donde $t_{\alpha/2}$ es el valor t con

$$v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

grados de libertad, que deja un área $\alpha/2$ a la derecha.

Nótese que el valor v anterior incluye variables aleatorias, y por ello representa una *estimación* de los grados de libertad. En las aplicaciones esta estimación no será un número entero, y por ello el analista lo debe redondear al siguiente entero.

Antes de ilustrar el intervalo de confianza anterior con un ejemplo, debemos señalar que todos los intervalos de confianza sobre $\mu_1 - \mu_2$ son de la misma forma general como los de una sola media; a saber, se pueden escribir como

$$\text{estimación puntual} \pm t_{\alpha/2} \hat{s.e.} (\text{estimación puntual})$$

o

$$\text{estimación puntual} \pm z_{\alpha/2} \text{s.e.} (\text{estimación puntual}).$$

Por ejemplo, en el caso donde $\sigma_1 = \sigma_2 = \sigma$, el error estándar estimado de $\bar{x}_1 - \bar{x}_2$ es $s_p \sqrt{1/n_1 + 1/n_2}$. Para el caso donde $\sigma_1^2 \neq \sigma_2^2$,

$$\hat{s.e.} (\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Ejemplo 9.8 El departamento de zoología del Instituto Politécnico y Universidad Estatal de Virginia llevó a cabo un estudio para estimar la diferencia en la cantidad de ortofósforo químico medido en dos estaciones diferentes del río James. El orto-

fósforo se mide en miligramos por litro. Se reunieron quince muestras de la estación 1 y 12 muestras de la estación 2. Las 15 muestras de la estación 1 tuvieron un contenido promedio de ortofósforo de 3.84 miligramos por litro y una desviación estándar de 3.07 miligramos por litro, mientras que las 12 muestras de la estación 2 tuvieron un contenido promedio de 1.49 miligramos por litro y una desviación estándar de 0.80 miligramos por litro. Encuentre un intervalo de confianza de 95% para la diferencia en el contenido promedio real de ortofósforo en estas dos estaciones, suponga que las observaciones vienen de poblaciones normales con varianzas diferentes.

SOLUCIÓN

Para la estación 1 tenemos $\bar{x}_1 = 3.84$, $s_1 = 3.07$, y $n_1 = 15$. Para la estación 2, $\bar{x}_2 = 1.49$, $s_2 = 0.80$, y $n_2 = 12$. Descamos encontrar un intervalo de confianza de 95% para $\mu_1 - \mu_2$. Como las varianzas poblacionales se suponen diferentes, sólo podemos encontrar un intervalo de confianza de 95% aproximado basado en la distribución t con v grados de libertad, donde

$$v = \frac{(3.07^2/15 + 0.80^2/12)^2}{[(3.07^2/15)^2/14] + [(0.80^2/12)^2/11]} = 16.3 \approx 16$$

Nuestra estimación puntual de $\mu_1 - \mu_2$ es

$$\bar{x}_1 - \bar{x}_2 = 3.84 - 1.49 = 2.35$$

Al usar $\alpha = 0.05$, encontramos en la tabla A.4 que $t_{0.025} = 2.120$ para $v = 16$ grados de libertad. Por tanto, el intervalo de confianza del 95% para $\mu_1 - \mu_2$ es

$$2.35 - 2.120 \sqrt{\frac{3.07^2}{15} + \frac{0.80^2}{12}} < \mu_1 - \mu_2 < 2.35 + 2.120 \sqrt{\frac{3.07^2}{15} + \frac{0.80^2}{12}}$$

que se simplifica a $0.60 < \mu_1 - \mu_2 < 4.10$. Por ello tenemos una confianza de 95% de que el intervalo de 0.60 a 4.10 miligramos por litro contiene la diferencia de los contenidos promedio reales de ortofósforo para estos dos lugares. ■

9.8 Observaciones pareadas

En esta sección consideraremos los procedimientos de estimación para la diferencia de dos medias cuando las muestras no son independientes y las varianzas de las dos poblaciones no necesariamente son iguales. La situación que se considera aquí tiene que ver con una situación experimental muy especial, a saber, la de *observaciones pareadas*. A diferencia de la situación que se describió antes, las condiciones de las dos poblaciones no se asignan de forma aleatoria a las unidades experimentales. Más bien, cada unidad experimental homogénea recibe ambas condiciones poblacionales; como resultado, cada unidad experimental tiene un par de observaciones, una para cada población. Por ejemplo, si realizamos una prueba de una nueva dieta con 15 individuos, el peso antes y después de llevar a cabo la dieta forman la información de nuestras dos muestras. Estas dos poblaciones son "antes" y "después" y la unidad experimental es el individuo. Obviamente las observaciones en un par tienen algo en común. Para determinar si la dieta es efectiva, consideramos las

diferencias d_1, d_2, \dots, d_n en las observaciones pareadas. Estas diferencias son los valores de una muestra aleatoria D_1, D_2, \dots, D_n de una población de diferencias que supondremos distribuidas normalmente con media $\mu_D = \mu_1 - \mu_2$ y varianza σ_D^2 . Estimamos σ_D^2 mediante s_D^2 , la varianza de las diferencias que constituyen nuestra muestra. El estimador puntual de μ_D está dado por \bar{D} .

Parear observaciones en un experimento es una estrategia que se puede emplear en muchos campos de aplicación. Se expondrá al lector este concepto en el material relativo a la prueba de hipótesis en el capítulo 10 y en los temas de diseño experimental en los capítulos 13 y 15. Al seleccionar unidades experimentales relativamente homogéneas (dentro de las unidades) y permitir que cada unidad experimente ambas condiciones poblacionales, se reduce la "varianza del error experimental" efectiva (en este caso σ_D^2). El lector puede visualizar que el i -ésimo par consiste en la medición

$$D_i = X_{1i} - X_{2i}.$$

Como las dos observaciones se toman de la unidad experimental de la muestra, no son independientes y, de hecho,

$$\text{Var } D_i = \text{Var}(X_{1i} - X_{2i}) = \sigma_1^2 + \sigma_2^2 - 2 \text{Cov}(X_{1i}, X_{2i}).$$

Ahora bien, de manera intuitiva, se espera que σ_D^2 se debe reducir debido a la similitud en la naturaleza de los "errores" de las dos observaciones dentro de una unidad experimental, y esto se realiza a través de la expresión anterior. En realidad se espera que si la unidad es homogénea, la covarianza es positiva. Como resultado, la ganancia en calidad del intervalo de confianza sobre el no pareado será mayor cuando haya homogeneidad dentro de las unidades y diferencias grandes conforme se vaya de una unidad a otra. Se debe tener en cuenta que el comportamiento del intervalo de confianza dependerá del error estándar de \bar{D} que es, por supuesto, σ_D/\sqrt{n} donde n es el número de pares. Como indicamos antes, la intención de parear es para reducir σ_D .

Al comparar el intervalo de confianza contra el de la situación sin parear, se hace evidente que hay un "intercambio" involucrado. Aunque en realidad el pareo debe reducir la varianza y por ello reducir el error estándar de la estimación puntual, los grados de libertad se reducen al reducir el problema a uno de una sola muestra. Como resultado, el punto $t_{n/2}$ unido al error estándar se ajusta en consecuencia. Por ello el pareamiento puede ser contraproducente. Éste en realidad sería el caso si uno experimenta sólo una reducción modesta en la varianza (a través de σ_D^2) mediante el pareamiento.

Otra ilustración de pareamiento podría involucrar la elección de n pares de sujetos donde cada par tiene una característica similar, como CI, edad, raza, etcétera; entonces para cada par se selecciona un miembro al azar para obtener un valor de X_1 , mientras que el otro proporciona el valor de X_2 . En este caso X_1 y X_2 pueden representar las calificaciones que obtienen dos individuos de CI igual cuando uno de los individuos se asigna al azar a un grupo que usa el sistema de clases convencional mientras que el otro individuo se asigna a un grupo que utiliza materiales programados.

Se puede establecer un intervalo de confianza de $(1 - \alpha)$ 100% para μ_D al escribir

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha,$$

donde $T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}}$ y $t_{\alpha/2}$, como antes, es un valor de la distribución t con $n - 1$ grados de libertad.

Es ahora un procedimiento de rutina reemplazar T , por su definición, en la desigualdad anterior y desarrollar los pasos matemáticos que conduzcan al siguiente intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_1 - \mu_2 = \mu_D$.

Intervalo de confianza para $\mu_D = \mu_1 - \mu_2$ para observaciones pareadas

Si \bar{d} y s_d son la media y la desviación estándar de las diferencias distribuidas normalmente de n pares aleatorios de mediciones, un intervalo de confianza de $(1 - \alpha)$ 100% para $\mu_D = \mu_1 - \mu_2$ es

$$\bar{d} - t_{\alpha/2} \frac{s_d}{\sqrt{n}} \leq \mu_D \leq \bar{d} + t_{\alpha/2} \frac{s_d}{\sqrt{n}},$$

donde $t_{\alpha/2}$ es el valor t con $v = n - 1$ grados de libertad, que deja un área de $\alpha/2$ a la derecha.

Ejemplo 9.9 Un estudio publicado en *Chemosphere* reporta los niveles de la dioxina TCDD de 20 veteranos de Vietnam residentes en Massachusetts que posiblemente se expusieron al Agente Naranja. El monto de los niveles de TCDD en plasma y en tejido adiposo fueron los siguientes:

Veterano	TCDD en plasma	Niveles de TCDD en tejido adiposo	d_i	Veterano	TCDD en plasma	Niveles de TCDD en tejido adiposo	d_i
1	2.5	4.9	-2.4	11	6.9	7.0	-0.1
2	3.1	5.9	-2.8	12	3.3	2.9	0.4
3	2.1	4.4	-2.3	13	4.6	4.6	0.0
4	3.5	6.9	-3.4	14	1.6	1.4	0.2
5	3.1	7.0	-3.9	15	7.2	7.7	-0.5
6	1.8	4.2	-2.4	16	1.8	1.1	0.7
7	6.0	10.0	-4.0	17	20.0	11.0	9.0
8	3.0	5.5	-2.5	18	2.0	2.5	-0.5
9	36.0	41.0	-5.0	19	2.5	2.3	0.2
10	4.7	4.4	0.3	20	4.1	2.5	1.6

Fuente: A. Scheeter, et al., "Partitioning of 2, 3, 7, 8-chlorinated dibenzo-*p*-dioxins and dibenzofurans between adipose tissue and plasma lipid of 20 Massachusetts Vietnam veterans", *Chemosphere*, Vol. 20, Nos. 7-9, 1990, pp. 954-955 (tablas I y II).

Encuentre un intervalo de confianza de 95% para $\mu_1 - \mu_2$, donde μ_1 y μ_2 representen la TCDD real en plasma y en tejido adiposo, respectivamente. Suponga que la distribución de las diferencias es aproximadamente normal.

SOLUCIÓN

Deseamos encontrar un intervalo de confianza de 95% para $\mu_1 - \mu_2$. Como las observaciones son pareadas, $\mu_1 - \mu_2 = \mu_D$. La estimación puntual de μ_D es $d = -0.87$, la desviación estándar S_d de las diferencias muestrales es

$$S_d = \sqrt{\frac{\sum (d_i - d)^2}{n-1}} = \sqrt{\frac{168.4220}{19}} = 2.9773.$$

Con el uso de $\alpha = 0.05$ encontramos en la tabla A.4 que $t_{0.025} = 2.093$ para $r = n - 1 = 19$ grados de libertad. Por tanto, el intervalo de confianza de 95% es

$$-0.8700 - (2.093) \left(\frac{2.9773}{\sqrt{20}} \right) < \mu_D < -0.8700 + (2.093) \left(\frac{2.9773}{\sqrt{20}} \right)$$

o simplemente $-2.2634 < \mu_D < -0.5234$, del cual podemos concluir que no hay diferencia significativa en el nivel medio de TCDD en plasma y el nivel medio de TCDD en tejido adiposo. ■

Ejercicios

- Una muestra aleatoria de tamaño $n_1 = 25$ que se toma de una población normal con una desviación estándar $\sigma_1 = 5$ tiene una media $\bar{x}_1 = 80$. Una segunda muestra aleatoria de tamaño $n_2 = 36$, que se toma de una población normal diferente con una desviación estándar $\sigma_2 = 3$, tiene una media $\bar{x}_2 = 75$. Encuentre un intervalo de confianza de 95% para $\mu_1 - \mu_2$.
- Se comparan las resistencias de dos clases de hilo. Cincuenta piezas de cada clase de hilo se prueban bajo condiciones similares. La marca A tiene una resistencia a la tracción promedio de 78.3 kilogramos con una desviación estándar de 5.6 kilogramos, mientras que la marca B tiene una resistencia a la tracción promedio de 87.2 kilogramos con una desviación estándar de 6.3 kilogramos. Construya un intervalo de confianza de 95% para la diferencia de las medias poblacionales.
- Se lleva a cabo un estudio para determinar si cierto tratamiento metálico tiene algún efecto sobre la cantidad de metal que se elimina en una operación de decapado. Se sumerge una muestra aleatoria de 100 piezas en un baño por 24 horas sin el tratamiento, lo que da un promedio de 12.2 milímetros eliminados de metal y una desviación estándar de 1.1 milímetros. Una segunda muestra de 200 piezas se somete al tratamiento, seguido de 24 horas de inmersión en el baño, lo que da como resultado una eliminación promedio de 9.1 milímetros de metal con una desviación estándar de 0.9 milímetros. Calcule una estimación del intervalo de confianza del 98% para la diferencia entre las medias de las poblaciones. ¿El tratamiento parece reducir la cantidad media de metal eliminado?
- En un proceso químico por lotes, se comparan los efectos de dos catalizadores sobre la potencia de la reacción del proceso. Se preparó una muestra de 12 lotes con el uso del catalizador 1 y se obtuvo una muestra de 10 lotes con el catalizador 2. Los 12 lotes para los que se utilizó el catalizador 1 dan un rendimiento promedio de 85 con una desviación estándar muestral de 4, y para la segunda muestra el promedio es 81 con una desviación estándar muestral de 5. Encuentre un intervalo de confianza de 90% para la diferencia entre las medias poblacionales, suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales.
- Los estudiantes pueden elegir entre un curso de física sin laboratorio de tres semestres-hora y un curso con laboratorio de cuatro semestres-hora. El examen escrito final es el mismo para cada sección. Si 12 estudiantes de la sección con laboratorio tiene una calificación promedio en el examen de 84 con una desviación estándar de 4, y 18 estudiantes de la sección sin laboratorio tienen una calificación promedio de 77 con una desviación estándar de 6, encuentre un intervalo de confianza de 99% para la diferencia entre las calificaciones promedio para los dos cursos. Suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales.
- En un estudio que se lleva a cabo en el Instituto Politécnico y Universidad Estatal de Virginia sobre el desarrollo de ectomycorrhizal, una relación simbiótica entre las raíces de los árboles y un hongo en la que se transfieren minerales del hongo a los árboles y ar-

cares de los árboles a los hongos, se plantan en un invernadero 20 robles rojos con el hongo *Pisolithus tinctorius*. Todos los arbolitos se plantan en el mismo tipo de suelo y reciben la misma cantidad de luz solar y agua. La mitad no recibe nitrógeno en el momento de plantarlos para servir como control y la otra mitad recibe 368 ppm de nitrógeno en forma de NaNO_3 . Los pesos de los tallos, que se registran en gramos, al final de 140 días se registran como sigue:

Sin nitrógeno	Con nitrógeno
0.32	0.26
0.53	0.43
0.28	0.47
0.37	0.49
0.47	0.52
0.43	0.75
0.36	0.79
0.42	0.86
0.38	0.62
0.43	0.46

Construya un intervalo de confianza del 95% para la diferencia en los pesos medios de los tallos entre los que no recibieron nitrógeno y los que recibieron 368 ppm de nitrógeno. Suponga que las poblaciones están distribuidas normalmente con varianzas iguales.

7. Los siguientes datos, registrados en días, representan el tiempo de recuperación para pacientes que se tratan al azar con uno de dos medicamentos para curar infecciones graves de la vejiga:

Medicamento 1	Medicamento 2
$n_1 = 14$	$n_2 = 16$
$\bar{x}_1 = 17$	$\bar{x}_2 = 19$
$s_1^2 = 1.5$	$s_2^2 = 1.8$

Encuentre un intervalo de confianza de 99% para la diferencia $\mu_2 - \mu_1$ en el tiempo promedio de recuperación para los dos medicamentos, suponga poblaciones normales con varianzas iguales.

8. Un experimento reportado en *Popular Science* compara las economías en combustible para dos tipos de camiones compactos a diesel equipados de forma similar. Supongamos que se utilizaron 12 camiones Volkswagen y 10 Toyota en pruebas de velocidad constante de 90 kilómetros por hora. Si los 12 Volkswagen promedian 16 kilómetros por litro con una desviación estándar de 1.0 kilómetro por litro y los 10 Toyota promedian 11 kilómetros por litro con una desviación estándar de 0.8 kilómetros

por litro, construya un intervalo de confianza de 90% para la diferencia entre los kilómetros promedio por litro de estos dos camiones compactos. Suponga que las distancias por litro para cada modelo de camión están distribuidas de forma aproximadamente normal con varianzas iguales.

9. Una compañía de taxis trata de decidir si comprar neumáticos de la marca A o de la B para su flota de taxis. Para estimar la diferencia de las dos marcas, se lleva a cabo un experimento utilizando 12 de cada marca. Los neumáticos se utilizan hasta que se gastan. Los resultados son

Marca A:	$\bar{x}_1 = 36,300$ kilómetros,
	$s_1 = 5000$ kilómetros.
Marca B:	$\bar{x}_2 = 38,100$ kilómetros
	$s_2 = 6100$ kilómetros.

Calcule un intervalo de confianza de 95% para $\mu_1 - \mu_2$, suponga que las poblaciones se distribuyen de forma aproximadamente normal. Puede no suponer que las varianzas son iguales.

10. Los siguientes datos representan los tiempos de duración de las películas que producen dos compañías cinematográficas.

Compañía	Tiempo (minutos)								
I	103	94	110	87	98				
II	97	82	123	92	175	88	118		

Calcule un intervalo de confianza de 90% para la diferencia entre los tiempos de duración promedio de las películas que producen las dos compañías. Suponga que las diferencias del tiempo de duración se distribuyen de forma aproximadamente normal con varianzas distintas.

11. El gobierno otorga fondos para los departamentos de agricultura de nueve universidades para probar las capacidades de rendimiento de dos nuevas variedades de trigo. Cada variedad se planta en parcelas de área igual en cada universidad y el rendimiento, en kilogramos por parcela, se registra como sigue:

Variedad	Universidad								
	1	2	3	4	5	6	7	8	9
1	38	23	35	41	44	29	37	31	38
2	45	25	31	38	50	33	36	40	43

Encuentre un intervalo de confianza de 95% para la diferencia media entre los rendimientos de las dos variedades, suponga que las diferencias de rendimiento se distribuyen de forma aproximadamente normal.

Explique por qué se necesita el pareamiento en este problema.

12. Con referencia al ejercicio 9, encuentre un intervalo de confianza del 99% para $\mu_1 - \mu_2$ si se asigna al azar un neumático de cada compañía a las ruedas traseras de ocho taxis y se registran las siguientes distancias, en kilómetros:

Taxi	Marca A	Marca B
1	34,400	36,700
2	45,500	46,800
3	36,700	37,700
4	32,000	31,100
5	48,400	47,800
6	32,800	36,400
7	38,100	38,900
8	30,100	31,500

Suponga que las diferencias de las distancias se distribuyen de forma aproximadamente normal.

13. A continuación se listan 10 de las 431 compañías estudiadas en *Fortune Magazine* (marzo de 1997). Se listan las utilidades totales para los 10 años anteriores a 1996 y también para 1996. Encuentre un intervalo de confianza de 95% para el cambio promedio en el porcentaje de utilidad de los inversionistas.

Compañía	Utilidad total para los inversionistas	
	1986-96	1996
Coca-Cola	29.8%	43.3%
Mirage-Resorts	27.9%	25.4%
Merck	22.1%	24.0%
Microsoft	44.5%	88.3%
Johnson & Johnson	22.2%	18.1%
Intel	43.8%	131.2%
Pfizer	21.7%	34.0%
Procter & Gamble	21.9%	32.1%
Berkshire Hathaway	28.3%	6.2%
S&P 500	11.8%	20.3%

9.9 Una sola muestra: estimación de una proporción

Un estimador puntual de la proporción p en un experimento binomial está dado por la estadística $\hat{P} = X/n$, donde X representa el número de éxitos en n pruebas. Por tanto, la proporción de la muestra $\hat{p} = x/n$ se utilizará como el estimador puntual del parámetro p .

Si no se espera que la proporción p desconocida esté demasiado cerca de cero o de 1, podemos establecer un intervalo de confianza para p al considerar la distribución muestral de \hat{P} . Al designar un fracaso en cada prueba binomial mediante el valor 0 y un éxito con el valor 1, el número de éxitos, x , se puede interpretar como la suma de n valores que consisten sólo de ceros y unos, y \hat{p} es sólo la media muestral de estos n valores. De aquí, por el teorema del límite central, para n suficientemente grande, \hat{P} está distribuida de forma aproximadamente normal con media

$$\mu_{\hat{p}} = E(\hat{P}) = E\left[\frac{X}{n}\right] = \frac{np}{n} = p$$

y varianza

$$\sigma_{\hat{p}}^2 = \sigma_{X/n}^2 = \frac{\sigma_x^2}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}$$

Por tanto, podemos asegurar que

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha,$$

donde

$$Z = \frac{\hat{P} - p}{\sqrt{pq/n}}$$

y $z_{\alpha/2}$ es el valor de la curva normal estándar sobre la cual encontramos un área de $\alpha/2$. Al sustituir para Z , escribimos

$$P\left(-z_{\alpha/2} < \frac{\hat{P} - p}{\sqrt{pq/n}} < z_{\alpha/2}\right) = 1 - \alpha$$

Al multiplicar cada término de la desigualdad por $\sqrt{pq/n}$, y después restar \hat{P} y multiplicar por -1 , obtenemos

$$P\left(\hat{P} - z_{\alpha/2}\sqrt{\frac{pq}{n}} < p < \hat{P} + z_{\alpha/2}\sqrt{\frac{pq}{n}}\right) = 1 - \alpha.$$

Es difícil manipular las desigualdades de modo que se obtenga un intervalo aleatorio cuyos puntos extremos sean independientes de p , el parámetro desconocido. Cuando n es grande, se introduce un error muy pequeño al sustituir la estimación puntual $\hat{p} = x/n$ por p bajo el signo del radical. Entonces podemos escribir

$$P\left(\hat{P} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{P} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) \approx 1 - \alpha.$$

Para nuestra muestra aleatoria particular de tamaño n , se calcula la proporción muestral $\hat{p} = x/n$ y se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% aproximado para p .

Intervalo de confianza para p de una muestra grande

Si \hat{p} es la proporción de éxitos en una muestra aleatoria de tamaño n , y $\hat{q} = 1 - \hat{p}$, un intervalo de confianza aproximado de $(1 - \alpha)$ 100% para el parámetro binomial p está dado por

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}},$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Cuando n es pequeña y la proporción desconocida p se considera cercana a 0 o a 1, el procedimiento del intervalo de confianza que se establece aquí no es confiable y, por tanto, no se debe utilizar. Para estar seguro, se debe requerir que $n\hat{p}$ o $n\hat{q}$ sea mayor que o igual a 5. El método para encontrar un intervalo de confianza para el parámetro binomial p también se aplica cuando la distribución binomial se utiliza para aproximar la distribución hipergeométrica; es decir, cuando n es pequeña en relación con N , como se ilustra en el ejemplo 9.10.

Ejemplo 9.10 En una muestra aleatoria de $n = 500$ familias que tienen televisores en la ciudad de Hamilton, Canadá, se encuentra que $x = 340$ están suscritas a HBO. Encuentre un intervalo de confianza de 95% para la proporción real de familias en esta ciudad que están suscritas a HBO.

SOLUCION

La estimación puntual de p es $\hat{p} = 340/500 = 0.68$. Con el uso de la tabla A.3, encontramos que $z_{0.025} = 1.96$. Por tanto, el intervalo de confianza de 95% para p es

$$0.68 - 1.96 \sqrt{\frac{(0.68)(0.32)}{500}} < p < 0.68 + 1.96 \sqrt{\frac{(0.68)(0.32)}{500}}$$

que se simplifica a $0.64 < p < 0.72$.

Si p es el valor central de un intervalo de confianza de $(1 - \alpha) 100\%$ entonces \hat{p} estima a p sin error. La mayor parte de las veces, sin embargo, \hat{p} no será exactamente igual a p y la estimación puntual es errónea. El tamaño de este error será la diferencia positiva que separa a p y \hat{p} , y podemos tener una confianza del $(1 - \alpha) 100\%$ de que esta diferencia no excederá $z_{\alpha/2} \sqrt{\hat{p}\hat{q}/n}$. Podemos ver esto fácilmente si dibujamos un diagrama de un intervalo de confianza típico como en la figura 9.6. ■

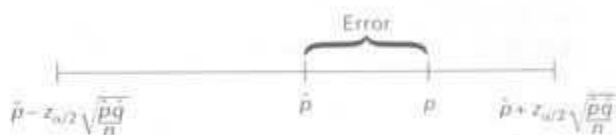


Figura 9.6 Error en la estimación de p por \hat{p} .

Teorema 9.3

Si \hat{p} se utiliza como una estimación de p , podemos tener una confianza del $(1 - \alpha) 100\%$ de que el error no excederá $z_{\alpha/2} \sqrt{\hat{p}\hat{q}/n}$.

En el ejemplo 9.10 tenemos una confianza de 95% de que la proporción de la muestra $\hat{p} = 0.68$ difiere de la proporción real p por una cantidad que no excede 0.04.

Determinemos ahora qué tan grande se requiere que sea una muestra para asegurar que el error al estimar p sea menor que una cantidad específica e . Por el teorema 7.3, esto significa que debemos elegir n de modo que $z_{\alpha/2} \sqrt{\hat{p}\hat{q}/n} = e$.

Teorema 9.4

Si \hat{p} se utiliza como estimación de p , podemos tener una confianza del $(1 - \alpha) 100\%$ de que el error será menor que una cantidad específica e cuando el tamaño de la muestra es aproximadamente

$$n = \frac{z_{\alpha/2}^2 \hat{p}\hat{q}}{e^2}$$

El teorema 9.4 es algo engañoso, pues debemos utilizar \hat{p} para determinar el tamaño n de la muestra, pero \hat{p} se calcula a partir de la muestra. Si se puede hacer una estimación cruda de p sin tomar una muestra, podríamos usar este valor para \hat{p} y después determinar n . A falta de tal estimación, podríamos tomar una muestra preliminar de tamaño $n \geq 30$ para proporcionar una estimación de p . Después, con el uso del teorema 9.4 podríamos determinar de forma aproximada cuántas observaciones se necesitan para proporcionar el grado de precisión que se desea. Una vez más, todos los valores fraccionarios de n se redondean al siguiente número entero.

Ejemplo 9.11 ¿Qué tan grande se requiere que sea una muestra en el ejemplo 9.10 si queremos tener 95% de confianza de que nuestra estimación de p está dentro de 0.02?

SOLUCIÓN

Tratemos a las 500 familias como una muestra preliminar que proporciona una estimación $\hat{p} = 0.68$. Entonces, por el teorema 9.4,

$$n = \frac{(1.96)^2 (0.68)(0.32)}{(0.02)^2} = 2090.$$

Por tanto, si basamos nuestra estimación de p sobre una muestra aleatoria de tamaño 2090, podemos tener una confianza de 95% de que nuestra proporción muestral no diferirá de la proporción real por más de 0.02. ■

De vez en cuando será poco práctico obtener una estimación de p que se utilice para determinar el tamaño muestral para un grado específico de confianza. Si esto sucede, se establece un límite superior para n al notar que $\hat{p}\hat{q} = \hat{p}(1 - \hat{p})$, que debe ser a lo más igual a $1/4$, puesto que \hat{p} debe estar entre 0 y 1. Este hecho se puede verificar al completar cuadrados. De aquí

$$\hat{p}(1 - \hat{p}) = -(\hat{p}^2 - \hat{p}) = \frac{1}{4} - (\hat{p}^2 - \hat{p} + \frac{1}{4}) = \frac{1}{4} - (\hat{p} - \frac{1}{2})^2,$$

que siempre es menor que $1/4$ excepto cuando $\hat{p} = 1/2$ y entonces $\hat{p}\hat{q} = 1/4$. Por tanto, si sustituimos $\hat{p} = 1/2$ en la fórmula para n del teorema 9.4, cuando, de hecho, p realmente difiere de $1/2$, entonces n se hará más grande de lo necesario para el grado de confianza específico y como resultado nuestro grado de confianza aumentará.

Teorema 9.5

Si \hat{p} se utiliza como estimación de p , podemos tener una confianza de **al menos** $(1 - \alpha)$ 100% de que el error no excederá una cantidad específica e cuando el tamaño de la muestra sea

$$n = \frac{z_{\alpha/2}^2}{4e^2}.$$

Ejemplo 9.12 ¿Qué tan grande se requiere que sea la muestra en el ejemplo 9.10 si queremos tener una confianza de al menos 95% de que nuestra estimación de p está dentro de 0.02?

SOLUCIÓN

A diferencia del ejemplo 9.11, debemos suponer ahora que no se ha tomado una muestra preliminar para proporcionar una estimación de p . En consecuencia, podemos tener al menos una confianza de 95% de que nuestra proporción de la muestra no diferirá de la proporción real por más de 0.02 si elegimos una muestra de tamaño

$$n = \frac{(1.96)^2}{(4)(0.02)^2} = 2401.$$

Al comparar los resultados de los ejemplos 9.11 y 9.12, vemos que la información con respecto a p , proporcionada por una muestra preliminar, o quizá a partir de la experiencia pasada, nos permite elegir una muestra más pequeña mientras mantenemos nuestro grado de precisión requerido. ■

9.10 Dos muestras: estimación de la diferencia entre dos proporciones

Considere el problema donde deseamos estimar la diferencia entre dos parámetros binomiales p_1 y p_2 . Por ejemplo, podemos hacer que p_1 sea la proporción de fumadores con cáncer de pulmón y p_2 la proporción de no fumadores con cáncer de pulmón. Nuestro problema, entonces, es estimar la diferencia entre estas dos proporciones. Primero, seleccionamos muestras aleatorias independientes de tamaño n_1 y n_2 a partir de las dos poblaciones binomiales con medias $n_1 p_1$ y $n_2 p_2$ y varianzas $n_1 p_1 q_1$ y $n_2 p_2 q_2$, respectivamente, después determinamos los números x_1 y x_2 de personas con cáncer de pulmón en cada ejemplo, y formar las proporciones $\hat{p}_1 = x_1/n_1$ y $\hat{p}_2 = x_2/n_2$. Un estimador puntual de la diferencia entre las dos proporciones, $p_1 - p_2$, está dado por la estadística $\hat{P}_1 - \hat{P}_2$. Por tanto, la diferencia de las proporciones muestrales, $\hat{p}_1 - \hat{p}_2$ se utilizará como la estimación puntual de $p_1 - p_2$.

Se puede establecer un intervalo de confianza para $p_1 - p_2$ al considerar la distribución muestral de $\hat{P}_1 - \hat{P}_2$. De la sección 9.9 sabemos que \hat{P}_1 y \hat{P}_2 están distribuidos cada uno de forma aproximadamente normal, con medias p_1 y p_2 y varianzas $p_1 q_1/n_1$ y $p_2 q_2/n_2$, respectivamente. Al elegir muestras independientes de las dos poblaciones, las variables \hat{P}_1 y \hat{P}_2 serán independientes, y entonces por la propiedad reproductiva de la distribución normal que se estableció en el teorema 7.11, concluimos que $\hat{P}_1 - \hat{P}_2$ está distribuida de forma aproximadamente normal con media

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 \quad \text{y varianza} \quad \sigma_{\hat{p}_1 - \hat{p}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

Por tanto, podemos asegurar que

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha,$$

donde $Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{(p_1 q_1/n_1) + (p_2 q_2/n_2)}}$ y $z_{\alpha/2}$ es un valor de la curva normal estándar sobre la cual encontramos un área de $\alpha/2$. Al sustituir para Z , escribimos

$$P\left[-z_{\alpha/2} < \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{(p_1 q_1/n_1) + (p_2 q_2/n_2)}} < z_{\alpha/2}\right] = 1 - \alpha$$

Después de llevar a cabo las manipulaciones matemáticas usuales, reemplazamos p_1 , p_2 , q_1 , y q_2 bajo el signo del radical por sus estimaciones $\hat{p}_1 = x_1/n_1$, $\hat{p}_2 = x_2/n_2$, $\hat{q}_1 = 1 - \hat{p}_1$, y $\hat{q}_2 = 1 - \hat{p}_2$, dado que $n_1 \hat{p}_1$, $n_1 \hat{q}_1$, $n_2 \hat{p}_2$, y $n_2 \hat{q}_2$ son

todas mayores que o iguales a 5, y se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% aproximado para $p_1 - p_2$.

Intervalo de confianza para $p_1 - p_2$ de una muestra grande

Si \hat{p}_1 y \hat{p}_2 son las proporciones de éxitos en muestras aleatorias de tamaño n_1 y n_2 , respectivamente, $\hat{q}_1 = 1 - \hat{p}_1$ y $\hat{q}_2 = 1 - \hat{p}_2$, un intervalo de confianza aproximado de $(1 - \alpha)$ 100% para la diferencia de dos parámetros binomiales $p_1 - p_2$, está dado por

$$\begin{aligned} (\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} &< p_1 - p_2 \\ &< (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}, \end{aligned}$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Ejemplo 9.13 Se considera cierto cambio en un proceso de fabricación de partes componentes. Se toman muestras del procedimiento existente y del nuevo para determinar si éste tiene como resultado una mejoría. Si se encuentra que 75 de 1500 artículos del procedimiento actual son defectuosos y 80 de 2000 artículos del procedimiento nuevo también lo son, encuentre un intervalo de confianza de 90% para la diferencia real en la fracción de defectuosos entre el proceso actual y el nuevo.

SOLUCIÓN

Sean p_1 y p_2 las proporciones reales de defectuosos para los procesos actual y nuevo, respectivamente. De aquí, $\hat{p}_1 = 75/1500 = 0.05$ y $\hat{p}_2 = 80/2000 = 0.04$, y la estimación puntual de $p_1 - p_2$ es

$$\hat{p}_1 - \hat{p}_2 = 0.05 - 0.04 = 0.01.$$

Con el uso de la tabla A.3, encontramos $z_{0.05} = 1.645$. Por tanto, al sustituir en esta fórmula obtenemos el intervalo de confianza de 90%

$$\begin{aligned} 0.01 - 1.645 \sqrt{\frac{(0.05)(0.95)}{1500} + \frac{(0.04)(0.96)}{2000}} &< p_1 - p_2 \\ &< 0.01 + 1.645 \sqrt{\frac{(0.05)(0.95)}{1500} + \frac{(0.04)(0.96)}{2000}}, \end{aligned}$$

que se simplifica a $-0.0017 < p_1 - p_2 < 0.0217$. Como el intervalo contiene el valor 0, no hay razón para creer que el nuevo procedimiento producirá una disminución significativa en la proporción de artículos defectuosos comparado con el método existente. ■

Ejercicios

1. (a) Se selecciona una muestra aleatoria de 200 votantes y se encuentra que 114 apoyan un convenio de anexión. Encuentre el intervalo de confianza de 96% para la fracción de la población votante que favorece el convenio.

(b) ¿Qué podemos asegurar con 96% de confianza acerca de la posible magnitud de nuestro error si estimamos que la fracción de votantes que favorecen la anexión es 0.57?

2. Un fabricante de reproductores de discos compactos utiliza un conjunto de pruebas amplias para evaluar la función eléctrica de su producto. Todos los reproductores de discos compactos deben pasar todas las pruebas antes de venderse. Una muestra aleatoria de 500 reproductores tiene como resultado 15 que fallan en una o más pruebas. Encuentre un intervalo de confianza de 90% para la proporción de los reproductores de discos compactos de la población que pasan todas las pruebas.

3. En una muestra aleatoria de 1000 casas en cierta ciudad, se encuentra que 228 se calientan con petróleo. Encuentre el intervalo de confianza de 99% para la proporción de casas en esta ciudad que se calientan con petróleo.

4. Calcule un intervalo de confianza de 98% para la proporción de artículos defectuosos en un proceso cuando se encuentra que una muestra de tamaño 100 da 8 defectuosos.

5. Se considera un nuevo sistema de lanzamiento de cohetes para el despliegue de cohetes pequeños de corto alcance. El sistema existente tiene $p = 0.8$ como probabilidad de lanzamiento exitoso. Se realiza una muestra de 40 lanzamientos experimentales con el nuevo sistema y 34 resultan exitosos.

(a) Construya un intervalo de confianza de 95% para p .
 (b) ¿Concluiría que el nuevo sistema es mejor?

6. Un genetista se interesa en la proporción de hombres africanos que tiene cierto trastorno sanguíneo menor. En una muestra aleatoria de 100 hombres africanos, se encuentra que 24 lo padecen.

(a) Calcule un intervalo de confianza de 99% para la proporción de hombres africanos que tienen este trastorno sanguíneo.

(b) ¿Qué se puede asegurar con 99% de confianza acerca de la posible magnitud de nuestro error si estimamos que la proporción de hombres africanos con este trastorno sanguíneo es 0.24?

7. (a) De acuerdo con un reporte del *Roanoke Times & World-News*, aproximadamente $2/3$ de los 1600

adultos encuestados vía telefónica dijeron que piensan que el programa del transbordador espacial es una buena inversión para el país. Encuentre un intervalo de confianza de 95% para la proporción de adultos estadounidenses que piensan que el programa del transbordador espacial es una buena inversión para el país.
 (b) ¿Qué podemos asegurar con una confianza de 95% acerca de la posible magnitud de nuestro error si estimamos que la proporción de adultos estadounidenses que piensan que el programa del transbordador espacial es una buena inversión es $2/3$?

8. En el artículo del periódico al que se hace referencia en el ejercicio 7, 32% de los 1600 adultos encuestados dijeron que el programa espacial estadounidense debe enfatizar la exploración científica. ¿Qué tan grande se necesita que sea la muestra de adultos en la encuesta si se desea tener una confianza de 95% de que el porcentaje estimado esté dentro de 2% del porcentaje real?

9. ¿Qué tan grande se requiere que sea la muestra en el ejercicio 1 si deseamos tener una confianza de 96% de que nuestra proporción de la muestra estará dentro del 0.02 de la fracción real de la población votante?

10. ¿Qué tan grande se requiere que sea la muestra en el ejercicio 3 si deseamos tener una confianza de 99% de que nuestra proporción de la muestra estará dentro del 0.05 de la proporción real de casas en esta ciudad que se calientan con petróleo?

11. ¿Qué tan grande se necesita la muestra en el ejercicio 4 si deseamos tener una confianza de 98% de que nuestra proporción de la muestra esté dentro del 0.05 de la proporción real de defectuosos?

12. Se lleva a cabo un estudio para estimar el porcentaje de ciudadanos de una ciudad que están a favor de tener su agua fluorada. ¿Qué tan grande se requiere que sea la muestra si se desea tener al menos una confianza de 95% de que nuestra estimación está dentro del 1% del porcentaje real?

13. La conjetura de un miembro del cuerpo docente del departamento de microbiología de la Escuela de Medicina Dental de la Universidad de Washington en St. Louis afirma que un par de tasas diarias de té verde o negro proporcionan suficiente flúor para evitar la caries en los dientes. ¿Qué tan grande se requiere que sea la muestra para estimar el porcentaje de habitantes de cierta ciudad que están a favor de tener su agua fluorada si se desea tener al menos el 99% de confianza de que la estimación está dentro de 1% del porcentaje real?

14. Se lleva a cabo un estudio para estimar la proporción de residentes de cierta ciudad y sus suburbios que están a favor de la construcción de una planta de energía nuclear. ¿Qué tan grande se requiere que sea la muestra si se desea tener al menos 95% de confianza de que la estimación está dentro del 0.04 de la proporción real de residentes de esta ciudad y sus suburbios que están a favor de la construcción de la planta de energía nuclear?

15. Cierta genetista se interesa en la proporción de hombres y mujeres en la población que tienen cierto trastorno sanguíneo menor. En una muestra aleatoria de 1000 hombres se encuentra que 250 lo padecen, mientras que 275 de 1000 mujeres examinadas parecen tener el trastorno. Calcule un intervalo de confianza del 95% para la diferencia entre la proporción de hombres y mujeres que padecen el trastorno sanguíneo.

16. Se encuestan diez escuelas de ingeniería en Estados Unidos. La muestra contiene 250 ingenieros eléctricos, 80 son mujeres; 175 ingenieros químicos, 40 son mujeres. Calcule un intervalo de confianza de 90% para la diferencia entre la proporción de mujeres en estos dos campos de la ingeniería. ¿Hay una diferencia significativa entre las dos proporciones?

17. Se lleva a cabo una prueba clínica para determinar si cierto tipo de inoculación tiene un efecto sobre la incidencia de cierta enfermedad. Se mantiene una muestra de 1000 ratas en un medio controlado por un periodo de un año y a 500 de éstas se les inoculó. Del grupo al que no se le dio la medicina, hubo 120 incidencias de la enfermedad, mientras que 98 del grupo inoculado la contrajo. Si p_1 es la probabilidad de incidencia de la

enfermedad en las ratas no inoculadas y p_2 es la probabilidad de incidencia después de recibir la medicina, calcule un intervalo de confianza de 90% para $p_1 - p_2$.

18. En un estudio, *Germination and Emergence of Broccoli*, que lleva a cabo el Departamento de Horticultura del Instituto Politécnico y Universidad Estatal de Virginia, un investigador encuentra que a 5 °C, 10 semillas de 20 germinaron, mientras que a 15 °C, 15 de 20 semillas germinaron. Calcule un intervalo de confianza del 95% para la diferencia entre la proporción de germinación en las dos diferentes temperaturas y decida si hay una diferencia significativa.

19. Una encuesta de 1000 estudiantes concluye que 274 eligen al equipo profesional de beisbol A como su equipo favorito. En 1991, se realizó la misma encuesta con 760 estudiantes. Concluyó que 240 de ellos también eligieron al equipo A como su favorito. Calcule un intervalo de confianza de 95% para la diferencia entre la proporción de estudiantes que favorecen al equipo A entre las dos encuestas. ¿Hay una diferencia significativa?

20. De acuerdo con *USA Today* (17 de marzo de 1997), las mujeres constituían 33.7% del personal editorial en las estaciones locales de televisión en 1990 y el 36.2% en 1994. Suponga que se contrataron 20 nuevos empleados para el personal editorial.

(a) Estime el número que habrían sido mujeres en cada año respectivamente.

(b) Calcule un intervalo de confianza de 95% para ver si hay evidencia de que la proporción de mujeres contratadas como personal editorial en 1994 fue mayor que la proporción contratada en 1990.

9.11 Una sola muestra: estimación de la varianza

Si se extrae una muestra de tamaño n de una población normal con varianza σ^2 , y se calcula la varianza muestral s^2 obtenemos un valor de la estadística S^2 . Esta varianza muestral calculada se usará como estimación puntual de σ^2 . Por ello la estadística S^2 se llama estimador de σ^2 .

Se puede establecer una estimación por intervalos de σ^2 mediante el uso de la estadística

$$X^2 = \frac{(n-1)S^2}{\sigma^2}.$$

De acuerdo con el teorema 8.4, la estadística X^2 tiene una distribución ji cuadrada con $n-1$ grados de libertad cuando las muestras se eligen de una población normal. Podemos escribir (véase la figura 9.7)

$$P(\chi_{n-1/2}^2 < X^2 < \chi_{n-1/2}^2) = 1 - \alpha,$$

donde $\chi^2_{1-\alpha/2}$ y $\chi^2_{\alpha/2}$ son valores de la distribución ji cuadrada con $n - 1$ grados de libertad, que dejan áreas de $1 - \alpha/2$ y $\alpha/2$, respectivamente, a la derecha. Al sustituir para X^2 , escribimos

$$P\left[\chi^2_{1-\alpha/2} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}\right] = 1 - \alpha.$$

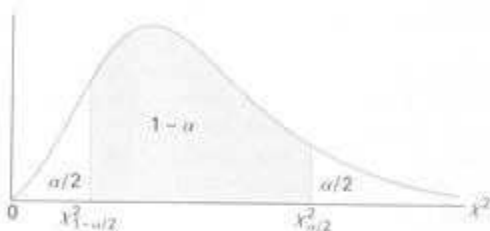


Figura 9.7 $P(\chi^2_{1-\alpha/2} < X^2 < \chi^2_{\alpha/2}) = 1 - \alpha$.

Al dividir cada término de la desigualdad entre $(n - 1)S^2$, y después invertir cada término (lo que cambia el sentido de las desigualdades), obtenemos

$$P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right] = 1 - \alpha.$$

Para nuestra muestra aleatoria particular de tamaño n , se calcula la varianza muestral s^2 y se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% para σ^2 .

Intervalo de confianza para σ^2

Si s^2 es la varianza de una muestra aleatoria de tamaño n de una población normal, un intervalo de confianza de $(1 - \alpha)$ 100% para σ^2 es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}},$$

donde $\chi^2_{\alpha/2}$ y $\chi^2_{1-\alpha/2}$ son valores χ^2 con $v = n - 1$ grados de libertad, que dejan áreas de $\alpha/2$ y $1 - \alpha/2$, respectivamente, a la derecha.

Un intervalo de confianza del $(1 - \alpha)$ 100% para σ se obtiene al tomar la raíz cuadrada de cada extremo del intervalo para σ^2 .

Ejemplo 9.14 Los siguientes son los pesos, en decagramos, de 10 paquetes de semillas de pasto distribuidas por cierta compañía: 46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2 y 46.0. Encuentre un intervalo de confianza de 95% para la varianza de todos los paquetes de semillas de pasto que distribuye esta compañía, suponga una población normal.

SOLUCIÓN

Primero encontramos

$$s^2 = \frac{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}{n(n-1)} = \frac{(10)(21,273.12) - (461.2)^2}{(10)(9)} = 0.286.$$

Para obtener un intervalo de confianza de 95%, elegimos $\alpha = 0.05$. Después, con el uso de la tabla A.5 con $\nu = 9$ grados de libertad, encontramos $\chi_{0.025}^2 = 19.023$ y $\chi_{0.975}^2 = 2.700$. Por tanto, el intervalo de confianza de 95% para σ^2 es

$$\frac{(9)(0.286)}{19.023} < \sigma^2 < \frac{(9)(0.286)}{2.700} \quad \text{o simplemente } 0.135 < \sigma^2 < 0.953. \quad \blacksquare$$

9.12 Dos muestras: estimación de la razón de dos varianzas

Una estimación puntual de la razón de dos varianzas poblacionales σ_1^2/σ_2^2 está dada por la razón s_1^2/s_2^2 de las varianzas muestrales. De aquí, la estadística S_1^2/S_2^2 se denomina estimador de σ_1^2/σ_2^2 .

Si σ_1^2 y σ_2^2 son las varianzas de poblaciones normales, podemos establecer una estimación por intervalos de σ_1^2/σ_2^2 mediante el uso de la estadística

$$F = \frac{\sigma_1^2 S_1^2}{\sigma_2^2 S_2^2}$$

De acuerdo con el teorema 8.8, la variable aleatoria F tiene una distribución F con $\nu_1 = n_1 - 1$ y $\nu_2 = n_2 - 1$ grados de libertad. Por tanto, podemos escribir (véase la figura 9.8)

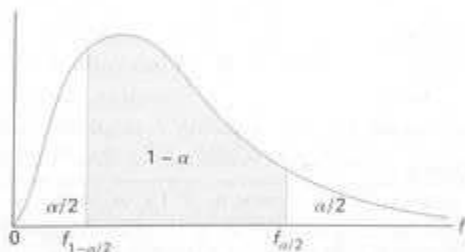


Figura 9.8 $P\{f_{1-\alpha/2}(\nu_1, \nu_2) < F < f_{\alpha/2}(\nu_1, \nu_2)\} = 1 - \alpha$.

$$P\{f_{1-\alpha/2}(\nu_1, \nu_2) < F < f_{\alpha/2}(\nu_1, \nu_2)\} = 1 - \alpha,$$

donde $f_{1-\alpha/2}(\nu_1, \nu_2)$ y $f_{\alpha/2}(\nu_1, \nu_2)$ son los valores de la distribución F con ν_1 y ν_2 grados de libertad, que dejan áreas de $1 - \alpha/2$ y $\alpha/2$, respectivamente, a la derecha. Al sustituir para F , escribimos

$$P\left\{f_{1-\alpha/2}(\nu_1, \nu_2) < \frac{\sigma_1^2 S_1^2}{\sigma_2^2 S_2^2} < f_{\alpha/2}(\nu_1, \nu_2)\right\} = 1 - \alpha.$$

Al multiplicar cada término en la desigualdad por S_2^2/S_1^2 , y después invertir cada término (de nuevo para cambiar el sentido de las desigualdades), obtenemos

$$P\left[\frac{S_1^2}{S_2^2} \frac{1}{f_{\alpha/2}(\nu_1, \nu_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{f_{1-\alpha/2}(\nu_1, \nu_2)}\right] = 1 - \alpha.$$

Los resultados del teorema 8.7 nos permiten reemplazar la cantidad $f_{1-\alpha/2}(\nu_1, \nu_2)$ por $1/f_{\alpha/2}(\nu_2, \nu_1)$. Por tanto,

$$P\left[\frac{S_1^2}{S_1^2 f_{\alpha/2}(\nu_1, \nu_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\alpha/2}(\nu_2, \nu_1)\right] = 1 - \alpha.$$

Para cualesquiera dos muestras aleatorias independientes de tamaño n_1 y n_2 que se seleccionan de dos poblaciones normales, la razón de las varianzas muestrales s_1^2/s_2^2 se calcula y se obtiene el siguiente intervalo de confianza de $(1 - \alpha)$ 100% para σ_1^2/σ_2^2 .

Intervalo de confianza para σ_1^2/σ_2^2

Si s_1^2 y s_2^2 son las varianzas de muestras independientes de tamaño n_1 y n_2 , respectivamente, de poblaciones normales, entonces un intervalo de confianza de $(1 - \alpha)$ 100% para σ_1^2/σ_2^2 es

$$\frac{s_1^2}{s_2^2} \frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\alpha/2}(v_2, v_1),$$

donde $f_{\alpha/2}(v_1, v_2)$ es un valor f con $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$ grados de libertad que deja un área de $\alpha/2$ a la derecha $f_{\alpha/2}(v_2, v_1)$ es un valor f similar con $v_2 = n_2 - 1$ y $v_1 = n_1 - 1$ grados de libertad.

Como en la sección 9.11, se obtiene un intervalo de confianza de $(1 - \alpha)$ 100% para σ_1/σ_2 al tomar la raíz cuadrada de cada extremo del intervalo para σ_1^2/σ_2^2 .

Ejemplo 9.15 En el ejemplo 9.8 de la página 258 se construyó un intervalo de confianza para la diferencia en el contenido medio de ortofósforo, que se mide en miligramos por litro, en dos estaciones sobre el río James mediante la suposición de que las varianzas normales de la población son diferentes. Justifique esta suposición mediante la construcción de un intervalo de confianza del 98% para σ_1^2/σ_2^2 y para σ_1/σ_2 , donde σ_1^2 y σ_2^2 son las varianzas poblacionales del contenido de ortofósforo en la estación 1 y en la estación 2, respectivamente.

SOLUCIÓN

Del ejemplo 9.8, tenemos $n_1 = 15$, $n_2 = 12$, $s_1 = 3.07$, y $s_2 = 0.80$. Para un intervalo de confianza de 98%, $\alpha = 0.02$. Al interpolar en la tabla A.6, encontramos $f_{0.01}(14, 11) = 4.30$ y $f_{0.01}(11, 14) = 3.87$. Por tanto, el intervalo de confianza de 98% para σ_1^2/σ_2^2 es

$$\frac{3.07^2}{0.80^2} \left(\frac{1}{4.30} \right) < \frac{\sigma_1^2}{\sigma_2^2} < \frac{3.07^2}{0.80^2} (3.87),$$

que se simplifica a $3.425 < \frac{\sigma_1^2}{\sigma_2^2} < 56.991$.

Al calcular las raíces cuadradas de los límites de confianza, encontramos que un intervalo de confianza de 98% para σ_1/σ_2 es

$$1.851 < \frac{\sigma_1}{\sigma_2} < 7.549.$$

Como este intervalo no permite la posibilidad de que σ_1/σ_2 sea igual a 1, es correcto suponer que $\sigma_1 \neq \sigma_2$ o $\sigma_1^2 \neq \sigma_2^2$ en el ejemplo 9.8.

Hasta este momento todos los intervalos de confianza presentados son de la forma

estimación puntual $\pm K$ s.e. (estimación puntual),

donde K es una constante (ya sea t o punto porcentual de la normal). Éste es el caso cuando el parámetro es una media, diferencia entre medias, proporción o diferencia entre proporciones. Sin embargo, esto no se aplica a varianzas y razones de varianzas. ■

Ejercicios

1. Un fabricante de baterías para automóvil afirma que sus baterías durarán, en promedio, tres años con una varianza de un año. Si cinco de estas baterías tienen duraciones de 1.9, 2.4, 3.0, 3.5 y 4.2 años, construya un intervalo de confianza de 95% para σ^2 y decida si la afirmación del fabricante de que $\sigma^2 = 1$ es válida. Suponga que la población de duraciones de las baterías se distribuye de forma aproximadamente normal.
2. Una muestra aleatoria de 20 estudiantes obtiene una media de $\bar{x} = 72$ y una varianza de $s^2 = 16$ en un examen de colocación en matemáticas. Suponga que las calificaciones se distribuyen normalmente y construya un intervalo de confianza de 98% para σ^2 .
3. Construya un intervalo de confianza de 95% para σ en el ejercicio 12 de la página 252.
4. Construya un intervalo de confianza de 99% para σ^2 en el ejercicio 13 de la página 252.
5. Construya un intervalo de confianza de 99% para σ en el ejercicio 14 de la página 252.
6. Construya un intervalo de confianza de 90% para σ^2 en el ejercicio 15 de la página 252.
7. Construya un intervalo de confianza de 98% para σ_1/σ_2 en el ejercicio 8 de la página 263, donde σ_1 y σ_2 son, respectivamente, las desviaciones estándar para las distancias que se obtienen por litro de combustible en los camiones compactos Volkswagen y Toyota.
8. Construya un intervalo de confianza de 90% para σ_1^2/σ_2^2 en el ejercicio 9 de la página 263. ¿Estamos justificados al suponer que $\sigma_1^2 = \sigma_2^2$ cuando construimos nuestro intervalo de confianza para $\mu_1 - \mu_2$?
9. Construya un intervalo de confianza del 90% para σ_1^2/σ_2^2 en el ejercicio 10 de la página 263. ¿Debimos suponer $\sigma_1^2 = \sigma_2^2$ al construir nuestro intervalo de confianza para $\mu_2 - \mu_1$?

9.13 Métodos bayesianos de estimación

Los métodos clásicos de estimación que estudiamos hasta ahora se basan sólo en información que proporciona la muestra aleatoria. Estos métodos interpretan esencialmente probabilidades como frecuencias relativas. Por ejemplo, para obtener un intervalo de confianza de 95% para μ , interpretamos el planteamiento

$$P(-1.96 < Z < 1.96) = 0.95$$

como que 95% de las veces en experimentos repetidos Z caerá entre -1.96 y 1.96 . Las probabilidades de este tipo que se pueden interpretar en el sentido de la frecuencia se denominarán como **probabilidades objetivas**. El enfoque bayesiano de los métodos estadísticos de estimación combina información muestral con otra información anterior disponible que pueda parecer pertinente.

Probabilidad subjetiva

Considere el problema de encontrar una estimación puntual del parámetro θ para la población $f(x; \theta)$. El enfoque clásico sería tomar una muestra aleatoria de tamaño n y sustituir la información que proporciona la muestra en el estimador apropiado o función de decisión. Así, en el caso de una población binomial $b(x; n, p)$ nuestra estimación de p , la proporción de éxitos, sería $\hat{p} = x/n$. Ahora bien, supon-

ga que se proporciona información adicional acerca de θ , a saber, que se sabe que varía de acuerdo con alguna distribución de probabilidad $f(\theta)$, que a menudo se denomina **distribución a priori**, con **media a priori** μ_0 y **varianza a priori** σ_0^2 . Es decir, suponemos ahora que θ es un valor de una variable aleatoria Θ con distribución de probabilidad $f(\theta)$ y deseamos estimar el valor particular θ para la población de la que seleccionamos nuestra muestra aleatoria. Las probabilidades que se asocian con esta distribución a priori se denominan **probabilidades subjetivas**, pues miden el grado de creencia de una persona en la posición de un parámetro. La persona utiliza su propia experiencia y conocimiento como la base para llegar a las probabilidades subjetivas dadas por la distribución a priori. Las técnicas bayesianas utilizan la distribución a priori $f(\theta)$ junto con la distribución conjunta de la muestra $f(x_1, x_2, \dots, x_n, \theta)$ para calcular la **distribución a posteriori** $f(\theta|x_1, x_2, \dots, x_n)$. La distribución a posteriori consiste en información de la distribución a priori subjetiva y de la distribución muestral objetiva y expresa nuestro grado de creencia en la posición del parámetro θ después que observamos la muestra.

En lo que sigue denotaremos con $f(x_1, x_2, \dots, x_n, \theta)$ la distribución de probabilidad conjunta de la muestra, condicional sobre el parámetro θ , en una situación donde θ es una variable aleatoria. La distribución conjunta de la muestra X_1, X_2, \dots, X_n y el parámetro Θ es entonces

$$f(x_1, x_2, \dots, x_n, \theta) = f(x_1, x_2, \dots, x_n | \theta) f(\theta),$$

de donde obtenemos con facilidad la distribución marginal

$$g(x_1, x_2, \dots, x_n) = \begin{cases} \sum_{\theta} f(x_1, x_2, \dots, x_n, \theta) & \text{(caso discreto)} \\ \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n, \theta) d\theta & \text{(caso continuo).} \end{cases}$$

De aquí la distribución a posteriori se puede escribir como

$$f(\theta|x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n, \theta)}{g(x_1, x_2, \dots, x_n)}.$$

Definición 9.3

La media de la distribución a posteriori $f(\theta|x_1, x_2, \dots, x_n)$, denotada por θ^* , se llama **estimación de Bayes** de θ .

Ejemplo 9.16 Supongamos que la distribución a priori para la proporción p de artículos defectuosos que produce una máquina es

p	0.1	0.2
$f(p)$	0.6	0.4

Encuentre la estimación de Bayes para la proporción de defectuosos que produce esta máquina si una muestra aleatoria de tamaño 2 da 1 artículo defectuoso.

SOLUCIÓN

Sea X el número de artículos defectuosos en nuestra muestra. Entonces la distribución de probabilidad para nuestra muestra es

$$f(x|p) = b(x; n, p) = \binom{2}{x} p^x q^{2-x}, \quad x = 0, 1, 2.$$

Si $p = 0.1$, la probabilidad de que la muestra aleatoria de tamaño 2 dé un artículo defectuoso es

$$f(1|0.1) = b(1; 2, 0.1) = \binom{2}{1} (0.1)(0.9) = 0.18.$$

De manera similar, cuando $p = 0.2$, encontramos que

$$f(1|0.2) = b(1; 2, 0.2) = \binom{2}{1} (0.2)(0.8) = 0.32.$$

Del hecho de que $f(x, p) = f(x|p)f(p)$, podemos elaborar la tabla

p	0.1	0.2
$f(1, p)$	0.108	0.128

de la que obtenemos

$$g(1) = \sum_p f(1, p) = 0.236.$$

Obtenemos la distribución a posteriori para la proporción de artículos defectuosos, cuando $x = 1$, a partir de la fórmula $f(p|x) = f(x, p)/g(x)$. De aquí tenemos

p	0.1	0.2
$f(p x=1)$	0.458	0.542

de donde obtenemos que la estimación de Bayes de p , que se denota con p^* , es

$$p^* = (0.1)(0.458) + (0.2)(0.542) = 0.1542.$$

Nótese que la estimación de Bayes es mucho más pequeña que el valor $1/2$ dado por la estimación clásica $\hat{p} = x/n$. ■

Ejemplo 9.17 Repita el ejemplo 9.16 con el uso de la distribución a priori uniforme $f(p) = 1, 0 < p < 1$.

SOLUCIÓN

Como antes, tenemos

$$f(x|p) = \binom{2}{x} p^x p^{2-x}, \quad x = 0, 1, 2.$$

Del hecho de que $f(x, p) = f(x|p)f(p)$, podemos escribir

$$f(1, p) = \binom{2}{1} pq = 2p(1-p), \quad 0 < p < 1$$

y entonces

$$g(1) = \int_0^1 2p(1-p) dp = \frac{1}{3}.$$

La distribución a posteriori para la proporción de artículos defectuosos, p , cuando $x = 1$, es entonces

$$f(p|x=1) = \frac{2p(1-p)}{\frac{1}{3}} = 6p(1-p), \quad 0 < p < 1,$$

de la que obtenemos que la estimación de Bayes de p es

$$p^* = 6 \int_0^1 p^2(1-p) dp = \frac{1}{2}.$$

En este caso vemos que la estimación de Bayes p^* y la estimación clásica \bar{p} son equivalentes.

Se puede construir un intervalo bayesiano de $(1 - \alpha)$ 100% para el parámetro θ al encontrar un intervalo centrado en la media a posteriori que contenga $(1 - \alpha)$ 100% de la probabilidad a posteriori. ■

Definición 9.4

El intervalo $a < \theta < b$ se llamará intervalo de Bayes de $(1 - \alpha)$ 100% para θ si

$$\int_a^b f(\theta|x_1, x_2, \dots, x_n) d\theta = \int_a^{\theta^*} f(\theta|x_1, x_2, \dots, x_n) d\theta = \frac{1 - \alpha}{2}.$$

Los métodos bayesianos de estimación que involucran a la media μ de una población normal se basan en el siguiente teorema.

Teorema 9.6

Si \bar{x} es la media de una muestra aleatoria de tamaño n de una población normal con varianza σ^2 conocida, y la distribución a priori de la media poblacional es una distribución normal con media μ_0 y varianza σ_0^2 , entonces la distribución a posteriori de la media poblacional también es una distribución normal con media μ^* y desviación estándar σ^* , donde

$$\mu^* = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2} \quad \text{y} \quad \sigma^* = \sqrt{\frac{\sigma_0^2\sigma^2}{n\sigma_0^2 + \sigma^2}}.$$

PRUEBA Al multiplicar la densidad de nuestra muestra

$$f(x_1, x_2, \dots, x_n|\mu) = \frac{1}{(2\pi)^{n/2} \cdot \sigma^n} \exp\left[-\left(\frac{1}{2}\right) \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2\right],$$

para $-\infty < x_i < \infty$ y $i = 1, 2, \dots, n$ por nuestra distribución a priori

$$f(\mu) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{-\left(\frac{1}{2}\right) \left[\frac{\mu - \mu_0}{\sigma_0}\right]^2\right\}, \quad -\infty < \mu < \infty,$$

obtenemos la densidad conjunta de la muestra aleatoria y la media de la población de la que se selecciona la muestra. Es decir,

$$f(x_1, x_2, \dots, x_n, \mu) = \frac{1}{(2\pi)^{(n+1)/2} \cdot \sigma^n \sigma_0} \times \exp \left\{ -\left(\frac{1}{2}\right) \left[\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2 + \left(\frac{\mu - \mu_0}{\sigma_0}\right)^2 \right] \right\}.$$

En la sección 8.4 establecimos la identidad

$$\sum_{i=1}^n (x_i - \mu)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2,$$

que nos permite escribir

$$f(x_1, x_2, \dots, x_n, \mu) = \frac{1}{(2\pi)^{(n+1)/2} \cdot \sigma^n \sigma_0} \exp \left[-\left(\frac{1}{2}\right) \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 \right] \times \exp \left\{ -\left(\frac{1}{2}\right) \left[\frac{n(\bar{x} - \mu)^2}{\sigma} \right]^2 + \left[\frac{(\mu - \mu_0)}{\sigma_0} \right]^2 \right\}.$$

Al completar cuadrados en el segundo exponente, podemos escribir la densidad conjunta de la muestra aleatoria y la media poblacional en la forma

$$f(x_1, x_2, \dots, x_n, \mu) = K \exp \left\{ -\left(\frac{1}{2}\right) \left[\frac{(\mu - \mu^*)}{\sigma^*} \right]^2 \right\},$$

donde

$$\mu^* = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}, \quad \sigma^* = \sqrt{\frac{\sigma_0^2\sigma^2}{n\sigma_0^2 + \sigma^2}}.$$

y K es una función de los valores muestrales y de los valores conocidos. La distribución marginal de la muestra es entonces

$$g(x_1, x_2, \dots, x_n) = K \sqrt{2\pi\sigma^*} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^*}} \exp \left\{ -\left(\frac{1}{2}\right) \left[\frac{(\mu - \mu^*)}{\sigma^*} \right]^2 \right\} d\mu \\ = K \sqrt{2\pi\sigma^*},$$

y la distribución a posteriori es

$$f(\mu|x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n, \mu)}{g(x_1, x_2, \dots, x_n)} \\ = \frac{1}{\sqrt{2\pi\sigma^*}} \exp \left\{ -\left(\frac{1}{2}\right) \left[\frac{(\mu - \mu^*)}{\sigma^*} \right]^2 \right\}, \quad -\infty < \mu < \infty,$$

que se identifica como distribución normal con media μ^* y desviación estándar σ^* , donde μ^* y σ^* se definen antes.

El teorema del límite central nos permite utilizar el teorema 9.6 también cuando seleccionamos las muestras aleatorias de tamaño $n \geq 30$ a partir de poblaciones no normales, y cuando la distribución a priori de la media es aproximadamente normal.

Para calcular μ^* y σ^* mediante las fórmulas del teorema 9.6 suponemos que σ se conoce. Como éste por lo general no es el caso, reemplazaremos σ^2 por la varianza muestral s^2 siempre que $n \geq 30$. La media a posteriori μ^* es la estimación de Bayes de la media poblacional μ , y se puede construir un **intervalo bayesiano** del $(1 - \alpha)$ 100% para μ mediante el cálculo del intervalo

$$\mu^* - z_{\alpha/2}\sigma^* < \mu < \mu^* + z_{\alpha/2}\sigma^*,$$

que se centra en la media a posteriori y contiene $(1 - \alpha)$ 100% de la probabilidad a posteriori. ■

Ejemplo 9.18 Una empresa eléctrica fabrica focos que tienen una duración que está distribuida de forma aproximadamente normal con una desviación estándar de 100 horas. Experiencia anterior nos conduce a creer que μ es un valor de una variable aleatoria normal con una media $\mu_0 = 800$ horas y una desviación estándar $\sigma_0 = 10$ horas. Si una muestra aleatoria de 25 focos tiene una duración promedio de 780 horas, encuentre un intervalo bayesiano de 95% para μ .

SOLUCIÓN

De acuerdo con el teorema 9.6, la distribución a posteriori de la media también es una distribución normal con media

$$\mu^* = \frac{(25)(780)(10)^2 + (800)(100)^2}{(25)(10)^2 + (100)^2} = 796$$

y desviación estándar

$$\sigma^* = \sqrt{\frac{(10)^2(100)^2}{(25)(10)^2 + (100)^2}} = \sqrt{80}.$$

El intervalo bayesiano de 95% para μ está dado entonces por

$$796 - 1.96\sqrt{80} < \mu < 796 + 1.96\sqrt{80}$$

o

$$778.5 < \mu < 813.5.$$

Si se ignora la información previa acerca de μ del ejemplo 9.18, podemos proceder como en la sección 9.4 y construir el intervalo de confianza clásico de 95%

$$780 - (1.96)\left(\frac{100}{\sqrt{25}}\right) < \mu < 780 + (1.96)\left(\frac{100}{\sqrt{25}}\right)$$

o

$$740.8 < \mu < 819.2.$$

que se observa que es más amplio que el intervalo bayesiano correspondiente. ■

Ejercicios

1. Estime la proporción de defectuosos que produce la máquina del ejemplo 9.16 si la muestra aleatoria de tamaño 2 da dos defectuosos.

2. Supongamos que la distribución a priori para la proporción p de bebidas de una máquina despachadora que las derrama es

p	0.05	0.10	0.15
$f(p)$	0.3	0.5	0.2

Si dos de las siguientes nueve bebidas de esta máquina se derraman, encuentre

- (a) la distribución posterior para la proporción p ;
 (b) la estimación de Bayes de p .

3. Repita el ejercicio 2 cuando una de las siguientes cuatro bebidas se derrama y la distribución uniforme a priori es

$$f(p) = 10, \quad 0.05 < p < 0.15,$$

4. El constructor de un nuevo complejo de condominios afirma que tres de cinco compradores preferirá un departamento de dos recámaras, mientras que su banquero afirma que sería más correcto decir que siete de 10 compradores preferirán una unidad de dos recámaras. En las predicciones a priori de este tipo el banquero ha sido dos veces más confiable que el constructor. Si 12 de los siguientes 15 condominios que se venden en este complejo son de dos recámaras, encuentre

- (a) las probabilidades a posteriori que se asocian con las afirmaciones del constructor y del banquero;
 (b) una estimación puntual de la proporción de compradores que prefieren un departamento de dos recámaras.

5. El tiempo en que se consume la primera etapa de un cohete es una variable aleatoria normal con una desviación estándar de 0.8 minutos. Suponga una distribución a priori normal para μ con una media de ocho minutos y una desviación estándar de 0.2 minutos. Si se lanzan 10 de estos cohetes y la primera etapa tiene un tiempo de consumo promedio de nueve minutos, encuentre un intervalo bayesiano de 95% para μ .

6. La utilidad diaria de una máquina despachadora de jugos que se coloca en un edificio de oficinas es un valor de una variable aleatoria normal con media μ y varianza σ^2 desconocidas. Por supuesto, la media va-

rá algo de edificio a edificio, y el distribuidor considera que estas utilidades promedio diarias se pueden describir mejor mediante una distribución normal con media $\mu_0 = \$30.00$ y desviación estándar $\sigma_0 = \$1.75$. Si una de estas máquinas despachadoras de jugo, que se coloca en cierto edificio, muestra una utilidad promedio diaria de $\bar{x} = \$24.90$ durante los primeros 30 días con una desviación estándar de $s = \$2.10$, encuentre

- (a) una estimación de Bayes de la utilidad promedio diaria real para este edificio;
 (b) un intervalo bayesiano de 96% de μ para este edificio;
 (c) la probabilidad de que la utilidad promedio diaria de la máquina en este edificio esté entre \$24.00 y \$26.00.

7. El departamento de matemáticas de una universidad grande diseña un examen de colocación para aplicarlo a los grupos de nuevo ingreso a primer año. Los miembros del departamento consideran que la calificación promedio para este examen variará de un grupo de primer año a otro. Esta variación de la calificación promedio del grupo se expresa de manera subjetiva mediante una distribución normal con una media $\mu_0 = 72$ y varianza $\sigma_0^2 = 5.76$.

- (a) ¿Qué probabilidad a priori de que la calificación promedio real caiga entre 71.8 y 73.4 asigna el departamento para los alumnos de los grupos de primer grado del siguiente año?
 (b) Si el examen se prueba en una muestra aleatoria de 100 estudiantes de primer grado del siguiente grupo de nuevo ingreso a primer grado que tiene como resultado una calificación promedio de 70 con una varianza de 64, construya un intervalo bayesiano de 95% para μ .
 (c) ¿Qué probabilidad a posteriori debe asignar el departamento al evento de la parte (a)?

8. Suponga que en el ejemplo 9.18 la empresa eléctrica no tiene suficiente información a priori con respecto a la duración media poblacional para ser capaz de suponer una distribución normal para μ . La empresa cree, sin embargo, que μ seguramente está entre 770 y 830 horas y considera que una aproximación bayesiana más realista sería suponer la distribución a priori

$$f(\mu) = \frac{1}{80}, \quad 770 < \mu < 830,$$

Si una muestra aleatoria de 25 focos da una vida promedio de 780 horas, siga los pasos de la demostración del teorema 9.6 para encontrar la distribución a posteriori

$$f(\mu | x_1, x_2, \dots, x_{25}).$$

9. Suponga que el tiempo de falla T para cierta bisagra es una variable aleatoria exponencial con densidad de probabilidad

$$f(t) = \theta e^{-\theta t}, \quad t > 0.$$

De cierta experiencia anterior nos inclinamos a pensar que θ es un valor de una variable aleatoria exponencial con densidad de probabilidad

$$f(\theta) = 2e^{-2\theta}, \quad \theta > 0.$$

Si tenemos una muestra de n observaciones de T , muestre que la distribución a posteriori de θ es una distribución gamma con parámetros

$$\alpha = n + 1 \quad \text{y} \quad \beta = \frac{1}{\left(\sum_{i=1}^n t_i + 2\right)}.$$

9.14 Estimación de probabilidad máxima

A menudo los estimadores de parámetros son los que recurren a la intuición. El estimador \bar{X} ciertamente parece razonable como estimador de una media poblacional μ . La virtud de S^2 como estimador de σ^2 se subraya en la discusión de estimadores insesgados de la sección 9.3. El estimador para un parámetro binomial p es simplemente una proporción de la muestra, que por supuesto es un promedio y recurre al sentido común. Pero hay muchas situaciones en las que no es del todo obvio cuál debería ser el estimador apropiado. Como resultado, el estudiante de estadística tiene mucho por aprender con respecto a las diferentes filosofías que producen diversos métodos de estimación. En esta sección tratamos el **método de probabilidad máxima**.

La estimación por probabilidad máxima representa una de las aproximaciones a la estimación más importantes en toda la inferencia estadística. No haremos un desarrollo completo del método. En su lugar, intentaremos comunicar la filosofía de la probabilidad máxima y la ilustraremos con ejemplos que la relacionan con otros problemas de estimación que se discuten en este capítulo.

Función de probabilidad

Como el nombre lo indica, el método de probabilidad máxima es aquel para el que se maximiza la *función de probabilidad*. La función de probabilidad se ilustra mejor a través del uso de un ejemplo con una distribución discreta y un solo parámetro. Sean X_1, X_2, \dots, X_n variables aleatorias independientes tomadas de una distribución de probabilidad representada por $f(X, \theta)$, donde θ es un solo parámetro de la distribución. Ahora bien

$$\begin{aligned} L(X_1, X_2, \dots, X_n; \theta) &= f(X_1, X_2, \dots, X_n; \theta) \\ &= f(X_1, \theta) \cdot f(X_2, \theta) \cdot \dots \cdot f(X_n, \theta) \end{aligned}$$

es la *distribución conjunta de las variables aleatorias*. Esto a menudo se denomina función de probabilidad. Sean x_1, x_2, \dots, x_n valores observados en una muestra. En el caso de una variable aleatoria discreta la interpretación es muy clara. La cantidad $L(x_1, x_2, \dots, x_n; \theta)$, *la probabilidad de la muestra*, es la siguiente probabilidad conjunta:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n).$$

En otras palabras es la probabilidad de obtener los valores muestrales x_1, x_2, \dots, x_n . Para el caso discreto el estimador de probabilidad máxima es el que tiene como resultado un valor máximo para esta probabilidad conjunta, o que maximiza la probabilidad de la muestra.

Considérese un ejemplo ficticio donde se inspeccionan tres artículos de una línea de ensamblaje. Los artículos se clasifican como defectuosos o no defectuosos y entonces se aplica el proceso de Bernoulli. La inspección de los tres artículos tiene como resultado dos artículos no defectuosos seguidos por uno defectuoso. Es de interés estimar p , la proporción de no defectuosos en el proceso. La probabilidad de la muestra para esta ilustración está dada por

$$p \cdot p \cdot q = p^2 q = p^2 - p^3,$$

donde $q = 1 - p$. La estimación de probabilidad máxima daría una estimación de p para la que se maximiza la probabilidad. Es claro que si diferenciamos la probabilidad con respecto a p , hacemos la derivada igual a cero, y la resolvemos, obtenemos el valor

$$\hat{p} = \frac{2}{3}$$

Ahora bien, por supuesto, en esta situación $\hat{p} = 2/3$ es la proporción muestral de defectuosos y por ello un estimador razonable de la probabilidad de un defectuoso. El lector debe intentar comprender que la filosofía de la estimación de probabilidad máxima se desprende de la noción de que el estimador razonable de un parámetro que se basa en información muestral es el valor del parámetro que produce la mayor probabilidad de obtener la muestra. Ésta es, en realidad, la interpretación para el caso discreto, pues la probabilidad es la probabilidad de observar de manera conjunta los valores en la muestra.

Ahora bien, mientras que la interpretación de la función de probabilidad como una probabilidad conjunta se reduce al caso discreto, la noción de probabilidad máxima se extiende a la estimación de parámetros de una distribución continua. Presentamos ahora una definición formal de la estimación de probabilidad máxima.

Definición 9.5

Dadas las observaciones independientes x_1, x_2, \dots, x_n de una función de densidad de probabilidad (caso continuo) o de una función de masa de probabilidad (caso discreto) $f(x, \theta)$ el estimador de probabilidad máxima $\hat{\theta}$ es el que maximiza la función de probabilidad

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta).$$

Muy a menudo conviene trabajar con el logaritmo natural de la función de probabilidad para encontrar el máximo de la misma. Considere el siguiente ejemplo acerca del parámetro μ de una distribución de Poisson.

Ejemplo 9.19 Considere una distribución de Poisson con función de masa de probabilidad

$$f(x, \mu) = \frac{e^{-\mu} \mu^x}{x!}, \quad x = 0, 1, 2, \dots$$

Suponga que se toma una muestra aleatoria x_1, x_2, \dots, x_n de la distribución. ¿Cuál es la estimación de probabilidad máxima de μ ?

SOLUCIÓN

La función de probabilidad es

$$L(x_1, x_2, \dots, x_n, \mu) = \prod_{i=1}^n f(x_i, \mu) = \frac{e^{-n\mu} \mu^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

Considere ahora

$$\ln L(x_1, x_2, \dots, x_n, \mu) = -n\mu + \sum_{i=1}^n x_i \ln \mu - \ln \prod_{i=1}^n x_i!$$

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n, \mu)}{\partial \mu} = -n + \sum_{i=1}^n \frac{x_i}{\mu}.$$

Al resolver para $\hat{\mu}$, el estimador de probabilidad máxima implica hacer la derivada igual a cero y resolver para el parámetro. De esta forma

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}.$$

Como μ es la media de la distribución de Poisson (capítulo 5), el promedio muestral en realidad parecería ser un estimador razonable. ■

El siguiente ejemplo muestra el uso del método de probabilidad máxima para encontrar estimaciones de dos parámetros. Simplemente encontramos los valores de los parámetros que maximizan (de manera conjunta) la función de probabilidad.

Ejemplo 9.20 Considere una muestra aleatoria x_1, x_2, \dots, x_n de una distribución normal $N(\mu, \sigma^2)$. Encuentre los estimadores de probabilidad máxima para μ y σ^2 .

SOLUCIÓN

La función de probabilidad para la distribución normal es

$$L(x_1, x_2, \dots, x_n, \mu, \sigma^2) = \frac{1}{(2\pi)^{n/2} (\sigma^2)^{n/2}} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2 \right].$$

Tomar logaritmos da

$$\ln L(x_1, x_2, \dots, x_n, \mu, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^2$$

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2}.$$

Al igualar la derivada a cero, obtenemos

$$\sum_{i=1}^n x_i - n\mu = 0.$$

De esta manera el estimador de probabilidad máxima está dado por

$$\hat{\mu} = \sum_{i=1}^n \frac{x_i}{n} = \bar{x}.$$

Este resultado es satisfactorio pues \bar{x} juega un importante papel en este capítulo como estimador puntual de μ . ■

Para el caso del estimador de σ^2 , el cálculo directo y la manipulación algebraica da

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}.$$

Resulta interesante notar la distinción entre el estimador de probabilidad máxima de σ^2 y el estimador insesgado S^2 que se desarrolló al principio de este capítulo. El numerador es idéntico, por supuesto, y el denominador son los "grados de libertad" $n - 1$ para el estimador insesgado y n para el estimador de probabilidad máxima. Los estimadores de probabilidad máxima no necesariamente gozan de la propiedad de ser insesgados. Sin embargo, los estimadores de probabilidad máxima tienen importantes propiedades asintóticas. Esto se discute de forma breve en la siguiente sección.

Ejemplo 9.21 Suponga que se utilizan 15 ratas en un estudio biomédico donde se inyecta a las ratas un medicamento contra el cáncer diseñado para aumentar su tasa de sobrevivencia. Los tiempos de sobrevivencia, en meses, son 14, 17, 27, 18, 12, 8, 22, 13, 19 y 12. Suponga que se aplica la distribución exponencial.

SOLUCIÓN

Del capítulo 6 sabemos que la función de densidad de probabilidad para la variable aleatoria exponencial x es

$$f(x, \lambda) = \frac{1}{\beta} e^{-x/\beta}, \quad x > 0$$

$$= 0, \quad \text{en cualquier otro caso,}$$

y por ello la probabilidad de los datos solos ($n = 10$).

$$L(x_1, x_2, \dots, x_{10}) = -10 \ln \beta - \frac{1}{\beta} \sum x_i.$$

$$\frac{\alpha \ln L}{\alpha \beta} = \frac{-10}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^{10} x_i$$

y así $\frac{\alpha \ln L}{\alpha \beta} = 0$ implica que

$$\hat{\beta} = \frac{\sum_{i=1}^{10} x_i}{10} = \frac{14 + 17 + 27 + 18 + 12 + 8 + 22 + 13 + 19 + 12}{10} = 16.2$$

Como resultado el estimador del parámetro β es el promedio muestral \bar{x} . ■

Comentarios adicionales con respecto a la estimación por probabilidad máxima

Una discusión completa de las propiedades de la estimación por probabilidad máxima está fuera del alcance de este libro y por lo general es un tema principal de un curso de teoría de inferencia estadística. El método de probabilidad máxima permite al analista hacer uso del conocimiento de la distribución para determinar un estimador apropiado. *El método de probabilidad máxima no se puede aplicar sin el conocimiento de la distribución subyacente.* Aprendimos en el ejemplo 9.20 que el estimador de probabilidad máxima no necesariamente es insesgado. El estimador de probabilidad máxima es insesgado en el límite; es decir, la magnitud del sesgo se aproxima a cero conforme la muestra crece más. Al principio de este capítulo se discutió la noción de eficiencia, que se vincula con la propiedad de varianza de un estimador. Los estimadores de probabilidad máxima poseen propiedades de varianza deseables en el límite. El lector debe consultar a Lehman o a Bickel y Doksum para más detalles.

Ejercicios

1. Suponga que hay n pruebas x_1, x_2, \dots, x_n de un proceso de Bernoulli con parámetro p , la probabilidad de un éxito. Es decir, la probabilidad de r éxitos está dada por $\binom{n}{r} p^r q^{n-r}$. Calcule el estimador de probabilidad máxima para el parámetro p .

2. Considere una muestra de x_1, x_2, \dots, x_n , observaciones de una distribución de Weibull con parámetros α y β y función de densidad

$$f(x) = \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}, \quad x > 0$$

$$= 0, \quad \text{en cualquier otro caso}$$

para $\alpha, \beta > 0$.

(a) Escriba la función de probabilidad.
 (b) Escriba las ecuaciones que al resolverse dan los estimadores de probabilidad máxima de α y β .

3. Considere la distribución logarítmica normal con la función de densidad dada en la sección 6.9. Suponga que tenemos una muestra x_1, x_2, \dots, x_n de una distribución logarítmica normal.

(a) Escriba la función de probabilidad.
 (b) Desarrolle los estimadores de probabilidad máxima de μ y σ^2 .

4. Considere las observaciones x_1, x_2, \dots, x_n de la distribución gamma que se discutió en la sección 6.7.
 (a) Escriba la función de probabilidad.
 (b) Escriba un conjunto de ecuaciones que cuando se resuelvan den los estimadores de probabilidad máxima de α y β .

5. Considere un experimento hipotético donde un hombre con un hongo utiliza un medicamento fungicida y se cura. Considere éste, entonces, como una muestra de uno de una distribución de Bernoulli con función de probabilidad

$$f(x) = p^x q^{1-x}, \quad x = 0, 1,$$

donde p es la probabilidad de un éxito (curación) y $q = 1 - p$. Ahora, por supuesto, la información muestral da $x = 1$. Escriba un desarrollo que muestre que $\hat{p} = 1.0$ es el estimador de probabilidad máxima de la probabilidad de curación.

Ejercicios de repaso

1. Considere dos estimadores de σ^2 en la muestra x_1, x_2, \dots, x_n que se extrae de una distribución normal con media μ y varianza σ^2 . Los estimadores son

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1} \quad \text{estimador insesgado}$$

$$\hat{\sigma}^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n} \quad \text{estimador de probabilidad máxima}$$

Discuta las propiedades de la varianza de estos dos estimadores.

2. Se afirma que una nueva dieta reducirá el peso de una persona en 4.5 kilogramos en promedio en un periodo de dos semanas. Los pesos de siete mujeres que siguieron esta dieta se registraron antes y después de un periodo de dos semanas.

	Mujer						
	1	2	3	4	5	6	7
Peso antes	58.5	60.3	61.7	69.0	64.0	62.6	56.7
Peso después	60.0	54.9	58.1	62.1	58.5	59.9	54.4

Pruebe la afirmación del fabricante mediante el cálculo de un intervalo de confianza de 95% para la diferencia media en el peso. Suponga que las diferencias de los pesos se distribuyen de forma aproximadamente normal.

3. De acuerdo con el *Roanoke Times* (16 de marzo de 1997), la empresa McDonald's vendió 42.1% del mercado de hamburguesas. Una muestra aleatoria de 75 hamburguesas vendidas tiene como resultado 28 de ellas vendidas por McDonald's. Utilice el material de la sección 9.9 para determinar si esta información apoya la afirmación del *Roanoke Times*.

4. Un gimnasio afirma que un nuevo programa de ejercicio reducirá la talla de la cintura de una persona en dos centímetros en promedio en un periodo de cinco días. Las tallas de cinturas de seis hombres que participaron en este programa de ejercicio se registraron antes y después del periodo de cinco días en la siguiente tabla:

	Hombre					
	1	2	3	4	5	6
Talla de cintura antes	90.4	95.5	98.7	115.9	104.0	85.6
Talla de cintura después	91.7	93.9	97.4	112.8	101.3	84.0

Mediante el cálculo de un intervalo de confianza del 95% para la reducción media de la talla de la cintura, determine si la afirmación del gimnasio es válida. Suponga que la distribución de las diferencias de tallas de cintura antes y después del programa es aproximadamente normal.

5. Se llevó a cabo un estudio en el Instituto Politécnico y Universidad Estatal de Virginia para determinar si el fuego se puede utilizar como una herramienta de control viable para aumentar la cantidad de forraje disponible para los venados durante los meses críticos a finales de invierno y principios de primavera. El calcio es un elemento que requieren las plantas y los animales. La cantidad que la planta toma y almacena está íntimamente correlacionada con la cantidad presente en el suelo. Se formuló la hipótesis de que el fuego puede cambiar los niveles de calcio presentes en el suelo y afectar de esta forma la cantidad disponible para los venados. Se seleccionó una superficie grande de tierra en el Fishburn Forest para un incendio controlado. Se tomaron muestras de suelo de 12 parcelas de igual área justo antes de la quema, y se analizaron para verificar el contenido de calcio. Los niveles de calcio después de la quema se analizaron de las mismas parcelas. Esos valores, en kilogramos por parcela, se presentan en la siguiente tabla:

Parcela	Nivel de calcio (kg/parcela)	
	Antes de la quema	Después de la quema
1	50	9
2	50	18
3	82	45
4	64	18
5	82	18
6	73	9
7	77	32
8	54	9
9	23	18
10	45	9
11	36	9
12	54	9

Construya un intervalo de confianza de 95% para la diferencia media en el nivel de calcio presente en el suelo antes y después del incendio controlado. Suponga que la distribución de las diferencias de los niveles de calcio es aproximadamente normal.

6. El departamento de ingeniería civil del Instituto Politécnico y Universidad Estatal de Virginia compara una técnica de ensaye modificada (M-5 hr) para recuperar coliformes fecales en charcos de agua de lluvia en un área urbana con la técnica del número más probable (MPN). Se colecta un total de 12 muestras de charcos y se analizan con las dos técnicas. Los conteos de coliformes por 100 mililitros se registran en la siguiente tabla:

Muestra	Conteo MPN	Conteo con M-5 hr
1	2300	2010
2	1200	930
3	450	400
4	210	436
5	270	4100
6	450	2090
7	154	219
8	179	169
9	192	194
10	230	174
11	340	274
12	194	183

Construya un intervalo de confianza de 90% para la diferencia en los conteos medios de coliformes fecales entre las técnicas M-5 hr y MPN. Suponga que las diferencias de conteos se distribuyen de forma aproximadamente normal.

7. Un antropólogo se interesa en la proporción de individuos de dos tribus indias con doble remolino de cabello en la parte occipital del cráneo. Suponga que se toman muestras independientes de cada una de las dos tribus, y que se encuentra que 24 de 100 individuos de la tribu A y 36 de 120 individuos de la tribu B poseen esta característica. Construya un intervalo de confianza de 95% para la diferencia $p_B - p_A$ entre las proporciones de estas dos tribus con remolinos de cabello en la parte occipital de la cabeza.

8. Se lleva a cabo un experimento para determinar si el acabado superficial tiene un efecto sobre el límite de aguante del acero. Una teoría existente dice que el pulido aumenta el límite de aguante promedio (flexión inversa). El experimento se realiza sobre acero al carbón al 0.4% con el uso de especímenes sin pulir y con pulido suave. El acabado sobre los especímenes con pulido suave se obtiene al pulir con lijas del número 0 y del 00. Los datos son los siguientes:

Límite de aguante (psi) para:	
Acero al carbón al 0.4% pulido	Acero al carbón al 0.4% sin pulir
85,500	82,600
91,900	82,400
89,400	81,700
84,000	79,500
89,900	79,400
78,700	69,800
87,500	79,900
83,100	83,400

Desde un punto de vista práctico, el pulido no debería tener ningún efecto sobre la desviación estándar del límite de aguante, que a partir de los resultados de numerosos experimentos del límite de aguante se sabe que es de 4000 psi. Encuentre un intervalo de confianza de 95% para la diferencia entre las medias poblacionales para los dos métodos, suponga que las poblaciones se distribuyen de forma aproximadamente normal.

9. Un fabricante de planchas eléctricas produce estos artículos en dos plantas. Ambas plantas tienen a mismo proveedor de partes pequeñas. Se puede tener un ahorro al comprar termostatos para la planta B de un proveedor local. Se compra un solo lote del proveedor local y se desea probar si estos nuevos termostatos son tan precisos como los anteriores. Los termostatos se prueban en planchas a 550°F, y las temperaturas reales se redondean al siguiente 0.1°F con un termopar. Los datos son los siguientes:

Proveedor nuevo (°F):						
530.3	559.3	549.4	544.0	551.7	566.3	549.9
556.9	536.7	558.8	538.8	543.3	559.1	555.0
538.6	551.1	565.4	554.9	550.0	554.9	554.7
536.1	569.1					
Proveedor anterior (°F):						
559.7	534.7	554.8	545.0	544.6	538.0	550.7
563.1	551.1	553.8	538.8	564.6	554.5	553.0
538.4	548.3	552.9	535.1	555.0	544.8	558.4
548.7	560.3					

Encuentre un intervalo de confianza de 95% para σ_1^2/σ_2^2 y para σ_1/σ_2 , donde σ_1^2 y σ_2^2 son las varianzas poblacionales de las lecturas de los termostatos del proveedor nuevo y del anterior, respectivamente.

10. Se afirma que la resistencia del alambre A es mayor que la resistencia del alambre B. Un experimento sobre los alambres muestra los siguientes resultados (en ohmios):

Alambre A	Alambre B
0.140	0.135
0.138	0.140
0.143	0.136
0.142	0.142
0.144	0.138
0.137	0.140

Suponga varianzas iguales, ¿qué conclusiones extrae? Justifique su respuesta.

11. Una forma alterna de estimación se lleva a cabo a través del método de momentos. El método impli-

ca igualar la media y varianza poblacionales a la correspondiente media muestral \bar{x} y varianza muestral s^2 y resolver para el parámetro, el resultado son los estimadores del momento. En el caso de un solo parámetro, sólo se utilizan las medias. Dé un argumento de que en el caso de la distribución de Poisson el estimador de probabilidad máxima y los estimadores de momento son iguales:

12. Especifique los estimadores del momento para μ y σ^2 para la distribución normal.

13. Especifique los estimadores del momento para μ y σ^2 para la distribución logarítmica normal.

14. Especifique los estimadores del momento para α y β para la distribución gamma.

Pruebas de hipótesis de una y dos muestras

10.1 Hipótesis estadísticas: conceptos generales

A menudo, el problema al que se enfrenta el científico o el ingeniero no es tanto la estimación de un parámetro poblacional como se discutió en el capítulo 9, sino más bien la formación de un procedimiento de decisión que se base en los datos que pueda producir una conclusión acerca de algún sistema científico. Por ejemplo, un investigador médico puede decidir sobre la base de evidencia experimental si beber café aumenta el riesgo de cáncer en humanos; un ingeniero puede tener que decidir sobre la base de datos muestrales si hay una diferencia entre la precisión de dos tipos de medidores; o un sociólogo puede desear reunir los datos apropiados que le permitan decidir si el tipo sanguíneo de una persona y el color de los ojos son variables independientes. En cada uno de estos casos el científico o el ingeniero *postula* o *conjetura* algo acerca de un sistema. Además, cada uno debe incluir el uso de datos experimentales y la toma de decisiones basadas en éstos. De manera formal, en cada caso, la conjetura se puede poner en forma de hipótesis estadística. Los procedimientos que conducen a la aceptación o rechazo de hipótesis estadísticas como éstas comprenden un área principal de la inferencia estadística. Primero, definamos con precisión lo que entendemos por **hipótesis estadística**.

Definición 10.1

Una **hipótesis estadística** es una aseveración o conjetura con respecto a una o más poblaciones.

La verdad o falsedad de una hipótesis estadística nunca se sabe con absoluta certidumbre a menos que examinemos toda la población. Esto, por supuesto, sería

poco práctico en la mayoría de las situaciones. En su lugar, tomamos una muestra aleatoria de la población de interés y utilizamos los datos contenidos en esta muestra para proporcionar evidencia que apoye o no la hipótesis. La evidencia de la muestra que es inconsistente con la hipótesis que se establece conduce al rechazo de ésta, mientras que la evidencia que la apoya conduce a su aceptación.

Debe quedar claro al lector que el diseño de un procedimiento de decisión se debe hacer con la idea en mente de la *probabilidad de una conclusión errónea*. Por ejemplo, suponga que la conjetura (hipótesis) que postula el ingeniero es que la fracción p de defectuosos en cierto proceso es 0.10. El experimento es la observación de una muestra aleatoria del producto en cuestión. Suponga que se prueban 100 artículos y se encuentra que 12 están defectuosos. Es razonable concluir que esta evidencia no rechaza la condición $p = 0.10$, y por ello puede conducir a la aceptación de la hipótesis. Sin embargo, tampoco rechaza $p = 0.12$ o quizá incluso $p = 0.15$. Como resultado, el lector se debe acostumbrar a comprender que **la aceptación de una hipótesis simplemente implica que los datos no dan suficiente evidencia para rechazarla. Por otro lado, el rechazo implica que la evidencia muestral la refuta.** Puesto de otra forma, **el rechazo significa que hay una pequeña probabilidad de obtener la información muestral observada cuando, de hecho, la hipótesis es verdadera.** Por ejemplo, en nuestra hipótesis de proporción de defectuosos, una muestra de 100 que revela 20 artículos defectuosos es ciertamente evidencia de rechazo. ¿Por qué? Si, en realidad, $p = 0.10$, la probabilidad de obtener 20 o más defectuosos es aproximadamente 0.0035. Con el pequeño riesgo resultante de una *conclusión errónea*, parecería seguro **rechazar la hipótesis** de que $p = 0.10$. En otras palabras, el rechazo de una hipótesis tiende a casi "excluir" la hipótesis. Por otro lado, es muy importante enfatizar que la aceptación o, más bien, la falla al rechazo no excluye otras posibilidades. Como resultado, *el analista de los datos establece una conclusión firme cuando se rechaza una hipótesis.*

El planteamiento formal de una hipótesis a menudo está influido por la estructura de la probabilidad de una conclusión errónea. Si el científico se interesa en *apoyar con fuerza* una opinión, desea llegar a la opinión en la forma de rechazo de una hipótesis. Si el investigador médico desea mostrar fuerte evidencia a favor de la opinión de que beber café aumenta el riesgo de contraer cáncer, la hipótesis a probar debe ser de la forma "no hay aumento en el riesgo de contraer cáncer como producto de beber café". Como resultado, la opinión se alcanza por medio de un rechazo. De manera similar, para apoyar la afirmación de que un tipo de medidores es más preciso que otro, el ingeniero prueba la hipótesis de que no hay diferencia en la precisión de los dos tipos de medidor.

Hipótesis nula y alternativa

La estructura de la prueba de hipótesis se formulará con el uso del término **hipótesis nula**. Éste se refiere a cualquier hipótesis que deseamos probar y se denota con H_0 . El rechazo de H_0 conduce a la aceptación de una **hipótesis alternativa**, que se denota con H_1 . Una hipótesis nula con respecto a un parámetro poblacional siempre se establecerá de modo que especifique un valor exacto del parámetro, mien-

tras que la hipótesis alternativa permite la posibilidad de varios valores. De aquí, si H_0 es la hipótesis nula $p = 0.5$ para una población binomial, la hipótesis alternativa H_1 sería una de las siguientes:

$$p > 0.5, \quad p < 0.5, \quad \text{o,} \quad p \neq 0.5.$$

10.2 Prueba de una hipótesis estadística

Para ilustrar los conceptos que se utilizan al probar una hipótesis estadística acerca de una población, considere el siguiente ejemplo. Se sabe que cierto tipo de vacuna contra el catarro sólo es efectiva en 25% después de un periodo de dos años. Para determinar si una vacuna nueva y algo más cara es superior al proporcionar protección contra el mismo virus durante un periodo más largo, suponga que se eligen 20 personas al azar y se inoculan. En un estudio real de este tipo los participantes que reciben la nueva vacuna pueden ascender a varios miles. El número 20 se utiliza aquí sólo para demostrar los pasos básicos para llevar a cabo una prueba estadística. Si más de ocho de los que reciben la nueva vacuna pasan el periodo de dos años sin contraer el virus, la nueva vacuna se considerará superior a la que se usa en la actualidad. El requisito de que el número exceda de ocho es algo arbitrario pero parece razonable, pues representa una ganancia modesta sobre las cinco personas que se esperaría que recibieran protección si las 20 personas se inocularan con la vacuna ya en uso. Probamos en esencia la hipótesis nula de que la vacuna nueva es igualmente efectiva después de un periodo de dos años como la que por lo general se usa ahora. La hipótesis alternativa es que la nueva vacuna es de hecho superior. Esto es equivalente a probar la hipótesis de que el parámetro binomial para la probabilidad de un éxito sobre una prueba dada es $p = 1/4$ contra la alternativa de que $p > 1/4$. Esto por lo general se escribe como sigue:

$$H_0: p = \frac{1}{4}$$

$$H_1: p > \frac{1}{4}$$

La **estadística de prueba** sobre la que se basa nuestra decisión es X , el número de individuos en nuestro grupo de prueba que reciben protección de la nueva vacuna por un periodo de al menos dos años. Los valores posibles de X , de 0 a 20, se dividen en dos grupos: los números menores que o iguales a ocho y los mayores que ocho. Todos los posibles valores mayores que ocho constituyen la **región crítica**, y todos los posibles valores menores que o iguales a ocho determinan la **región de aceptación**. El último número que observamos al pasar de la región de aceptación a la región crítica se llama **valor crítico**. En nuestra ilustración el valor crítico es el número ocho. Por tanto, si $x > 8$, rechazamos H_0 a favor de la hipótesis alternativa H_1 . Si $x \leq 8$, aceptamos H_0 . Este criterio de aceptación se ilustra en la figura 10.1.

El procedimiento de decisión recién descrito podría conducir a cualquiera de dos conclusiones erróneas. Por ejemplo, la nueva vacuna puede no ser mejor que la que se usa actualmente y, para este grupo particular de individuos seleccionado de

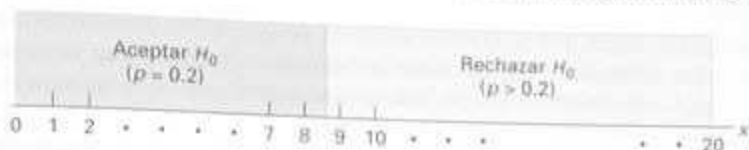


Figura 10.1 Criterio de decisión para probar $p = 0.2$ contra $p > 0.2$.

forma aleatoria, más de ocho pasan el periodo de dos años sin contraer el virus. Cometeríamos un error al rechazar H_0 a favor de H_1 cuando, de hecho, H_0 es verdadera. Tal error se llama **error tipo I**.

Definición 10.2

El rechazo de la hipótesis nula cuando es verdadera se llama **error tipo I**.

Una segunda clase de error se comete si ocho o menos del grupo pasan el periodo de dos años de forma exitosa y concluimos que la nueva vacuna no es mejor cuando realmente lo es. En este caso aceptaríamos H_0 cuando es falsa. Éste se llama **error tipo II**.

Definición 10.3

La aceptación de la hipótesis nula cuando es falsa se llama **error tipo II**.

Al probar cualquier hipótesis estadística, hay cuatro situaciones posibles que determinan si nuestra decisión es correcta o errónea. Estas cuatro situaciones se resumen en la tabla 10.1.

Tabla 10.1 Situaciones posibles al probar una hipótesis estadística

	H_0 es verdadera	H_0 es falsa
Aceptar H_0	Decisión correcta	Error tipo II
Rechazar H_0	Error tipo I	Decisión correcta

La probabilidad de cometer un error tipo I, también llamada **nivel de significancia**, se denota con la letra griega α . En nuestra ilustración, un error tipo I ocurrirá cuando más de ocho individuos pasen el periodo de dos años sin contraer el virus al usar la nueva vacuna que en realidad equivale a la que está en uso. De aquí, si X es el número de individuos que permanecen libres del virus por al menos dos años,

$$\begin{aligned} \alpha &= P(\text{error tipo I}) = P\left(X > 8 \text{ cuando } p = \frac{1}{4}\right) = \sum_{x=9}^{20} b\left(x; 20, \frac{1}{4}\right) \\ &= 1 - \sum_{x=0}^8 b\left(x; 20, \frac{1}{4}\right) = 1 - 0.9591 = 0.0409. \end{aligned}$$

Decimos que la hipótesis nula, $p = 1/4$, se prueba al nivel de significancia $\alpha = 0.0409$. Algunas veces el nivel de significancia se llama tamaño de la región crítica. Una región crítica de tamaño 0.0409 es muy pequeña y por tanto es poco práctica.

bable que se cometa un error de tipo I. En consecuencia, sería poco probable que más de ocho individuos permanezcan inmunes a un virus por un periodo de dos años mediante el uso de una vacuna nueva que es en esencia equivalente a la que existe ahora en el mercado.

La probabilidad de cometer un error tipo II, que se denota por β , es imposible de calcular a menos que tengamos una hipótesis alternativa específica. Si probamos la hipótesis nula $p = 1/4$ contra la hipótesis alternativa $p = 1/2$, entonces seremos capaces de calcular la probabilidad de aceptar H_0 cuando es falsa. Simplemente encontramos la probabilidad de obtener ocho o menos en el grupo que sobrepasan el periodo de dos años cuando $p = 1/2$. En este caso

$$\beta = P(\text{error tipo II}) = P\left(X \leq 8 \text{ cuando } p = \frac{1}{2}\right) = \sum_{x=0}^8 b\left(x; 20, \frac{1}{2}\right) = 0.2517.$$

Ésta es una probabilidad más bien alta, lo que indica un procedimiento de prueba en el que es muy probable que rechacemos la nueva vacuna cuando, de hecho, es superior a la que está en uso. Idealmente, preferiríamos utilizar un procedimiento de prueba en el que los errores tipo I y tipo II son pequeños.

Es posible que el director del programa de prueba esté dispuesto a cometer un error tipo II si la vacuna más cara no es significativamente superior. De hecho la única ocasión en la que desea estar prevenido contra un error tipo II es cuando el valor real de p es al menos 0.7. Si $p = 0.7$, este procedimiento de prueba da

$$\beta = P(\text{error tipo II}) = P(X \leq 8 \text{ cuando } p = 0.7) = \sum_{x=0}^8 b(x; 20, 0.7) = 0.0051.$$

Con una probabilidad tan pequeña de cometer un error tipo II, es improbable en extremo que se rechace la nueva vacuna cuando tiene una efectividad de 70% después de un periodo de dos años. Conforme la hipótesis alternativa se aproxima a la unidad, el valor de β tiende a cero.

Supongamos que el director del programa de prueba no está dispuesto a cometer un error tipo II cuando la hipótesis alternativa $p = 1/2$ es verdadera aunque se encuentre que la probabilidad de tal error es $\beta = 0.2517$. Una reducción de β siempre es posible al aumentar el tamaño de la región crítica. Por ejemplo, considere que sucede a los valores de α y β cuando cambiamos nuestro valor crítico a siete de modo que todos los valores mayores que siete caigan en la región crítica y aquellos menores que o iguales a siete caigan en la región de aceptación. Ahora bien, al probar $p = 1/4$ contra la hipótesis alternativa $p = 1/2$, encontramos que

$$\alpha = \sum_{x=8}^{20} b\left(x; 20, \frac{1}{4}\right) = 1 - \sum_{x=0}^7 b\left(x; 20, \frac{1}{4}\right) = 1 - 0.8982 = 0.1018$$

y

$$\beta = \sum_{x=0}^7 b\left(x; 20, \frac{1}{2}\right) = 0.1316.$$

Al adoptar un nuevo procedimiento de decisión, reducimos la probabilidad de cometer un error tipo II a costa de aumentar la probabilidad de cometer un error tipo I. Para un tamaño muestral fijo, una disminución en la probabilidad de un

error por lo general tendrá como resultado un aumento en la probabilidad del otro error. Por fortuna, la probabilidad de cometer ambos tipos de error se puede reducir al aumentar el tamaño de la muestra. Considere el mismo problema mediante el uso de una muestra aleatoria de 100 individuos. Si más de 36 del grupo sobrepasan el periodo de dos años, rechazaremos la hipótesis nula $p = 1/4$ y aceptaremos la hipótesis alternativa $p > 1/4$. El valor crítico ahora es 36. Todos los valores posibles por arriba de 36 constituyen la región crítica y todos los valores posibles menores que o iguales a 36 caen en la región de aceptación.

Para determinar la probabilidad de cometer un error tipo I, utilizaremos la aproximación de la curva normal con

$$\mu = np = (100)\left(\frac{1}{4}\right) = 25 \quad \text{y} \quad \sigma = \sqrt{npq} = \sqrt{(100)\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)} = 4.33.$$

Con referencia a la figura 10.2, necesitamos el área bajo la curva normal a la derecha de $x = 36.5$. El valor z correspondiente es

$$z = \frac{36.5 - 25}{4.33} = 2.66.$$

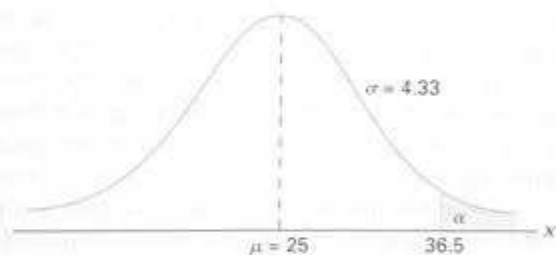


Figura 10.2 Probabilidad de un error tipo I.

De la tabla A.3 encontramos que

$$\begin{aligned} \alpha &= p(\text{error tipo I}) = P\left(X > 36 \text{ cuando } p = \frac{1}{4}\right) = P(Z > 2.66) \\ &= 1 - P(Z < 2.66) = 1 - 0.9961 = 0.0039. \end{aligned}$$

Si H_0 es falsa y el verdadero valor de H_1 es $p = 1/2$, podemos determinar la probabilidad de un error tipo II con el uso de la aproximación a la curva normal con

$$\mu = np = (100)\left(\frac{1}{2}\right) = 50 \quad \text{y} \quad \sigma = \sqrt{npq} = \sqrt{(100)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5.$$

La probabilidad de caer en la región de aceptación cuando H_1 es verdadera está dada por el área de la región sombreada a la izquierda de $x = 36.5$ en la figura 10.3. El valor z que corresponde a $x = 36.5$ es

$$z = \frac{36.5 - 50}{5} = -2.7.$$

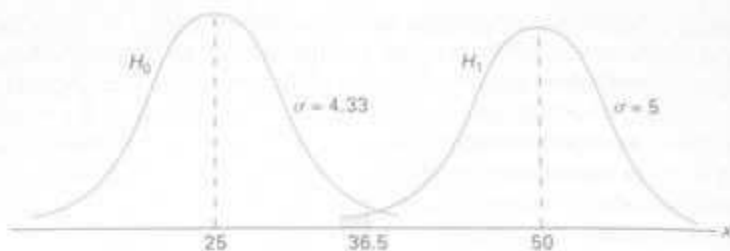


Figura 10.3 Probabilidad de un error tipo II.

Por tanto,

$$\beta = P(\text{error tipo II}) = P\left(X \leq 36 \text{ cuando } p = \frac{1}{2}\right) = P(Z < -2.7) = 0.0035.$$

Obviamente, los errores tipo I y tipo II rara vez ocurren si el experimento consiste en 100 individuos.

La ilustración anterior subraya la estrategia del científico en la prueba de hipótesis. Después de que se establecen las hipótesis nula y alternativa, es importante considerar la sensibilidad del procedimiento de prueba. Por esto queremos decir que debe haber una determinación, para un α fija, de un valor razonable para la probabilidad de aceptar de manera errónea H_0 (es decir, el valor de β) cuando la situación real representa alguna *desviación importante de H_0* . Por lo general se puede determinar el valor del tamaño de la muestra para el que hay un balance razonable entre α y el valor de β que se calcula de esta manera. El problema de la vacuna es un ejemplo.

Los conceptos que aquí se discuten para una población discreta se pueden aplicar bien de igual forma a poblaciones continuas. Considere la hipótesis nula de que el peso promedio de estudiantes hombres en cierta universidad es 68 kilogramos contra la hipótesis alternativa de que es diferente de 68. Es decir, deseamos probar

$$H_0: \mu = 68,$$

$$H_1: \mu \neq 68.$$

La hipótesis alternativa nos permite la posibilidad de que $\mu < 68$ o $\mu > 68$.

Una media muestral que cae cerca del valor hipotético de 68 se consideraría como evidencia en favor de H_0 . Por otro lado, una media muestral considerablemente menor que o mayor que 68 sería una evidencia de inconsistencia de H_0 y por tanto favorecería a H_1 . La media muestral es la estadística de prueba en este caso. Una región crítica para la estadística de prueba se puede elegir de manera arbitraria como los dos intervalos $\bar{x} < 67$ y $\bar{x} > 69$. La región de aceptación será entonces el intervalo $67 \leq \bar{x} \leq 69$. Este criterio de decisión se ilustra en la figura 10.4.



Figura 10.4 Probabilidad de un error tipo II.

Utilicemos ahora el criterio de decisión de la figura 10.4 para calcular las probabilidades de cometer errores tipo I y tipo II cuando se prueba la hipótesis nula $\mu = 68$ kilogramos contra la alternativa $\mu \neq 68$ kilogramos para la población continua de los pesos de los estudiantes.

Suponga que la desviación estándar de la población de pesos es $\sigma = 3.6$. Para muestras grandes podemos sustituir s por σ si no se dispone de ninguna otra estimación de σ . Nuestra estadística de decisión, que se basa en una muestra aleatoria de tamaño $n = 36$, será \bar{X} , el estimador más eficiente de μ . Del teorema del límite central, sabemos que la distribución muestral de \bar{X} es aproximadamente normal con desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 3.6/6 = 0.6$.

La probabilidad de cometer un error tipo I, o el nivel de significancia de nuestra prueba, es igual a la suma de las áreas sombreadas en cada cola de la distribución en la figura 10.5. Por tanto,

$$\alpha = P(\bar{X} < 67 \text{ cuando } \mu = 68) + P(\bar{X} > 69 \text{ cuando } \mu = 68).$$

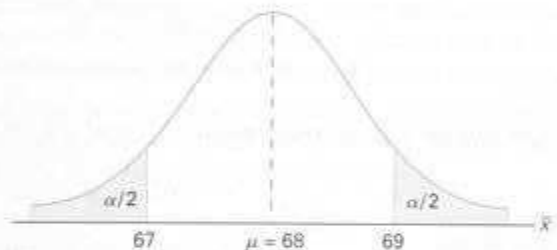


Figura 10.5 Región crítica para probar $\mu = 68$ contra $\mu \neq 68$.

Los valores z correspondientes a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando H_0 es verdadera son

$$z_1 = \frac{67 - 68}{0.6} = -1.67 \quad \text{y} \quad z_2 = \frac{69 - 68}{0.6} = 1.67.$$

Por tanto,

$$\alpha = P(Z < -1.67) + P(Z > 1.67) = 2P(Z < -1.67) = 0.0950.$$

De esta manera 9.5% de todas las muestras de tamaño 36 nos conducirían a rechazar $\mu = 68$ kilogramos cuando ésta es verdadera. Para reducir α , tenemos que escoger entre aumentar el tamaño de la muestra o ampliar la región de aceptación. Suponga que aumentamos el tamaño de la muestra a $n = 64$. Entonces $\sigma_{\bar{X}} = 3.6/8 = 0.45$. Ahora bien,

$$z_1 = \frac{67 - 68}{0.45} = -2.22 \quad \text{y} \quad z_2 = \frac{69 - 68}{0.45} = 2.22.$$

De aquí

$$\alpha = P(Z < -2.22) + P(Z > 2.22) = 2P(Z < -2.22) = 0.0264.$$

La reducción en α no es suficiente por sí misma para garantizar un buen procedimiento de prueba. Debemos evaluar β para varias hipótesis alternativas que

consideremos se deben aceptar si son verdaderas. Por tanto, es importante rechazar H_0 cuando la media real es algún valor $\mu \geq 70$ o $\mu \leq 66$, entonces la probabilidad de cometer un error tipo II se debe calcular y examinar para las alternativas $\mu = 66$ y $\mu = 70$. Debido a la simetría, sólo es necesario considerar la probabilidad de aceptar la hipótesis nula $\mu = 68$ cuando la alternativa $\mu = 70$ es verdadera. Resultará un error tipo II cuando la media muestral \bar{x} caiga entre 67 y 69 cuando H_1 sea verdadera. Por tanto, con referencia a la figura 10.6, encontramos que

$$\beta = P(67 \leq \bar{X} \leq 69 \text{ cuando } \mu = 70).$$

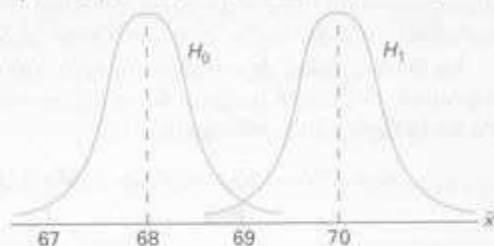


Figura 10.6 Error tipo II para probar $\mu = 68$ contra $\mu = 70$.

Los valores z que corresponden a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando H_1 es verdadera son

$$z_1 = \frac{67 - 70}{0.45} = -6.67 \quad \text{y} \quad z_2 = \frac{69 - 70}{0.45} = -2.22.$$

Por tanto,

$$\begin{aligned} \beta &= P(-6.67 < Z < -2.22) = P(Z < -2.22) - P(Z < -6.67) \\ &= 0.0132 - 0.0000 = 0.0132. \end{aligned}$$

Si el valor real de μ es la alternativa $\mu = 66$, el valor de β nuevamente será 0.0132. Para todos los valores posibles de $\mu < 66$ o $\mu > 70$, el valor de β será incluso más pequeño cuando $n = 64$, y en consecuencia habrá poca oportunidad de aceptar H_0 cuando sea falsa.

La probabilidad de cometer un error tipo II aumenta rápidamente cuando el valor real de μ se aproxima al valor hipotético, pero no es igual a éste. Por supuesto, esta es por lo general la situación donde podemos cometer un error tipo II. Por ejemplo, si la hipótesis alternativa $\mu = 68.5$ es verdadera, podemos cometer un error tipo II al concluir que la respuesta verdadera es $\mu = 68$. La probabilidad de cometer tal error será alta cuando $n = 64$. Con referencia a la figura 10.7, tenemos

$$\beta = P(67 \leq \bar{X} \leq 69 \text{ cuando } \mu = 68.5).$$

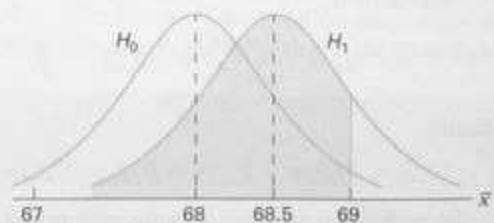


Figura 10.7 Error tipo II para probar $\mu = 68$ contra $\mu = 68.5$.

Los valores z correspondientes a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando $\mu = 68.5$ son

$$z_1 = \frac{67 - 68.5}{0.45} = -3.33 \quad \text{y} \quad z_2 = \frac{69 - 68.5}{0.45} = 1.11.$$

Por tanto,

$$\begin{aligned} \beta &= P(-3.33 < Z < 1.11) = P(Z < 1.11) - P(Z < -3.33) \\ &= 0.8665 - 0.0004 = 0.8661. \end{aligned}$$

Los ejemplos anteriores ilustran las propiedades importantes siguientes:

1. Los errores tipo I y tipo II están relacionados. Una disminución en la probabilidad de uno por lo general tiene como resultado un aumento en la probabilidad del otro.
2. El tamaño de la región crítica, y por tanto la probabilidad de cometer un error tipo I, siempre se puede reducir al ajustar el o los valores críticos.
3. Un aumento en el tamaño muestral n reducirá α y β de forma simultánea.
4. Si la hipótesis nula es falsa, β es un máximo cuando el valor real de un parámetro se aproxima al valor hipotético. Entre más grande sea la distancia entre el valor real y el valor hipotético, será menor β .

Un concepto muy importante que se relaciona con las probabilidades de error es la noción de potencia de una prueba.

Definición 10.4

La **potencia** de una prueba es la probabilidad de rechazar H_0 dado que una alternativa específica es verdadera.

La potencia de una prueba se puede calcular como $1 - \beta$. A menudo diferentes tipos de pruebas se comparan al contrastar propiedades de potencia. Considere la ilustración anterior en la que probamos $H_0: \mu = 68$ y $H_1: \mu \neq 68$. Como antes, suponga que nos interesamos en evaluar la sensibilidad de la prueba. La prueba está gobernada por la regla de que la aceptamos si $67 \leq \bar{x} \leq 69$. Buscamos la capacidad de la prueba para rechazar de manera apropiada H_0 cuando en realidad $\mu = 68.5$. Vimos que la probabilidad de un error tipo II está dada por $\beta = 0.8661$. De esta manera la **potencia** de la prueba es $1 - 0.8661 = 0.1339$. En cierto sentido, la potencia es una medida más sucinta de cuán sensible es la prueba para "detectar diferencias" entre una media de 68 y 68.5. En este caso, si μ es realmente 68.5, la prueba como se describe *rechazará de forma apropiada H_0 sólo 13.39% de las veces*. Como resultado, la prueba no sería buena si es importante que el analista tenga una oportunidad razonable de distinguir realmente entre una media de 68.0 (que especifica H_0) y una media de 68.5. De lo anterior, es claro que para producir una potencia deseable (digamos, mayor que 0.8), se debe aumentar α o aumentar el tamaño de la muestra.

En las secciones anteriores de este capítulo, la mayor parte del texto sobre prueba de hipótesis gira alrededor de las bases y definiciones. En las secciones que siguen seremos más específicos y clasificaremos a las hipótesis en categorías así co-

mo discutiremos pruebas de hipótesis sobre varios parámetros de interés. Comenzamos formulando la distinción entre hipótesis unilaterales y bilaterales.

10.3 Pruebas de una y dos colas

Una prueba de cualquier hipótesis estadística, donde la alternativa es **unilateral**, como

$$\begin{array}{ll} H_0: \theta = \theta_0, & H_0: \theta = \theta_0, \\ H_1: \theta > \theta_0, & \text{o quizá} \quad H_1: \theta < \theta_0, \end{array}$$

se denomina **prueba de una sola cola**.

En la sección 10.2 se hizo referencia a la **estadística de prueba** para una hipótesis. Por lo general, la región crítica para la hipótesis alternativa $\theta > \theta_0$ yace en la cola derecha de la distribución de la estadística de prueba, mientras que la región crítica para la hipótesis alternativa $\theta < \theta_0$ yace por completo en la cola izquierda. En cierto sentido, el símbolo de desigualdad apunta en la dirección donde se encuentra la región crítica. En el experimento de la vacuna de la sección 10.2 se utilizó una prueba de una sola cola para probar la hipótesis $p = 1/4$ contra la alternativa unilateral $p > 1/4$ para la distribución binomial. La región crítica de una sola cola por lo general es obvia. Para una mejor comprensión el lector debe visualizar el comportamiento de la estadística de prueba y notar la *señal* obvia que produciría evidencia que apoye la hipótesis alternativa.

Una prueba de cualquier hipótesis alternativa donde la alternativa es **bilateral**, como

$$\begin{array}{l} H_0: \theta = \theta_0, \\ H_1: \theta \neq \theta_0 \end{array}$$

se llama **prueba de dos colas**, pues la región crítica se divide en dos partes, que a menudo tienen probabilidades iguales que se colocan en cada cola de la distribución de la estadística de prueba. La hipótesis alternativa $\theta \neq \theta_0$ establece que $\theta < \theta_0$ o $\theta > \theta_0$. Una prueba de dos colas se utilizó para probar la hipótesis nula $\mu = 68$ kilogramos contra la alternativa bilateral $\mu \neq 68$ kilogramos para la población continua de pesos de estudiantes en la sección 10.2.

La hipótesis nula, H_0 , siempre se establecerá con el uso del signo de igualdad para que se especifique un solo valor. De esta manera se puede controlar la probabilidad de cometer un error tipo I. Si se establece una prueba de una cola o de dos colas dependerá de la conclusión que se extraiga si H_0 se rechaza. La posición de la región crítica se puede determinar sólo después que se establece H_1 . Por ejemplo, al probar una medicina nueva, se establece la hipótesis de que no es mejor que las medicinas similares actualmente en el mercado y se prueba ésta contra la hipótesis alternativa de que la medicina nueva es superior. Tal hipótesis alternativa tendrá como resultado de una prueba de una sola cola con la región crítica en la cola derecha. Sin embargo, si deseamos comparar una nueva técnica de enseñanza con el procedimiento

to convencional de salón, la hipótesis alternativa debe permitir que la nueva aproximación sea inferior o superior al procedimiento convencional. Por ello, la prueba es de dos colas con la región crítica dividida en partes iguales de modo que caiga en los extremos de las colas izquierda y derecha de la distribución de nuestra estadística.

Son deseables ciertos principios para determinar cuál hipótesis se establecerá como H_0 y cuál como H_1 . Primero, lea el problema con mucho cuidado y determine la afirmación que desea probar. Si la afirmación sugiere una sola dirección como *mayor que*, *menor que*, *superior a*, *inferior a*, etcétera, entonces H_1 se debe establecer con el uso del símbolo de desigualdad ($<$ o $>$) que corresponda a la dirección sugerida. Si, por ejemplo, al probar una nueva medicina deseamos mostrar una fuerte evidencia de que *más de 30%* de las personas tendrán ayuda, de inmediato escribimos $H_1: p > 0.3$ y entonces la hipótesis nula se escribe como $H_0: p = 0.3$. Si la afirmación sugiere una dirección compuesta (igualdad y dirección) como *al menos mayor o igual que*, *a lo más, no mayor que*, etcétera, entonces toda esta dirección compuesta (\leq o \geq) se expresa como H_0 , pero con el uso únicamente del signo igual, y H_1 se da en la dirección opuesta. Finalmente, si no se sugiere ninguna dirección en la afirmación, entonces H_1 se establece con el signo de *diferente* (\neq).

Ejemplo 10.1 Un fabricante de cierta marca de cereal de arroz afirma que el contenido promedio de grasa saturada no excede 1.5 gramos. Establezca las hipótesis nula y alternativa a utilizar para probar esta afirmación y determinar dónde se localiza la región crítica.

SOLUCIÓN

La afirmación del fabricante se debe rechazar sólo si μ es mayor que 1.5 miligramos y se debe aceptar si μ es menor o igual que 1.5 miligramos. Como la hipótesis nula siempre especifica un solo valor del parámetro, probamos

$$H_0: \mu = 1.5,$$

$$H_1: \mu > 1.5.$$

Aunque establecemos la hipótesis nula con un signo igual, se entiende que incluye cualquier valor no especificado por la hipótesis alternativa. En consecuencia, la aceptación de H_0 no implica que μ es exactamente igual a 1.5 miligramos sino más bien que no tenemos suficiente evidencia que favorezca a H_1 . Como tenemos una prueba de una cola, el símbolo mayor que indica que la región crítica yace por completo en la cola derecha de la distribución de nuestra estadística de prueba \bar{X} . ■

Ejemplo 10.2 Un agente de bienes raíces afirma que 60% de todas las residencias privadas que se construyen actualmente son casas de tres recámaras. Para probar esta afirmación, se inspecciona una muestra grande de residencias nuevas; la proporción de estas casas con tres recámaras se registra y se utiliza como estadística de prueba. Establezca las hipótesis nula y alternativa a utilizarse en esta prueba y determine la posición de la región crítica.

SOLUCIÓN

Si la estadística de prueba fuera considerablemente mayor o menor que $p = 0.6$, rechazaríamos la afirmación del agente. De aquí debemos establecer la hipótesis

$$H_0: p = 0.6,$$

$$H_1: p \neq 0.6.$$

La hipótesis alternativa implica una prueba de dos colas con la región crítica dividida por igual en ambas colas de la distribución de P , nuestra estadística de prueba.

10.4 Uso de valores P para la toma de decisiones

Al probar hipótesis en las que la estadística de prueba es discreta, la región crítica se puede elegir de manera arbitraria y determinar su tamaño. Si α es demasiado grande, se puede reducir al hacer un ajuste en el valor crítico. Puede ser necesario aumentar el tamaño de la muestra para compensar la disminución que ocurre de manera automática en la potencia de la prueba.

Por generaciones enteras de análisis estadístico, se ha hecho costumbre elegir un α de 0.05 o 0.01 y seleccionar la región crítica en consecuencia. Entonces, por supuesto, el rechazo o no rechazo estricto de H_0 dependerá de esa región crítica. Por ejemplo, si la prueba es de dos colas y α se fija al nivel de significancia de 0.05 y la estadística de prueba involucra, digamos, la distribución normal estándar, entonces se observa un valor z de los datos y la región crítica es

$$z > 1.96, \quad z < -1.96,$$

donde el valor 1.96 se encuentra como $z_{0.025}$ en la tabla A.3. Un valor de z en la región crítica sugiere el planteamiento: "El valor de la estadística de prueba es significativo." Podemos traducir esto al lenguaje del usuario. Por ejemplo, si la hipótesis está dada por

$$H_0: \mu = 10,$$

$$H_1: \mu \neq 10,$$

se puede decir: "La media difiere de manera significativa del valor 10."

Esta preselección de un nivel de significancia α tiene sus raíces en la filosofía de que se debe controlar el riesgo máximo de cometer un error tipo I. Sin embargo, este enfoque no explica los valores de estadísticas de prueba que están "cerca" a la región crítica. Suponga que, por ejemplo, en la ilustración con $H_0: \mu = 10$; $H_1: \mu \neq 10$, se observa un valor $z = 1.87$; estrictamente hablando, con $\alpha = 0.05$ el valor no es significativo. Pero el riesgo de cometer un error tipo I si se rechaza H_0 en este caso difícilmente se podría considerar severo. De hecho, en un escenario de dos colas se puede cuantificar el riesgo como

$$P = 2P(z > 1.87 \text{ cuando } \mu = 10) = 2(0.0307) = 0.0614.$$

Como resultado, la probabilidad de obtener un valor z tan grande o mayor (en magnitud) que 1.87 cuando de hecho $\mu = 10$ es 0.0614. Aunque esta evidencia contra H_0 no es tan fuerte como la que resultaría de un rechazo en un nivel $\alpha = 0.05$, es información importante para el usuario. En realidad, el uso continuo de $\alpha = 0.05$ o 0.01 sólo es un resultado de lo que los estándares han establecido por generaciones. En la estadística aplicada los usuarios han adoptado de forma extensa la aproximación del valor P . La aproximación se diseña para dar al usuario una alternativa (en términos de una probabilidad) a la simple conclusión de "rechazo" o "no rechazo". El

cálculo del valor P también da al usuario información importante cuando el valor z cae por completo dentro de la región crítica ordinaria. Por ejemplo, si z es 2.73, es informativo para el usuario observar que

$$P = 2(0.0032) = 0.0064$$

y de esta forma el valor z es significativo a un nivel considerablemente menor que 0.05. Es importante saber que bajo la condición de H_0 , un valor de $z = 2.73$ es un evento en extremo raro. A saber, un valor al menos tan grande en magnitud sólo ocurriría 64 veces en 10,000 experimentos.

Una manera muy simple de explicar un valor P de forma gráfica es considerar dos muestras distintas prematuramente. Suponga que se consideran dos materiales para cubrir un tipo particular de metal a fin de evitar la corrosión. Se obtienen especímenes y se cubre un grupo con el material 1 y un grupo con el material 2. Los tamaños muestrales son $n_1 = n_2 = 10$ para cada muestra y la corrosión se mide en porcentaje de área superficial afectada. La hipótesis es que las muestras provienen de distribuciones comunes con media $\mu = 10$. Supongamos que la varianza poblacional es 1.0. Entonces probamos que

$$H_0: \mu_1 = \mu_2 = 10.$$

Representemos en la figura 10.8 una gráfica de puntos de los datos; los datos se colocan en la distribución que establece la hipótesis nula. Parece claro ahora que los datos rechazan la hipótesis nula. Pero ¿cómo se puede resumir esto en un número? El valor P se puede ver simplemente como la probabilidad de obtener este conjunto de datos *dado* que las muestras provienen de la distribución que se representa. Es claro que esta probabilidad es bastante pequeña, digamos, ¡0.00000001! De esta manera el pequeño valor P rechaza claramente H_0 , y la conclusión es que las medias poblacionales son significativamente diferentes.

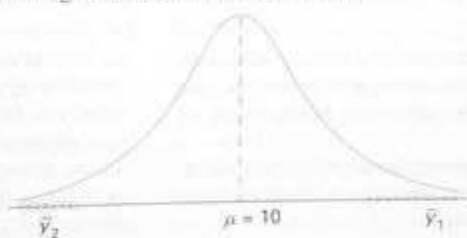


Figura 10.8 Datos probablemente generados de poblaciones que tienen dos medias diferentes.

La aproximación del valor P como ayuda en la toma de decisiones es bastante natural pues casi todos los paquetes de computadora que proporcionan el cálculo de pruebas de hipótesis entregan valores P junto con valores de la estadística de prueba apropiada. La siguiente es una definición formal de un valor P .

Definición 10.5

Un valor P es el nivel (de significancia) más bajo en el que el valor observado de la estadística de prueba es significativo.

Resulta apropiado en este momento resumir los procedimientos para la prueba de hipótesis. Esto puede servir como fundamento en el que se basen los casos es-

peciales de las secciones siguientes. Para este resumen, supongamos que la hipótesis es $H_0: \theta = \theta_0$.

1. Establecer la hipótesis nula H_0 que $\theta = \theta_0$.
2. Elegir una hipótesis alternativa apropiada H_1 a partir de una de las alternativas $\theta < \theta_0$, $\theta > \theta_0$, o $\theta \neq \theta_0$.
3. Elegir un nivel de significancia de tamaño α .
4. Seleccionar la estadística de prueba apropiada y establecer la región crítica. (Si la decisión se basa en un valor P , no es necesario establecer la región crítica.)
5. Calcular el valor de la estadística de prueba a partir de los datos de la muestra.
6. Decisión: Rechazar H_0 si la estadística de prueba tiene un valor en la región crítica (o si el valor P calculado es menor o igual que el nivel de significancia α que se desea); en cualquier otro caso, no rechazar H_0 .

El lector se debe dar cuenta que las conclusiones que extrae el analista pueden resultar afectadas por los valores P calculados. En otras palabras, se puede no tener en mente un nivel α preseleccionado y así sacar conclusiones que se basen en la información que proporciona el valor P . Como se indicó al principio, éste es el enfoque que a menudo se adopta en situaciones de la vida real.

Ejercicios

1. Suponga que un alergólogo desea probar la hipótesis de que al menos 30% del público es alérgico a algunos productos de queso. Explique cómo el alergólogo podría cometer
 - (a) un error tipo I;
 - (b) un error tipo II.
2. Una socióloga se interesa en la eficacia de un curso de entrenamiento diseñado para lograr que más conductores utilicen los cinturones de seguridad en los automóviles.
 - (a) ¿Qué hipótesis prueba si comete un error del tipo I al concluir de manera errónea que el curso de entrenamiento no es efectivo?
 - (b) ¿Qué hipótesis prueba si comete un error del tipo II al concluir de forma errónea que el curso de entrenamiento es efectivo?
3. Se acusa a una empresa de discriminación en sus prácticas de contratación.
 - (a) ¿Qué hipótesis se prueba si un jurado comete un error tipo I al encontrar culpable a la empresa?
 - (b) ¿Qué hipótesis se prueba si un jurado comete un error tipo II al encontrar que la empresa es culpable?
4. Se estima que la proporción de adultos que viven en una pequeña ciudad que son graduados universitarios es $p = 0.6$. Para probar esta hipótesis se selecciona una muestra aleatoria de 15 adultos. Si el número de graduados en nuestra muestra es cualquier número de 6 a 12, aceptaremos la hipótesis nula de que $p = 0.6$; en caso contrario, concluiremos que $p \neq 0.6$.
 - (a) Evalúe α con la suposición de que $p = 0.6$. Utilice la distribución binomial.
 - (b) Evalúe β para las alternativas $p = 0.5$ y $p = 0.7$.
 - (c) ¿Es éste un buen procedimiento de prueba?
5. Repita el ejercicio 4 cuando se seleccionan 200 adultos y la región de aceptación se define como $110 \leq x \leq 130$ donde x es el número de graduados universitarios en nuestra muestra. Utilice la aproximación normal.
6. Un fabricante de telas considera que la proporción de órdenes de materia prima que llega tarde es $p = 0.6$. Si una muestra aleatoria de 10 órdenes muestra que tres o menos llegan tarde, la hipótesis de que $p = 0.6$ se debería rechazar a favor de la alternativa $p < 0.6$. Utilice la distribución binomial.
 - (a) Encuentre la probabilidad de cometer un error tipo I si la proporción verdadera es $p = 0.6$.
 - (b) Encuentre la probabilidad de cometer un error tipo II para la alternativa $p = 0.3$, $p = 0.4$, y $p = 0.5$.
7. Repita el ejercicio 6 cuando se seleccionan 50 órdenes, y se define la región crítica como $x \leq 24$, donde x es el número de órdenes en nuestra muestra que llegan tarde. Utilice la aproximación normal.
8. Una tintorería afirma que un nuevo removedor de manchas eliminará más de 70% de las manchas en las

que se aplica. Para verificar esta afirmación, el removedor de manchas se utilizará sobre 12 manchas que se eligen al azar. Si menos de 11 de las manchas se eliminan, aceptaremos la hipótesis nula que $p = 0.7$; en cualquier otro caso, concluiremos que $p > 0.7$.

- (a) Evalúe α , suponga que $p = 0.7$.
 (b) Evalúe β para la alternativa $p = 0.9$.

9. Repita el ejercicio 8 cuando se tratan 100 manchas y la región crítica se define como $x > 82$, donde x es el número de manchas que se eliminan.

10. En la publicación *Relief from Arthritis* de Thorsons Publishers, Ltd., John E. Croft afirma que más de 40% de las personas que sufren de artritis ósea obtienen un alivio mensurable de un ingrediente que produce una especie particular de mejillón que se encuentra en la costa de Nueva Zelanda. Para probar esta afirmación, el extracto de mejillón se administra a un grupo de siete pacientes con artritis ósea. Si tres o más de los pacientes obtienen alivio, aceptaremos la hipótesis nula $p = 0.4$; de otro modo, concluiremos que $p < 0.4$.

- (a) Evalúe α , suponga que $p = 0.4$.
 (b) Evalúe β para la alternativa $p = 0.3$.

11. Repita el ejercicio 10 cuando se administra el extracto de mejillón a 70 pacientes y la región crítica se define como $x < 24$, donde x es el número de pacientes con artritis ósea que obtienen alivio.

12. Se pregunta a una muestra aleatoria de 400 votantes en cierta ciudad si están a favor de un impuesto adicional de 4% sobre la venta de gasolina para proporcionar ingresos que se necesitan con urgencia para la reparación de calles. Si más de 220 pero menos de 260 favorecen el impuesto sobre ventas, concluiremos que 60% de los votantes lo apoyan.

- (a) Encuentre la probabilidad de cometer un error tipo I si 60% de los votantes están a favor del aumento de impuestos.
 (b) ¿Cuál es la probabilidad de cometer un error tipo II al utilizar este procedimiento de prueba si en realidad sólo 48% de los votantes está a favor del impuesto adicional a la gasolina?

13. Suponga que, en el ejercicio 12, concluímos que 60% de los votantes está a favor del impuesto a la venta de gasolina si más de 214 pero menos de 266 votantes de nuestra muestra lo favorece. Muestre que esta nueva región de aceptación tiene como resultado un valor más pequeño para α a costa de aumentar β .

14. Un fabricante desarrolla un nuevo sedal para pesca, que afirma tiene una resistencia media a la ruptura de 15 kilogramos con una desviación estándar

de 0,5 kilogramos. Para probar la hipótesis de que $\mu = 15$ kilogramos contra la alternativa de que $\mu < 15$ kilogramos, se prueba una muestra aleatoria de 50 sedales. La región crítica se define como $\bar{x} < 14.9$. Suponga que $\sigma = 15$ es correcta.

- (a) Encuentre la probabilidad de cometer un error tipo I cuando H_0 es verdadera.
 (b) Evalúe β para las alternativas $\mu = 14.8$ y $\mu = 14.9$ kilogramos.

15. Una máquina de refrescos en un restaurante de carnes asadas se ajusta de modo que la cantidad de bebida que sirva esté distribuida de forma aproximadamente normal con una media de 200 mililitros y una desviación estándar de 15 mililitros. La máquina se verifica periódicamente con una muestra de nueve bebidas y con el cálculo del contenido promedio. Si \bar{x} cae en el intervalo $191 < \bar{x} < 209$, se considera que la máquina opera de manera satisfactoria; de otro modo, concluimos que $\mu \neq 200$ mililitros.

- (a) Encuentre la probabilidad de cometer un error tipo I cuando $\mu = 200$ mililitros.
 (b) Encuentre la probabilidad de cometer un error tipo II cuando $\mu = 215$ mililitros.

16. Repita el ejercicio 15 para muestras de tamaño $n = 25$. Utilice la misma región crítica.

17. Se desarrolla una nueva cura para cierto tipo de cemento que tiene como resultado un coeficiente de compresión de 5000 kilogramos por centímetro cuadrado y una desviación estándar de 120. Para probar la hipótesis de que $\mu = 5000$ contra la alternativa de que $\mu < 5000$, se prueba una muestra aleatoria de 50 piezas de cemento. La región crítica se define como $\bar{x} < 4970$.

- (a) Encuentre la probabilidad de cometer un error tipo I cuando H_0 es verdadera.
 (b) Evalúe β para la alternativa $\mu = 4970$ y $\mu = 4960$.

18. Si graficamos las probabilidades de aceptación de H_0 que corresponden a diversas alternativas para μ (incluido el valor especificado por H_0) y conectamos todos los puntos mediante una curva suave, obtenemos la **curva característica de operación** del criterio de prueba, o simplemente curva CO. Nótese que la probabilidad de aceptación de H_0 cuando es verdadera es simplemente $1 - \alpha$. Las curvas características de operación se utilizan ampliamente en aplicaciones industriales para proporcionar una muestra visual de los méritos del criterio de prueba. Con referencia al ejercicio 15, encuentre las probabilidades de aceptación de H_0 para los siguientes nueve valores de μ y grafique la curva CO: 184, 188, 192, 196, 200, 204, 208, 212, y 216.

10.5 Una sola muestra: pruebas con respecto a una sola media (varianza conocida)

En esta sección consideramos de manera formal pruebas de hipótesis en una sola media poblacional. Muchas de las ilustraciones de las secciones anteriores incluyen pruebas sobre la media, por lo que el lector ya debe tener una idea de algunos de los detalles que aquí se señalan. Primero debemos describir las suposiciones en las que se basa el experimento. El modelo para la situación subyacente se centra alrededor de un experimento con X_1, X_2, \dots, X_n que representan una muestra aleatoria de una distribución con media μ y varianza $\sigma^2 > 0$. Considere primero la hipótesis

$$H_0: \mu = \mu_0,$$

$$H_1: \mu \neq \mu_0.$$

La estadística de prueba apropiada se debe basar en la variable aleatoria \bar{X} . En el capítulo 8 se introdujo el teorema del límite central, el cual establece en esencia que sin importar la distribución de X , la variable aleatoria \bar{X} tiene una distribución aproximadamente normal con media μ y varianza σ^2/n para tamaños de muestras razonablemente grandes. De esta manera, $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}}^2 = \sigma^2/n$. Podemos determinar entonces una región crítica basada en el promedio muestral calculado, \bar{x} . Debe quedar claro ya al lector que habrá una región crítica de dos colas para la prueba.

Es conveniente estandarizar \bar{X} e incluir de manera formal la variable aleatoria normal estándar Z , donde

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

Sabemos que bajo H_0 , es decir, si $\mu = \mu_0$, entonces $(\bar{X} - \mu_0)/\sigma/\sqrt{n}$ tiene una distribución $N(0, 1)$ y por tanto se puede utilizar la expresión

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1 - \alpha$$

para escribir una región de aceptación apropiada. El lector debe tener en mente que, formalmente, la región crítica se diseña para controlar α , la probabilidad de cometer un error tipo I. Debe ser obvio que se necesita una señal de evidencia de dos colas para apoyar H_1 . Así, dado un valor calculado \bar{x} , la prueba formal implica rechazar H_0 si la estadística de prueba calculada

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_{\alpha/2} \quad \text{o} \quad z < -z_{\alpha/2}.$$

Si $-z_{\alpha/2} < z < z_{\alpha/2}$, no se rechaza H_0 . El rechazo de H_0 , por supuesto, implica la aceptación de la hipótesis alternativa $\mu \neq \mu_0$. Con esta definición de la región crítica

tica debe quedar claro que habrá la probabilidad α de rechazar H_0 (que cae en la región crítica) cuando, en realidad, $\mu = \mu_0$.

Aunque es más fácil entender la región crítica escrita en términos de z , escribimos la misma región crítica en términos del promedio calculado \bar{x} . Lo siguiente se puede escribir como un procedimiento de decisión idéntico:

$$\text{rechazar } H_0 \text{ si } \bar{x} > b \text{ o } \bar{x} < a,$$

donde

$$a = \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad b = \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

De aquí, para un nivel de significancia α , los valores críticos de la variable aleatoria z y \bar{x} se representan en la figura 10.9.

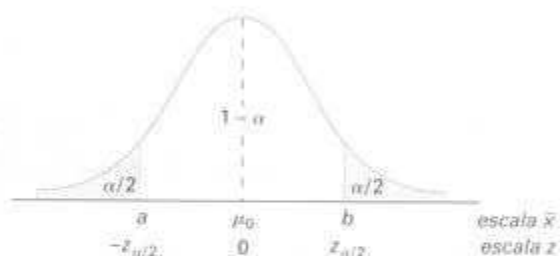


Figura 10.9 Región crítica para la hipótesis alternativa $\mu \neq \mu_0$.

Las pruebas de hipótesis unilaterales sobre la media involucran la misma estadística que se describe en el caso bilateral. La diferencia, por supuesto, es que la región crítica sólo está en una cola de la distribución normal estándar. Como resultado, por ejemplo, supongamos que buscamos probar

$$H_0: \mu = \mu_0,$$

$$H_1: \mu > \mu_0.$$

La señal que favorece a H_1 proviene de valores grandes de z . De esta forma el rechazo de H_0 resulta cuando $z > z_\alpha$. Obviamente, si la alternativa es $H_1: \mu < \mu_0$, la región crítica está por completo en la cola inferior y así el rechazo resulta de $z < -z_\alpha$.

Los siguientes dos ejemplos ilustran pruebas de medias para el caso en el que se conoce σ .

Ejemplo 10.3 Una muestra aleatoria de 100 muertes registradas en Estados Unidos el año pasado muestra una vida promedio de 71.8 años. Suponga una desviación estándar poblacional de 8.9 años, ¿esto parece indicar que la vida media hoy en día es mayor que 70 años? Utilice un nivel de significancia de 0.05.

SOLUCIÓN

1. $H_0: \mu = 70$ años.
2. $H_1: \mu > 70$ años.

- $\alpha = 0.05$.
- Región crítica: $z > 1.645$, donde $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$.
- Cálculos: $\bar{x} = 71.8$ años, $\sigma = 8.9$ años y $z = \frac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$.
- Decisión: rechazar H_0 y concluir que la vida media hoy día es mayor que 70 años.

En el ejemplo 10.3 el valor P que corresponde a $z = 2.02$ está dado por el área de la región sombreada en la figura 10.10.

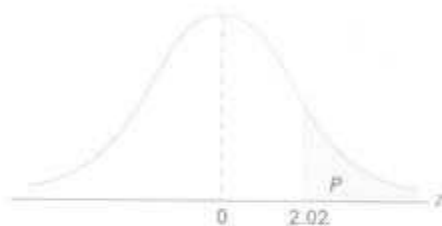


Figura 10.10 Valor P para el ejemplo 10.3.

Con el uso de la tabla A.3, tenemos

$$P = P(Z > 2.02) = 0.0217.$$

Como resultado, la evidencia a favor de H_1 es incluso más fuerte que la sugerida por un nivel de significancia del 0.05. ■

Ejemplo 10.4 Un fabricante de equipo deportivo desarrolla un nuevo sedal sintético que afirma tiene una resistencia media a la tensión de ocho kilogramos con una desviación estándar de 0.5 kilogramos. Pruebe la hipótesis $\mu = 8$ kilogramos contra la alternativa $\mu \neq 8$ kilogramos si se prueba una muestra aleatoria de 50 sedales y se encuentra que tiene una resistencia media a la tensión de 7.8 kilogramos. Utilice un nivel de significancia de 0.01.

SOLUCIÓN

- $H_0: \mu = 8$ kilogramos.
- $H_1: \mu \neq 8$ kilogramos.
- $\alpha = 0.01$.
- Región crítica: $z < -2.575$ y $z > 2.575$, donde $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$.
- Cálculos: $\bar{x} = 7.8$ kilogramos, $n = 50$, y de aquí $z = \frac{7.8 - 8}{0.5/\sqrt{50}} = -2.83$.
- Decisión: rechazar H_0 y concluir que la resistencia promedio a la tensión *no es* igual a ocho sino que, de hecho, es menor que ocho kilogramos.

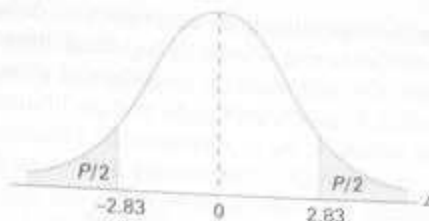


Figura 10.11 Valor P para el ejemplo 10.4.

Como la prueba en el ejemplo 10.4 es de dos colas, el valor de P que se desea es dos veces el área de la región sombreada de la figura 10.11 a la izquierda de $z = -2.83$. Por tanto, con el uso de la tabla A.3, tenemos:

$$P = P(|Z| > 2.83) = 2P(Z < -2.83) = 0.0046,$$

que nos permite rechazar la hipótesis nula $\mu = 8$ kilogramos en un nivel de significancia menor que 0.01. ■

10.6 Relación con la estimación del intervalo de confianza

El lector ya se habrá dado cuenta que el enfoque de la prueba de hipótesis para la inferencia estadística en este capítulo está relacionado muy de cerca con el enfoque del intervalo de confianza del capítulo 9. La estimación del intervalo de confianza incluye el cálculo de límites para los que es "razonable" que el parámetro en cuestión se encuentre dentro de ellos. Para el caso de una sola media poblacional μ con σ^2 conocida, la estructura de la prueba de hipótesis y de la estimación del intervalo de confianza se basan en la variable aleatoria

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

Resulta que la prueba de $H_0: \mu = \mu_0$ contra $H_1: \mu \neq \mu_0$ a un nivel de significancia α es equivalente a calcular un intervalo de confianza de $100(1 - \alpha)\%$ sobre μ y rechazar H_0 si μ_0 no está dentro del intervalo de confianza. Si μ_0 está dentro del intervalo de confianza, la hipótesis no se rechaza. La equivalencia es muy intuitiva y bastante simple de ilustrar. Recuerde que con un valor observado \bar{x} el fracaso a rechazar H_0 a un nivel de significancia α implica que

$$-z_{\alpha/2} \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha/2},$$

que es equivalente a

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

La equivalencia del intervalo de confianza a la prueba de hipótesis se extiende a las diferencias entre dos medias, varianzas, razones de varianzas, etcétera. Como

resultado, el estudiante de estadística no debe considerar la estimación del intervalo de confianza y la prueba de hipótesis como formas separadas de la inferencia estadística. De esta manera, considere el ejemplo 9.2. El intervalo de confianza de 95% sobre la media está dado por los límites [2.50, 2.70]. Así, con la misma información muestral, no se rechazará una hipótesis bilateral sobre μ que incluya cualquier valor hipotético entre 2.50 y 2.70. A medida que regresemos a diferentes áreas de la prueba de hipótesis, la equivalencia con la estimación del intervalo de confianza se seguirá aplicando.

10.7 Una sola muestra: pruebas sobre una sola media (varianza desconocida)

Ciertamente sospechamos que las pruebas sobre una media poblacional μ con σ^2 desconocida, como la estimación del intervalo de confianza, debe incluir el uso de la distribución t de Student. Estrictamente hablando, la aplicación de la t de Student para los intervalos de confianza y para la prueba de hipótesis se desarrolla bajo las suposiciones siguientes. Las variables aleatorias X_1, X_2, \dots, X_n representan una muestra aleatoria de una distribución normal con μ y σ^2 desconocidas. Entonces la variable aleatoria $\sqrt{n}(\bar{X} - \mu)/S$ tiene una distribución t de Student con $n - 1$ grados de libertad. La estructura de la prueba es idéntica a la del caso con σ conocida, con la excepción de que el valor σ en la estadística de prueba se reemplaza por la estimación de S calculada y la distribución normal estándar se reemplaza con una distribución t . Como resultado, para la hipótesis bilateral

$$H_0: \mu = \mu_0,$$

$$H_1: \mu \neq \mu_0$$

el rechazo de H_0 en un nivel de significancia α resulta cuando una estadística t calculada

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

excede $t_{\alpha/2, n-1}$ o es menor que $-t_{\alpha/2, n-1}$. El lector debe recordar de los capítulos 8 y 9 que la distribución t es simétrica alrededor del valor cero. Así, esta región crítica de dos colas se aplica de forma similar a la del caso de σ conocida. Para la hipótesis bilateral en un nivel de significancia α , se aplican las regiones críticas de dos colas. Para $H_1: \mu > \mu_0$, el rechazo resulta cuando $t > t_{\alpha, n-1}$. Para $H_1: \mu < \mu_0$, la región crítica está dada por $t < -t_{\alpha, n-1}$.

Ejemplo 10.5 El Instituto Eléctrico Edison publica cifras del número anual de kilowatt-hora que gastan varios aparatos electrodomésticos. Se afirma que una aspiradora gasta un promedio de 46 kilowatt-hora al año. Si una muestra aleatoria de 12 hogares que se incluye en un estudio planeado indica que las aspiradoras gastan un promedio de 42 kilowatt-hora al año con una desviación estándar de 11.9 kilowatt-hora, ¿esto sugiere en un nivel de significancia de 0.05 que las aspiradoras gastan, en promedio, menos de 46 kilowatt-hora anualmente? Suponga que la población de kilowatt-hora es normal.

SOLUCIÓN

1. $H_0: \mu = 46$ kilowatt-hora.
2. $H_1: \mu < 46$ kilowatt-hora.
3. $\alpha = 0.05$.
4. Región crítica: $t < -1.796$, donde $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ con $\nu = 11$ grados de libertad.
5. Cálculos: $\bar{x} = 42$ kilowatt-hora, $s = 11.9$ kilowatt-hora, y $n = 12$. De aquí

$$t = \frac{42 - 46}{11.9/\sqrt{12}} = -1.16, \quad P = P(T < -1.16) = 0.135.$$
6. Decisión: no rechazar H_0 y concluir que el número promedio de kilowatt-hora que gastan al año las aspiradoras domésticas no es significativamente menor que 46. ■

Comentario sobre la prueba T de una sola muestra

Es probable que el lector note que se mantiene la equivalencia de la prueba t de dos colas para una sola media y el cálculo de un intervalo de confianza sobre μ con σ reemplazada por s . Entonces, considere el ejemplo 9.4. Esencialmente, podemos ver ese cálculo como uno en el que encontramos todos los valores de μ_0 , el volumen medio hipotético de contenedores de ácido sulfúrico, para los que la hipótesis $H_0: \mu = \mu_0$ no se rechazará con $\alpha = 0.05$. Nuevamente, esto es consistente con el planteamiento: "Con base en la información muestral, los valores del volumen medio de la población entre 9.74 y 10.26 litros son razonables."

Vale la pena en este momento enfatizar algunos comentarios con respecto a la suposición de normalidad. Indicamos que cuando se conoce σ , el teorema del límite central permite el uso de una estadística de prueba o de un intervalo de confianza que se base en Z , la variable aleatoria normal estándar. Hablando estrictamente, por supuesto, el teorema del límite central y por ello el uso de la normal estándar no se aplica a menos que se conozca σ . Ahora bien, en el capítulo ocho, se da el desarrollo de la distribución t . En ese momento se estableció que la normalidad sobre X_1, X_2, \dots, X_n era una suposición básica. De esta manera, *en sentido estricto*, las tablas de la t de Student de puntos porcentuales para pruebas o intervalos de confianza no se deben utilizar a menos que se sepa que la muestra proviene de una población normal. En la práctica, σ rara vez se puede suponer conocida. Sin embargo, se puede disponer de una buena estimación a partir de experimentos anteriores. Muchos libros de estadística sugieren que se puede reemplazar con seguridad σ por s en la estadística de prueba

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

cuando $n \geq 30$ y aún utilizar las tablas Z para la región crítica apropiada. Aquí, la implicación es que en realidad se recurre al teorema del límite central y se cuenta con el hecho de que $s = \sigma$. Obviamente cuando se hace esto el resultado se debe ver

como aproximación. De esta manera un valor P calculado (de la distribución Z) de 0.15 puede ser 0.12 o quizá 0.17, o un intervalo de confianza calculado puede ser un intervalo de confianza de 93% en lugar de un intervalo de 95% como se desea. Ahora bien, ¿qué sucede con las situaciones donde $n \leq 30$? El usuario no puede confiar en que s es cercana a σ , y a fin de tomar en cuenta la inexactitud de la estimación, el intervalo de confianza debe ser más amplio o el valor crítico de magnitud más grande. Los puntos porcentuales de la distribución t realizan esto pero son correctos sólo cuando la muestra proviene de una distribución normal.

Para muestras pequeñas, a menudo es difícil detectar desviaciones de una distribución normal. (Las pruebas de la bondad de ajuste se presentan en una de las últimas secciones de este capítulo). Para distribuciones en forma de campana de las variables aleatorias X_1, X_2, \dots, X_n , el uso de la distribución t para pruebas o intervalos de confianza es probable que sea bastante buena. Cuando haya duda, el usuario debe recurrir a los procedimientos no paramétricos que se presentan en el capítulo 16.

Resultados de computadora comentados para pruebas T de una sola muestra

Debe ser de interés para el lector ver resultados de computadora comentados que muestren el resultado de una prueba t de una sola muestra. Suponga que un ingeniero se interesa en probar el sesgamiento en un medidor de pH. Se reúnen datos de una sustancia neutra ($\text{pH} = 7.0$). Se toma una muestra de las mediciones y los datos son los siguientes:

7.07 7.00 7.10 6.97 7.00
7.03 7.01 7.01 6.98 7.08

Es, entonces, de interés probar

$$H_0: \mu = 7.0,$$

$$H_1: \mu \neq 7.0.$$

En esta ilustración utilizamos el paquete de computadora MINITAB para ilustrar el análisis del conjunto de datos anterior. Nótese los componentes clave del impreso que se muestra en la figura 10.12. Por supuesto, la media $\bar{y} = 7.0250$, STDEV es simplemente la desviación estándar de la muestra $s = 0.440$, y SE MEAN es el error estándar estimado de la media y se calcula como $s/\sqrt{n} = 0.0139$. El valor t es la razón

$$(7.0250 - 7)/(0.0139) = 1.80.$$

```

H0: mu = 7 vs. H1: mu != 7
pH-meter
7.07 7.10 7.00 7.01 6.98 7.00 6.97 7.03 7.01 7.08
MTB > ttest mu = 7 'pH-meter'
TEST OF MU = 7.0000 VS MU N.E. 7.0000
      N    MEAN    STDEV  SE MEAN      T    P VALUE
pH-meter 10 7.0250 0.4400 0.0139  1.80    0.11

```

Figura 10.12 Impreso de MINITAB para la prueba t de una sola muestra para el medidor de pH.

El valor P de 0.11 sugiere resultados no concluyentes. No hay un fuerte rechazo de H_0 (con base en una α de 0.05 o de 0.10), en realidad aún no se puede concluir con certeza que el medidor de pH está insesgado. Nótese que el tamaño de la muestra de 10 es más bien pequeño. Un aumento en el tamaño de la muestra (quizá otro experimento) puede resolver las cosas. En la sección 10.10 aparece una discusión con respecto al tamaño apropiado de la muestra.

10.8 Dos muestras: pruebas sobre dos medias

El lector ya llegó a comprender la relación entre pruebas e intervalos de confianza y puede confiar por completo en los detalles que proporciona el material sobre el intervalo de confianza del capítulo 9. Las pruebas con respecto a dos medias representan un conjunto de herramientas analíticas muy importantes para el científico o el ingeniero. El procedimiento experimental es muy parecido al que se describe en la sección 9.7. Se extraen dos muestras aleatorias independientes de tamaño n_1 y n_2 , respectivamente, de dos poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 . Sabemos que la variable aleatoria

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

tiene una distribución normal estándar. Suponemos aquí que n_1 y n_2 son suficientemente grandes por lo que se aplica el teorema del límite central. Por supuesto, si las dos poblaciones son normales, la estadística anterior tiene una distribución normal estándar aun para n_1 y n_2 pequeñas. Obviamente, si podemos suponer que $\sigma_1 = \sigma_2 = \sigma$, la estadística anterior se reduce a

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma \sqrt{1/n_1 + 1/n_2}}$$

Las dos estadísticas anteriores sirven como base para el desarrollo de los procedimientos de prueba que incluyen dos medias. La equivalencia con el intervalo de confianza y la facilidad de la transición del caso de pruebas sobre una sola media hacen que esto sea simple.

La hipótesis bilateral sobre dos medias se puede escribir con bastante generalidad como

$$H_0: \mu_1 - \mu_2 = d_0.$$

Obviamente, la alternativa puede ser bilateral o unilateral. De nuevo, la distribución que se utiliza es la distribución de la estadística de prueba bajo H_0 . Los valores \bar{x}_1 y \bar{x}_2 se calculan y para σ_1 y σ_2 conocidas, la estadística de prueba está dada por

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

con una región crítica de dos colas en el caso de una alternativa bilateral. Es decir, rechazar H_0 a favor de $H_1: \mu_1 - \mu_2 \neq d_0$ si $z > z_{\alpha/2}$ o $z < -z_{\alpha/2}$. Las regiones críticas de una cola se utilizan en el caso de alternativas unilaterales. El lector debe estudiar, como antes, la estadística de prueba y estar satisfecho de que para, digamos, $H_1: \mu_1 - \mu_2 > d_0$, la señal que favorece a H_1 provenga de valores grandes de z . De esta manera se aplica la región crítica de la cola superior.

Varianzas desconocidas

Las situaciones que más prevalecen que implican pruebas sobre dos medias son las que tienen varianzas desconocidas. Si el científico involucrado está dispuesto a suponer que ambas distribuciones son normales y que $\sigma_1 = \sigma_2 = \sigma$, se puede utilizar la prueba *t* combinada (a menudo llamada prueba *t* de dos muestras). La estadística de prueba (véase la sección 9.7) está dada por el siguiente procedimiento de prueba.

Prueba *T* combinada de dos muestras:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{1/n_1 + 1/n_2}}$$

donde

$$s_p^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

Se incluye la distribución *t* y la hipótesis bilateral *no se rechaza* cuando

$$-t_{\alpha/2, n_1 + n_2 - 2} < t < t_{\alpha/2, n_1 + n_2 - 2}$$

Recuerde del material del capítulo 9 que los grados de libertad para la distribución *t* son un resultado de la combinación de la información de las dos muestras para estimar σ^2 . Las alternativas unilaterales sugieren regiones críticas unilaterales, como era de esperarse. Por ejemplo, para $H_1: \mu_1 - \mu_2 > d_0$, se rechaza $H_0: \mu_1 - \mu_2 = d_0$ cuando $t > t_{\alpha, n_1 + n_2 - 2}$.

Ejemplo 10.6 Se lleva a cabo un experimento para comparar el desgaste por abrasivo de dos diferentes materiales laminados. Se prueban 12 piezas del material 1 mediante la exposición de cada pieza a una máquina para medir el desgaste. Diez piezas del material 2 se prueban de manera similar. En cada caso, se mide la profundidad del desgaste. Las muestras del material 1 dan un desgaste promedio (codificado) de 85 unidades con una desviación estándar muestral de 4, mientras que las muestras del material 2 dan un promedio de 81 y una desviación estándar muestral de 5. ¿Podemos concluir con un nivel de significancia del 0.05 que el desgaste abrasivo del material 1 excede el del material 2 en más de 2 unidades? Suponga que las poblaciones son aproximadamente normales con varianzas iguales.

SOLUCIÓN

Representemos con μ_1 y μ_2 las medias poblacionales del desgaste abrasivo para el material 1 y el material 2, respectivamente.

1. $H_0: \mu_1 - \mu_2 = 2$.
2. $H_1: \mu_1 - \mu_2 > 2$.
3. $\alpha = 0.05$.

4. Región crítica: $t > 1.725$, donde $t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{1/n_1 + 1/n_2}}$ con $v = 20$ grados de libertad.

5. Cálculos:

$$\begin{aligned}\bar{x}_1 &= 85, & s_1 &= 4, & n_1 &= 12, \\ \bar{x}_2 &= 81, & s_2 &= 5, & n_2 &= 10.\end{aligned}$$

De aquí

$$s_p = \sqrt{\frac{(11)(16) + (9)(25)}{12 + 10 - 2}} = 4.478,$$

$$t = \frac{(85 - 81) - 2}{4.478 \sqrt{(1/12) + (1/10)}} = 1.04,$$

$$P = P(T > 1.04) = 0.16.$$

6. Decisión: no rechazar H_0 . Somos incapaces de concluir que el desgaste abrasivo del material 1 excede el del material 2 en más de 2 unidades. ■

Varianzas desconocidas pero diferentes

Existen situaciones donde el analista **no** es capaz de suponer que $\sigma_1 = \sigma_2$. Recuerde del capítulo 9 que, si las poblaciones son normales, la estadística

$$t' = \frac{(\bar{X} - \bar{X}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

tiene una distribución t aproximada con grados de libertad aproximados

$$v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

Como resultado el procedimiento de prueba es *no rechazar H_0* cuando

$$-t_{\alpha/2, v} \leq t' < t_{\alpha/2, v}$$

con v dado como antes. De nuevo, como en el caso de la prueba t combinada, las alternativas unilaterales sugieren regiones críticas unilaterales.

Observaciones pareadas

Cuando el estudiante de estadística estudia la prueba t de dos muestras o el intervalo de confianza sobre la diferencia entre medias, se debe dar cuenta que algunas nociones elementales que se tratan en el diseño experimental se vuelven relevantes y se deben considerar. Recuerde la discusión sobre las unidades experimentales en el capítulo 9, donde se sugiere en ese momento que la condición de las dos poblaciones (a menudo denominadas los dos tratamientos) se deben asignar de manera aleatoria a las unidades experimentales. Esto se realiza para evitar resultados sesgados debido a diferencias sistemáticas entre las unidades experimentales. En otras palabras, en términos de la jerga en pruebas de hipótesis, es importante que la diferencia significativa que se encuentra (o que no se encuentra) entre las medias se deba a las diferentes condiciones de las poblaciones y no debido a las unidades

experimentales en el estudio. Por ejemplo, considere el ejercicio 6, sección 9.8. Los 20 retoños juegan el papel de unidades experimentales. Diez de ellas se tratan con nitrógeno y 10 sin nitrógeno. Puede ser muy importante que esta asignación al tratamiento "con nitrógeno" y "sin nitrógeno" sea aleatoria para asegurar que las diferencias sistemáticas entre los retoños no interfieran con una comparación válida entre las medias.

En el ejemplo 10.6, el tiempo de medición es la elección más probable de la unidad experimental. Las 22 piezas de material se deben medir en orden aleatorio. Necesitamos protegernos contra la posibilidad de que las mediciones del desgaste que se realicen casi al mismo tiempo puedan tender a dar resultados similares. No se esperan diferencias sistemáticas (no aleatorias) en las unidades experimentales. Sin embargo, las asignaciones aleatorias protegen contra el problema.

Referencias a la planeación de experimentos, aleatorización, elección del tamaño de la muestra, etcétera, continuarán influyendo en gran parte del desarrollo en los capítulos 13, 14 y 15. Cualquier científico o ingeniero cuyo interés descansa en el análisis de datos reales debe estudiar este material. La prueba t combinada se amplía en el capítulo 13 para cubrir más de dos medias.

La prueba de dos medias se puede llevar a cabo cuando los datos están en la forma de observaciones pareadas como se discutió en el capítulo 9. En esta estructura de pareamiento, las condiciones de las dos poblaciones (tratamientos) se asignan de forma aleatoria dentro de unidades homogéneas. El cálculo del intervalo de confianza para $\mu_1 - \mu_2$ en la situación con observaciones pareadas se basa en la variable aleatoria

$$T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}},$$

donde \bar{D} y S_D son variables aleatorias que representan la media muestral y las desviaciones estándar de las diferencias de las observaciones en las unidades experimentales. Como en el caso de la *prueba t combinada*, la suposición es que las observaciones de cada población son normales. Este problema de dos muestras se reduce en esencia a un problema de una sola muestra mediante el uso de las diferencias calculadas d_1, d_2, \dots, d_n . De esta manera la hipótesis se reduce a

$$H_0: \mu_D = d_0.$$

La estadística de prueba calculada está dada entonces por

$$t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}.$$

Las regiones críticas se construyen con el uso de la distribución t con $n - 1$ grados de libertad.

Ejemplo 10.7 En un estudio conducido en el Departamento de Silvicultura y Fauna del Instituto Politécnico y Universidad Estatal de Virginia, J. A. Wesson examinó la influencia del fármaco *succinylcholine* sobre los niveles de circulación de

andrógenos en la sangre. Se obtuvieron muestras de sangre de venados salvajes vía la vena yugular inmediatamente después de una inyección intramuscular de succinylcholine con un rifle de caza con dardos. Los venados se sangraron nuevamente aproximadamente 30 minutos después de la inyección y después se liberaron. Los niveles de andrógenos al momento de la captura y 30 minutos después, medidos en nanogramos por mililitro (ng/ml), para 15 venados son los siguientes:

Venado	Al momento de la inyección	Andrógenos (ng/ml) 30 minutos después de la inyección	d_i
1	2.76	7.02	4.26
2	5.18	3.10	-2.08
3	2.68	5.44	2.76
4	3.05	3.99	0.94
5	4.10	5.21	1.11
6	7.05	10.26	3.21
7	6.60	13.91	7.31
8	4.79	18.53	13.74
9	7.39	7.91	0.52
10	7.30	4.85	-2.45
11	11.78	11.10	-0.68
12	3.90	3.74	-0.16
13	26.00	94.03	68.03
14	67.48	94.03	26.55
15	17.04	41.70	24.66

Suponga que las poblaciones de andrógenos al momento de la inyección y 30 minutos después se distribuyen normalmente, pruebe a un nivel de significancia de 0.05 si las concentraciones de andrógenos se alteran después de 30 minutos de encierro.

SOLUCIÓN

Sean μ_1 y μ_2 la concentración promedio de andrógenos al momento de la inyección y 30 minutos después, respectivamente. Procedemos como sigue:

1. $H_0: \mu_1 = \mu_2$ o $\mu_D = \mu_1 - \mu_2 = 0$.
2. $H_1: \mu_1 \neq \mu_2$ o $\mu_D = \mu_1 - \mu_2 \neq 0$.
3. $\alpha = 0.05$.

4. Región crítica: $t < -2.145$ y $t > 2.145$, donde $t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$ con $v = 14$ grados de libertad.

5. Cálculos: la media muestral y la desviación estándar para las d_i son

$$\bar{d} = 9.848 \quad \text{y} \quad s_d = 18.474.$$

Por tanto,

$$t = \frac{9.848 - 0}{18.474/\sqrt{15}} = 2.06.$$

6. Aunque la estadística t no es significativa al nivel 0.05,

$$P = P(|T| > 2.06) = 0.06.$$

Como resultado, existe alguna evidencia de que hay una diferencia en los niveles medios circulantes de andrógenos.

En el caso de observaciones pareadas, es importante que no haya interacción entre los tratamientos y las unidades experimentales. Esto se discutió en el capítulo 9 en el desarrollo de intervalos de confianza. La suposición de no interacción implica que el efecto de la unidad experimental, o pareada, es el mismo para cada uno de los dos tratamientos. En el ejemplo 10.7, suponemos que el efecto en el venado es el mismo para las dos condiciones bajo estudio, a saber "en el momento de la inyección" y 30 minutos después de la inyección. ■

Resultados de computadora comentados para la prueba *T* pareada

FiLa figura 10.13 muestra un impreso de computadora del SAS para una prueba *t* pareada con el uso de los datos del ejemplo 10.7. Nótese que la apariencia del impreso es la de una prueba *t* de una sola muestra y, por supuesto, esto es exactamente lo que se llevó a cabo pues la prueba busca determinar si d es significativamente diferente de cero.

Analysis Variable : DIFF Difference in Levels of Androgens				
N	Mean	Std Error	T	Prob> T
15	9.8480000	4.7698699	2.0646265	0.0580

Figura 10.13 Impreso del SAS de la prueba *t* pareada para los datos del ejemplo 10.7.

Resumen de los procedimientos de prueba

Dado que completamos el desarrollo formal de pruebas sobre medias poblacionales, ofrecemos la tabla 10.2, que resume el procedimiento de prueba para los casos de una sola media y de dos medias. Nótese el procedimiento aproximado cuando las distribuciones son normales y las varianzas se desconocen pero se suponen iguales. Esta estadística se introdujo en el capítulo 9.

10.9 Elección del tamaño de la muestra para probar medias

En la sección 10.2 demostramos como el analista puede explotar las relaciones entre el tamaño de la muestra, el nivel de significancia α y la potencia de la prueba para alcanzar cierto estándar de calidad. En la mayor parte de las circunstancias prácticas el experimento se debe planear con la elección de un tamaño muestral que se realiza antes del proceso de recolección de datos si es posible. El tamaño de la muestra por lo general se hace para lograr una buena potencia para una α fija y una alternativa específica fija. Esta alternativa fija puede ser en la forma de $\mu = \mu_0$ en el caso de una hipótesis que incluya una sola media o $\mu_1 = \mu_2$ en el caso de un problema que involucre dos medias. Los casos específicos serán ilustrativos.

Tabla 10.2 Pruebas relacionadas con medias

H_0	Valor de la estadística de prueba	H_1	Región crítica
$\mu = \mu_0$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ conocida	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ y $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n - 1$, σ desconocida	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ y $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$ σ_1 y σ_2 conocidas	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ y $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{(1/n_1) + (1/n_2)}}$ $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ pero desconocidas $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ y $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t' = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$ $v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2 + (s_2^2/n_2)^2}$ $\sigma_1 \neq \sigma_2$ y desconocidas	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t' < -t_\alpha$ $t' > t_\alpha$ $t' < -t_{\alpha/2}$ y $t' > t_{\alpha/2}$
$\mu_D = d_0$	$t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$; $v = n - 1$, observaciones apareadas	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ y $t > t_{\alpha/2}$

Suponga que deseamos probar la hipótesis

$$H_0: \mu = \mu_0,$$

$$H_1: \mu \geq \mu_0,$$

con un nivel de significancia α cuando la varianza σ^2 se conoce. Para una alternativa específica, digamos $\mu = \mu_0 + \delta$, se muestra, en la figura 10.14, que la potencia de nuestra prueba es

$$1 - \beta = P(X > a \text{ cuando } \mu = \mu_0 + \delta).$$

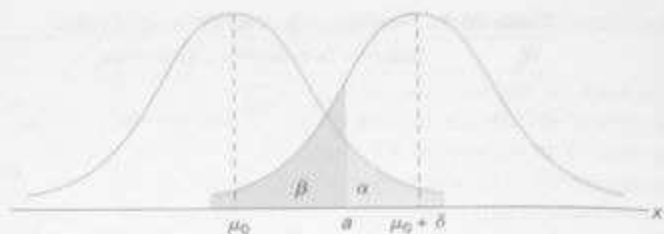


Figura 10.14 Prova de $\mu = \mu_0$ contra $\mu = \mu_0 + \delta$.

Por tanto,

$$\begin{aligned} \beta &= P(\bar{X} < a \text{ cuando } \mu = \mu_0 + \delta) \\ &= P\left[\frac{\bar{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}} < \frac{a - (\mu_0 + \delta)}{\sigma/\sqrt{n}} \text{ cuando } \mu = \mu_0 + \delta\right]. \end{aligned}$$

Bajo la hipótesis alternativa $\mu = \mu_0 + \delta$, la estadística

$$\frac{\bar{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}}$$

es la variable normal estándar Z . Por tanto,

$$\beta = P\left(Z < \frac{a - \mu_0}{\sigma/\sqrt{n}} - \frac{\delta}{\sigma/\sqrt{n}}\right) = P\left(Z < z_\alpha - \frac{\delta}{\sigma/\sqrt{n}}\right),$$

de donde concluimos que

$$-z_\beta = z_\alpha - \frac{\delta\sqrt{n}}{\sigma}$$

y de aquí

Selección del tamaño de la muestra

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2},$$

resultado que también es verdadero cuando la hipótesis alternativa es $\mu < \mu_0$.

En el caso de una prueba de dos colas obtenemos la potencia $1 - \beta$ para una alternativa específica cuando

$$n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2}.$$

Ejemplo 10.8 Suponga que deseamos probar la hipótesis

$$H_0: \mu = 68 \text{ kilogramos,}$$

$$H_1: \mu > 68 \text{ kilogramos}$$

para los pesos de estudiantes varones en cierta universidad con el uso de un nivel de significancia $\alpha = 0.05$ cuando se sabe que $\sigma = 5$. Encuentre el tamaño muestral que se requiere si la potencia de nuestra prueba debe ser 0.95 cuando la media real es 69 kilogramos.

SOLUCIÓN

Como $\alpha = \beta = 0.05$, tenemos $z_\alpha = z_\beta = 1.645$. Para la alternativa $\mu = 69$, tomamos $\delta = 1$ y entonces

$$n = \frac{(1.645 + 1.645)^2(25)}{1} = 270.6.$$

Por tanto, se requieren 271 observaciones si la prueba debe rechazar la hipótesis nula 95% de las veces cuando, de hecho, μ es tan grande como 69 kilogramos.

Se puede utilizar un procedimiento similar para determinar el tamaño de la muestra $n = n_1 = n_2$ que se requiere para una potencia específica de la prueba en la que se comparan dos medias poblacionales. Por ejemplo, suponga que deseamos probar la hipótesis

$$H_0: \mu_1 - \mu_2 = d_0,$$

$$H_1: \mu_1 - \mu_2 \neq d_0,$$

cuando se conocen σ_1 y σ_2 . Para una alternativa específica, digamos $\mu_1 - \mu_2 = d_0 + \delta$, en la figura 10.15 se muestra que la potencia de nuestra prueba es

$$1 - \beta = P(|\bar{X}_1 - \bar{X}_2| > a \text{ cuando } \mu_1 - \mu_2 = d_0 + \delta).$$

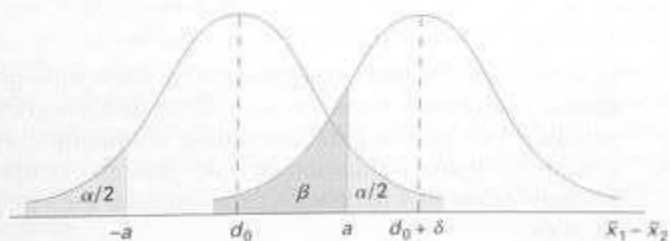


Figura 10.15 Prueba de $\mu_1 - \mu_2 = d_0$ contra $\mu_1 - \mu_2 = d_0 + \delta$.

Por tanto,

$$\begin{aligned} \beta &= P(-a < \bar{X}_1 - \bar{X}_2 < a \text{ cuando } \mu_1 - \mu_2 = d_0 + \delta) \\ &= P\left[\frac{-a - (d_0 + \delta)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < \frac{\bar{X}_1 - \bar{X}_2 - (d_0 + \delta)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \right. \\ &\quad \left. < \frac{a - (d_0 + \delta)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \text{ cuando } \mu_1 - \mu_2 = d_0 + \delta\right]. \end{aligned}$$

Bajo la hipótesis alternativa $\mu_1 - \mu_2 = d_0 + \delta$, la estadística

$$\frac{\bar{X}_1 - \bar{X}_2 - (d_0 + \delta)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

es la variable normal estándar Z . Ahora bien, al escribir

$$-z_{\alpha/2} = \frac{-a - d_0}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \quad \text{y} \quad z_{\alpha/2} = \frac{a - d_0}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

tenemos

$$\beta = P\left[-z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < Z < z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}\right],$$

de donde concluimos que

$$-z_{\beta} = z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \quad \text{y de aquí} \quad n = \frac{(z_{\alpha/2} + z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}.$$

Para la prueba de una sola cola, la expresión para el tamaño requerido de la muestra cuando $n = n_1 = n_2$ es

Elección del tamaño de la muestra

$$n = \frac{(z_{\alpha} + z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{\delta^2}.$$

Cuando se desconoce la varianza poblacional (o varianzas en la situación de dos muestras), la elección del tamaño de la muestra no es directa. Al probar la hipótesis $\mu = \mu_0$ cuando el valor real es $\mu = \mu_0 + \delta$, la estadística

$$\frac{\bar{X} - (\mu_0 + \delta)}{S/\sqrt{n}}$$

no sigue la distribución t , como podría esperarse, sino que en su lugar sigue la **distribución t no central**. Sin embargo, existen tablas o gráficas que se basan en la distribución t no central para determinar el tamaño apropiado de la muestra si se dispone de alguna estimación de σ o si δ es un múltiplo de σ . La tabla A.8 da los tamaños muestrales necesarios para controlar los valores de α y β para varios valores de

$$\Delta = \frac{|\delta|}{\sigma} = \frac{|\mu - \mu_0|}{\sigma}$$

para pruebas de una y de dos colas. En el caso de la prueba t de dos muestras en la que se desconocen las varianzas pero se suponen iguales, obtenemos los tamaños muestrales $n = n_1 = n_2$ necesarios para controlar los valores de α y β para varios valores de

$$\Delta = \frac{|\delta|}{\sigma} = \frac{|\mu_1 - \mu_2 - d_0|}{\sigma} \quad \text{de la tabla A.9.} \quad \blacksquare$$

Ejemplo 10.9 Al comparar el comportamiento de dos catalizadores sobre el efecto del rendimiento de una reacción, se realiza una prueba t de dos muestras con

$\alpha = 0,05$. Las varianzas de los rendimientos se consideran las mismas para los dos catalizadores. ¿De qué tamaño se necesita una muestra para cada catalizador para probar la hipótesis

$$H_0: \mu_1 = \mu_2,$$

$$H_1: \mu_1 \neq \mu_2$$

si es esencial detectar una diferencia de $0,8\sigma$ entre los catalizadores con probabilidad 0,9?

SOLUCIÓN

De la tabla A.9, con $\alpha = 0,05$ para una prueba de dos colas, $\beta = 0,1$, y

$$\Delta = \frac{|0,8\sigma|}{\sigma} = 0,8,$$

encontramos que el tamaño de la muestra que se requiere es $n = 34$.

Se enfatiza que en situaciones prácticas puede ser difícil forzar a un científico o a un ingeniero a hacer un compromiso sobre la información de la que se puede encontrar un valor de Δ . Se recuerda al lector que el valor Δ cuantifica la clase de diferencia entre las medias que el científico considera importantes; es decir, una diferencia que se considere *significativa* desde un punto de vista científico, no estadístico. El ejemplo 10.9 ilustra cómo se hace a menudo esta elección, a saber, mediante la selección de una fracción de σ . Obviamente, si el tamaño de la muestra se basa en una elección de $|\delta|$ que es una pequeña fracción de σ , el tamaño muestral que resulta puede ser bastante grande comparado con lo que el estudio permite. ■

10.10 Métodos gráficos para comparar medias

En el capítulo 3 se pone una considerable atención hacia la presentación de datos en forma gráfica. Los diagramas de tronco y hojas, en el capítulo 8, y las gráficas de caja y extensión, gráficas de cuantiles y gráficas normales cuantil-cuantil se utilizan para proporcionar una "imagen" y resumir así un conjunto de datos experimentales. Muchos paquetes de software computacional producen representaciones gráficas. A medida que procedamos con otras formas de análisis de datos (por ejemplo, el análisis de regresión y el análisis de varianza), los métodos gráficos se hacen aún más informativos.

Las ayudas gráficas que se utilizan junto con la prueba de hipótesis no se usan como un reemplazo del procedimiento de prueba. En realidad, el valor de la estadística de prueba indica el tipo apropiado de evidencia en apoyo de H_0 o H_1 . Sin embargo, una representación como imagen proporciona una buena ilustración y a menudo es un mejor comunicador de evidencia para el beneficiario del análisis. Además, una imagen con frecuencia dejará claro por qué se encontró una diferencia significativa. La falla de una suposición importante se puede descubrir mediante un resumen de tipo gráfico.

Para la comparación de las medias, las gráficas de caja y extensión de lado a lado proporcionan una presentación reveladora. El lector debe recordar que estas gráficas muestran el percentil 25, el percentil 75 y la mediana en un conjunto de datos. Además, las extensiones muestran los extremos en un conjunto de datos. Con-

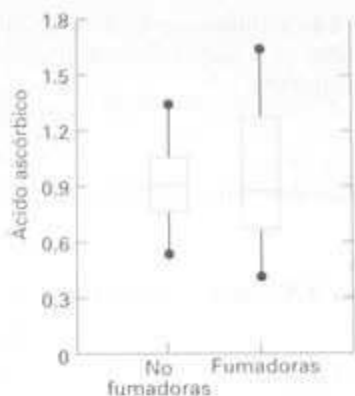


Figura 10.16 Gráfica múltiple de caja y extensión para el ácido ascórbico en plasma en fumadoras y no fumadoras.

sidere el ejercicio 22 que sigue a esta sección. Se miden los niveles en plasma de ácido ascórbico en dos grupos de mujeres embarazadas, fumadoras y no fumadoras. La figura 10.16 muestra las gráficas de caja y extensión para ambos grupos de mujeres. Dos cosas son muy evidentes. Al tomar en cuenta la variabilidad, parece haber una diferencia insignificante en las medias muestrales. Además, la variabilidad en los dos grupos parece ser algo diferente. Por supuesto, el analista debe tener en mente las diferencias más bien considerables entre los tamaños muestrales en este caso.

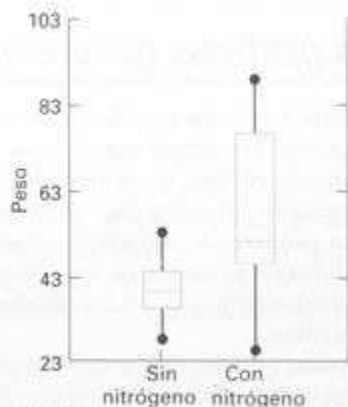


Figura 10.17 Gráfica múltiple de caja y extensión para los datos de los retoños.

Considere el ejercicio 6 que sigue de la sección 9.8. La figura 10.17 muestra la gráfica múltiple de caja y extensión para los datos de 10 retoños, la mitad con nitrógeno y la otra mitad sin nitrógeno. La gráfica revela una variabilidad más pequeña para el grupo que no contiene nitrógeno. Además, la falta de traslape de las cajas sugiere una diferencia significativa entre los pesos medios de los tallos de los dos grupos. Parecería que la presencia de nitrógeno aumenta el peso de los tallos y quizá aumenta la variabilidad en los pesos.

No existen ciertas reglas empíricas respecto a cuando las gráficas de caja y extensión proporcionan evidencia de diferencias significativas entre las medias. Sin

embargo, una pauta aproximada es que si la línea del percentil 25 para una muestra excede la línea de la mediana de la otra muestra, hay fuerte evidencia de una diferencia entre las medias.

Se hará más énfasis en los métodos gráficos en un estudio de caso de la vida real que se muestra más adelante en este capítulo.

Resultados de computadora comentados para una prueba T de dos muestras

Considere los datos del ejercicio 6 de la sección 9.8, donde se reunieron los datos de los retoños en condiciones con nitrógeno y sin nitrógeno. Probar

$$H_0: \mu_{\text{NIT}} = \mu_{\text{NON}}$$

$$H_1: \mu_{\text{NIT}} > \mu_{\text{NON}}$$

donde las medias poblacionales indican los pesos medios. La figura 10.18 es un impreso de computadora comentado del paquete SAS. Nótese que las desviaciones estándar y el error estándar se muestran para ambas muestras. Se da la estadística t bajo las suposiciones de "varianza igual" y de "varianza diferente". De la gráfica de caja y extensión de la figura 10.17 en realidad parecería que se viola la suposición de varianza igual. Un valor P de 0.0229 sugiere una conclusión de medias diferentes. Esto coincide con la información de diagnóstico que se da en la figura 10.17. Incidentalmente, nótese que t y t' son iguales en este caso, pues $n_1 = n_2$.

TTEST PROCEDURE				
Variable: WEIGHT				
MINERAL	N	Mean	Std Dev	Std Error
no nitrogen	10	0.34900000	0.07279347	0.02301932
nitrogen	10	0.56500000	0.18674106	0.05905271
Variances:		T	DF	Prob > T
Unequal	-2.6191	11.7	0.0229	
Equal	-2.6191	18.0	0.0174	
For HO: Variances are equal, F = 6.58 DF = (9,9)				
Prob > F = 0.0096				

Figura 10.18 Impreso del SAS para la prueba t de dos muestras.

Ejercicios

1. Una empresa eléctrica fabrica focos que tienen una duración que se distribuye de forma aproximadamente normal con una media de 800 horas y una desviación estándar de 40 horas. Pruebe la hipótesis de que $\mu = 800$ horas contra la alternativa $\mu \neq 800$ horas si una muestra aleatoria de 30 focos tiene una duración promedio de 788 horas. Utilice un nivel de significancia de 0.04.

2. Una muestra aleatoria de 64 bolsas de palomitas de maíz con queso cheddar pesan, en promedio, 5.23 onzas con una desviación estándar de 0.24 onzas.

Pruebe la hipótesis de que $\mu = 5.5$ onzas contra la hipótesis alternativa, $\mu < 5.5$ onzas en el nivel de significancia de 0.05.

3. En un informe de investigación de Richard H. Weindruch de la Escuela de Medicina de la UCLA, se afirma que los ratones con una vida promedio de 32 meses vivirán hasta alrededor de 40 meses de edad cuando 40% de las calorías en su comida se reemplacen por vitaminas y proteínas. ¿Hay alguna razón para creer que $\mu < 40$ si 64 ratones que se sujetan a esta dieta tienen una vida promedio de 38 meses con una

desviación estándar de 5.8 meses? Utilice un nivel de significancia del 0.025.

4. La estatura promedio de mujeres en el grupo de primer año de cierta universidad es de 162.5 centímetros con una desviación estándar de 6.9 centímetros. ¿Hay alguna razón para creer que hay un cambio en la estatura promedio si una muestra aleatoria de 50 mujeres en el grupo actual de primer año tiene una altura promedio de 165.2 centímetros? Utilice un valor P en su conclusión.

5. Se afirma que un automóvil se maneja en promedio más de 20,000 kilómetros por año. Para probar esta afirmación, se pide a una muestra de 100 propietarios de automóviles que lleven un registro de los kilómetros que viajen. ¿Está de acuerdo con esta afirmación si la muestra aleatoria muestra un promedio de 23,500 kilómetros y una desviación estándar de 3900 kilómetros? Utilice un valor P en su conclusión.

6. En el boletín de la Asociación Americana del Corazón, *Hypertension*, investigadores reportan que los individuos que practican la meditación trascendental (MT) bajan su presión sanguínea de manera significativa. Si una muestra aleatoria de 225 hombres practicantes de MT meditan 8.5 horas a la semana con una desviación estándar de 2.25 horas, ¿esto sugiere que, en promedio, los hombres que utilizan la MT meditan más de ocho horas a la semana? Cite un valor P en su conclusión.

7. Pruebe la hipótesis de que el contenido promedio de los envases de un lubricante particular es de 10 litros si los contenidos de una muestra aleatoria de 10 envases son 10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3 y 9.8 litros. Utilice un nivel de significancia de 0.01 y suponga que la distribución del contenido es normal.

8. De acuerdo con un estudio dietético una ingesta alta de sodio se puede relacionar con úlceras, cáncer de estómago y migraña. El requerimiento humano de sal es de sólo 220 miligramos por día, el cual se rebasa en la mayoría de las porciones individuales de cereales listos para comerse. Si una muestra aleatoria de 20 porciones similares de Special K tienen un contenido medio de 244 miligramos de sodio y una desviación estándar de 24.5 miligramos, ¿esto sugiere, en el nivel de significancia del 0.05, que el contenido promedio de sodio para porciones individuales de Special K es mayor que 220 miligramos? Suponga que la distribución de contenidos de sodio es normal.

9. Un estudio de la Universidad de Colorado en Boulder muestra que correr aumenta el porcentaje de la tasa metabólica de descanso (RMR) en mujeres an-

cianas. La RMR promedio de 30 ancianas corredoras fue 34.0 por ciento más alta que la RMR promedio de 30 ancianas sedentarias y las desviaciones estándar fueron 10.5 y 10.2 por ciento, respectivamente. ¿Hay un aumento significativo en la RMR de las corredoras con respecto a las sedentarias? Suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales. Utilice un valor P en sus conclusiones.

10. De acuerdo con el *Chemical Engineering* una propiedad importante de la fibra es su absorción de agua. Se encuentra que el porcentaje promedio de absorción de 25 piezas de algodón que se seleccionan al azar es 20 con una desviación estándar de 1.5. Una muestra aleatoria de 25 piezas de acetato dan un porcentaje promedio de 12 con una desviación estándar de 1.25. Hay fuerte evidencia de que el porcentaje promedio de absorción de la población para la fibra de algodón es significativamente más alta que la media para el acetato. Suponga que el porcentaje de absorción se distribuye de forma aproximadamente normal y que las varianzas de la población en el porcentaje de absorción para las dos fibras son las mismas. Utilice un nivel de significancia de 0.05.

11. La experiencia pasada indica que el tiempo para que los estudiantes de último año de preparatoria terminen un examen estandarizado es una variable aleatoria normal con una media de 35 minutos. Si a una muestra aleatoria de 20 estudiantes de último año de preparatoria le toma un promedio de 33.1 minutos completar esta prueba con una desviación estándar de 4.3 minutos, pruebe la hipótesis de que $\mu = 35$ minutos contra la alternativa de que $\mu < 35$ minutos con un nivel de significancia de 0.025.

12. Una muestra aleatoria de tamaño $n_1 = 25$, que se toma de una población normal con una desviación estándar $\sigma_1 = 5.2$, tiene una media $\bar{x}_1 = 81$. Una segunda muestra aleatoria de tamaño $n_2 = 36$, que se toma de una población normal diferente con una desviación estándar $\sigma_2 = 3.4$, tiene una media $\bar{x}_2 = 76$. Pruebe la hipótesis de que $\mu_1 = \mu_2$ contra la alternativa $\mu_1 \neq \mu_2$. Cite un valor P en su conclusión.

13. Un fabricante afirma que la resistencia a la tracción promedio del hilo A excede la resistencia a la tracción promedio del hilo B en al menos 12 kilogramos. Para probar esta afirmación, se prueban 50 piezas de cada tipo de hilo bajo condiciones similares. El hilo tipo A tiene una resistencia a la tracción promedio de 86.7 kilogramos con una desviación estándar de 6.28 kilogramos, mientras que el hilo tipo B tiene

una resistencia promedio a la tracción de 77.8 kilogramos con una desviación estándar de 5.61 kilogramos. Pruebe la afirmación del fabricante con el uso de un nivel de significancia de 0.05.

14. El *Amstat News* (febrero de 1997) lista los salarios medios de profesores asociados de estadística en instituciones de investigación y en escuelas de ciencias sociales y en otras instituciones en Estados Unidos. Suponga que una muestra de 200 profesores asociados de instituciones de investigación tienen un salario promedio de \$51,750 al año con una desviación estándar de \$5000. Suponga también una muestra de 200 profesores asociados de otros tipos de institución que tienen un salario promedio de \$47,500 con una desviación estándar de \$5000. Pruebe la hipótesis de que el salario medio de profesores asociados en instituciones de investigación es \$2000 más alto que los de los de otras instituciones. Utilice un nivel de significancia del 0.01.

15. Se lleva a cabo un estudio para ver si el aumento de la concentración de sustrato tiene un efecto apreciable sobre la velocidad de una reacción química. Con una concentración de sustrato de 1.5 moles por litro, la reacción se realizó 15 veces con una velocidad promedio de 7.5 micromoles por 30 minutos y una desviación estándar de 1.5. Con una concentración de sustrato de 2.0 moles por litro, se realizan 12 reacciones, que dan una velocidad promedio de 8.8 micromoles por 30 minutos y una desviación estándar muestral de 1.2. ¿Hay alguna razón para creer que este incremento en la concentración de sustrato ocasiona un aumento en la velocidad media por más de 0.5 micromoles por 30 minutos? Utilice un nivel de significancia de 0.01 y suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales.

16. Se realiza un estudio para determinar si los temas de la materia en un curso de física se comprenden mejor cuando un laboratorio constituye parte del curso. Se seleccionan estudiantes al azar para que participen en un curso de tres semestres-hora sin laboratorio o en un curso de cuatro semestres-hora con laboratorio. En la sección con laboratorio 11 estudiantes tuvieron una calificación promedio de 85 con una desviación estándar de 4.7, y en la sección sin laboratorio 17 estudiantes tuvieron una calificación promedio de 79 con una desviación estándar de 6.1. ¿Diría que el curso con laboratorio aumenta la calificación promedio hasta 8 puntos? Utilice un valor P en su conclusión y suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales.

17. Para encontrar si un nuevo suero detiene la leucemia, se seleccionan nueve ratones, todos con una etapa avanzada de la enfermedad. Cinco ratones reciben el tratamiento y cuatro no. Los tiempos de sobrevivencia, en años, a partir del momento en que comienza el experimento son los siguientes:

Con tratamiento	2.1	5.3	1.4	4.6	0.9
Sin tratamiento	1.9	0.5	2.8	3.1	

¿Se puede decir en el nivel de significancia del 0.05 que el suero es efectivo? Suponga que las dos distribuciones se distribuyen de forma normal con varianzas iguales.

18. Una compañía armadora de automóviles grande trata de decidir si compra llantas de la marca A o de la B para sus modelos nuevos. Se lleva a cabo un experimento, para ayudar a llegar a una decisión, en el que se usan 12 llantas de cada marca. Las llantas se utilizan hasta que se acaban. Los resultados son

$$\begin{aligned} \text{Marca A: } & \bar{x}_1 = 37,900 \text{ kilómetros,} \\ & s_1 = 5100 \text{ kilómetros.} \\ \text{Marca B: } & \bar{x}_2 = 39,800 \text{ kilómetros,} \\ & s_2 = 5900 \text{ kilómetros.} \end{aligned}$$

Pruebe la hipótesis de que no hay diferencia en las dos marcas de llantas con un nivel de significancia de 0.05. Suponga que las poblaciones se distribuyen de forma aproximadamente normal con varianzas iguales.

19. En el ejercicio 8 de la página 263, pruebe la hipótesis de que los camiones compactos Volkswagen, en promedio, exceden a los camiones compactos Toyota equipados de forma similar por cuatro kilómetros por litro. Utilice un nivel de significancia del 0.10.

20. Un investigador de la UCLA afirma que la vida promedio de un ratón se puede prolongar ocho meses más cuando las calorías en su comida se reducen en aproximadamente 40% desde el momento en que se destetan. Las dietas restringidas se enriquecen a niveles normales con vitaminas y proteína. Suponga que se alimenta una muestra aleatoria de 10 ratones con una dieta normal y tiene una vida promedio de 32.1 meses con una desviación estándar de 3.2 meses, mientras que una muestra aleatoria de 15 ratones se alimenta con la dieta restringida y viven un promedio de 37.6 meses con una desviación estándar de 2.8 meses. Pruebe con un nivel de significancia de 0.05 la hipótesis de que la vida promedio de los ratones con esta dieta restringida aumenta ocho meses contra la alternativa de que el aumento es menor que ocho meses. Suponga que las distribuciones de las vidas con

las dietas regular y restringida son aproximadamente normales con varianzas iguales.

21. Los siguientes datos representan los tiempos de duración de películas producidas por dos compañías cinematográficas:

Compañía	Tiempo (minutos)					
1	102	86	98	109	92	
2	81	165	97	134	92	87 114

Pruebe la hipótesis de que el tiempo de duración promedio de las películas producidas por la compañía 2 excede el tiempo promedio de duración de las que produce la compañía 1 en 10 minutos, contra la alternativa unilateral de que la diferencia es de más de 10 minutos. Utilice un nivel de significancia de 0.1 y suponga que las distribuciones de los tiempos son aproximadamente normales con varianzas iguales.

22. En un estudio realizado en el Instituto Politécnico y Universidad Estatal de Virginia, se compararon los niveles de ácido ascórbico en plasma en mujeres embarazadas fumadoras contra no fumadoras. Se seleccionaron para el estudio 32 mujeres en los últimos tres meses de embarazo, libres de padecimientos importantes y con edades entre 15 y 32 años. Antes de tomar las muestras de 20 ml de sangre, se pidió a las participantes ir en ayunas, no consumir sus complementos vitamínicos y evitar comidas con alto contenido de ácido ascórbico. De las muestras de sangre se determinaron los siguientes valores en miligramos por 100 mililitros de ácido ascórbico en plasma de cada mujer:

Valores de ácido ascórbico en plasma

No fumadoras	Fumadoras	
0.97	1.16	0.48
0.72	0.86	0.71
1.00	0.85	0.98
0.81	0.58	0.68
0.62	0.57	1.18
1.32	0.64	1.36
1.24	0.98	0.78
0.99	1.09	1.64
0.90	0.92	
0.74	0.78	
0.88	1.24	
0.94	1.18	

¿Existe suficiente evidencia para concluir que hay una diferencia entre los niveles de ácido ascórbico en plasma entre fumadoras y no fumadoras? Suponga

que los dos conjuntos de datos provienen de poblaciones normales con varianzas diferentes. Utilice un valor P .

23. El Departamento de Zoología del Instituto Politécnico y Universidad Estatal de Virginia lleva a cabo un estudio para determinar si hay una diferencia significativa en la densidad de organismos en dos estaciones diferentes colocadas en Cedar Run, un cauce secundario que se localiza en la cuenca del río Roanoke. El drenaje de una planta de tratamiento de aguas negras y el sobreflujo del estanque de sedimentación de la Federal Mogul Corporation entran al flujo cerca del nacimiento del río. Los siguientes datos dan las medidas de densidad, en número de organismos por metro cuadrado, en las dos diferentes estaciones colectoras:

Número de organismos por metro cuadrado			
Estación 1		Estación 2	
5030	4980	2800	2810
13,700	11,910	4670	1330
10,730	8130	6890	3320
11,400	26,850	7720	1230
860	17,660	7030	2130
2200	22,800	7330	2190
4250	1130		
15,040	1690		

¿Podemos concluir, con un nivel de significancia de 0.05, que las densidades promedio en las dos estaciones son iguales? Suponga que las observaciones provienen de poblaciones normales con varianzas diferentes.

24. Cinco muestras de una sustancia ferrosa se usan para determinar si hay una diferencia entre un análisis químico de laboratorio y un análisis de fluorescencia de rayos X del contenido de hierro. Cada muestra se divide en dos submuestras y se aplican los dos tipos de análisis. A continuación se presentan los datos codificados que muestran los análisis de contenido de hierro:

Análisis	Muestra				
	1	2	3	4	5
Rayos X	2.0	2.0	2.3	2.1	2.4
Químico	2.2	1.9	2.5	2.3	2.4

Suponga que las poblaciones son normales, pruebe con un nivel de significancia de 0.05 si los dos métodos de análisis dan, en promedio, el mismo resultado.

25. El administrador de una compañía de taxis trata de decidir si el uso de llantas radiales en lugar de llantas regulares de cinturón mejora la economía de combustible. Se equipan 12 autos con llantas radiales y se manejan por un recorrido de prueba preestablecido. Sin cambiar de conductores, los mismos autos se equipan con llantas comunes con cinturón y se manejan otra vez por el recorrido de prueba. El consumo de gasolina, en kilómetros por litro, se registró como sigue:

Auto	Kilómetros por litro	
	Llantas radiales	Llantas con cinturón
1	4.2	4.1
2	4.7	4.9
3	6.6	6.2
4	7.0	6.9
5	6.7	6.8
6	4.5	4.4
7	5.7	5.7
8	6.0	5.8
9	7.4	6.9
10	4.9	4.7
11	6.1	6.0
12	5.2	4.9

¿Podemos concluir que los autos equipados con llantas radiales dan una economía de combustible mejor que los equipados con llantas de cinturón? Suponga que las poblaciones se distribuyen normalmente. Utilice un valor P en su conclusión.

26. En el ejercicio 2 de la página 287, utilice la distribución t para probar la hipótesis de que la dieta reduce el peso de una persona en 4.5 kilogramos en promedio contra la hipótesis alternativa de que la diferencia media en peso es menor que 4.5 kilogramos. Utilice un valor P .

27. De acuerdo con reportes publicados, el ejercicio bajo condiciones de fatiga altera los mecanismos que gobiernan el comportamiento. Se lleva a cabo un experimento en el que se usan 15 estudiantes universitarios entrenados para realizar un movimiento horizontal continuo del brazo, de derecha a izquierda, desde un microconmutador hasta una barrera, golpeando sobre la barrera en coincidencia con la llegada de una manecilla del reloj a la posición de las seis en punto. Se registra el valor absoluto de la diferencia entre el tiempo, en milisegundos, que toma golpear sobre la barrera y el tiempo para que la manecilla alcance la posición de las seis en punto (50

ms). Cada participante ejecuta la tarea cinco veces bajo condiciones sin fatiga y con fatiga, y se registraron las sumas de las diferencias absolutas para las cinco ejecuciones como sigue:

Sujeto	Diferencias absolutas de tiempo (ms)	
	Sin fatiga	Con fatiga
1	158	91
2	92	59
3	65	215
4	98	226
5	33	223
6	89	91
7	148	92
8	58	177
9	142	134
10	117	116
11	74	153
12	66	219
13	109	143
14	57	164
15	85	100

Un aumento en las diferencias medias absolutas de tiempo cuando la tarea se ejecuta bajo condiciones de fatiga apoyaría la afirmación de que el ejercicio bajo condiciones de fatiga altera el mecanismo que gobierna el comportamiento. Con la suposición de que las poblaciones se distribuyen normalmente, pruebe esta afirmación.

28. En un estudio realizado por el Departamento de Nutrición Humana y Alimentos del Instituto Politécnico y Universidad Estatal de Virginia se registraron los siguientes datos acerca de la comparación de residuos de ácido sórbico, en partes por millón, en jamón inmediatamente después de sumergirlo en una solución de ácido y después de 60 días de almacenamiento:

Rebanada	Residuos de ácido sórbico en jamón	
	Antes del almacenamiento	Después del almacenamiento
1	224	116
2	270	96
3	400	239
4	444	329
5	590	437
6	660	597
7	1400	689
8	680	576

Si se supone que las poblaciones se distribuyen normalmente, ¿hay suficiente evidencia, al nivel de significancia de 0.05, para decir que la duración del almacenamiento influye en las concentraciones residuales de ácido sórbico?

29. ¿Qué tan grande se requiere que sea la muestra en el ejercicio 2 si la potencia de nuestra prueba debe ser 0.90 cuando la media real es 5.20? Suponga que $\sigma = 0.24$.

30. Si la distribución del tiempo de vida en el ejercicio 3 es aproximadamente normal, ¿qué tan grande se requiere que sea una muestra a fin de que la probabilidad de cometer un error tipo II sea 0.1 cuando la media real es 35.9 meses? Suponga que $\sigma = 5.8$ meses.

31. ¿Qué tan grande se requiere que sea la muestra en el ejercicio 4 si la potencia de nuestra prueba debe ser 0.95 cuando las estaturas promedio reales difieren de 162.5 en 3.1 centímetros?

32. ¿Qué tan grandes deben ser las muestras en el ejercicio 13 si la potencia de nuestra prueba debe ser 0.95 cuando la diferencia real entre los tipos de hilo A y B es ocho kilogramos?

33. ¿Qué tan grande se requiere que sea la muestra en el ejercicio 6 si la potencia de nuestra prueba será 0.8 cuando el tiempo medio real de meditación exceda el valor hipotético por 1.2σ ? Utilice $\alpha = 0.05$.

34. Se considera una prueba t de nivel $\alpha = 0.05$ para probar

$$H_0: \mu = 14,$$

$$H_1: \mu = 14,$$

¿Qué tamaño de la muestra se necesita a fin de que la probabilidad sea 0.1 o aceptar de manera errónea H_0 cuando la media poblacional real difiera de 14 en 0.5? Estimamos a partir de una muestra preliminar que σ es 1.25.

35. Se llevó a cabo un estudio en el Departamento de Medicina Veterinaria del Instituto Politécnico y Universidad Estatal de Virginia para determinar si la "resistencia" de una herida de incisión quirúrgica resulta afectada por la temperatura del bisturí. Se utilizaron ocho perros en el experimento. La incisión se realizó en el abdomen de los animales. Se aplicó una incisión "caliente" y una "fría" en cada perro y se midió la resistencia. Los datos que resultaron aparecen abajo.

(a) Escriba una hipótesis apropiada para determinar si hay una diferencia significativa en la resistencia entre las incisiones caliente y fría.

(b) Pruebe la hipótesis mediante el uso de una prueba t pareada. Utilice un valor P en su conclusión.

Perro Bisturí Resistencia

1	Caliente	5,120
1	Frío	8,200
2	Caliente	10,000
2	Frío	8,600
3	Caliente	10,000
3	Frío	9,200
4	Caliente	10,000
4	Frío	6,200
5	Caliente	10,000
5	Frío	10,000
6	Caliente	7,900
6	Frío	5,200
7	Caliente	510
7	Frío	885
8	Caliente	1,020
8	Frío	460

36. Se utilizaron nueve sujetos en un experimento para determinar si una atmósfera que implica la exposición a monóxido de carbono tiene un impacto sobre la capacidad de respiración. Los datos fueron recolectados por el personal del Departamento de Salud y Educación Física del Instituto Politécnico y Universidad Estatal de Virginia. Los datos se analizaron en el Centro de Consulta Estadística del mismo instituto. Los sujetos se colocaron en cámaras de respiración, una de las cuales contenía una alta concentración de CO . Se realizaron varias mediciones de respiración para cada sujeto en cada cámara. Los sujetos se colocaron en las cámaras de respiración en una secuencia aleatoria. Los siguientes datos dan la frecuencia respiratoria en número de respiraciones por minuto.

Sujeto	Con CO	Sin CO
1	30	30
2	45	40
3	26	25
4	25	23
5	34	30
6	51	49
7	46	41
8	32	35
9	30	28

Realice una prueba unilateral de la hipótesis de que la frecuencia respiratoria media es la misma para los dos ambientes. Utilice $\alpha = 0.05$.

10.11 Una muestra: prueba sobre una sola proporción

Las pruebas de hipótesis que se relacionan con proporciones se requieren en muchas áreas. El político se interesa ciertamente en conocer qué fracción de votantes lo favorecerá en la siguiente elección. Todas las empresas fabricantes se preocupan por la proporción de artículos defectuosos cuando se realiza un embarque. El jugador depende de un conocimiento de la proporción de resultados que considera favorables.

Consideraremos el problema de probar la hipótesis de que la proporción de éxitos en un experimento binomial es igual a algún valor específico. Es decir, probaremos la hipótesis nula H_0 que $p = p_0$, donde p es el parámetro de la distribución binomial. La hipótesis alternativa puede ser una de las alternativas unilaterales o bilaterales usuales:

$$p < p_0, p > p_0, \quad \text{o} \quad p \neq p_0.$$

La variable aleatoria apropiada sobre la que basamos nuestro criterio de decisión es la variable aleatoria binomial X , aunque también podríamos usar sólo la estadística $\hat{p} = X/n$. Los valores de X que están lejos de la media $\mu = np_0$ conducirán al rechazo de la hipótesis nula. Como X es una variable binomial discreta, es poco probable que se pueda establecer una región crítica cuyo tamaño sea *exactamente* igual a un valor prescrito de α . Por esta razón es preferible, al tratar con muestras pequeñas, basar nuestra decisión en valores P . Para probar la hipótesis

$$H_0: p = p_0,$$

$$H_1: p < p_0$$

utilizamos la distribución binomial para calcular el valor P

$$P = P(X \leq x \text{ cuando } p = p_0).$$

El valor x es el número de éxitos en nuestra muestra de tamaño n . Si este valor P es menor que o igual a α , nuestra prueba es significativa en el nivel α y rechazamos H_0 a favor de H_1 . De manera similar, para probar la hipótesis

$$H_0: p = p_0,$$

$$H_1: p > p_0,$$

en el nivel de significancia α , calculamos

$$P = P(X \geq x \text{ cuando } p = p_0)$$

y rechazamos H_0 a favor de H_1 si este valor P es menor que o igual a α . Finalmente, para probar la hipótesis

$$H_0: p = p_0,$$

$$H_1: p \neq p_0,$$

al nivel de significancia α , calculamos

$P = 2P(X \leq x \text{ cuando } p = p_0)$ si $x < np_0$ o $P = 2P(X \geq x \text{ cuando } p = p_0)$ si $x > np_0$ y se rechaza H_0 a favor de H_1 si el valor P calculado es menor o igual a α .

Los pasos para probar una hipótesis nula acerca de una proporción contra varias alternativas mediante el uso de las probabilidades binomiales de la tabla A.1 son los siguientes:

Prueba de una proporción: muestras pequeñas

1. H_0 : $p = p_0$.
2. H_1 : Las alternativas son $p < p_0$, $p > p_0$, o $p \neq p_0$.
3. Elegir un nivel de significancia igual a α .
4. Estadística de prueba: variable binomial X con $p = p_0$.
5. Cálculos: Encontrar x , el número de éxitos, y calcular el valor P apropiado.
6. Decisión: Extraer las conclusiones apropiadas basadas en el valor P .

Ejemplo 10.10 Un constructor afirma que se instalan bombas de calor en 70% de todas las casas que se construyen hoy en día en la ciudad de Richmond. ¿Estaría de acuerdo con esta afirmación si una investigación de casas nuevas en esta ciudad muestra que 8 de 15 tienen instaladas bombas de calor? Utilice un nivel de significancia de 0.10.

SOLUCIÓN

1. H_0 : $p = 0.7$.
2. H_1 : $p \neq 0.7$.
3. $\alpha = 0.10$.
4. Estadística de prueba: Variable binomial X con $p = 0.7$ y $n = 15$.
5. Cálculos: $x = 8$ y $np_0 = (15)(0.7) = 10.5$. Por tanto, de la tabla A.1, el valor P calculado es

$$P = 2P(X \leq 8 \text{ cuando } p = 0.7) = 2 \sum_{r=0}^8 b(x; 15, 0.7) \\ = 0.2622 > 0.10.$$

6. Decisión: No rechazar H_0 . Concluir que no hay razón suficiente para dudar de la afirmación del constructor.

En la sección 5.3, vimos que las probabilidades binomiales se obtienen de la fórmula binomial real o de la tabla A.1 cuando n es pequeña. Para n grande, se requieren procedimientos de aproximación. Cuando el valor hipotético p_0 es muy cercano a 0 o a 1, se puede utilizar la distribución de Poisson, con parámetro $\mu = np_0$. Sin embargo, la aproximación de la curva normal, con parámetros $\mu = np_0$ y $\sigma^2 = np_0q_0$, por lo general se prefiere para n grande y es muy precisa mientras que p_0 no esté extremadamente cerca de 0 o de 1. Si utilizamos la aproximación normal, el valor z para probar $p = p_0$ está dado por

$$z = \frac{x - np_0}{\sqrt{np_0q_0}}$$

que es un valor de la variable normal estándar Z . De aquí que, para una prueba de dos colas al nivel de significancia α , la región crítica es $z < -z_{\alpha/2}$ y $z > z_{\alpha/2}$. Para la alternativa unilateral $p < p_0$, la región crítica es $z < -z_{\alpha}$, y para la alternativa $p > p_0$, la región crítica es $z > z_{\alpha}$. ■

Ejemplo 10.11 Una medicina que se prescribe comúnmente para aliviar la tensión nerviosa se considera que es efectivo en 60%. Resultados experimentales con una nueva medicina que se administra a una muestra aleatoria de 100 adultos que padecen de tensión nerviosa muestran que 70 tuvieron alivio, ¿Esta es evidencia suficiente para concluir que la nueva medicina es superior a la que se prescribe actualmente? Utilice un nivel de significancia de 0,05.

SOLUCIÓN

1. $H_0: p = 0,6$.
2. $H_1: p > 0,6$.
3. $\alpha = 0,05$.
4. Región crítica: $z > 1,645$.
5. Cálculos: $x = 70, n = 100, np_0 = (100)(0,6) = 60$, y

$$z = \frac{70 - 60}{\sqrt{(100)(0,6)(0,4)}} = 2,04, \quad P = P(Z > 2,04) < 0,025.$$

6. Decisión: Rechazar H_0 y concluir que la nueva medicina es superior. ■

10.12 Dos muestras: pruebas sobre dos proporciones

A menudo existen situaciones donde deseamos probar la hipótesis de que dos proporciones son iguales. Por ejemplo, podemos tratar de mostrar evidencia de que la proporción de doctores que son pediatras en un estado es igual a la proporción de pediatras en otro estado. Una persona puede decidir dejar de fumar sólo si se convence de que la proporción de fumadores con cáncer de pulmón excede la proporción de no fumadores con cáncer de pulmón.

En general, deseamos probar la hipótesis nula de que dos proporciones, o parámetros binomiales, son iguales. Es decir, probaremos que $p_1 = p_2$ contra una de las alternativas $p_1 < p_2$, $p_1 > p_2$, o $p_1 \neq p_2$. Por supuesto, esto es equivalente a probar la hipótesis nula $p_1 - p_2 = 0$ contra una de las alternativas $p_1 - p_2 < 0$, $p_1 - p_2 > 0$, o $p_1 - p_2 \neq 0$. La estadística sobre la que basamos nuestra decisión es la variable aleatoria $\hat{P}_1 - \hat{P}_2$. Se seleccionan al azar muestras independientes de tamaño n_1 y n_2 de dos poblaciones binomiales y se calcula la proporción de éxitos \hat{P}_1 y \hat{P}_2 para las dos muestras.

En nuestra construcción de intervalos de confianza para p_1 y p_2 , señalamos, para n_1 y n_2 suficientemente grandes, que el estimador puntual \hat{P}_1 menos \hat{P}_2 y varianza

$$\mu_{\hat{P}_1 - \hat{P}_2} = p_1 - p_2 \quad \text{y varianza } \sigma_{\hat{P}_1 - \hat{P}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}.$$

Por tanto, nuestra aceptación y regiones críticas se pueden establecer mediante el uso de la variable normal estándar

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{(p_1 q_1 / n_1) + (p_2 q_2 / n_2)}}.$$

Cuando H_0 es verdadera, podemos sustituir $p_1 = p_2 = p$ y $q_1 = q_2 = q$ (donde p y q son los valores comunes) en la fórmula anterior para Z y obtener la forma

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq[(1/n_1) + (1/n_2)]}}$$

Para calcular un valor de Z , sin embargo, debemos estimar los parámetros p y q que aparecen en el radical. Al combinar los datos de ambas muestras, la **estimación combinada de la proporción p** es

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2},$$

donde x_1 y x_2 son el número de éxitos en cada una de las dos muestras. Al sustituir \hat{p} por p y $\hat{q} = 1 - \hat{p}$ por q , el **valor z para probar $p_1 = p_2$** se determina a partir de la fórmula

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}[(1/n_1) + (1/n_2)]}}$$

Las regiones críticas para las hipótesis alternativas apropiadas se establecen como antes mediante el uso de puntos críticos de la curva normal estándar. De aquí, para la alternativa $p_1 \neq p_2$ con un nivel de significancia α , la región crítica es $z < -z_{\alpha/2}$ y $z > z_{\alpha/2}$. Para una prueba donde la alternativa es $p_1 < p_2$, la región crítica es $z < -z_\alpha$, y cuando la alternativa es $p_1 > p_2$, la región crítica es $z > z_\alpha$.

Ejemplo 10.12 Se tomará el voto entre los residentes de una ciudad y el condado circundante para determinar si se debe construir una planta química propuesta. El lugar de construcción está dentro de los límites de la ciudad y por esta razón muchos votantes del condado consideran que la propuesta pasará debido a la gran proporción de votantes que favorecen la construcción. Para determinar si hay una diferencia significativa en la proporción de votantes de la ciudad y votantes del condado que favorecen la propuesta, se realiza una encuesta. Si 120 de 200 votantes de la ciudad favorecen la propuesta y 240 de 500 residentes del condado también lo hacen, ¿estaría de acuerdo en que la proporción de votantes de la ciudad que favorecen la propuesta es más alto que la proporción de votantes del condado? Utilice un nivel de significancia de 0.025.

SOLUCIÓN

Sean p_1 y p_2 las proporciones reales de votantes en la ciudad y el condado, respectivamente, que favorecen la propuesta.

1. $H_0: p_1 = p_2$.
2. $H_1: p_1 > p_2$.
3. $\alpha = 0.025$.
4. Región crítica: $z > 1.96$.
5. Cálculos:

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{120}{200} = 0.60$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{240}{500} = 0.48$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{120 + 240}{200 + 500} = 0.51.$$

Por tanto,

$$z = \frac{0.60 - 0.48}{\sqrt{(0.51)(0.49)\left[\frac{1}{200} + \frac{1}{500}\right]}} = 2.9.$$

$$P = P(Z > 2.9) = 0.0019.$$

6. Decisión: Rechazar H_0 y estar de acuerdo en que la proporción de votantes de la ciudad a favor de la propuesta es más alta que la proporción de votantes del condado. ■

Ejercicios

1. Un experto en mercadotecnia de una compañía fabricante de pasta considera que 40% de los amantes de la pasta prefieren la lasagna. Si nueve de 20 amantes de la pasta eligen lasagna sobre otras pastas, ¿qué se puede concluir acerca de la afirmación del experto? Utilice un nivel de significancia de 0.05.

2. Suponga que, en el pasado, 40% de todos los adultos favorecían la pena capital. ¿Tenemos razón para creer que la proporción de adultos que hoy en día favorecen la pena capital ha aumentado si, en una muestra aleatoria de 15 adultos, ocho están a favor de la pena capital? Utilice un nivel de significancia de 0.05.

3. Se lanza 20 veces una moneda, con un resultado de cinco caras. ¿Esta es suficiente evidencia para rechazar la hipótesis de que la moneda está balanceada a favor de la alternativa de que las caras ocurren menos de 50% de las veces? Cite un valor P .

4. Se cree que al menos 60% de los residentes de cierta área favorecen una demanda de anexión de una ciudad vecina. ¿Qué conclusión extraería si sólo 110 en una muestra de 200 votantes están a favor de la demanda? Utilice un nivel de significancia de 0.05.

5. Una compañía petrolera afirma que un quinto de las casas en cierta ciudad se calientan con petróleo. ¿Tenemos razón en dudar de esta afirmación si, en una muestra aleatoria de 1000 casas en esta ciudad, se encuentra que 136 se calientan con petróleo? Utilice un nivel de significancia de 0.01.

6. En cierta universidad se estima que a lo más 25% de los estudiantes van en bicicleta a la escuela. ¿Esta parece ser una estimación válida si, en una muestra aleatoria de 90 estudiantes universitarios, se encuen-

tra que 28 van en bicicleta a la escuela? Utilice un nivel de significancia de 0.05.

7. Se considera un nuevo dispositivo de radar para cierto sistema de misiles de defensa. El sistema se verifica mediante la experimentación con aeronaves reales en las que se simula una situación de *muerde* o de *no muerde*. Si en 300 pruebas ocurren 259 muertes, acepte o rechace, en un nivel de significancia de 0.04, la afirmación de que la probabilidad de una muerte con el sistema nuevo no excede la probabilidad de 0.8 del sistema existente.

8. En un experimento de laboratorio controlado, científicos de la Universidad de Minnesota descubren que 25% de cierta camada de ratas sujetas a una dieta de 20% de grano de café y luego forzadas a consumir un poderoso químico causante de cáncer desarrollaron tumores cancerosos. ¿Tendríamos razones para creer que la proporción de ratas que desarrolla tumores cuando se sujeta a esta dieta aumenta si el experimento se repite y 16 de 48 ratas desarrollan tumores? Utilice un nivel de significancia de 0.05.

9. En un estudio para estimar la proporción de residentes de cierta ciudad y sus suburbios que están a favor de la construcción de una planta de energía nuclear, se encuentra que 63 de 100 residentes urbanos están a favor de la construcción mientras que sólo 59 de 125 residentes suburbanos la favorecen. ¿Hay una diferencia significativa entre la proporción de residentes urbanos y suburbanos que favorecen la construcción de la planta nuclear? Use un valor P .

10. En un estudio sobre la fertilidad de mujeres casadas conducido por Martin O'Connell y Carolyn C. Rogers para la Oficina de Censos en 1979, se seleccio-

naron al azar dos grupos de esposas con edades de 25 a 29 años sin hijos y a cada mujer se le preguntó si planeaba tener un hijo. Se seleccionó un grupo entre las mujeres con menos de dos años de casadas y otro entre las que tenían cinco años de casadas. Suponga que 240 de 300 con menos de dos años de casadas planean tener algún día un hijo comparadas con 288 de las 400 con cinco años de casadas. ¿Podemos concluir que la proporción de mujeres con menos de dos años de casadas que planean tener hijos es significativamente más alta que la proporción con cinco años de casadas? Use un valor P .

11. Una comunidad urbana quiere mostrar que la incidencia de cáncer de seno es más alta que en una área rural vecina. (Se encontró que los niveles de PCB son más altos en el suelo de la comunidad urba-

na.) Si se encuentra que 20 de 200 mujeres adultas en la comunidad urbana tienen cáncer de pecho y 10 de 150 mujeres adultas en la comunidad rural tienen cáncer de seno, ¿podemos concluir con un nivel de significancia de 0.06 que este tipo de cáncer prevalece más en la comunidad urbana?

12. En un invierno con epidemia de gripe, una compañía farmacéutica bien conocida estudió 2000 bebés para determinar si la nueva medicina de la compañía era efectiva después de dos días. Entre 120 bebés que tenían gripe y se les administró la medicina, 29 se curaron dentro de dos días. Entre 280 bebés que tenían gripe pero que no recibieron la medicina, 56 se curaron dentro de dos días. ¿Hay alguna indicación significativa que apoye la afirmación de la compañía de la efectividad de la medicina?

10.13 Pruebas referentes a varianzas de una y dos muestras

En esta sección nos referimos a la prueba de hipótesis relacionadas con varianzas o desviaciones estándar poblacionales. Las pruebas de una y dos muestras sobre varianzas en realidad no son difíciles de motivar. Los ingenieros y los científicos constantemente se enfrentan a estudios en los que se les requiere demostrar que las mediciones involucradas con productos o procesos caen dentro de las especificaciones que fijan los consumidores. Las especificaciones a menudo se cumplen si la varianza del proceso es suficientemente pequeña. La atención también se concentra en experimentos comparativos entre métodos o procesos donde la reproducibilidad o variabilidad inherentes se deben comparar de manera formal. Además, a menudo se aplica una prueba que compara dos varianzas antes de llevar a cabo una prueba t sobre dos medias. El objetivo es determinar si se viola la suposición de varianzas iguales.

Consideremos primero el problema de probar la hipótesis nula H_0 de que la varianza poblacional σ^2 es igual a un valor específico σ_0^2 contra una de las alternativas comunes $\sigma^2 < \sigma_0^2$, $\sigma^2 > \sigma_0^2$, o $\sigma^2 \neq \sigma_0^2$. La estadística apropiada sobre la que basamos nuestra decisión es la misma estadística χ^2 cuadrada del teorema 8.4 que se utiliza en el capítulo 9 para construir un intervalo de confianza para σ^2 . Por tanto, si suponemos que la distribución de la población que se muestrea es normal, el valor de χ^2 cuadrada para probar $\sigma^2 = \sigma_0^2$ está dado por

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2},$$

donde n es el tamaño de la muestra, s^2 es la varianza muestral y σ_0^2 es el valor de σ^2 dado por la hipótesis nula. Si H_0 es verdadera, χ^2 es un valor de la distribución χ^2 cuadrada con $v = n - 1$ grados de libertad. De aquí que, para una prueba de dos colas en el nivel de significancia α , la región crítica es $\chi^2 < \chi_{1-\alpha/2}^2$ y $\chi^2 > \chi_{\alpha/2}^2$. Para la alternativa unilateral $\sigma^2 < \sigma_0^2$, la región crítica es $\chi^2 < \chi_{1-\alpha}^2$, y para la alternativa unilateral $\sigma^2 > \sigma_0^2$, la región crítica es $\chi^2 > \chi_{\alpha}^2$.

Robustez de la normalidad

El lector puede percibir que varias pruebas dependen, al menos en teoría, de la suposición de normalidad. En general, muchos procedimientos en estadística aplicada tienen fundamentos teóricos que dependen de la distribución normal. Estos procedimientos varían en el grado de su dependencia de la suposición de la normalidad. Un procedimiento que es razonablemente insensible a la suposición se denomina **procedimiento robusto**; es decir, robusto para la normalidad. La prueba χ^2 sobre una sola varianza es nada robusta hacia la normalidad; es decir, el éxito práctico del procedimiento depende de la normalidad. Como resultado el valor P calculado puede ser apreciablemente diferente del valor P real si la población muestreada no es normal. En realidad, es bastante factible que un valor P estadísticamente significativo puede no ser una verdadera señal de $H_1: \sigma \neq \sigma_0$ sino, más bien, que un valor significativo puede ser un resultado de la violación de las suposiciones de normalidad. Por tanto, el analista se debe aproximar al uso de esta prueba χ^2 particular con precaución.

Ejemplo 10.13 Un fabricante de baterías para auto afirma que la duración de sus baterías se distribuye de forma aproximadamente normal con una desviación estándar igual a 0.9 años. Si una muestra aleatoria de 10 de tales baterías tiene una desviación estándar de 1.2 años, ¿considera que $\sigma > 0.9$ años? Utilice un nivel de significancia de 0.05.

SOLUCIÓN

1. $H_0: \sigma^2 = 0.81$.
2. $H_1: \sigma^2 > 0.81$.
3. $\alpha = 0.05$.
4. Región crítica: de la figura 10.19 vemos que la hipótesis nula se rechaza cuando $\chi^2 > 16.919$, donde $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$ con $v = 9$ grados de libertad.

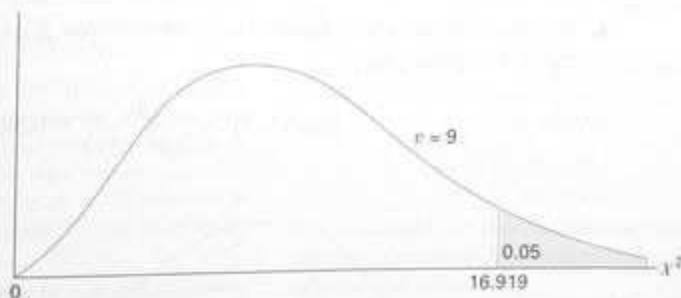


Figura 10.19 Región crítica para la hipótesis alternativa $\sigma^2 > 81$.

5. Cálculos: $s^2 = 1.44$, $n = 10$, y

$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16.0, \quad P \approx 0.07.$$

6. Decisión: la estadística χ^2 no es significativa en el nivel 0.05. Sin embargo, hay alguna evidencia de que $\sigma > 0.9$.

Consideremos ahora el problema de probar la igualdad de las varianzas σ_1^2 y σ_2^2 de dos poblaciones. Es decir, probaremos la hipótesis nula H_0 de que $\sigma_1^2 = \sigma_2^2$ contra una de las alternativas usuales

$$\sigma_1^2 < \sigma_2^2, \sigma_1^2 > \sigma_2^2, \text{ o } \sigma_1^2 \neq \sigma_2^2.$$

Para muestras aleatorias independientes de tamaño n_1 y n_2 , respectivamente, de las dos poblaciones, el valor f para probar $\sigma_1^2 = \sigma_2^2$ es la razón

$$f = \frac{s_1^2}{s_2^2},$$

donde s_1^2 y s_2^2 son las varianzas calculadas de las dos muestras. Si las dos poblaciones se distribuyen de forma aproximadamente normal y la hipótesis nula es verdadera, de acuerdo con el teorema 8.8 la razón $f = s_1^2/s_2^2$ es un valor de la distribución F con $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$ grados de libertad. Por tanto, las regiones críticas de tamaño α que corresponden a las alternativas unilaterales $\sigma_1^2 < \sigma_2^2$ y $\sigma_1^2 > \sigma_2^2$ son, respectivamente, $f < f_{1-\alpha}(v_1, v_2)$ y $f > f_{\alpha}(v_1, v_2)$. Para la alternativa bilateral $\sigma_1^2 \neq \sigma_2^2$, la región crítica es $f < f_{1-\alpha/2}(v_1, v_2)$ y $f > f_{\alpha/2}(v_1, v_2)$. ■

Ejemplo 10.14 Al probar la diferencia en el desgaste abrasivo de los dos materiales del ejemplo 10.6, supusimos que las dos varianzas poblacionales desconocidas eran iguales. ¿Se justifica esta suposición? Utilice un nivel de significancia de 0.10.

SOLUCIÓN

Sean σ_1^2 y σ_2^2 las varianzas poblacionales para el desgaste abrasivo del material 1 y del material 2, respectivamente.

1. H_0 : $\sigma_1^2 = \sigma_2^2$.
2. H_1 : $\sigma_1^2 \neq \sigma_2^2$.
3. $\alpha = 0.10$.
4. Región crítica: de la figura 10.20, vemos que $f_{0.05}(11, 9) = 3.11$, mediante el uso del teorema 8.7,

$$f_{0.95}(11, 9) = \frac{1}{f_{0.05}(9, 11)} = 0.34.$$

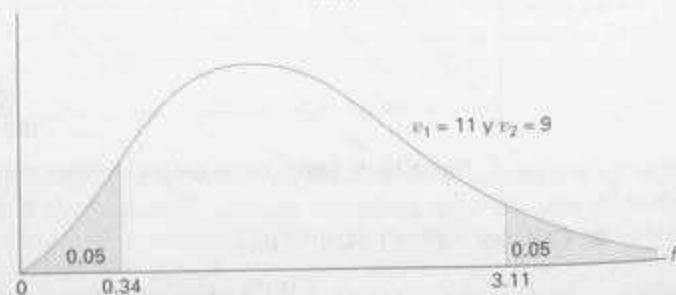


Figura 10.20 Región crítica para la alternativa $\sigma_1^2 \neq \sigma_2^2$.

Por tanto, la hipótesis nula se rechaza cuando $f < 0.34$ o $f > 3.11$, donde $f = s_1^2/s_2^2$ con $v_1 = 11$ y $v_2 = 9$ grados de libertad.

5. Cálculos: $s_1^2 = 16$, $s_2^2 = 25$, y por ello $f = \frac{16}{25} = 0.64$.
6. Decisión: no rechazar H_0 . Concluir que no hay suficiente evidencia de que las varianzas difieran. ■

Prueba F para probar varianzas en el SAS

La figura 10.18 muestra una prueba t de dos muestras donde se comparan dos medias, como ejercicio, con los datos de los retoños. La gráfica de caja y extensión en la figura 10.17 de la página 324 sugiere que las varianzas no son homogéneas, y por ello la estadística t' y su valor P correspondiente son relevantes. Nótese también que el impreso muestra la estadística F para $H_0: \sigma_1 = \sigma_2$ con un valor P de 0.0098, evidencia adicional de que se debe esperar más variabilidad cuando se usa nitrógeno comparado con la condición sin nitrógeno.

Ejercicios

1. Se sabe que el volumen de los envases de un lubricante particular se distribuye normalmente con una varianza de 0.03 litros. Pruebe la hipótesis de que $\sigma^2 = 0.03$ contra la alternativa de que $\sigma^2 \neq 0.03$ para la muestra aleatoria de 10 envases del ejercicio 7 de la página 326. Use un nivel de significancia de 0.01.

2. Experiencia anterior indica que el tiempo que se requiere para que los estudiantes de último año de preparatoria completen una prueba estandarizada es una variable aleatoria normal con una desviación estándar de seis minutos. Pruebe la hipótesis de que $\sigma = 6$ contra la alternativa de que $\sigma < 6$ si una muestra aleatoria de 20 estudiantes del último año de preparatoria tiene una desviación estándar $s = 4.51$. Utilice un nivel de significancia de 0.05.

3. Se deben supervisar las aflotoxinas producidas por moho en cosechas de cacahuate en Virginia. Una muestra de 64 lotes de cacahuate revela niveles de 24.17 ppm, en promedio, con una varianza de 4.25 ppm. Pruebe la hipótesis de que $\sigma^2 = 4.2$ ppm con la alternativa de que $\sigma^2 \neq 4.2$ ppm. Utilice un nivel de significancia de 0.05.

4. Datos históricos indican que la cantidad de dinero que contribuyeron los residentes trabajadores de una ciudad grande a un escuadrón de rescate voluntario es una variable aleatoria normal con una desviación estándar de \$1.40. Se sugiere que las contribuciones al escuadrón de rescate sólo de los empleados del departamento de sanidad son mucho más variables. Si las contribuciones de una muestra aleatoria de 12 emplea-

dos del departamento de sanidad tienen una desviación estándar de \$1.75, ¿podemos concluir con un nivel de significancia de 0.01 que la desviación estándar de las contribuciones de todos los trabajadores de sanidad es mayor que la de todos los trabajadores que viven en esta ciudad?

5. Se dice que una máquina despachadora de refrescos está fuera de control si la varianza de los contenidos excede 1.15 decilitros. Si una muestra aleatoria de 25 bebidas de esta máquina tiene una varianza de 2.03 decilitros, ¿esto indica con un nivel de significancia de 0.05 que la máquina está fuera de control? Suponga que los contenidos se distribuyen de forma aproximadamente normal.

6. Prueba de $\sigma^2 = \sigma_0^2$ para una muestra grande: Cuando $n \geq 30$ podemos probar la hipótesis nula de que $\sigma^2 = \sigma_0^2$, o que, al calcular

$$z = \frac{s^2 - \sigma_0^2}{\sigma_0^2 / \sqrt{2n}}$$

que es un valor de una variable aleatoria cuya distribución de muestreo es aproximadamente la distribución normal estándar.

(a) Con referencia al ejemplo 10.3, pruebe con un nivel de significancia de 0.05 si $\sigma = 7.5$ años contra la alternativa de que $\sigma \neq 7.5$ años.

(b) Se sospecha que la varianza de la distribución de distancias en kilómetros logrados con cinco litros de combustible por un modelo nuevo de automóvil equipado con un motor diesel es menor que la varian-

za de la distribución de distancias lograda por el mismo modelo equipado con un motor a gasolina de seis cilindros, que se sabe es $\sigma^2 = 6.25$. Si 72 recorridos de prueba en el modelo diesel tienen una varianza de 4.41, ¿podemos concluir con un nivel de significancia de 0.02 que la varianza de las distancias alcanzadas por el modelo diesel es menor que la del modelo de gasolina?

7. Se lleva a cabo un estudio para comparar la longitud de tiempo entre hombres y mujeres para ensamblar cierto producto. Experiencia pasada indica que la distribución de los tiempos para hombres y mujeres es aproximadamente normal pero que la varianza de los tiempos para las mujeres es menor que para los hombres. Una muestra aleatoria de tiempos para 11 hombres y 14 mujeres produce los siguientes datos:

Hombres	Mujeres
$n_1 = 11$	$n_2 = 14$
$s_1 = 6.1$	$s_2 = 5.3$

Pruebe la hipótesis de que $\sigma_1^2 = \sigma_2^2$ contra la alternativa de que $\sigma_1^2 > \sigma_2^2$. Utilice un nivel de significancia de 0.01.

8. En el ejercicio 23 de la página 328, pruebe la hipótesis al nivel de significancia de 0.05 de que $\sigma_1^2 = \sigma_2^2$ contra la alternativa de que $\sigma_1^2 \neq \sigma_2^2$, donde σ_1^2 y σ_2^2 son las varianzas para el número de organismos por metro cuadrado en los dos diferentes lugares de Cedar Run.

9. Con referencia al ejercicio 18 de la página 327, pruebe la hipótesis de que $\sigma_1 = \sigma_2$ contra la alternativa de que $\sigma_1 < \sigma_2$, donde σ_1 y σ_2 son las desviaciones estándar de las distancias que se obtienen por las llantas marca A y marca B, respectivamente. Utilice un nivel de significancia de 0.05.

10. Con referencia al ejercicio 21 de la página 328, pruebe la hipótesis de que $\sigma_1^2 = \sigma_2^2$ contra la alternativa de que $\sigma_1^2 \neq \sigma_2^2$, donde σ_1^2 y σ_2^2 son las varianzas para los tiempos de duración de películas producidas por la compañía 1 y la compañía 2, respectivamente. Utilice un nivel de significancia de 0.10.

11. Se comparan dos tipos de instrumentos para medir la cantidad de monóxido de azufre en la atmósfera en un experimento de contaminación del aire. Se desea determinar si los dos tipos de instrumentos dan mediciones que tengan la misma variabilidad. Se registran las siguientes lecturas para los dos instrumentos:

Monóxido de azufre

Instrumento A	Instrumento B
0.86	0.87
0.82	0.74
0.75	0.63
0.61	0.55
0.89	0.76
0.64	0.70
0.81	0.69
0.68	0.57
0.65	0.53

Suponga que las poblaciones de mediciones se distribuyen de forma aproximadamente normal, pruebe la hipótesis de que $\sigma_A = \sigma_B$ contra la alternativa de que $\sigma_A \neq \sigma_B$. Use un valor P .

12. Se lleva a cabo un experimento para comparar el contenido de alcohol en una salsa de soya en dos líneas de producción diferentes. La producción se supervisa ocho veces al día. Los datos son los que aquí se muestran.

Línea de producción 1:

0.48 0.39 0.42 0.52 0.40 0.48 0.52 0.52

Línea de producción 2:

0.38 0.37 0.39 0.41 0.38 0.39 0.40 0.39

Suponga que ambas poblaciones son normales. Se sospecha que la línea de producción uno no produce con la consistencia de la línea dos en términos de contenido de alcohol. Pruebe la hipótesis de que $\sigma_1 = \sigma_2$ contra la alternativa $\sigma_1 \neq \sigma_2$. Utilice un valor P .

13. Se sabe que las emisiones de hidrocarburos disminuyeron de forma dramática durante la década de 1980. Se lleva a cabo un estudio para comparar las emisiones de hidrocarburos a velocidad estacionaria en partes por millón (ppm), para automóviles de 1980 y 1990. Se seleccionaron al azar 20 autos de cada modelo y se registraron sus niveles de emisión de hidrocarburos. Los datos son:

modelos 1980:

141 359 247 940 882 494 306 210 105 880
200 223 188 940 241 190 300 435 241 380

modelos 1990:

140 160 20 20 223 60 20 95 360 70
220 400 217 58 235 380 200 175 85 65

Pruebe la hipótesis de que $\sigma_1 = \sigma_2$ contra la alternativa de que $\sigma_1 \neq \sigma_2$. Suponga que ambas poblaciones son normales.

10.14 Prueba de la bondad de ajuste

A lo largo de este capítulo nos ocupamos de la prueba de hipótesis estadísticas acerca de parámetros de una sola población como μ , σ^2 , y p . Ahora consideraremos una prueba para determinar si una población tiene una distribución teórica específica. La prueba se basa en qué tan buen ajuste tenemos entre la frecuencia de ocurrencia de las observaciones en una muestra observada y las frecuencias esperadas que se obtienen a partir de la distribución hipotética.

Para ilustrar, considere el lanzamiento de un dado. Elaboramos la hipótesis de que el dado es legal, lo que equivale a probar la hipótesis de que la distribución de resultados es la distribución uniforme discreta

$$f(x) = \frac{1}{6}, \quad x = 1, 2, \dots, 6.$$

Suponga que el dado se lanza 120 veces y que se registra cada resultado. Teóricamente, si el dado está balanceado, esperaríamos que cada cara ocurriera 20 veces. Los resultados se dan en la tabla 10.3. Al comparar las frecuencias observadas con las frecuencias esperadas correspondientes, debemos decidir si es posible que estas discrepancias ocurran como resultado de fluctuaciones del muestreo y de que el dado está balanceado o que éste no es legal y de que la distribución de resultados no es uniforme. Es práctica común referirse a cada resultado posible de un experimento como una celda. De aquí, en nuestra ilustración, tenemos seis celdas. La estadística apropiada en la que basamos nuestro criterio de decisión para un experimento que incluye k celdas se define mediante el siguiente teorema.

Teorema 10.1

Una **prueba de la bondad de ajuste** entre las frecuencias observadas y esperadas se basa en la cantidad

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i},$$

donde χ^2 es un valor de una variable aleatoria cuya distribución muestral se aproxima muy de cerca con la distribución ji cuadrada con $v = k - 1$ grados de libertad. Los símbolos o_i y e_i representan las frecuencias observada y esperada, respectivamente, para la i -ésima celda.

El número de grados de libertad que se asocia con la distribución ji cuadrada que aquí se utiliza es igual a $k - 1$, pues hay sólo $k - 1$ frecuencias de celdas libremente determinadas. Es decir, una vez que se determinan las frecuencias de $k - 1$ celdas queda determinada la frecuencia para la k -ésima celda.

Si las frecuencias observadas están cerca de las frecuencias esperadas correspondientes, el valor χ^2 será pequeño, lo que indica un buen ajuste. Si las frecuencias observadas difieren de manera considerable de las frecuencias esperadas, el valor χ^2 será grande y el ajuste es pobre. Un buen ajuste conduce a la aceptación de H_0 , mientras que un ajuste pobre conduce a su rechazo. La región crítica caerá, por tanto, en la cola derecha de la distribución ji cuadrada. Para un nivel de significancia igual a α , encontramos el valor crítico χ_α^2 de la tabla A.5, y entonces $\chi^2 \geq \chi_\alpha^2$.

Tabla 10.3 Frecuencias observadas y esperadas de 120 lanzamientos de un dado

Cara:	1	2	3	4	5	6
Observada	20	22	17	18	19	24
Esperada	20	20	20	20	20	20

constituye la región crítica. El criterio de decisión que aquí se describe no se debe usar a menos que cada una de las frecuencias esperadas sea al menos igual a 5. Esta restricción puede requerir la combinación de celdas adyacentes lo que tiene como resultado una reducción en el número de grados de libertad.

De la tabla 10.3, encontramos que el valor χ^2 es:

$$\begin{aligned}\chi^2 &= \frac{(20 - 20)^2}{20} + \frac{(22 - 20)^2}{20} + \frac{(17 - 20)^2}{20} \\ &\quad + \frac{(18 - 20)^2}{20} + \frac{(19 - 20)^2}{20} + \frac{(24 - 20)^2}{20} = 1.7.\end{aligned}$$

Con el uso de la tabla A.5, encontramos $\chi_{0.05}^2 = 11.070$ para $v = 5$ grados de libertad. Como 1.7 es menor que el valor crítico, no se rechaza H_0 . Concluimos que no hay suficiente evidencia de que el dado no está balanceado.

Como segunda ilustración probemos la hipótesis de que la distribución de frecuencia de las duraciones de baterías dadas en la tabla 3.4 se puede aproximar mediante una distribución normal con media $\mu = 3.5$ y desviación estándar $\sigma = 0.7$. Las frecuencias esperadas para las siete clases (celdas), que se listan en la tabla 10.4, se obtienen al calcular las áreas bajo la curva normal hipotética que caen entre los diversos límites de clase.

Por ejemplo, los valores z que corresponden a los límites de la cuarta clase son

$$z_1 = \frac{2.95 - 3.5}{0.7} = -0.79 \quad \text{y} \quad z_2 = \frac{3.45 - 3.5}{0.7} = -0.07.$$

De la tabla A.3 encontramos que el área entre $z_1 = -0.79$ y $z_2 = -0.07$ es

$$\begin{aligned}\text{área} &= P(-0.79 < Z < -0.07) = P(Z < -0.07) - P(Z < -0.79) \\ &= 0.4721 - 0.2148 = 0.2573.\end{aligned}$$

De aquí, la frecuencia esperada para la cuarta clase es

$$e_4 = (0.2573)(40) = 10.3.$$

Se acostumbra redondear estas frecuencias a un decimal.

La frecuencia esperada para el primer intervalo de clase se obtiene al utilizar el área total bajo la curva normal a la izquierda del límite 1.95. Para el último intervalo de clase, usamos el área total a la derecha del límite 4.45. Todas las demás frecuencias esperadas se determinan mediante el método que se describe para la cuarta clase. Nótese que combinamos clases adyacentes en la tabla 10.4, donde las frecuencias esperadas son menores que cinco. En consecuencia, el número total de intervalos se reduce de siete a cuatro, lo que tiene como resultado $v = 3$ grados de libertad. El valor χ^2 está dado entonces por

$$\chi^2 = \frac{(7 - 8.5)^2}{8.5} + \frac{(15 - 10.3)^2}{10.3} + \frac{(10 - 10.7)^2}{10.7} + \frac{(8 - 10.5)^2}{10.5} = 3.05.$$

Tabla 10.4 Frecuencias observadas y esperadas de las duraciones de las baterías suponiendo normalidad

Límites de clase	o_i	e_i
1.45-1.95	2	0.5
1.95-2.45	1	
2.45-2.95	4	
2.95-3.45	15	10.3
3.45-3.95	10	10.7
3.95-4.45	5	7.0
4.45-4.95	3	

Como el valor χ^2 calculado es menor que $\chi_{0.05}^2 = 7.815$ para 3 grados de libertad, no tenemos razón para rechazar la hipótesis nula y concluimos que la distribución normal con $\mu = 3.5$ y $\sigma = 0.7$ proporciona un buen ajuste para la distribución de duraciones de las baterías.

La prueba de bondad de ajuste ji cuadrada es un recurso importante, en particular dado que muchos procedimientos estadísticos en la práctica dependen, en un sentido teórico, de la suposición que los datos reunidos provienen de un tipo de distribución específico. Como ya vimos, la *suposición de normalidad* se hace con bastante frecuencia. En los capítulos que siguen continuaremos haciendo suposiciones de normalidad a fin de proporcionar una base teórica para ciertas pruebas e intervalos de confianza.

En la literatura hay pruebas que son más poderosas que la prueba ji cuadrada para probar la normalidad. Una de tales pruebas es la **prueba de Geary**. Esta prueba se basa en una estadística muy simple que es una razón de dos estimadores de la desviación estándar poblacional σ . Suponga que se toma una muestra aleatoria X_1, X_2, \dots, X_n de una distribución normal, $N(\mu, \sigma)$. Considere la razón

$$u = \frac{\sqrt{\pi/2} \sum_{i=1}^n |X_i - \bar{X}|/n}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2/n}}$$

El lector debe reconocer que el denominador es un estimador razonable de σ si la distribución es normal o no. El numerador es un buen estimador de σ si la distribución es normal pero puede sobrestimar o subestimar σ cuando hay desviaciones de la normalidad. Así los valores de U que difieren considerablemente de 1.0 representan la señal de que la hipótesis de normalidad se debe rechazar.

Para muestras grandes una prueba razonable se basa en la normalidad aproximada de U . La estadística de prueba es entonces una estandarización de U , dada por

$$Z = \frac{U - 1}{0.2661/\sqrt{n}}$$

Por supuesto, el procedimiento de prueba incluye la región crítica bilateral. Calculamos un valor de z a partir de los datos y no rechazamos la hipótesis de normalidad cuando

$$-z_{\alpha/2} < z < z_{\alpha/2}$$

En la bibliografía se cita un artículo que trata de la prueba de Geary.

10.15 Prueba de independencia (datos categóricos)

El procedimiento de prueba de ji cuadrada que se presenta en la sección 10.14 también se puede usar para probar la hipótesis de independencia de dos variables de clasificación. Suponga que deseamos determinar si las opiniones de los votantes residentes del estado de Illinois con respecto a una nueva reforma de impuestos son independientes de sus niveles de ingreso. Una muestra aleatoria de 1000 votantes registrados del estado de Illinois se clasifican de acuerdo con su posición en las categorías de ingreso bajo, medio o alto y si están a favor o no de la nueva reforma de impuestos. Las frecuencias observadas se presentan en la tabla 10.5, que se conoce como **tabla de contingencia**.

Tabla 10.5 Tabla de contingencia 2×3

Reforma de impuestos	Nivel de ingreso			Total
	Bajo	Medio	Alto	
A favor	182	213	203	598
En contra	154	138	110	402
Total	336	351	313	1000

Una tabla de contingencia con r renglones y c columnas se denomina tabla $r \times c$ (" $r \times c$ " se lee " r por c "). Los totales de renglones y columnas en la tabla 10.5 se denominan **frecuencias marginales**. Nuestra decisión de aceptar o rechazar la hipótesis nula, H_0 , de independencia entre la opinión de un votante con respecto a la nueva reforma de impuestos y su nivel de ingreso se basa en qué tan buen ajuste tenemos entre las frecuencias observadas en cada una de las seis celdas de la tabla 10.5 y las frecuencias que esperaríamos para cada celda bajo la suposición de que H_0 es verdadera. Para encontrar estas frecuencias esperadas, definamos los siguientes eventos:

- L : Una persona seleccionada está en el nivel de ingresos bajo.
- M : Una persona seleccionada está en el nivel de ingresos medio.
- H : Una persona seleccionada está en el nivel de ingresos alto.
- F : Una persona seleccionada está a favor de la nueva reforma de impuestos.
- A : Una persona seleccionada está en contra de la nueva reforma de impuestos.

Con el uso de las frecuencias marginales, podemos listar las siguientes estimaciones de probabilidad:

$$P(L) = \frac{336}{1000}, \quad P(M) = \frac{351}{1000}, \quad P(H) = \frac{313}{1000},$$

$$P(F) = \frac{598}{1000}, \quad P(A) = \frac{402}{1000}.$$

Ahora bien, si H_0 es verdadera y las dos variables son independientes, debemos tener

$$P(L \cap F) = P(L)P(F) = \left(\frac{336}{1000}\right)\left(\frac{598}{1000}\right),$$

$$P(L \cap A) = P(L)P(A) = \left(\frac{336}{1000}\right)\left(\frac{402}{1000}\right),$$

$$P(M \cap F) = P(M)P(F) = \left(\frac{351}{1000}\right)\left(\frac{598}{1000}\right).$$

$$P(M \cap A) = P(M)P(A) = \left(\frac{351}{1000}\right)\left(\frac{402}{1000}\right).$$

$$P(H \cap F) = P(H)P(F) = \left(\frac{313}{1000}\right)\left(\frac{598}{1000}\right).$$

$$P(H \cap A) = P(H)P(A) = \left(\frac{313}{1000}\right)\left(\frac{402}{1000}\right).$$

Las frecuencias esperadas se obtienen al multiplicar cada probabilidad de una celda por el número total de observaciones. Como antes, redondeamos estas frecuencias a un decimal. Así, se estima que el número esperado de votantes de bajo ingreso en nuestra muestra que favorecen la nueva reforma fiscal es

$$\left(\frac{336}{1000}\right)\left(\frac{598}{1000}\right)(1000) = \frac{(336)(598)}{1000} = 200.9$$

cuando H_0 es verdadera. La regla general para obtener la frecuencia esperada de cualquier celda está dada por la siguiente fórmula:

$$\text{frecuencia esperada} = \frac{(\text{total de la columna}) \times (\text{total del renglón})}{\text{gran total}}$$

En la tabla 10.6 la frecuencia esperada para cada celda se registra entre paréntesis a un lado del valor observado real. Nótese que las frecuencias esperadas en

Tabla 10.6 Frecuencias observadas y esperadas

Reforma fiscal	Nivel de ingreso			Total
	Bajo	Medio	Alto	
A favor	182 (200.9)	213 (209.9)	203 (187.2)	598
En contra	154 (135.1)	138 (141.1)	110 (125.8)	402
Total	336	351	313	1000

cualquier renglón o columna se suman al total marginal apropiado. En nuestro ejemplo necesitamos calcular sólo las dos frecuencias esperadas en el renglón superior de la tabla 10.6 y después encontrar las otras por sustracción. El número de grados de libertad asociados con la prueba ji cuadrada que aquí se usa es igual al número de frecuencias de celdas que se pueden llenar libremente cuando se nos dan los totales marginales y el gran total, y en este ejemplo este número es 2. Una fórmula simple que proporciona el número correcto de grados de libertad es

$$v = (r - 1)(c - 1).$$

De aquí, para nuestro ejemplo, $v = (2 - 1)(3 - 1) = 2$ grados de libertad.

Para probar la hipótesis nula de independencia, usamos el criterio de decisión siguiente:

Prueba de independencia

Calcular

$$\chi^2 = \sum_r \frac{(o_r - e_r)^2}{e_r}$$

donde la suma se extiende a todas las celdas rc en la tabla de contingencia $r \times c$. Si $\chi^2 > \chi_{\alpha}^2$ con $v = (r - 1)(c - 1)$ grados de libertad, rechazar la hipótesis nula de independencia al nivel de significancia α ; en cualquier otro caso, aceptar la hipótesis nula.

Al aplicar este criterio a nuestro ejemplo, encontramos que

$$\begin{aligned} \chi^2 &= \frac{(182 - 200.9)^2}{200.9} + \frac{(213 - 209.9)^2}{209.9} + \frac{(203 - 187.2)^2}{187.2} \\ &+ \frac{(154 - 135.1)^2}{135.1} + \frac{(138 - 141.1)^2}{141.1} + \frac{(110 - 125.8)^2}{125.8} \\ &= 7.85, \end{aligned}$$

$$P = 0.02.$$

De la tabla A.5 encontramos que $\chi_{0.05}^2 = 5.991$ para $v = (2 - 1)(3 - 1) = 2$ grados de libertad. La hipótesis nula se rechaza. Concluimos que la opinión de un votante con respecto a la nueva reforma fiscal y su nivel de ingresos no son independientes.

Es importante recordar que la estadística sobre la que basamos nuestra decisión tiene una distribución que sólo se aproxima por la distribución ji cuadrada. Los valores χ^2 calculados dependen de las frecuencias de las celdas y en consecuencia son discretas. La distribución ji cuadrada continua parece aproximar muy bien a la distribución de muestreo discreta de χ^2 , dado que el número de grados de libertad es mayor que 1. En una tabla de contingencia de 2×2 , donde sólo tenemos 1 grado de libertad, se aplica una correlación llamada **corrección de Yates para continuidad**. La fórmula corregida se vuelve entonces

$$\chi^2(\text{corregida}) = \sum_r \frac{(|o_r - e_r| - 0.5)^2}{e_r}$$

Si las frecuencias de celdas esperadas son grandes, los resultados corregidos y sin corrección son casi los mismos. Cuando las frecuencias esperadas están entre 5 y 10, se debe aplicar la corrección de Yates. Para frecuencias esperadas menores que 5, se debe utilizar la prueba exacta de Fisher-Irwin. Una discusión de esta prueba se puede encontrar en *Basic Concepts of Probability and Statistics* de Hodges and Lehmann (véase la bibliografía). La prueba de Fisher-Irwin se puede evitar, sin embargo, mediante la elección de una muestra grande.

10.16 Prueba de homogeneidad

Cuando probamos la independencia en la sección 10.15, se seleccionó una muestra aleatoria de 1000 votantes y los totales de renglón y columna para nuestra tabla de contingencia se determinaron al azar. Otro tipo de problema para el que se aplica el método de la sección 10.15 es uno en el que se predeterminan los totales de renglón y columna. Suponga, por ejemplo, que decidimos de antemano seleccionar 200 demócratas, 150 republicanos y 150 independientes de los votantes del estado de Carolina del Norte y registrar si favorecen una propuesta de ley del aborto, están en contra o están indecisos. Las respuestas observadas se dan en la tabla 10.7.

Ahora bien, en lugar de probar la independencia, probamos la hipótesis de que las proporciones de población dentro de cada renglón son las mismas. Es decir, probamos la hipótesis de que las proporciones de demócratas, republicanos e independientes que favorecen la ley sobre el aborto son las mismas; las proporciones de cada afiliación política contra la ley son las mismas; y las proporciones de cada afiliación política que están indecisos son las mismas. Básicamente nos interesamos en determinar si las tres categorías de votantes son **homogéneas** con respecto a sus opiniones acerca de la propuesta de ley del aborto. Tal prueba se llama prueba de homogeneidad.

Al suponer homogeneidad, de nuevo encontramos las frecuencias esperadas de las celdas al multiplicar los totales de renglón y columna correspondientes y después dividir entre el gran total. El análisis entonces continúa mediante el uso de la misma estadística χ^2 cuadrada como antes. Ilustramos este proceso en el siguiente ejemplo para los datos de la tabla 10.7.

Tabla 10.7 Tabla de contingencia 3×3

Ley del aborto	Afiliación política			Total
	Demócrata	Republicano	Independiente	
A favor	82	70	62	214
En contra	93	62	67	222
Indecisos	<u>25</u>	<u>18</u>	<u>21</u>	<u>64</u>
Total	200	150	150	500

Ejemplo 10.15 Con referencia a los datos de la tabla 10.7, probar la hipótesis de que las opiniones con respecto a la ley del aborto propuesta son las mismas dentro de cada afiliación política. Utilice un nivel de significancia de 0.05.

SOLUCIÓN

- H_0 : para cada opinión las proporciones de demócratas, republicanos e independientes son las mismas.
- H_1 : para al menos una opinión las proporciones de demócratas, republicanos e independientes no son las mismas.
- $\alpha = 0.05$.
- Región crítica: $\chi^2 > 9.488$ con $\nu = 4$ grados de libertad.
- Cálculos: Al usar la fórmula de las frecuencias de celdas esperadas de la página 345 necesitamos calcular las cuatro frecuencias de celdas. Todas las demás

frecuencias se encuentran por sustracción. Las frecuencias de celdas observadas y esperadas se muestran en la tabla 10.8.

Tabla 10.8 Frecuencias observadas y esperadas

Ley del aborto	Afiliación política			Total
	Demócrata	Republicano	Independiente	
A favor	82 (85.6)	70 (64.2)	62 (64.2)	214
En contra	93 (88.8)	62 (66.6)	67 (66.6)	222
Indecisos	25 (25.6)	18 (19.2)	21 (19.2)	64
Total	200	150	150	500

Ahora bien,

$$\begin{aligned} \chi^2 &= \frac{(82 - 85.6)^2}{85.6} + \frac{(70 - 64.2)^2}{64.2} + \frac{(62 - 64.2)^2}{64.2} \\ &+ \frac{(93 - 88.8)^2}{88.8} + \frac{(62 - 66.6)^2}{66.6} + \frac{(67 - 66.6)^2}{66.6} \\ &+ \frac{(25 - 25.6)^2}{25.6} + \frac{(18 - 19.2)^2}{19.2} + \frac{(21 - 19.2)^2}{19.2} \\ &= 1.53. \end{aligned}$$

6. Decisión: No rechazar H_0 . No hay suficiente evidencia para concluir que la proporción de demócratas, republicanos e independientes difieren para cada opinión establecida. ■

10.17 Prueba para varias proporciones

La estadística ji cuadrada para probar la homogeneidad también se aplica cuando se prueba la hipótesis de que k parámetros binomiales tienen el mismo valor. Esta es, por tanto, una extensión de la prueba que se presenta en la sección 10.12 para determinar diferencias entre dos proporciones a una prueba para determinar diferencias entre k proporciones. Por ello nos interesamos en probar la hipótesis nula

$$H_0: p_1 = p_2 = \dots = p_k$$

contra la hipótesis alternativa, H_1 , de que las proporciones de la población *no son todas iguales*. Para ejecutar esta prueba, primero observamos muestras aleatorias independientes de tamaño n_1, n_2, \dots, n_k de las k poblaciones y acomodamos los datos como en la tabla de contingencia $2 \times k$, tabla 10.9.

Tabla 10.9 k muestras binomiales independientes

Muestra:	1	2	...	k
Éxitos	x_1	x_2	...	x_k
Fracasos	$n_1 - x_1$	$n_2 - x_2$...	$n_k - x_k$

Según si los tamaños de las muestras aleatorias se predeterminaron u ocurrieron al azar, el procedimiento de prueba es idéntico a la prueba de homogeneidad o a la

prueba de independencia. Por tanto, las frecuencias de celdas esperadas se calculan como antes y se sustituyen junto con las frecuencias observadas en la estadística ji cuadrada

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i}$$

con

$$v = (2 - 1)(k - 1) = k - 1$$

grados de libertad. Al seleccionar la región crítica apropiada de la cola superior de la forma $\chi^2 > \chi_{\alpha}^2$, se puede llegar ahora a una decisión con respecto a H_0 .

Ejemplo 10.16 En un estudio de un taller, se reúne un conjunto de datos para determinar si la proporción de defectuosos producida por los trabajadores es la misma para el turno matutino, vespertino o nocturno. Se reunieron los datos siguientes:

Turno:	Matutino	Vespertino	Nocturno
Defectuosos	45	55	70
No defectuosos	905	890	870

Utilice un nivel de significancia de 0.025 para determinar si la proporción de defectuosos es la misma para los tres turnos.

SOLUCIÓN

Representemos con p_1 , p_2 , y p_3 la proporción real de defectuosos para los turnos matutino, vespertino y nocturno, respectivamente.

1. $H_0: p_1 = p_2 = p_3$.
2. $H_1: p_1, p_2, \text{ y } p_3$ no son todas iguales.
3. $\alpha = 0.025$.
4. Región crítica: $\chi^2 > 7.378$ para $v = 2$ grados de libertad.
5. Cálculos: en correspondencia con las frecuencias observadas $o_1 = 45$ y $o_2 = 55$, encontramos

$$e_1 = \frac{(950)(170)}{2835} = 57.0 \quad \text{y} \quad e_2 = \frac{(945)(170)}{2835} = 56.7.$$

Tabla 10.10 Frecuencias esperadas y observadas

Turno:	Matutino	Vespertino	Nocturno	Total
Defectuosos	45 (57.0)	55 (56.7)	70 (56.3)	170
No defectuosos	905 (893.0)	890 (888.3)	870 (883.7)	2665
Total	950	945	940	2835

Todas las demás frecuencias esperadas se encuentran por sustracción y se muestran en la tabla 10.10. Ahora bien,

$$\chi^2 = \frac{(45 - 57.0)^2}{57.0} + \frac{(55 - 56.7)^2}{56.7} + \frac{(70 - 56.3)^2}{56.3} \\ + \frac{(905 - 893.0)^2}{893.0} + \frac{(890 - 888.3)^2}{888.3} + \frac{(870 - 883.7)^2}{883.7} = 6.29$$

$$P = 0.04$$

6. Decisión: no rechazamos H_0 con $\alpha = 0.025$. No obstante, con el anterior valor P calculado, ciertamente sería riesgoso concluir que la proporción de defectuosos producidos es la misma para todos los turnos. ■

10.18 Estudio de dos muestras

En esta sección consideramos un estudio donde mostramos un análisis completo con el uso del análisis gráfico y el formal junto con resultados de computadora comentados y conclusiones. En un estudio de análisis de datos que llevó a cabo el personal del Centro de Consulta Estadística del Tecnológico de Virginia, se compararon dos materiales diferentes, digamos aleación A y aleación B , en términos de la resistencia de ruptura. La aleación B es más cara, pero en realidad se debe adoptar si se puede demostrar que es más fuerte que la aleación A . Se debe tomar en cuenta la consistencia del rendimiento de las dos aleaciones.

Se seleccionaron muestras aleatorias de vigas de cada aleación y la resistencia se midió en una deflexión de 0,001 pulgadas cuando se aplicó una fuerza fija en ambos extremos de la viga. Se utilizaron 20 especímenes para cada una de las dos aleaciones. Los datos son los siguientes:

Aleación A			Aleación B		
88	82	87	75	81	80
79	85	90	77	78	81
84	88	83	86	78	77
89	80	81	84	82	78
81	85		80	80	
83	87		78	76	
82	80		83	85	
79	78		76	79	

Es importante que el ingeniero compare las dos aleaciones. La preocupación es la resistencia promedio y la reproducibilidad. Interesa determinar si hay una violación severa de la suposición de normalidad que requieren las pruebas t y F . Las figuras 10.21 y 10.22 son gráficas de cuantil-cuantil normal de las muestras para las dos aleaciones. No parece haber ninguna violación seria de la suposición de normalidad. Además, la figura 10.23 muestra dos gráficos de caja y extensión en la misma gráfica. Los gráficos de caja y extensión sugieren que no hay una diferencia apreciable de la variabilidad en la deflexión para las dos aleaciones. Sin embargo, parece sugerir que la media de la

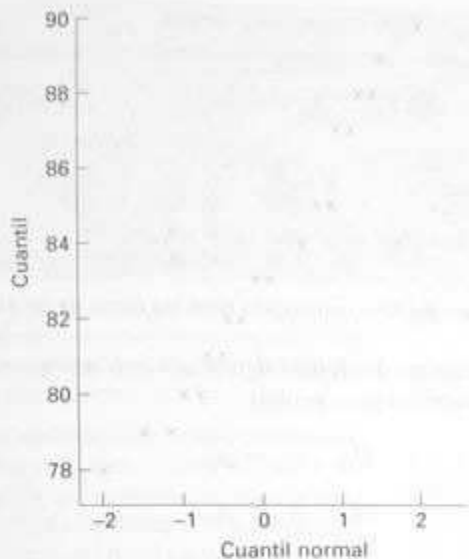


Figura 10.21 Gráfica de cuantil-cuantil normal de los datos para la aleación A.

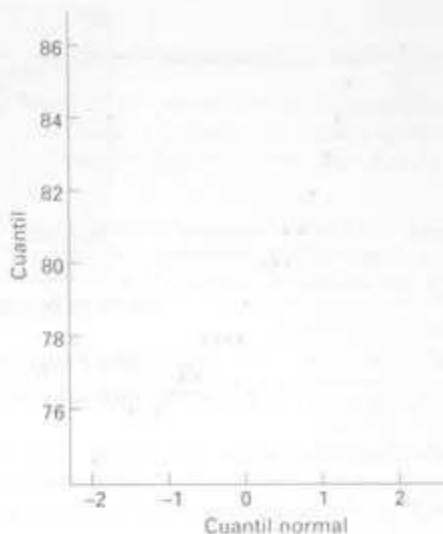


Figura 10.22 Gráfica de cuantil-cuantil normal de los datos para la aleación B.

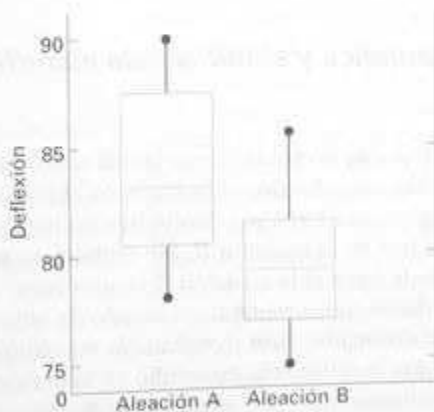


Figura 10.23 Gráficos de caja y extensión para ambas aleaciones.

aleación B es significativamente más pequeña, lo que sugiere al menos gráficamente que la aleación B es más fuerte. Las medias muestrales y las desviaciones estándar son

$$\bar{y}_A = 83.55, \quad s_A = 3.663,$$

$$\bar{y}_B = 79.70, \quad s_B = 3.097.$$

El impreso del SAS para el PROC prueba t se muestra en la figura 10.24.

t-Test Procedure				
ALLOY	N	Mean	Std Dev	Std Error
1	20	63.55000000	3.44311630	0.61909771
2	20	79.70000000	3.09666753	0.69244036
Variances:		T	DF	Prob> T
Unequal		3.5895	37.0	0.0010
Equal		3.5895	38.0	0.0009
For HO: Variances are equal, F = 1.40 DF = (19, 19)				
Prob > F = 0.4709				

Figura 10.24 Impreso del SAS comentado para los datos de las aleaciones.

La prueba F sugiere ninguna diferencia significativa en las varianzas ($P = 0.4709$) y la estadística t de dos muestras para probar

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_B > \mu_A$$

($t = 3.5895$, $P = 0.0009$) rechaza H_0 en favor de H_1 y de esta manera confirma lo que sugiere la información gráfica. Aquí utilizamos la prueba t que reúne las varianzas de dos muestras a la luz de los resultados de la prueba F . Con base en este análisis la adopción de la aleación B sería adecuada.

Significancia estadística y significancia científica o en ingeniería

Mientras que el estadístico se puede sentir muy a gusto con los resultados de la comparación entre las dos aleaciones en el estudio anterior, permanece un dilema para el ingeniero. El análisis demuestra un mejoramiento estadísticamente significativo con el uso de la aleación B . Sin embargo, ¿se encuentra que la diferencia realmente vale la pena si la aleación B es más cara? Esta ilustración resalta un asunto muy importante que a menudo es dejado de lado por los estadísticos y el analista de datos: *la distinción entre significancia estadística y significancia científica o en ingeniería*. Aquí la diferencia promedio en la deflexión es $\bar{y}_A - \bar{y}_B = 0.00385$ pulgadas. En un análisis completo el ingeniero debe determinar si la diferencia es suficiente para justificar el costo extra en el largo plazo.

Éste es un asunto económico y de ingeniería. El lector debe comprender que una diferencia estadísticamente significativa sólo implica que la diferencia en las medias muestrales que se encuentra en los datos con dificultad puede ocurrir al azar. Esto no implica que la diferencia en las medias poblacionales sea profunda o particularmente significativa en el contexto del problema. Por ejemplo, en la sección 10.7 se utiliza un impreso de computadora para mostrar evidencia de que un medidor de pH está, de hecho, sesgado. Es decir, esto no demuestra un pH medio de 7.00 para el material en el que se probó. Pero la variabilidad entre las observaciones en la muestra es muy pequeña. El ingeniero puede decidir que las desviaciones pequeñas de 7.0 hacen al medidor de pH bastante adecuado.

Ejercicios

1. Se lanza 180 veces un dado con los siguientes resultados:

x	1	2	3	4	5	6
f	28	36	36	30	27	23

¿Es un dado balanceado? Utilice un nivel de significancia de 0.01.

2. En 100 lanzamientos de una moneda se observan 63 caras y 37 cruces. ¿Es una moneda balanceada? Utilice un nivel de significancia de 0.05.

3. Se supone que una máquina mezcla cacahuates, avellanas, anacardos y pacanas a razón de 5:2:2:1. Se encuentra que una lata que contiene 500 de estas nueces mezcladas tiene 269 cacahuates, 112 avellanas, 74 anacardos y 45 pacanas. Al nivel de significancia de 0.05, pruebe la hipótesis de que la máquina mezcla las nueces a una razón de 5:2:2:1.

4. Las calificaciones de un curso de estadística para un semestre particular fueron las siguientes:

Calificación	A	B	C	D	F
f	14	18	32	20	16

Pruebe la hipótesis, al nivel de significancia de 0.05, de que la distribución de calificaciones es uniforme.

5. Se extraen tres cartas de una baraja ordinaria, con reemplazo, y se registra el número Y de espadas. Después de repetir el experimento 64 veces, se registran los siguientes resultados:

y	0	1	2	3
f	21	31	12	0

Pruebe la hipótesis, con un nivel de significancia de 0.01, de que los datos registrados se pueden ajustar mediante una distribución binomial $b(y; 3, 1/4)$, $y = 0, 1, 2, 3$.

6. Se seleccionan tres canicas de una urna que contiene cinco canicas rojas y tres verdes. Después de registrar el número X de canicas rojas, las canicas se reemplazan en la urna y el experimento se repite 112 veces. Los resultados que se obtienen son los siguientes:

x	0	1	2	3
f	1	31	55	25

Pruebe la hipótesis con un nivel de significancia de 0.05, de que los datos registrados se pueden ajustar con la distribución hipergeométrica $h(x; 8, 3, 5)$, $x = 0, 1, 2, 3$.

7. Se lanza una moneda hasta que sale una cara y se registra el número de lanzamientos X . Después de repetir el experimento 256 veces, obtenemos los siguientes resultados:

x	1	2	3	4	5	6	7	8
f	136	60	34	12	9	1	3	1

Pruebe la hipótesis con un nivel de significancia de 0.05 de que la distribución observada de X se puede ajustar por una distribución geométrica $g(x; 1/2)$, $x = 1, 2, 3, \dots$

8. Repita el ejercicio 5 con el uso de un conjunto nuevo de datos obtenidos al llevar a cabo realmente el experimento que se describe 64 veces.

9. Repita el ejercicio 7 con el uso de un nuevo conjunto de datos obtenidos al realizar el experimento que se describe 256 veces.

10. En el ejercicio 1 de la página 68, pruebe la bondad de ajuste entre las frecuencias de clase que se observan y las frecuencias esperadas correspondientes de una distribución normal con $\mu = 65$ y $\sigma = 21$, utilice un nivel de significancia de 0.05.

11. En el ejercicio 5 de la página 69, pruebe la bondad del ajuste entre las frecuencias de clase que se observan y las frecuencias esperadas correspondientes de una distribución normal con $\mu = 1.8$ y $\sigma = 0.4$, utilice un nivel de significancia de 0.01.

12. En un experimento para estudiar la dependencia de la hipertensión de los hábitos de fumar, se tomaron los siguientes datos de 180 individuos:

	No fumadores	Fumadores moderados	Fumadores empedernidos
Con hipertensión	21	36	30
Sin hipertensión	48	26	19

Pruebe la hipótesis de que la presencia o ausencia de hipertensión es independiente de los hábitos de fumar. Utilice un nivel de significancia de 0.05.

13. Una muestra aleatoria de 90 adultos se clasifica de acuerdo a su género y al número de horas que pasan viendo la televisión durante una semana:

	Sexo	
	Masculino	Femenino
Más de 25 horas	15	29
Menos de 25 horas	27	19

Utilice un nivel de significancia de 0.01 y pruebe la hipótesis de que el tiempo que pasan viendo televisión es independiente de si el televidente es hombre o mujer.

14. Una muestra aleatoria de 200 hombres casados, todos retirados, se clasifica de acuerdo con la educación y el número de hijos:

Educación	Número de hijos		
	0-1	2-3	Over 3
Elemental	14	37	32
Secundaria	19	42	17
Universidad	12	17	10

Pruebe la hipótesis, con un nivel de significancia de 0.05, de que el tamaño de la familia es independiente del nivel de instrucción del padre.

15. Un criminólogo realizó una investigación para determinar si la incidencia de ciertos tipos de crímenes varían de una parte a otra en una ciudad grande. Los crímenes particulares de interés son asalto, robo, hurto y homicidio. La siguiente tabla muestra el número de crímenes cometidos en cuatro áreas de la ciudad durante el año pasado.

Distrito	Tipo de crimen			
	Asalto	Robo	Hurto	Homicidio
1	162	118	451	18
2	310	196	996	25
3	258	193	458	10
4	280	175	390	19

¿Podemos concluir a partir de estos datos con un nivel de significancia de 0.01 que la ocurrencia de estos tipos de crimen es dependiente del distrito de la ciudad?

16. La enfermería de una universidad realiza un experimento para determinar el grado de alivio que proporcionan tres remedios para la tos. Cada remedio para la tos se trata en 50 estudiantes y se registran los siguientes datos:

	Remedio para la tos		
	NyQuil	Robitussin	Triaminic
Sin alivio	11	13	9
Cierto alivio	32	28	27
Alivio completo	7	9	14

Pruebe la hipótesis, con un nivel de significancia de 0.05, de que los tres remedios para la tos son igualmente efectivos.

17. Para determinar las posiciones actuales acerca de las oraciones en escuelas públicas se llevó a cabo una investigación en cuatro condados de Virginia. La siguiente tabla da las posiciones de 200 padres del condado de Craig, 150 padres del condado de Giles, 100 padres de Franklin y 100 del Montgomery:

Posición	Condado			
	Craig	Giles	Franklin	Montgomery
A favor	65	66	40	34
En contra	42	30	33	42
Sin opinión	93	54	27	24

Al nivel de significancia de 0.01, pruebe la homogeneidad de las posiciones entre los cuatro condados con respecto a las oraciones en escuelas públicas.

18. De acuerdo con un estudio de la Universidad Johns Hopkins publicado en el *American Journal of Public Health*, las viudas viven más que los viudos. Considere los siguientes datos de sobrevivencia de 100 viudas y 100 viudos después de la muerte del cónyuge:

Años vividos	Viuda	Viudo
Menos de 5	25	39
de 5 a 10	42	40
Más de 10	33	21

¿Podemos concluir con un nivel de significancia de 0.05 que las proporciones de viudas y viudos son iguales con respecto a los diferentes periodos que un cónyuge sobrevive a la muerte de su compañero?

19. Las siguientes respuestas respecto al estándar de vida al momento de una encuesta de opinión independiente de 1000 familias contra un año antes parece estar de acuerdo con los resultados de un estudio publicado en *Across the Board* (junio de 1981). (Ver tabla en la siguiente página):

Pruebe la hipótesis con un nivel de significancia de 0.05 de que las proporciones de familias dentro de cada estándar de vida son las mismas para cada uno de los cuatro periodos.

Estándar de vida

Periodo	Algo mejor	Igual	No tan bueno	Total
1980: Enero	72	144	84	300
Mayo	63	135	102	300
Sept.	47	100	53	200
1981: Enero	40	105	55	200

20. Se lleva a cabo un estudio en Indiana, Kentucky y Ohio para determinar la posición de los votantes con respecto al transporte escolar. Una encuesta de 200 votantes de cada uno de estos estados da los siguientes resultados:

Posición del votante:

Estado	Apoya	No apoya	Indeciso
Indiana	82	97	21
Kentucky	107	66	27
Ohio	93	74	33

Al nivel de significancia de 0.025, pruebe la hipótesis nula de que las proporciones de votantes dentro de cada categoría de posición son las mismas para cada uno de los tres estados.

21. Se lleva a cabo una investigación en dos ciudades de Virginia para determinar el sentimiento de los votantes hacia los candidatos en una elección próxima. Se seleccionan 500 votantes al azar de cada ciudad y se registran los siguientes datos:

Sentimiento del votante	Ciudad	
	Richmond	Norfolk
A favor de A	204	225
A favor de B	211	198
Indeciso	85	77

Con un nivel de significancia de 0.05, pruebe la hipótesis nula de que las proporciones de votantes que favorecen al candidato A, al candidato B o están indecisos son las mismas para cada ciudad.

22. En un estudio para estimar la proporción de esposas que de manera regular ven telenovelas, se encuentra que 52 de 200 esposas en Denver, 31 de 150 en Phoenix y 37 de 150 en Rochester ven al menos una telenovela. Utilice un nivel de significancia de 0.05 para probar la hipótesis de que no hay diferencia entre las proporciones reales de esposas que ven telenovelas en estas tres ciudades.

Ejercicios de repaso

1. Un genetista se interesa en la proporción de hombres y mujeres de una población que tiene cierto trastorno sanguíneo menor. En una muestra aleatoria de 100 hombres, se encuentra que 31 lo padecen, mientras que sólo 24 de 100 mujeres parecen tener el trastorno. ¿Podemos concluir con un nivel de significancia de 0.01 que la proporción de hombres en la población con este trastorno sanguíneo es significativamente mayor que la proporción de mujeres afectadas?

2. Considere la situación del ejercicio 36 de la página 330. El consumo de oxígeno en ml/kg/min también se midió en los nueve sujetos.

Sujeto	Con CO	Sin CO
1	26.46	25.41
2	17.46	22.53
3	16.32	16.32
4	20.19	27.48
5	19.84	24.97

Sujeto	Con CO	Sin CO
6	20.65	21.77
7	28.21	28.17
8	33.94	32.02
9	29.32	28.96

Se conjetura que el consumo de oxígeno debe ser mayor en un ambiente relativamente libre de CO. Realice una prueba de significancia y discuta la conjetura.

3. Establezca las hipótesis nula y alternativa a utilizarse en la prueba de las siguientes afirmaciones y determine de manera general dónde se localiza la región crítica:

- La caída de nieve promedio en el lago George durante el mes de febrero es 21.8 centímetros.
- No más de 20% del cuerpo de profesores en la universidad local contribuye a un fondo anual.
- En promedio, los niños asisten a la escuela dentro de 6.2 kilómetros de sus casas en un suburbio de St. Louis.

(d) Al menos 70% de los autos nuevos del siguiente año caerán en la categoría de compactos y semicompactos.

(e) La proporción de votantes que favorecen al actual en la próxima elección es 0.58.

(f) El filete rib-eye promedio en el restaurante Loughorn Steak es de al menos 340 gramos.

4. Se realiza un estudio para determinar si más italianos que americanos prefieren la champaña blanca a la rosa en las bodas. De los 300 italianos que se seleccionaron al azar, 72 prefieren champaña blanca, y de los 400 americanos seleccionados, 70 prefieren champaña blanca en lugar de la rosa. ¿Podemos concluir que una proporción más grande de italianos que de americanos prefiere champaña blanca en las bodas? Utilice un nivel de significancia de 0.05.

5. En un conjunto de datos analizados por el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia se pidió a un grupo de sujetos completar cierta tarea en la computadora. La respuesta medida fue el tiempo de terminación. El propósito del experimento fue probar un grupo de herramientas de ayuda desarrolladas por el Departamento de Ciencias Computacionales del mismo instituto. Hubo 10 sujetos involucrados. Con una asignación al azar, a cinco se les dio un procedimiento estándar con lenguaje Fortran para completar la tarea. A los otros cinco se les pidió realizar la tarea con el uso de las herramientas de ayuda. A continuación se presentan los datos de los tiempos de terminación de las tareas.

Grupo 1 (procedimiento estándar)	Grupo 2 (herramienta de ayuda)
161	132
169	162
174	134
158	138
163	133

Suponga que las varianzas poblacionales son las mismas para los dos grupos; apoye o rechace la conjetura de que las herramientas de ayuda aumentan la velocidad con la que se lleva a cabo la tarea.

6. Establezca las hipótesis nula y alternativa a utilizar en la prueba de las siguientes afirmaciones, y determine de manera general dónde se ubica la región crítica:

(a) A lo más 20% de la cosecha de trigo del próximo año se exportará a la Unión Soviética.

(b) En promedio, las amas de casa americanas beben tres tazas de café al día.

(c) La proporción de graduados en Virginia este año que se especializan en ciencias sociales es al menos 0.15.

(d) La donación promedio a la Asociación Americana del Pulmón es no más de \$10.

(e) Los residentes del suburbio Richmond recorren, en promedio, 15 kilómetros hasta su lugar de trabajo.

7. Si se selecciona una lata que contiene 500 nueces al azar de cada uno de tres diferentes distribuidores de nueces surtidas y contienen, respectivamente, 345, 313 y 359 cacahuates en cada una de las latas, ¿podemos concluir con un nivel de significancia de 0.01 que las nueces surtidas de los tres distribuidores contienen proporciones iguales de cacahuates?

8. Valor z para probar $p_1 - p_2 = d_0$. Para probar la hipótesis nula H_0 de que $p_1 - p_2 = d_0$, donde $d_0 \neq 0$, basamos nuestra decisión en

$$z = \frac{\hat{p}_1 - \hat{p}_2 - d_0}{\sqrt{\hat{p}_1 \hat{q}_1 / n_1 + \hat{p}_2 \hat{q}_2 / n_2}}$$

que es un valor de una variable aleatoria cuya distribución aproxima la distribución normal estándar en tanto que n_1 y n_2 sean grandes.

Con referencia al ejemplo 10.12 de la página 334, pruebe la hipótesis de que el porcentaje de votantes de la ciudad que favorecen la construcción de la planta química no excederá el porcentaje de votantes del condado por más de 3%. Utilice un valor P en su conclusión.

9. Se realiza un estudio para determinar si hay una diferencia entre las proporciones de padres en los estados de Maryland, Virginia, Georgia y Alabama que están a favor de colocar Biblias en las escuelas primarias. En la siguiente tabla se registran las respuestas de 100 padres seleccionados al azar en cada uno de estos estados:

Preferencia	Estado			
	Maryland	Virginia	Georgia	Alabama
Si	65	71	78	82
No	35	29	22	18

¿Podemos concluir que las proporciones de padres que están a favor de colocar Biblias en las escuelas son las mismas para estos cuatro estados? Utilice un nivel de significancia de 0.025.

10. Se lleva a cabo un estudio en el Centro de Medicina Veterinaria Equina del Colegio Regional de Virginia-Maryland para determinar si la ejecución de cierto tipo de cirugía en caballos jóvenes tiene algún efecto en ciertas clases de células sanguíneas en el

animal. Se toman muestras del fluido de cada uno de seis potros antes y después de la cirugía. Se analizan las muestras para el número de leucogramos WBC posoperatorios. También se realiza una medición de leucogramos WBC preoperatorios. Utilice una prueba t de una muestra pareada para determinar si hay un cambio significativo en los leucogramos WBC con la cirugía.

Potro	Precirugía*	Postcirugía*
1	10.80	10.60
2	12.90	16.60
3	9.59	17.20
4	8.81	14.00
5	12.00	10.60
6	6.07	8.60

*Todos los valores $\times 10^{-3}$.

11. Se lleva a cabo un estudio en el Departamento de Salud y Educación Física del Instituto Politécnico y Universidad Estatal de Virginia para determinar si ocho semanas de entrenamiento realmente reducen los niveles de colesterol en los participantes. A un grupo de tratamiento que consiste en 15 personas se les dan conferencias dos veces a la semana de cómo reducir su nivel de colesterol. Otro grupo de 18 personas de edad similar se selecciona al azar como grupo de control. Se registran los niveles de colesterol de todos los participantes al final del programa de ocho semanas y se listan a continuación:

Grupo con tratamiento:

129	131	154	172	115	126	175	191
122	238	159	156	176	175	126	

Grupo de control:

151	132	196	195	188	198	187	168	115
165	137	208	133	217	191	193	140	146

¿Podemos concluir, con un nivel de significancia de 5%, que el nivel de colesterol promedio se reduce como consecuencia del programa?

12. En un estudio que realiza el Departamento de Ingeniería Mecánica y que analiza el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia, se comparan las varillas de acero que proveen dos compañías diferentes. Se fabrican diez resortes de muestra con las varillas proporcionadas por cada compañía y se estudia "la capacidad de rebote". Los datos son los siguientes:

Compañía A:

9.3	8.8	6.8	8.7	8.5	6.7	8.0	6.5	9.2	7.0
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Compañía B:

11.0	9.8	9.9	10.2	10.1	9.7	11.0	11.1	10.2	9.6
------	-----	-----	------	------	-----	------	------	------	-----

¿Puede concluir que casi no hay diferencia entre las varillas de acero proporcionadas por las dos compañías? Utilice un valor P para llegar a su conclusión.

13. En un estudio que conduce el Centro de Recursos Acuáticos y que analiza el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia, se comparan dos plantas de tratamiento de aguas residuales. La planta A se ubica donde el ingreso medio de los hogares está por abajo de \$22,000 al año, y la planta B se ubica donde el ingreso medio de los hogares está por arriba de \$60,000 anuales. La cantidad de agua residual que trata cada planta (miles de galones/día) se muestrea de forma aleatoria durante 10 días. Los datos son los siguientes:

Planta A:

21	19	20	23	22	28	32	19	13	18
----	----	----	----	----	----	----	----	----	----

Planta B:

20	39	24	33	30	28	30	22	33	24
----	----	----	----	----	----	----	----	----	----

¿Podemos concluir, con un nivel de significancia de 5%, que la cantidad promedio de agua residual tratada en la vecindad de altos ingresos es mayor que la del área de bajos ingresos?

14. Los siguientes datos muestran el número de defectos en 100,000 líneas de código en un tipo particular de programa de software hecho en Estados Unidos y Japón. ¿Hay suficiente evidencia para afirmar que existe una diferencia significativa entre los programas de los dos países?

Estados Unidos	48	39	42	52	40	48	52	52
Japón	54	48	52	55	43	46	48	52
	50	48	42	40	43	48	50	46
	38	38	36	40	40	48	48	45

15. Estudios muestran que la concentración de PCB es mucho más alta en tejido maligno de pecho que en tejido normal de pecho. Si un estudio de 50 mujeres con cáncer de pecho revela una concentración promedio de PCB de 22.8×10^{-4} gramos, con una desviación estándar de 4.8×10^{-4} gramos, ¿la concentración media de PCB es menor que 24×10^{-4} gramos?

Regresión lineal simple y correlación

11.1 Introducción a la regresión lineal

A menudo, en la práctica, se requiere resolver problemas que incluyen conjuntos de variables cuando se sabe que existen algunas relaciones inherentes entre ellas. Por ejemplo, en una situación industrial se puede saber que el contenido de alquitrán en el flujo saliente de un proceso químico se relaciona con la temperatura de entrada. Puede ser de interés desarrollar un método de predicción; es decir, un procedimiento para estimar el contenido de alquitrán para varios niveles de temperatura de entrada a partir de la información experimental. El aspecto estadístico del problema se convierte entonces en lograr la mejor estimación de la relación entre las variables.

Gran parte de lo presentado y que se discutirá depende de la suposición del modelo. Por supuesto, el tipo de justificación intuitiva que señala la necesidad del uso de un análisis de regresión lineal simple es una **gráfica de los datos**. La gráfica que se muestra en la figura 11.1 de la página 360 en realidad sugiere que el modelo es lineal. Continuamos con un mayor desarrollo y discusión respecto a la elección del modelo. Sin embargo, siempre se deben graficar los datos.

Para este ejemplo y para la mayor parte de las aplicaciones hay una distinción clara entre las variables en lo que respecta a su papel en el proceso experimental. Muy a menudo existe una sola *variable dependiente* o respuesta Y , que no se controla en el experimento. Esta respuesta depende de una o más **variables de regresión independientes**, digamos x_1, x_2, \dots, x_k , que se miden con un error insignificante y a menudo realmente se controlan en el experimento. De esta manera las variables independientes x_1, x_2, \dots, x_k *no* son variables aleatorias y por tanto no tienen propiedades de distribución. En el ejemplo que se cita al principio, la temperatura de

entrada es la variable independiente o variable de regresión x y el contenido de alquitrán es la respuesta Y . La relación que se ajusta a un conjunto de datos experimentales se caracteriza por una ecuación de predicción que se denomina **ecuación de regresión**. En el caso de una sola Y y una sola x , la situación se convierte en una regresión de Y sobre x . Para k variables independientes, hablamos en términos de una regresión de Y sobre x_1, x_2, \dots, x_k . Un ingeniero químico se puede preocupar, de hecho, por la cantidad de hidrógeno que se pierde de muestras de un metal particular cuando el material se almacena. En este caso puede haber dos entradas, el tiempo de almacenamiento x_1 en horas y la temperatura de almacenamiento x_2 en grados centígrados. La respuesta sería entonces la pérdida de hidrógeno Y en partes por millón.

En este capítulo nos ocupamos del tema de la **regresión lineal simple**, sólo para el caso de una sola variable de regresión. Para el caso de más de una variable de regresión, se remite al lector al capítulo 12. Denotemos una muestra aleatoria de tamaño n con el conjunto $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Si se toman muestras adicionales mediante el uso de exactamente los mismos valores de x , debemos esperar que varíen los valores y . De aquí el valor y , en el par ordenado (x, y) es un valor de alguna variable aleatoria Y . Por conveniencia definimos $Y|x$ como la variable aleatoria Y que corresponde a un valor fijo x y denotamos su media y varianzas con $\mu_{Y|x}$ y $\sigma_{Y|x}^2$, respectivamente. Es claro entonces que, si $x = x_c$, el símbolo $Y|x_c$ representa la variable aleatoria Y_c con media $\mu_{Y|x_c}$ y varianzas $\sigma_{Y|x_c}^2$.

El término **regresión lineal** implica que $\mu_{Y|x}$ se relaciona linealmente con x mediante la **ecuación de regresión de población**

$$\mu_{Y|x} = \alpha + \beta x,$$

donde los **coeficientes de regresión** α y β son parámetros a estimar a partir de los datos muestrales. Al denotar sus estimaciones con a y b , respectivamente, podemos estimar $\mu_{Y|x}$ con \hat{y} a partir de la regresión de la muestra o la **línea de regresión ajustada**

$$\hat{y} = a + bx,$$

donde las estimaciones a y b representan la intersección y y la pendiente, respectivamente. El símbolo \hat{y} se utiliza aquí para distinguir entre el valor estimado o predicho dado por la línea de regresión de la muestra y un valor experimental real observado y para algún valor de x .

Uno de los problemas más desafiantes que enfrenta el campo del control de la contaminación del agua lo presenta la industria del curtido de pieles. Los desechos de las curtidurías son químicamente complejos. Se caracterizan por los altos valores de demanda bioquímica de oxígeno, sólidos volátiles y otras medidas de contaminación. Considérense los datos experimentales de la tabla 11.1, que se obtuvieron de 33 muestras de desechos químicamente tratados en un estudio que realizó el Instituto Politécnico y Universidad Estatal de Virginia. Se registraron las lecturas de x , la reducción porcentual de sólidos totales, y y , la reducción porcentual en la demanda química de oxígeno para las 33 muestras.

Los datos de la tabla 11.1 se grafican en la figura 11.1, que muestra un **diagrama de dispersión**. De la inspección de este diagrama de dispersión se ve que los puntos cercanos siguen una línea recta, lo que indica que la suposición de linealidad entre las dos variables parece ser razonable.

Tabla 11.1 Mediciones de sólidos y demanda química de oxígeno

Reducción de sólidos, x (%)	Demanda química de oxígeno, y (%)	Reducción de sólidos, x (%)	Demanda química de oxígeno, y (%)
3	5	36	34
7	11	37	36
11	21	38	38
15	16	39	37
18	16	39	36
27	28	39	45
29	27	40	39
30	25	41	41
30	35	42	40
31	30	42	44
31	40	43	37
32	32	44	44
33	34	45	46
33	32	46	46
34	34	47	49
36	37	50	51
36	38		

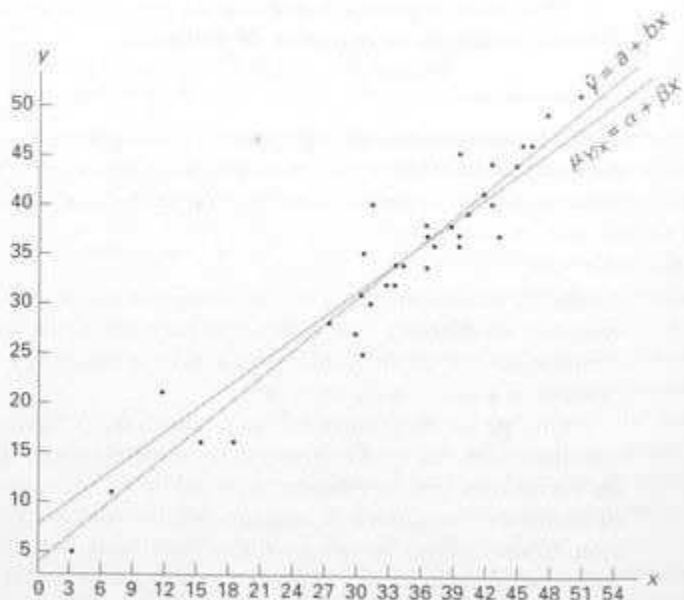


Figura 11.1 Diagrama de dispersión con líneas de regresión.

En el diagrama de dispersión de la figura 11.1 se muestran una línea de regresión ajustada y una línea de regresión hipotética real. Se espera que la concordancia entre la línea muestral y la línea hipotética desconocida sea buena cuando se disponga de una cantidad grande de datos.

En la siguiente sección desarrollamos procedimientos para encontrar estimaciones de los coeficientes de regresión α y β a fin de que se pueda utilizar la ecuación de regresión para predecir o estimar la respuesta media o una respuesta individual para un valor específico de la variable independiente x .

11.2 Regresión lineal simple

En el caso de la regresión lineal simple donde hay una sola variable de regresión independiente x y una sola variable aleatoria dependiente Y , los datos se pueden representar mediante los pares de observaciones $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Es conveniente utilizar los conceptos de la sección anterior para definir cada variable aleatoria $Y_i = Y|x_i$ por medio de un **modelo estadístico**. Si postulamos que todas las medias $\mu_{Y|x_i}$ caen en una línea recta, cada Y_i se puede describir con el **modelo de regresión lineal simple**.

$$Y_i = \mu_{Y|x_i} + E_i = \alpha + \beta x_i + E_i,$$

donde el error aleatorio E_i , el error del modelo, necesariamente debe tener una media de cero. Cada observación (x_i, y_i) en nuestra muestra satisface la ecuación

$$y_i = \alpha + \beta x_i + \varepsilon_i,$$

donde ε_i es el valor que toma E_i cuando Y_i toma el valor y_i . La ecuación anterior se puede ver como el modelo para una sola observación y_i .

De manera similar, con el uso de la línea de regresión estimada o ajustada

$$\hat{y} = a + bx,$$

cada par de observaciones satisface la relación

$$y_i = a + bx_i + e_i,$$

donde $e_i = y_i - \hat{y}_i$ se denomina **residuo** y describe el error en el ajuste del modelo en el i -ésimo punto de los datos. La diferencia entre e_i y ε_i se muestra con claridad en la figura 11.2.

La figura 11.1 representa el ajuste lineal para este conjunto de datos, a saber, $\hat{y} = a + bx$, y la línea refleja el modelo $\mu_{Y|x} = \alpha + \beta x$. Ahora bien, por supuesto, α y β son parámetros desconocidos. La línea ajustada es una estimación de la línea que produce el modelo estadístico. Tenga en cuenta que la línea $\mu_{Y|x} = \alpha + \beta x$ no

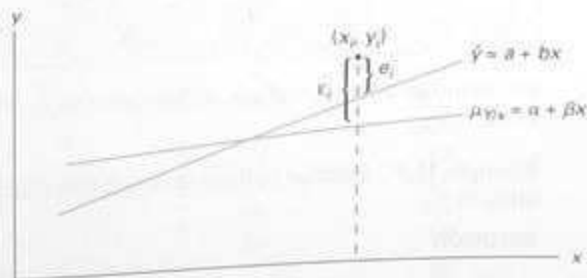


Figura 11.2 Comparación de e_i con el residuo ε_i .

se conoce sino que más bien es una noción conceptual simple de cómo se generaron los datos en el proceso científico. Como resultado, la realización de E_i , a saber, ε_i , en realidad nunca se observa. Sin embargo, se observa el residuo e_i .

Método de mínimos cuadrados

Encontraremos a y b , las estimaciones de α y β , de modo que la suma de los cuadrados de los residuos sea mínima. A menudo la suma de los cuadrados de los residuos se llama suma de cuadrados de los errores alrededor de la línea de regresión y se denota con SSE. Este procedimiento de minimización para estimar los parámetros se llama **método de mínimos cuadrados**. Por ello, encontraremos a y b de modo que se minimice

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2.$$

Al diferenciar SSE con respecto a a y b , tenemos

$$\frac{\partial(SSE)}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i), \quad \frac{\partial(SSE)}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bx_i)x_i.$$

Al hacer las derivadas parciales iguales a cero y reacomodar los términos, obtenemos las ecuaciones (denominadas **ecuaciones normales**)

$$na + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \quad a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i,$$

que podemos resolver de manera simultánea para obtener fórmulas de cálculo para a y b .

Estimación de los coeficientes de regresión

Dada la muestra $\{(x_i, y_i); i = 1, 2, \dots, n\}$, las estimaciones por mínimos cuadrados a y b de los coeficientes de regresión α y β se calculan a partir de las fórmulas

$$b = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

y

$$a = \frac{\sum_{i=1}^n y_i - b \sum_{i=1}^n x_i}{n} = \bar{y} - b\bar{x}.$$

El cálculo de a y b , con el uso de los datos de la tabla 11.1, se ilustran con el siguiente ejemplo.

Ejemplo 11.1 Estimar la línea de regresión para los datos de contaminación de la tabla 11.1.

SOLUCIÓN

$$\sum_{i=1}^{33} x_i = 1104, \quad \sum_{i=1}^{33} y_i = 1124, \quad \sum_{i=1}^{33} x_i y_i = 41,355, \quad \sum_{i=1}^{33} x_i^2 = 41,086.$$

Por tanto

$$\hat{b} = \frac{(33)(41,355) - (1104)(1124)}{(33)(41,086) - (1104)^2} = 0.903643$$

y

$$\hat{a} = \frac{1124 - (0.903643)(1104)}{33} = 3.829633.$$

De esta manera la línea de regresión estimada está dada por

$$\hat{y} = 3.8296 + 0.9036x,$$

donde los coeficientes se redondean a cuatro decimales. Al sustituir dos valores de x en esta ecuación, digamos $x = 3$ y $x = 50$, obtenemos los valores de predicción $\hat{y} = 6.5$ y $\hat{y} = 49$. La línea de regresión de la muestra en la figura 11.1 se dibuja al conectar estos dos puntos con una línea recta.

Con el uso de la línea de regresión del ejemplo 11.1, predecimos una reducción de 31% en la demanda química de oxígeno cuando la reducción en sólidos totales es 30%. La reducción de 31% en la demanda química de oxígeno se puede interpretar como una estimación de la media poblacional $\mu_{Y|30}$ o como una estimación de una observación nueva cuando la reducción en sólidos totales es 30%. Tales estimaciones, sin embargo, están sujetas a error. Incluso cuando el experimento se controla de modo que la reducción en sólidos totales es 30%, no es probable que midamos una reducción en la demanda química de oxígeno exactamente igual a 31%. De hecho, los datos originales que se registran en la tabla 11.1 muestran que las mediciones de 25% y 35% se registraron para la reducción en la demanda de oxígeno cuando la reducción en sólidos totales se mantuvo a 30%. ■

Ejercicios

1. Se llevó a cabo un estudio en el Instituto Politécnico y Universidad Estatal de Virginia para determinar si ciertas medidas de resistencia estática del brazo tienen alguna influencia en las características de "elevación dinámica" de un individuo. Se sometieron a pruebas de resistencia a 25 individuos y después se les pidió realizar una prueba de levantamiento de pesas en la que el peso se debía levantar en forma dinámica por arriba de la cabeza. Los datos son los siguientes:

Individuo	Resistencia del brazo, x	Levantamiento dinámico, y	Individuo	Resistencia del brazo, x	Levantamiento dinámico, y
1	17.3	71.7	14	29.6	78.3
2	19.3	48.3	15	29.9	60.0
3	19.5	88.3	16	29.9	71.7
4	19.7	75.0	17	30.3	85.0
5	22.9	91.7	18	31.3	85.0
6	23.1	100.0	19	36.0	88.3
7	26.4	73.3	20	39.5	100.0
8	26.8	65.0	21	40.4	100.0
9	27.6	75.0	22	44.3	100.0
10	28.1	88.3	23	44.6	91.7
11	28.2	68.3	24	50.4	100.0
12	28.7	96.7	25	55.9	71.7
13	29.0	76.7			

(a) Estime α y β para la curva de regresión lineal $\mu_{Y|x} = \alpha + \beta x$.

(b) Encuentre una estimación puntual de $\mu_{Y|30}$.

2. Las calificaciones de un grupo de nueve estudiantes en un reporte del examen parcial de mitad del trimestre (x) y del examen final (y) son las siguientes:

x	77	50	71	72	81	94	96	99	67
y	82	66	78	34	47	85	99	99	68

(a) Estime la línea de regresión lineal.

(b) Estime la calificación del examen final de un estudiante que obtiene 85 en el reporte del examen parcial de mitad del trimestre.

3. Se realiza un estudio sobre la cantidad de azúcar transformada en cierto proceso a varias temperaturas. Los datos se recolectan y se registran como sigue:

Temperatura, x Azúcar transformada, y

1.0	8.1
1.1	7.8
1.2	8.5
1.3	9.8
1.4	9.5
1.5	8.9
1.6	8.6
1.7	10.2
1.8	9.3
1.9	9.2
2.0	10.5

(a) Estime la línea de regresión lineal.

(b) Estime la cantidad media de azúcar transformada que se produce cuando la temperatura codificada es 1.75.

4. En cierto tipo de espécimen metálico de prueba, se sabe que la resistencia normal sobre un espécimen está relacionada funcionalmente con la resistencia de corte. El siguiente es un conjunto de datos experimentales codificados sobre las dos variables:

Resistencia normal, x Resistencia de corte, y

26.8	26.5
25.4	27.3
28.9	24.2
23.6	27.1
27.7	23.6
23.9	25.9
24.7	26.3
28.1	22.5
26.9	21.7
27.4	21.4
22.6	25.8
25.6	24.9

(a) Estime la línea de regresión $\mu_{y|x} = a + bx$.
 (b) Estime la resistencia de corte para una resistencia normal de 24.5 kilogramos por centímetro cuadrado.

5. Las cantidades de un compuesto químico y , que se disuelve en 100 gramos de agua a varias temperaturas, x , se registran como sigue:

x (°C)	y (gramos)		
0	8	6	8
15	12	10	14
30	25	21	24
45	31	33	28
60	44	39	42
75	48	51	44

(a) Encuentre la ecuación de la línea de regresión.
 (b) Grafique la línea en un diagrama de dispersión.
 (c) Estime la cantidad de compuesto químico que se disolverá en 100 gramos de agua a 50°C.

6. A todos los estudiantes de primer año en un pequeño colegio se les aplica un examen de clasificación en matemáticas. Al estudiante que obtiene una calificación por abajo de 35 se le niega la admisión al curso regular de matemáticas y se le coloca en un grupo de regularización. Las calificaciones del examen de clasificación y las calificaciones finales de 20 estudiantes que toman el curso regular se registran a continuación:

Examen de Clasificación del curso Examen de Clasificación del curso

50	53	90	54
35	41	80	91
35	61	60	48
40	56	60	71
55	68	60	71
65	36	40	47
35	11	55	53
60	70	50	68
90	79	65	57
35	59	50	79

(a) Grafique un diagrama de dispersión.
 (b) Encuentre la ecuación de la línea de regresión para predecir las calificaciones del curso a partir de las calificaciones en el examen de clasificación.
 (c) Grafique la línea en el diagrama de dispersión.
 (d) Si 60 es la calificación mínima para aprobar, ¿por debajo de qué calificación del examen de clasificación se debe negar en el futuro la admisión a este curso?

7. Un comerciante al menudeo lleva a cabo un estudio para determinar la relación entre los gastos semanales de publicidad y las ventas. Se registran los siguientes datos:

Costos de publicidad (\$)	Ventas (\$)
40	385
20	400
25	395
20	365
30	475
50	440
40	490
20	420
50	560
40	525
25	480
50	510

- (a) Grafique un diagrama de dispersión.
 (b) Encuentre la ecuación de la línea de regresión para predecir las ventas semanales a partir de los gastos de publicidad.
 (c) Estime las ventas semanales cuando los costos de publicidad son \$35.

8. Los siguientes datos se reunieron para determinar la relación entre presión y la correspondiente lectura en la escala con el propósito de calibración.

Presión, x (lb/pulgada cuadrada)	Lectura en la escala, y
10	13
10	18
10	16
10	15
10	20
50	86
50	90
50	88
50	88
50	92

- (a) Encuentre la ecuación de la línea de regresión.
 (b) El propósito de la calibración en esta aplicación es estimar la presión a partir de una lectura en la es-

cala observada. Estime la presión para una lectura en la escala de 54 con el uso de $\hat{x} = (54 - a)/b$.

9. Un estudio de la cantidad de precipitación pluvial y la cantidad de contaminación eliminada del aire produce los datos siguientes:

Precipitación pluvial diaria, x (0.01 cm)	Partículas eliminadas, y (mcg/m cúbico)
4.3	126
4.5	121
5.9	116
5.6	118
6.1	114
5.2	118
3.8	132
2.1	141
7.5	108

- (a) Encuentre la ecuación de la línea de regresión para predecir las partículas que se eliminan a partir de la cantidad de precipitación pluvial diaria.
 (b) Estime la cantidad de partículas eliminadas cuando la precipitación pluvial diaria es $x = 4.8$ unidades.
 10. Los siguientes datos son los precios de venta z de cierta marca y modelo de automóvil usado de w años:

w (años)	z (dólares)
1	6350
2	5695
2	5750
3	5395
5	4985
5	4895

(a) Ajuste una curva de la forma $\mu_{z|w} = \gamma\delta^w$ por medio de la ecuación de regresión no lineal de la muestra $\hat{z} = c\delta^w$. SUGERENCIA: escriba

$$\ln \hat{z} = \ln c + (\ln \delta)w \\ = a + bw.$$

11.3 Propiedades de los estimadores de mínimos cuadrados

Además de las suposiciones de que el término de error en el modelo

$$Y_i = \alpha + \beta x_i + E_i$$

es una variable aleatoria con media cero, suponga que hacemos la suposición adicional de que cada E_i tiene la misma varianza σ^2 y que E_1, E_2, \dots, E_n son independientes.

dientes entre las corridas en el experimento. Con estas suposiciones sobre las E_i tenemos un procedimiento para encontrar las medias y varianzas para los estimadores de α y β .

Es importante recordar que nuestros valores de a y b , que se basan en una muestra dada de n observaciones, son sólo estimaciones de los parámetros reales α y β . Si el experimento se repite varias veces, cada vez con el uso de los mismos valores fijos de x , las estimaciones de α y β que resultan muy probablemente diferirán de un experimento a otro. Estas estimaciones diferentes se pueden ver como valores que toman las variables aleatorias A y B .

Como los valores de x permanecen fijos, los valores de A y B dependen de las variaciones en los valores de y o, de forma más precisa, de los valores de las variables aleatorias Y_1, Y_2, \dots, Y_n . Las suposiciones de distribución sobre las E_i implican que las $Y_i, i = 1, 2, \dots, n$, también se distribuyen de forma independiente, con media $\mu_{Y_i} = \alpha + \beta x_i$ y varianzas σ^2 ; iguales; es decir,

$$\sigma_{Y_i}^2 = \sigma^2 \quad \text{para } i = 1, 2, \dots, n.$$

Media y varianza de estimadores

En lo que sigue mostramos que el estimador B es insesgado para β y demostramos las varianzas de A y B . Esto iniciará una serie de desarrollos que conducen a la prueba de hipótesis y a la estimación del intervalo de confianza sobre la intersección y la pendiente.

Como el estimador

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

es de la forma $\sum_{i=1}^n a_i Y_i$, donde

$$a_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad i = 1, 2, \dots, n,$$

podemos concluir del corolario 2 del teorema 4.7 que

$$\mu_B = E(B) = \frac{\sum_{i=1}^n (x_i - \bar{x})E(Y_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(\alpha + \beta x_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta,$$

y después, con el corolario 3 del teorema 4.10,

$$\sigma_B^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma_{Y_i}^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Se puede mostrar (ejercicio 1 de la página 375) que la variable aleatoria A tiene la media

$$\mu_A = \alpha \quad \text{y varianza} \quad \sigma_A^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2.$$

De los resultados anteriores, es evidente que los estimadores de mínimos cuadrados para α y β son estimadores insesgados.

Partición de variabilidad total

Para ser capaces de extraer inferencias sobre α y β es necesario llegar a una estimación del parámetro σ^2 que aparece en las dos fórmulas anteriores de la varianza para A y B . El parámetro σ^2 , varianza del error del modelo, refleja la variación aleatoria o la variación del error experimental alrededor de la línea de regresión. En gran parte de lo que sigue será ventajoso utilizar la notación

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2, \quad S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2, \quad S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

Ahora podemos escribir la suma de cuadrados del error como sigue:

$$\begin{aligned} SSE &= \sum_{i=1}^n (y_i - a - bx_i)^2 = \sum_{i=1}^n [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2b \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + b^2 \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= S_{yy} - 2bS_{xy} + b^2S_{xx} = S_{yy} - bS_{xy}, \end{aligned}$$

el paso final se sigue del hecho que $b = S_{xy}/S_{xx}$.

Teorema 11.1

Un estimador insesgado de σ^2 es

$$s^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2} = \frac{S_{yy} - bS_{xy}}{n-2}.$$

La demostración del teorema 11.1 se deja como ejercicio (véase el ejercicio de repaso 7).

11.4 Inferencias acerca de los coeficientes de regresión

Aparte de sólo estimar la relación lineal entre x y Y con propósitos de predicción, el experimentador se puede interesar en extraer ciertas inferencias acerca de la pendiente y la intersección. A fin de realizar la prueba de hipótesis y la construcción de

intervalos de confianza para α y β , se debe estar dispuesto a hacer la suposición adicional de que cada E_i , $i = 1, 2, \dots, n$, se distribuye de forma normal. Esta suposición implica que Y_1, Y_2, \dots, Y_n también se distribuyen normalmente, cada una con distribución de probabilidad $n(y_i; \alpha + \beta x_i, \sigma)$. Como A y B son funciones lineales de variables normales independientes, podemos deducir ahora del teorema 7.11 que A y B se distribuyen normalmente con distribuciones de probabilidad $n(\alpha; \alpha, \sigma_\alpha)$ y $n(\beta; \beta, \sigma_\beta)$, respectivamente.

Sucede que bajo la suposición de normalidad, un resultado muy análogo al dado en el teorema 8.4 nos permite concluir que $(n - 2)S^2/\sigma^2$ es una variable ji cuadrada con $n - 2$ grados de libertad, independiente de la variable aleatoria B . El teorema 8.5 nos asegura entonces que la estadística

$$T = \frac{(B - \beta)/(\sigma/\sqrt{S_{xx}})}{S/\sigma} = \frac{B - \beta}{S/\sqrt{S_{xx}}}$$

tiene una distribución t con $n - 2$ grados de libertad. La estadística T se puede usar para construir un intervalo de confianza de $(1 - \alpha)$ 100% para el coeficiente β .

Intervalo de confianza para β

Un intervalo de confianza de $(1 - \alpha)$ 100% para el parámetro β en la línea de regresión $\mu_{Y|x} = \alpha + \beta x$ es

$$b - \frac{t_{\alpha/2} s}{\sqrt{S_{xx}}} < \beta < b + \frac{t_{\alpha/2} s}{\sqrt{S_{xx}}}$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Ejemplo 11.2 Encuentre un intervalo de confianza de 95% para β en la línea de regresión $\mu_{Y|x} = \alpha + \beta x$, con base en los datos de contaminación de la tabla 11.1.

SOLUCIÓN

En el ejemplo 11.1 encontramos que

$$S_{xx} = 4152.18; \quad S_{yy} = 3752.09.$$

Además, encontramos que $S_{xy} = 3713.88$. Recuérdese que $b = 0.903643$. De aquí que,

$$s^2 = \frac{S_{yy} - bS_{xy}}{n - 2} = \frac{3713.88 - (0.903643)(3752.09)}{31} = 10.4299.$$

Por tanto, al sacar la raíz cuadrada obtenemos $s = 3.2295$. Con el uso de la tabla A.4, encontramos $t_{0.025} = 2.045$ para 31 grados de libertad. Por tanto, un intervalo de confianza de 95% para β es

$$0.903643 - \frac{(2.045)(3.2295)}{\sqrt{4152.18}} < \beta < 0.903643 + \frac{(2.045)(3.2295)}{\sqrt{4152.18}},$$

lo que se simplifica a

$$0.8011 < \beta < 1.0061. \quad \blacksquare$$

Prueba de hipótesis sobre la pendiente

Para probar la hipótesis nula H_0 de que $\beta = \beta_0$ contra una alternativa apropiada, de nuevo utilizamos la distribución t con $n - 2$ grados de libertad para establecer una región crítica y después basar nuestra decisión en el valor de

$$t = \frac{b - \beta_0}{s / \sqrt{S_{xx}}}$$

El método se ilustra con el siguiente ejemplo.

Ejemplo 11.3 Con el uso del valor estimado $b = 0.903643$ del ejemplo 11.1, pruebe la hipótesis de que $\beta = 1.0$ contra la alternativa de que $\beta < 1.0$.

SOLUCIÓN

$$H_0: \beta = 1.0,$$

$$H_1: \beta < 1.0,$$

$$t = \frac{0.903643 - 1.0}{3.2295 / \sqrt{4152.18}} = -1.92,$$

con $n - 2 = 31$ grados de libertad ($P = 0.03$).

Decisión: el valor t es significativo en el nivel 0.03, lo que sugiere fuerte evidencia de que $\beta < 1.0$.

Una prueba t importante sobre la pendiente es la prueba de la hipótesis

$$H_0: \beta = 0,$$

$$H_1: \beta \neq 0.$$

Cuando la hipótesis nula no se rechaza la conclusión es que no hay una relación lineal significativa entre $E(y)$ y la variable independiente x . La gráfica de los datos para el ejemplo 11.1 sugiere que existe una relación lineal. Sin embargo, en algunas aplicaciones en las que σ^2 es grande y por ello se presenta considerable "ruido" en los datos, una gráfica, aunque útil, puede no producir información clara para el investigador. El rechazo de la H_0 anterior implica que existe una regresión lineal significativa. Esta hipótesis se retoma en la sección 11.7 donde presentamos el uso del análisis de varianza en el contexto de la regresión lineal simple.

La figura 11.3 despliega un impreso del MINITAB que muestra la prueba t para

$$H_0: \beta = 0,$$

$$H_1: \beta \neq 0$$

para los datos del ejemplo 11.1. Nótese el coeficiente de regresión, el error estándar, el valor t y el valor P . Se rechaza la hipótesis nula. Claramente hay una relación lineal significativa entre la demanda química media de oxígeno y la reducción de sólidos. Nótese que la estadística t se calcula como

$$t = \frac{\text{coeficiente}}{\text{error estándar}} = \frac{b}{s/\sqrt{S_{xx}}} \quad \blacksquare$$

Regression Analysis

The regression equation is
 COD = 3.83 + 0.904 Per_Red

Predictor	Coef	StDev	T	P
Constant	3.830	1.768	2.17	0.038
Per_Red	0.90364	0.05012	18.03	0.000

S = 3.230 R-Sq = 91.3% R-Sq(adj) = 91.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	3390.6	3390.6	325.08	0.000
Error	31	323.3	10.4		
Total	32	3713.9			

Figura 11.3 Imprese del MINITAB para la prueba t para los datos del ejemplo 11.1

Inferencia estadística de la intersección

Los intervalos de confianza y la prueba de hipótesis sobre el coeficiente α se pueden establecer a partir del hecho de que A también se distribuye normalmente. No es difícil mostrar que

$$T = \frac{A - \alpha}{S \sqrt{\sum_{i=1}^n x_i^2 / n S_{xx}}}$$

tiene una distribución t con $n - 2$ grados de libertad de lo que podemos construir un intervalo de confianza de $(1 - \alpha)$ 100% para α .

Un intervalo de confianza de $(1 - \alpha)$ 100% para el parámetro α en la línea de regresión $\mu_{Y|x} = \alpha + \beta x$ es

$$a - \frac{t_{\alpha/2} s \sqrt{\sum_{i=1}^n x_i^2}}{\sqrt{n S_{xx}}} < \alpha < a + \frac{t_{\alpha/2} s \sqrt{\sum_{i=1}^n x_i^2}}{\sqrt{n S_{xx}}}$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Nótese que el símbolo α se utiliza aquí de dos formas sin relación alguna, primero como el nivel de significancia y después como la intersección de la línea de regresión.

Ejemplo 11.4 Encuentre un intervalo de confianza de 95% para α en la línea de regresión $\mu_{Y|x} = \alpha + \beta x$, con base en los datos de la tabla 11.1.

SOLUCIÓN

En los ejemplos 11.1 y 11.2 encontramos que

$$S_{xx} = 4152.18 \quad \text{y} \quad s = 3.2295.$$

Del ejemplo 11.1 tenemos que

$$\sum_{i=1}^n x_i^2 = 41,086 \quad \text{y} \quad a = 3.829633.$$

Con el uso de la tabla A.4, encontramos $t_{0.025} = 2.045$ para 31 grados de libertad. Por tanto, un intervalo de confianza de 95% para α es

$$3.829633 - \frac{(2.045)(3.2295)\sqrt{41,086}}{\sqrt{(33)(4152.18)}} < \alpha < 3.829633 + \frac{(2.045)(3.2295)\sqrt{41,086}}{\sqrt{(33)(4152.18)}}$$

que se simplifica a $0.2131 < \alpha < 7.4461$.

Para probar la hipótesis nula H_0 de que $\alpha = \alpha_0$ contra una alternativa adecuada podemos utilizar la distribución t con $n - 2$ grados de libertad para establecer una región crítica y después basar nuestra decisión en el valor de

$$t = \frac{a - \alpha_0}{s \sqrt{\sum_{i=1}^n x_i^2 / n S_{xx}}}$$

Ejemplo 11.5 Con el uso del valor estimado $\alpha = 3.829640$ del ejemplo 11.1, pruebe la hipótesis de que $\alpha = 0$ en el nivel de significancia de 0.05 contra la alternativa de que $\alpha \neq 0$.

SOLUCIÓN

$$H_0: \alpha = 0,$$

$$H_1: \alpha \neq 0.$$

$$t = \frac{3.829633 - 0}{3.2295 \sqrt{41,086 / (33)(4152.18)}} = 2.17$$

con 31 grados de libertad. De esta manera $P = 0.02$ y concluimos que $\alpha \neq 0$. ■

11.5 Predicción

Existen varias razones para construir una regresión lineal. Una, por supuesto, es predecir valores de respuesta a uno o más valores de la variable independiente. En esta sección nos enfocamos en los errores asociados con la predicción.

La ecuación $\hat{y} = a + bx$ se puede usar para predecir o estimar la **respuesta media** $\mu_{Y|x_0}$ en $x = x_0$, donde x_0 no es necesariamente uno de los valores preseleccionados, o se puede utilizar para predecir un solo valor y_0 de la variable Y_0 cuando $x = x_0$. Esperaríamos que el error de predicción fuese más alto en el caso de un so-

lo valor predicho que en el caso donde se predice una media. Esto, entonces, afectará el ancho de nuestros intervalos para los valores que se predicen.

Suponga que el experimentador desea construir un intervalo de confianza para $\mu_{Y|x_0}$. Utilizaremos el estimador puntual $\hat{Y}_0 = A + Bx_0$ para estimar $\mu_{Y|x_0} = \alpha + \beta x_0$. Se puede mostrar que la distribución muestral de \hat{Y}_0 es normal con media

$$\mu_{\hat{Y}_0} = E(\hat{Y}_0) = E(A + Bx_0) = \alpha + \beta x_0 = \mu_{Y|x_0},$$

y varianza

$$\sigma_{\hat{Y}_0}^2 = \sigma_{A+Bx_0}^2 = \sigma_{\hat{Y}, B(x_0-x)}^2 = \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right],$$

la última se sigue del hecho que $\text{Cov}(\hat{Y}, B) = 0$ (véase el ejercicio 2 de la página 375). De esta forma el intervalo de confianza de $(1 - \alpha) 100\%$ sobre la respuesta media $\mu_{Y|x_0}$ se puede construir a partir de la estadística

$$T = \frac{\hat{Y}_0 - \mu_{Y|x_0}}{S \sqrt{(1/n) + [(x_0 - \bar{x})^2/S_{xx}]}}$$

que tiene una distribución t con $n - 2$ grados de libertad.

Intervalo de confianza para $\mu_{Y|x_0}$

Un intervalo de confianza de $(1 - \alpha) 100\%$ para la respuesta media $\mu_{Y|x_0}$ es

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} < \mu_{Y|x_0} < \hat{y}_0 + t_{\alpha/2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}},$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Ejemplo 11.6 Con el uso de los datos de la tabla 11.1, construya los límites de confianza de 95% para la respuesta media $\mu_{Y|20}$.

SOLUCIÓN

De la ecuación de regresión encontramos para $x_0 = 20$, digamos,

$$\hat{y}_0 = 3.829633 + (0.903643)(20) = 21.9025.$$

Además, $\bar{x} = 33.4545$, $S_{xx} = 4152.18$, $s = 3.2295$, y $t_{0.025} = 2.045$ para 31 grados de libertad. Por tanto, un intervalo de confianza de 95% para $\mu_{Y|20}$ es

$$\begin{aligned} 21.9025 - (2.045)(3.2295) \sqrt{\frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} &< \mu_{Y|20} \\ &< 21.9025 + (2.045)(3.2295) \sqrt{\frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} \end{aligned}$$

o simplemente $20.1071 < \mu_{Y|20} < 23.6979$.

Al repetir los cálculos anteriores para cada uno de varios valores diferentes de x_0 se pueden obtener los límites de confianza correspondientes sobre cada $\mu_{Y|x_0}$. La

figura 11.4 muestra los puntos de datos, la línea de regresión estimada y los límites de confianza superior e inferior sobre la media de $Y|x$.

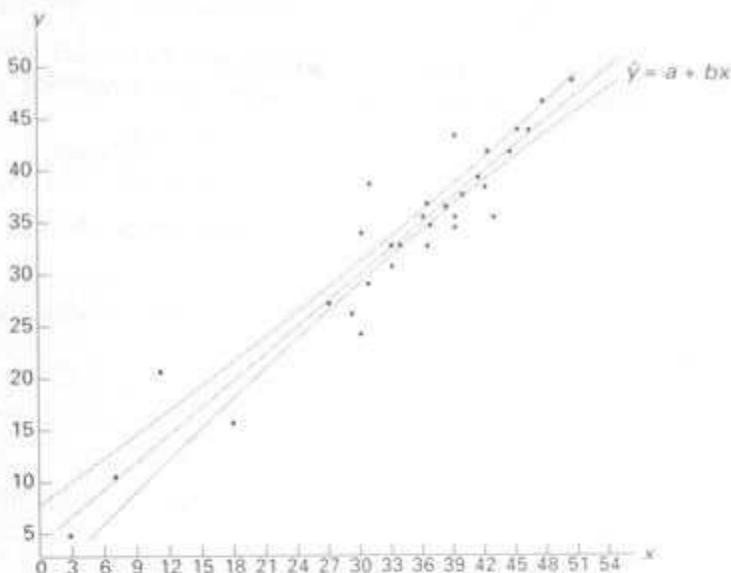


Figura 11.4 Límites de confianza para el valor medio de $Y|x$.

En el ejemplo 11.6 tenemos una confianza de 95% de que la demanda química de oxígeno de la población está entre 20,1071 y 23,6979%. ■

Intervalo de predicción

Otro tipo de intervalo que a menudo se mal interpreta y se confunde con el dado para $\mu_{Y|x}$ es el intervalo de predicción para una respuesta futura observada. Realmente, en muchos casos el intervalo de predicción es más relevante para el científico o para el ingeniero que el intervalo de confianza sobre la media. En el ejemplo del contenido de alquitrán-temperatura de entrada, que se cita en la sección 11.1, ciertamente habría interés no sólo en estimar el contenido medio de alquitrán a una temperatura específica sino también en construir un intervalo que refleje el error al predecir una cantidad futura observada de contenido de alquitrán a la temperatura dada.

Para obtener un **intervalo de predicción** para cualquier valor único y_0 de la variable Y_0 es necesario estimar la varianza de las diferencias entre las ordenadas \hat{y}_0 que se obtienen de las líneas de regresión calculadas en un muestreo repetido cuando $x = y_0$, y la ordenada real correspondiente y_0 . Podemos pensar en la diferencia $\hat{y}_0 - y_0$ como un valor de la variable aleatoria $\hat{Y}_0 - Y_0$, cuya distribución muestral se puede mostrar que es normal con media

$$\mu_{\hat{Y}_0 - Y_0} = E(\hat{Y}_0 - Y_0) = E[A + Bx_0 - (\alpha + \beta x_0 + E_0)] = 0,$$

y varianza

$$\sigma_{\hat{Y}_i - Y_i}^2 = \sigma_{A + Bx_i - Y_i}^2 = \sigma_{Y_i + B(x_i - \bar{x}) - Y_i}^2 = \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right].$$

De esta manera el intervalo de predicción de $(1 - \alpha)$ 100% para un solo valor pronosticado y_0 se puede construir a partir de la estadística

$$T = \frac{\hat{Y}_0 - Y_0}{S \sqrt{1 + (1/n) + [(x_0 - \bar{x})^2/S_{xx}]}}$$

que tiene una distribución t con $n - 2$ grados de libertad.

Intervalo de predicción para y_0

Un intervalo de predicción de $(1 - \alpha)$ 100% para una sola respuesta y_0 está dado por

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} < y_0 < \hat{y}_0 + t_{\alpha/2} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - 2$ grados de libertad.

Es claro que existe una distinción entre el concepto de un intervalo de confianza y el intervalo de predicción que se describe arriba. La interpretación del intervalo de confianza es idéntica a la que se describe para todos los intervalos de confianza sobre los parámetros de población que se presentan a lo largo del libro. En realidad, $\mu_{Y|x_0}$ es un parámetro poblacional. El intervalo de predicción calculado, sin embargo, representa un intervalo que tiene una probabilidad igual a $1 - \alpha$ de contener no un parámetro sino un valor futuro y_0 de la variable aleatoria Y_0 .

Ejemplo 11.7 Utilice los datos de la tabla 11.1 para construir un intervalo de predicción de 95% para y_0 cuando $x_0 = 20\%$.

SOLUCIÓN

Tenemos $n = 33$, $x_0 = 20$, $\bar{x} = 33.4545$, $\hat{y}_0 = 21.9025$, $S_{xx} = 4152.18$, $s = 3.2295$, y $t_{0.025} = 2.045$ para 31 grados de libertad. Por tanto, un intervalo de predicción de 95% para y_0 es

$$\begin{aligned} 21.9025 - (2.045)(3.2295) \sqrt{1 + \frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} < y_0 \\ < 21.9025 + (2.045)(3.2295) \sqrt{1 + \frac{1}{33} + \frac{(20 - 33.4545)^2}{4152.18}} \end{aligned}$$

que se simplifica a $15.0585 < y_0 < 28.7464$.

La figura 11.5 muestra otra gráfica de los datos de la demanda química de oxígeno, con el intervalo de confianza sobre la respuesta media y el intervalo de predicción sobre una respuesta individual graficados. La gráfica refleja un intervalo mucho más estrecho alrededor de la línea de regresión en el caso de la respuesta media. ■

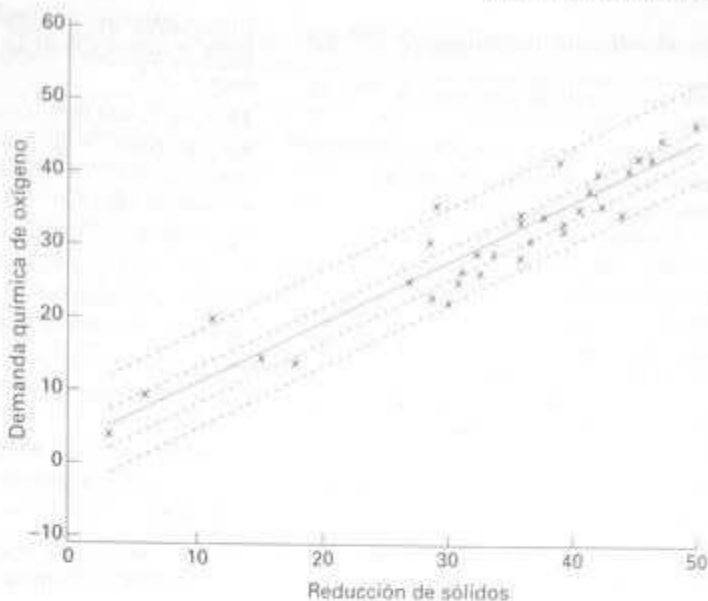


Figura 11.5 Intervalos de confianza y de predicción para los datos de demanda química de oxígeno.

Ejercicios

1. Suponga que las E_i son normales, independientes con medias cero y varianzas común σ^2 , muestre que A , el estimador de mínimos cuadrados de α en $\mu_{Y_i} = \alpha + \beta x_i$, se distribuye normalmente con media α y varianzas

$$\sigma_A^2 = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \sigma^2.$$

2. Para un modelo de regresión lineal simple

$$Y_i = \alpha + \beta x_i + E_i, \quad i = 1, 2, \dots, n,$$

donde las E_i son independientes y se distribuyen de forma normal con medias cero y varianzas σ^2 , iguales, muestre que \bar{Y} y

$$B = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

tienen covarianza cero.

3. Con referencia al ejercicio 1 de la página 363,

(a) evalúe s^2 ;

(b) pruebe la hipótesis de que $\beta = 0$ contra la alternativa de que $\beta \neq 0$ en el nivel de significancia 0.05 e interprete la decisión que resulta.

4. Con referencia al ejercicio 2 de la página 364,

(a) evalúe s^2 ;

(b) construya un intervalo de confianza de 95% para α ;

(c) construya un intervalo de confianza de 95% para β .

5. Con referencia al ejercicio 3 de la página 364,

(a) evalúe s^2 ;

(b) construya un intervalo de confianza de 95% para α ;

(c) construya un intervalo de confianza de 95% para β .

6. Con referencia al ejercicio 4 de la página 364,

(a) evalúe s^2 ;

(b) construya un intervalo de confianza de 99% para α ;

(c) construya un intervalo de confianza de 99% para β .

7. Con referencia al ejercicio 5 de la página 364,

(a) evalúe s^2 ;

(b) construya un intervalo de confianza de 99% para α .

(c) construya un intervalo de confianza de 99% para β .

8. Pruebe la hipótesis de que $\alpha = 10$ en el ejercicio 6 de la página 364 contra la alternativa de que $\alpha < 10$. Utilice un nivel de significancia de 0.05.

9. Pruebe la hipótesis de que $\beta = 6$ en el ejercicio 7 de la página 364 contra la alternativa de que $\beta < 6$. Utilice un nivel de significancia de 0.025.

10. Con el valor de s^2 que se encontró en el ejercicio 4(a), construya un intervalo de confianza de 95% para $\mu_{Y|30}$ en el ejercicio 2 de la página 364.

11. Con referencia al ejercicio 4 de la página 364, use el valor de s^2 que se encontró en el ejercicio 6(a) para calcular

(a) un intervalo de confianza de 95% para la resistencia de ruptura media cuando $x = 24.5$;

15. Considere la regresión del rendimiento de ciertos automóviles, medido en millas por galón (mpg) sobre su peso en libras (wt). Los datos son del *Consumer Reports* (abril de 1997). Parte de los resultados del SAS para el procedimiento se muestra en el siguiente impreso.

(b) un intervalo de predicción de 95% para un solo valor pronosticado de la resistencia de ruptura cuando $x = 24.5$.

12. Con el uso de s^2 que se encontró en el ejercicio 5(a), grafique la línea de regresión y las bandas de confianza del 95% para la respuesta media $\mu_{Y|3}$ para el dato del ejercicio 3 de la página 364.

13. Con el uso del valor de s^2 que se encontró en el ejercicio 5(a), construya un intervalo de confianza de 95% para la cantidad de azúcar transformada que corresponde $x = 1.6$ en el ejercicio 3 de la página 364.

14. Con referencia al ejercicio 5 de la página 364, use el valor de s^2 que se encuentra en el ejercicio 7(a) para calcular

(a) un intervalo de confianza de 99% para la cantidad promedio de químico que se disolverá en 100 gramos de agua a 50°C;

Root MSE	1.68743	R-square	0.9852
Dep Mean	23.79412	Adj R-sq	0.9023
C.V.	7.09181		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter=0	Prob > T
INTERCEP	1	45.901182	1.35291776	34.667	0.0001
WT	1	-0.007430	0.00042493	-17.484	0.0001

MODEL	WT	Dep Var MPG	Predict Value	Std Err Predict	95% Mean		95% Predict		Residual
					Lower	Upper	Lower	Upper	
GMV Safari	4520	15.0000	13.3195	0.665	11.9643	14.6747	9.6248	17.0142	1.4805
Gen Metro	3065	29.0000	31.5591	0.530	30.4794	32.6389	27.9563	35.1819	-2.0591
Honda Civic	2440	31.0000	28.7730	0.406	27.9460	29.6000	25.2377	32.3083	2.2270
Hyundai Accent	2290	28.0000	29.8875	0.493	28.9647	30.8102	26.3286	33.4463	-1.8875
Infiniti I30	3195	23.0000	23.1637	0.292	22.5697	23.7577	19.6756	26.6518	-0.1637
Isuzu Oasis	3480	21.0000	21.0463	0.329	20.3755	21.7171	17.5443	24.5463	-0.0463
Jeep Grand Char	4090	15.0000	16.5142	0.507	15.4814	17.5471	12.9252	20.1032	-1.5142
Land Rover Disc	4635	13.0000	13.2081	0.671	11.9411	14.6750	9.5090	16.9071	-0.2081
Lexus ES300	3390	22.0000	21.7149	0.333	21.0776	22.3522	16.2192	25.2107	0.2851
Lincoln Cont	3930	18.0000	17.7030	0.453	16.7804	18.6255	14.1441	21.2618	0.2710

(a) Estime el rendimiento para un vehículo que pesa 4000 libras.

(b) Suponga que los ingenieros de Honda afirman que, en promedio, el modelo Civic (o cualquier otro modelo de auto que pese 2440 libras) rinde más de 30 mpg. ¿Con base en los resultados del análisis de regresión creería en la afirmación? ¿Por qué sí o por qué no?

(c) Los ingenieros de diseño del Lexus ES300 fijan como objetivo 18 mpg como ideal para su modelo (o para cualquier otro modelo que pese 3390 libras) aunque se espera que se experimente alguna variación. ¿Es probable que este valor meta sea realista? Discuta.

11.6 Elección de un modelo de regresión

Gran parte de lo presentado hasta ahora sobre la regresión que involucra una sola variable independiente depende de la suposición de que el modelo que se elige es correcto, la presunción de que $\mu_{y|x}$ se relaciona con x linealmente en los parámetros. En realidad, no esperaríamos que la predicción de la respuesta fuera buena si hay varias variables independientes, que no se consideran en el modelo, y que afectan la respuesta y varían en el sistema. Además, la predicción sería en realidad inadecuada si la verdadera estructura que relaciona $\mu_{y|x}$ con x es en extremo no lineal en el rango de las variables que se consideran.

A menudo se utiliza el modelo de regresión lineal más simple aunque se sabe que el modelo es algo diferente del lineal o que la estructura real se desconoce. Este enfoque por lo general es sólido, en particular cuando el rango de x es pequeño. De esta manera el modelo que se usa se convierte en una función de aproximación que se espera sea una representación adecuada de la imagen real en la región de interés. Se debe notar, sin embargo, el efecto de un modelo inadecuado sobre los resultados presentados hasta el momento. Por ejemplo, si el modelo real, desconocido para el experimentador, es lineal en más de una x , digamos,

$$\mu_{y|x_1, x_2} = \alpha + \beta x_1 + \gamma x_2,$$

entonces la estimación ordinaria de mínimos cuadrados $b = S_{yy}/S_{xx}$, calculada al considerar sólo x_1 en el experimento, es, en circunstancias generales, una estimación sesgada del coeficiente β , donde el sesgo es una función del coeficiente adicional γ (véase el ejercicio 3 de la página 385). Además, la estimación s^2 para σ^2 está sesgada debido a la variable adicional.

11.7 Procedimiento del análisis de varianza

Con frecuencia el problema de analizar la calidad de la línea de regresión estimada se maneja mediante un enfoque de **análisis de varianza**: un procedimiento por el que la variación total en la variable dependiente se subdivide en componentes significativos que se observan y tratan de forma sistemática. El análisis de varianza, que se expone en el capítulo 13, es un recurso poderoso que se utiliza en muchas aplicaciones.

Suponga que tenemos n puntos de datos experimentales en la forma acostumbrada (x_i, y_i) y que se estima la línea de regresión. En nuestra estimación de σ^2 en la sección 11.3 establecimos la identidad

$$S_{yy} = bS_{xy} + SSE$$

Una formulación alternativa y quizá más informativa es

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Por lo que logramos una partición de la **suma total corregida de cuadrados de y** en dos componentes que deben reflejar un significado particular para el experimentador. Indicaremos esta partición de manera simbólica como

$$SST = SSR + SSE.$$

El primer componente de la derecha, SSR , se llama **suma de cuadrados de regresión** y refleja la cantidad de variación en los valores **y explicados por el modelo**, en este caso la línea recta postulada. El segundo componente es la ya familiar suma de cuadrados del error, que refleja la variación alrededor de la línea de regresión.

Suponga que nos interesa probar la hipótesis

$$H_0: \beta = 0,$$

$$H_1: \beta \neq 0,$$

donde la hipótesis nula dice en esencia que el modelo es $\mu_{Y|x} = \alpha$. Es decir, la variación en Y resulta del azar o de las fluctuaciones aleatorias que son independientes de los valores de x . Bajo las condiciones de esta hipótesis nula se puede mostrar que SSR/σ^2 y SSE/σ^2 son valores de variables ji cuadrada independientes con 1 y $n - 2$ grados de libertad, respectivamente, y entonces por el teorema 7.12 se sigue que SST/σ^2 también es un valor de una variable ji cuadrada con $n - 1$ grados de libertad. Para probar la hipótesis anterior, calculamos

$$f = \frac{SSR/1}{SSE/(n-2)} = \frac{SSR}{s^2}$$

y rechazamos H_0 al nivel de significancia α cuando $f > f_{\alpha}(1, n - 2)$.

Los cálculos por lo general se resumen por medio de una **tabla de análisis de varianza**, como se indica en la tabla 11.2. Se acostumbra hacer referencia a las diversas sumas de cuadrados divididas entre sus respectivos grados de libertad como los **cuadrados medios**.

Tabla 11.2 Análisis de varianza para probar $\beta = 0$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Regresión	SSR	1	SSR	SSR/s^2
Error	SSE	$n - 2$	$s^2 = \frac{SSE}{n - 2}$	
Total	SST	$n - 1$		

Cuando se rechaza la hipótesis nula, es decir, cuando la estadística F calculada excede el valor crítico $f_{\alpha}(1, n - 2)$, concluimos que hay una cantidad significativa de variación en la respuesta explicada por el modelo que se postula, la función de línea recta. Si la estadística F está en la región de aceptación, concluimos que los datos no reflejaron evidencia suficiente para apoyar el modelo postulado.

En la sección 11.4 se da un procedimiento con el que la estadística

$$T = \frac{B - \beta_0}{S/\sqrt{S_{xx}}}$$

se utiliza para probar la hipótesis

$$H_0: \beta = \beta_0,$$

$$H_1: \beta \neq \beta_0,$$

donde T sigue la distribución t con $n - 2$ grados de libertad. La hipótesis se rechaza si $|t| > t_{\alpha/2}$ para un nivel de significancia α . Es de interés notar que en el caso especial en el que probamos

$$H_0: \beta = 0,$$

$$H_1: \beta \neq 0,$$

el valor de nuestra estadística T se convierte en

$$t = \frac{b}{s/\sqrt{S_{xx}}},$$

y la hipótesis bajo consideración es idéntica a la que se prueba en la tabla 11.2. A saber, la hipótesis nula establece que la variación en la respuesta se debe sólo al azar. El análisis de varianza usa la distribución F en lugar de la distribución t . Para la alternativa bilateral, los dos procedimientos son idénticos. Esto lo podemos ver al escribir

$$t^2 = \frac{b^2 S_{xx}}{s^2} = \frac{b S_{xy}}{s^2} = \frac{SSR}{s^2},$$

que es idéntica al valor f que se utiliza en el análisis de varianza. La relación básica entre la distribución t con v grados de libertad y la distribución F con 1 y v grados de libertad es

$$t_{\alpha/2}^2 = f_{\alpha}(1, v)$$

Por supuesto, la prueba t permite la prueba contra una alternativa unilateral mientras que la prueba F se restringe a probar contra una alternativa bilateral.

Resultados de computadora comentados para regresión lineal simple

Considere nuevamente los datos de demanda química de oxígeno de la tabla 11.1. La figura 11.6 muestra un impreso de computadora comentado. En este caso se ilustra el software MINITAB PC. La columna razón t indica las pruebas para hipótesis nulas de valores cero sobre el parámetro. El término "ajuste" denota valores \hat{y} , a menudo llamados valores **ajustados**. El término "stdev.fit" se usa para calcular intervalos de confianza sobre la respuesta media. El elemento R^2 se calcula como $(SSR/SST) \times 100$ y significa la proporción de la variación en y explicada por la regresión de línea recta. Discutimos la estadística R^2 con más detalle en el capítulo 12. También se muestran intervalos de confianza sobre la respuesta media e intervalos de predicción sobre una nueva observación.

11.8 Prueba para la linealidad de la regresión: datos con observaciones repetidas

Para ciertas clases de situaciones experimentales el investigador tiene la capacidad de obtener observaciones repetidas sobre la respuesta para cada valor de x . Aunque no es necesario tener estas repeticiones a fin de estimar α y β , no obstante es-

The regression equation is
 $y = 3.63 + 0.904x$

Predictor	Coef	Stdev	t-ratio	P
Constant	3.630	1.768	2.07	0.038
X	0.90364	0.05032	16.03	0.000

n = 33 0 S-sq = 91.38

Analysis of Variance

SOURCE	DF	SS	MS	F	P
Regression	1	3390.4	3390.4	325.06	0.000
Error	31	323.9	10.4		
Total	32	3713.9			

Obs.	X	Y	Fit	Stdev. Fit	Residual
1	3.0	5.000	6.541	1.627	-1.541
2	7.0	11.000	10.155	1.440	0.845
3	11.0	21.000	13.770	1.258	7.230
4	15.0	16.000	17.384	1.082	-1.384
5	18.0	16.000	20.999	0.957	-4.999
6	22.0	28.000	28.228	0.649	-0.228
7	29.0	27.000	30.035	0.605	-3.035
8	30.0	25.000	30.939	0.588	-5.939
9	30.0	35.000	30.939	0.588	4.061
10	31.0	30.000	31.843	0.575	-1.843
11	31.0	40.000	31.843	0.575	8.157
12	32.0	32.000	32.746	0.567	-0.746
13	33.0	34.000	33.650	0.563	0.350
14	33.0	32.000	33.650	0.563	-1.650
15	34.0	34.000	34.554	0.563	-0.554
16	36.0	37.000	36.361	0.576	0.639
17	36.0	30.000	36.361	0.576	-6.361
18	36.0	34.000	36.361	0.576	-2.361
19	37.0	36.000	37.264	0.590	-1.264
20	36.0	38.000	38.168	0.607	-0.168
21	39.0	37.000	39.072	0.627	-2.072
22	39.0	36.000	39.072	0.627	-3.072
23	39.0	45.000	39.072	0.627	5.928
24	40.0	39.000	39.975	0.651	-0.975
25	43.0	41.000	40.879	0.678	-0.879
26	42.0	40.000	41.783	0.707	-1.783
27	42.0	44.000	41.783	0.707	2.217
28	43.0	37.000	42.686	0.738	-5.686
29	44.0	44.000	43.590	0.772	0.410
30	45.0	46.000	44.494	0.807	1.506
31	46.0	46.000	45.397	0.843	0.603
32	47.0	49.000	46.301	0.881	2.699
33	50.0	51.000	49.012	1.002	1.988

Fit	Stdev. Fit	95% C.I.	95% P.I.
6.541	1.627	(3.222, 9.859)	(-0.836, 13.917)
10.155	1.440	(7.217, 13.093)	(2.942, 17.369)
13.770	1.258	(11.203, 16.336)	(6.699, 20.840)
17.384	1.082	(15.176, 19.592)	(10.436, 24.333)
20.999	0.957	(18.143, 22.048)	(13.224, 28.767)
28.228	0.649	(26.905, 29.551)	(21.508, 34.948)
30.035	0.605	(28.801, 31.269)	(23.332, 36.736)
30.939	0.588	(29.739, 32.139)	(24.242, 37.636)
30.939	0.588	(29.739, 32.139)	(24.242, 37.636)
31.843	0.575	(30.669, 33.017)	(25.153, 38.535)
31.843	0.575	(30.669, 33.017)	(25.153, 38.535)
32.746	0.567	(31.590, 33.903)	(26.057, 39.435)
33.650	0.563	(32.502, 34.798)	(26.962, 40.337)
33.650	0.563	(32.502, 34.798)	(26.962, 40.337)
34.554	0.563	(33.405, 35.702)	(27.866, 41.241)
36.361	0.576	(35.185, 37.537)	(29.668, 43.053)
36.361	0.576	(35.185, 37.537)	(29.668, 43.053)
36.361	0.576	(35.185, 37.537)	(29.668, 43.053)
37.264	0.590	(36.082, 38.447)	(30.567, 43.962)
38.168	0.607	(36.931, 39.406)	(31.465, 44.872)

Figura 11.6 Impreso del MINITAB de la regresión lineal simple para los datos de la demanda química de oxígeno.

[continuación]

39.072	0.627	(37.792, 40.351)	(32.360, 45.763)
39.072	0.627	(37.792, 40.351)	(32.360, 45.763)
39.072	0.627	(37.792, 40.351)	(32.360, 45.763)
39.975	0.651	(38.646, 43.303)	(33.255, 46.696)
40.679	0.676	(39.497, 42.261)	(34.147, 47.611)
41.763	0.707	(40.341, 43.224)	(35.039, 48.527)
41.763	0.707	(40.341, 43.224)	(35.039, 48.527)
42.666	0.738	(41.180, 44.192)	(35.928, 49.444)
43.590	0.772	(42.016, 45.164)	(36.816, 50.364)
44.494	0.807	(42.848, 46.139)	(37.703, 51.284)
45.397	0.843	(43.677, 47.118)	(38.588, 52.206)
46.301	0.881	(44.503, 48.099)	(39.472, 53.130)
49.012	1.002	(46.966, 51.056)	(42.314, 55.910)

Figura 11.6 [conclusión]

to permite al experimentador obtener información cuantitativa con respecto a lo apropiado del modelo. De hecho, si se generan observaciones repetidas, el experimentador puede hacer una prueba de significancia como ayuda en la determinación de si el modelo es adecuado o no.

Seleccionemos una muestra aleatoria de n observaciones con el uso de k valores distintos de x , digamos x_1, x_2, \dots, x_k , tal que la muestra contenga n_1 valores observados de la variable aleatoria Y_1 que corresponden a x_1 , n_2 valores observados de Y_2 que corresponden a x_2 , ..., n_k valores observados de Y_k que corresponden a x_k .

Necesariamente $n = \sum_{i=1}^k n_i$. Definimos

y_{ij} = valor j -ésimo de la variable aleatoria Y_i ,

$$y_{i.} = T_{i.} = \sum_{j=1}^{n_i} y_{ij}$$

$$\bar{y}_{i.} = \frac{T_{i.}}{n_i}$$

De aquí que, si se hacen $n_4 = 3$, mediciones de Y correspondientes a $x = x_4$, indicáramos estas observaciones con y_{41} , y_{42} , y y_{43} . Entonces

$$T_{4.} = y_{41} + y_{42} + y_{43}$$

Concepto de falta de ajuste

La suma de los cuadrados del error consiste en dos partes: la cantidad debida a la variación entre los valores de Y dentro de los valores dados de x y un componente que normalmente se llama contribución de la **falta de ajuste**. El primer componente refleja la variación sólo aleatoria o **error experimental puro**, mientras que el segundo componente es una medida de la variación sistemática ocasionada por los términos de orden superior. En nuestro caso estos son términos en x diferentes de la contribución lineal o de primer orden. Nótese que al elegir un modelo lineal en la contribución lineal o de primer orden. Nótese que al elegir un modelo lineal en la contribución lineal o de primer orden no existe y por ello nuestra suma

de cuadrados del error se debe por completo a errores aleatorios. Si éste debe ser el caso, entonces $s^2 = SSE/(n - 2)$ es una estimación insesgada de σ^2 . Sin embargo, si el modelo no se ajusta de manera adecuada a los datos, entonces la suma de cuadrados del error se infla y produce una estimación sesgada de σ^2 . Aunque el modelo se ajuste a los datos o no, siempre se puede obtener una estimación insesgada de σ^2 cuando tenemos observaciones repetidas simplemente al calcular

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n_i - 1}; \quad i = 1, 2, \dots, k,$$

para cada uno de los k valores distintos de x y después combinando estas varianzas tenemos

$$s^2 = \frac{\sum_{i=1}^k (n_i - 1)s_i^2}{n - k} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{n - k}$$

El numerador de s^2 es una medida del error experimental puro. Un procedimiento de cálculo para separar la suma de cuadrados del error en los dos componentes que representan el error puro y la falta de ajuste es el siguiente:

1. Calcular la suma de cuadrados de los errores puros.

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2.$$

Esta suma de cuadrados tiene $n - k$ grados de libertad asociados y el cuadrado medio que resulta es nuestra estimación insesgada s^2 de σ^2 .

2. Restar la suma de cuadrados del error puro de la suma de cuadrados del error, SSE , y con ello obtener la suma de cuadrados debida a la falta de ajuste. Los grados de libertad para la falta de ajuste también se obtienen al restar simplemente

$$(n - 2) - (n - k) = k - 2.$$

Los cálculos que se requieren para probar hipótesis en un problema de regresión con mediciones repetidas de la respuesta se pueden resumir como se muestra en la tabla 11.3.

Tabla 11.3 Análisis de varianza para probar la linealidad de la regresión

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada
Regresión	SSR	1	SSR	$\frac{SSR}{s^2}$
Error	SSE	$n - 2$		
Falta de ajuste	$\left\{ \begin{array}{l} SSE - SSE(\text{puro}) \\ SSE(\text{puro}) \end{array} \right.$	$\left\{ \begin{array}{l} k - 2 \\ n - k \end{array} \right.$	$\frac{SSE - SSE(\text{puro})}{k - 2}$ $s^2 = \frac{SSE(\text{puro})}{n - k}$	$\frac{SSE - SSE(\text{puro})}{s^2(k - 2)}$
Total	SST	$n - 1$		

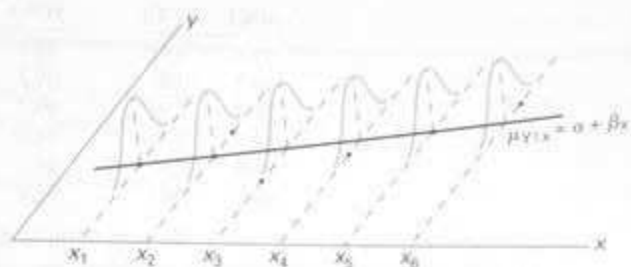


Figura 11.7 Modelo lineal correcto sin componente de falta de ajuste.

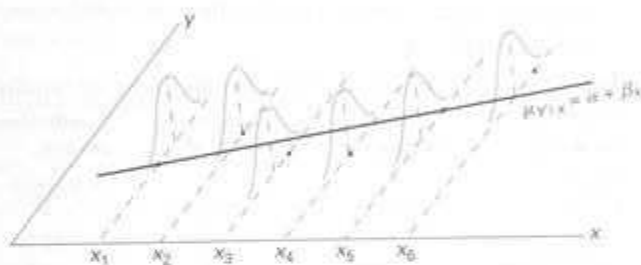


Figura 11.8 Modelo lineal incorrecto con componente de falta de ajuste.

Las figuras 11.7 y 11.8 enseñan los puntos muestrales para las situaciones del “modelo correcto” y del “modelo incorrecto”. En la figura 11.7, donde $\mu_{Y|X}$ cae en una línea recta, no hay falta de ajuste cuando se supone un modelo lineal, por lo que la variación muestral alrededor de la línea de regresión es un error puro que resulta de la variación que ocurre entre observaciones repetidas. En la figura 11.8, donde $\mu_{Y|X}$ claramente no cae en una línea recta, la falta de ajuste por elegir erróneamente un modelo lineal explica en gran parte la variación alrededor de la línea de regresión además del error puro.

El concepto de falta de ajuste es en extremo importante en aplicaciones del análisis de regresión. De hecho, la necesidad de construir o diseñar un experimento que tenga en consideración la falta de ajuste se hace más crítica conforme el problema y el mecanismo básico involucrados se tornan más complicados. Con seguridad, no siempre se puede estar seguro que la estructura que se postula, en este caso el modelo de regresión lineal, es correcto o incluso una representación adecuada. El siguiente ejemplo muestra cómo la suma de cuadrados del error se divide en los dos componentes que representan el error puro y la falta de ajuste. Lo adecuado del modelo se prueba al nivel de significancia α mediante la comparación del cuadrado medio de la falta de ajuste dividido entre s^2 con $f_{\alpha}(k-2, n-k)$.

Ejemplo 11.8 A continuación se registran las observaciones del rendimiento de una reacción química a varias temperaturas:

$y(\%)$	$x(^{\circ}\text{C})$	$y(\%)$	$x(^{\circ}\text{C})$
77.4	150	88.9	250
76.7	150	89.2	250
78.2	150	89.7	250
84.1	200	94.8	300
84.5	200	94.7	300
83.7	200	95.9	300

Estime el modelo lineal $\mu_{y|x} = \alpha + \beta x$ y pruebe la falta de ajuste.

Los resultados de los cálculos se muestran en la tabla 11.4. Conclusión: La división de la variación total revela de esta manera una variación significativa explicada por el modelo lineal y un monto insignificante de la variación debida a la falta de ajuste. Por ello, los datos experimentales no parecen sugerir la necesidad de considerar términos mayores que los de primer orden en el modelo, y no se rechaza la hipótesis nula. ■

Tabla 11.4 Análisis de varianza para los datos de rendimiento-temperatura.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada	Valores P
Regresión	509.1050	1	509.1050	1531.60	<0.0001
Error	4.0117	10			
Falta de ajuste	1.3517	2	0.6758	2.03	0.19
Error puro	2.6600	8	0.3325		
Total	513.1167	11			

Impreso de computadora comentado para la prueba de falta de ajuste

La figura 11.9 es un impreso de computadora comentado del análisis de los datos del ejemplo 11.8. Es resultado es un impreso del SAS. Nótese el "LOF" con 2 grados de libertad, que representa la contribución cuadrática y cúbica al modelo, y el valor P de 0.22, lo que sugiere que el modelo lineal (de primer orden) es adecuado.

General Linear Models Procedure					
Dependent Variable: Y Percent Yield					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	510.4888889	170.1522222	511.74	0.0001
Error	8	2.6600000	0.3325000		
Corrected Total	11	513.1168889			
		R-Square	C.V.	Root MSE	Y Mean
		0.994816	0.866751	0.5766281	86.483333
Source	DF	Type III SS	Mean Square	F Value	Pr > F
X	1	509.2508889	509.2508889	1631.58	0.0001
LOF	2	1.2060000	0.6030000	1.81	0.2241

Figura 11.9 Impreso del SAS que muestra el análisis de los datos del ejemplo 11.8.

Ejercicios

1. (a) Encuentre la estimación de mínimos cuadrados para el parámetro β en la ecuación lineal $\mu_{Y|x} = \beta x$.
 (b) Estime la línea de regresión que pasa a través del origen para los siguientes datos:

x	0.5	1.5	3.2	4.2	5.1	6.5
y	1.3	3.4	6.7	8.0	10.0	13.2

2. Suponga que en el ejercicio 1 no se conoce si la regresión real debe pasar por el origen. Estime el modelo lineal $\mu_{Y|x} = \alpha + \beta x$ y pruebe la hipótesis de que $\alpha = 0$ en el nivel de significancia de 0.10 contra la alternativa de que $\alpha \neq 0$.

3. Suponga que un experimentador postula un modelo del tipo

$$Y_i = \alpha + \beta x_{ij} + E_{ij}, \quad i = 1, 2, \dots, n,$$

cuando de hecho una variable adicional, digamos x_{2i} , también contribuye linealmente a la respuesta. El modelo real está dado entonces por

$$Y_i = \alpha + \beta x_{1i} + E_{1i}, \quad i = 1, 2, \dots, n.$$

Calcule el valor esperado del estimador

$$B = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1) Y_i}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}$$

4. Utilice un procedimiento de análisis de varianza para probar la hipótesis de que $\beta = 0$ contra la hipótesis alternativa de que $\beta \neq 0$ en el ejercicio 3 de la página 364 en el nivel de significancia de 0.05.

5. Se utilizan compuestos de organofosfato (OP) como pesticidas. Sin embargo, es importante estudiar sus efectos en especies expuestas a ellos. En el estudio de laboratorio, *Some Effects of Organophosphate Pesticides on Wildlife Species*, del Departamento de Industria Pesquera y Fauna del Instituto Politécnico y Universidad Estatal de Virginia, se llevó a cabo un experimento en el que se administraron diferentes dosis de un pesticida OP particular a cinco grupos de cinco ratones (*peromyscus leucopus*). Los 25 ratones eran hembras de edad y condiciones similares. Un grupo no recibió ningún químico. La respuesta básica y fue una medición de la actividad cerebral. Se postuló que la actividad cerebral disminuiría con un aumento en la dosis de OP. Los datos son los siguientes:

Animal	Dosis, x (mg/kg de peso corporal)	Actividad, y (moles/litro/min)
1	0.0	10.9
2	0.0	10.6
3	0.0	10.8
4	0.0	9.8
5	0.0	9.0
6	2.3	11.0
7	2.3	11.3
8	2.3	9.9
9	2.3	9.2
10	2.3	10.1
11	4.6	10.6
12	4.6	10.4
13	4.6	8.8
14	4.6	11.1
15	4.6	8.4
16	9.2	9.7
17	9.2	7.8
18	9.2	9.0
19	9.2	8.2
20	9.2	2.3
21	18.4	2.9
22	18.4	2.2
23	18.4	3.4
24	18.4	5.4
25	18.4	8.2

- (a) Con el uso del modelo

$$Y_i = \alpha + \beta x_i + E_i, \quad i = 1, 2, \dots, 25,$$

encuentre las estimaciones de mínimos cuadrados de α y β .

(b) Construya una tabla de análisis de varianza en la que la falta de ajuste y el error puro estén separados. Determine si la falta de ajuste es significativa en el nivel 0.05. Interprete los resultados.

6. Pruebe la linealidad de la regresión en el ejercicio 5 de la página 364. Use un nivel de significancia de 0.05.

7. Pruebe la linealidad de la regresión en el ejercicio 6 de la página 364.

8. La ganancia del transistor en un dispositivo de circuitos integrados entre emisor y colector (hFE) se relaciona con dos variables [Myers y Montgomery (1995)], que se pueden controlar en el proceso de de-

posición, tiempo de transmisión de emisor (x_1 , en minutos) y dosis del emisor (x_2 , en iones $\times 10^{14}$). Se observaron 14 muestras después de la deposición y los datos que resultaron se muestran en la tabla siguiente. Consideraremos modelos de regresión lineal con el uso de la ganancia como la respuesta y el tiempo de transmisión de emisor o la dosis de emisor como variables de regresión.

Tabla 11.5 Datos de ganancia del transistor y

Observación	x_1 (tiempo de transmisión, min)	x_2 (dosis, iones $\times 10^{14}$)	y (ganancia o hFE)
1	195	4.00	1004
2	255	4.00	1636
3	195	4.60	852
4	255	4.60	1506
5	225	4.20	1272
6	225	4.10	1270
7	225	4.60	1269
8	195	4.30	903
9	255	4.30	1555
10	225	4.00	1260
11	225	4.70	1146
12	225	4.30	1276
13	225	4.72	1225
14	340	4.30	1321

(a) Determine si el tiempo de transmisión de emisor influye en la ganancia con una relación lineal. Es decir, pruebe $H_0: \beta_1 = 0$, donde β_1 es la pendiente de la variable de regresión.

(b) Realice una prueba de falta de ajuste para determinar si la relación lineal es adecuada. Extraiga conclusiones.

(c) Determine si la dosis del emisor influye en la ganancia con una relación lineal. ¿Cuál variable de regresión es el mejor pronosticador de la ganancia?

9. Los siguientes datos son el resultado de una investigación del efecto de la temperatura de reacción x sobre el porcentaje de conversión de un proceso químico y . [Véase Myers y Montgomery (1995)]. Ajuste una regresión lineal simple y utilice una prueba de falta de ajuste para determinar si el modelo es adecuado.

Observación	Corrida	Temp. ($^{\circ}\text{C}$)	
		x	y
1	4	200	43
2	12	250	78
3	11	200	69
4	5	250	73
5	6	189.65	48
6	7	260.35	78
7	1	225	65
8	3	225	74
9	8	225	76
10	10	225	79
11	9	225	83
12	2	225	81

10. A menudo se utiliza el tratamiento térmico para carburar partes metálicas como los engranes. El espesor de la capa carburada se considera una característica importante del engrane y contribuye a la confiabilidad general de la parte. Debido a la naturaleza crítica de esta característica, se lleva a cabo una prueba de laboratorio en cada carga del horno. La prueba es destructiva, donde una parte real se corta de forma transversal y se remoja en un químico durante cierto tiempo. Esta prueba implica correr un análisis de carbón sobre la superficie del paso de engranaje (parte superior de los dientes del engrane) y la raíz del engrane (entre los dientes). Los siguientes datos son los resultados del análisis de carbón del paso de engranaje para 19 partes.

Tiempo de remojo	Paso de engranaje	Tiempo de remojo	Paso de engranaje
0.58	0.013	1.17	0.021
0.66	0.016	1.17	0.019
0.66	0.015	1.17	0.021
0.66	0.016	1.20	0.025
0.66	0.015	2.00	0.025
0.66	0.016	2.00	0.026
1.00	0.014	2.20	0.024
1.17	0.021	2.20	0.025
1.17	0.018	2.20	0.024
1.17	0.019		

(a) Ajuste una regresión lineal simple que relacione el análisis de carbón en el paso de engranaje y contra el tiempo de remojo. Pruebe $H_0: \beta_1 = 0$.

(b) Si la hipótesis de la parte (a) se rechaza, determine si el modelo lineal es adecuado.

11.9 Gráficas y transformaciones de datos

En este capítulo tratamos la construcción de modelos de regresión donde hay una variable independiente o de regresión. Además, suponemos, a través de la formulación del modelo, que x y y entran al modelo de *forma lineal*. A menudo se aconseja trabajar con un modelo alternativo en el que x y y (o ambas) entran de forma no lineal. Se puede indicar una *transformación* de los datos debido a las consideraciones teóricas inherentes al estudio científico, o una simple gráfica de los datos puede sugerir la necesidad de *reexpresar* las variables en el modelo. La necesidad de llevar a cabo una transformación es bastante simple de diagnosticar en el caso de la regresión lineal simple debido a que las gráficas bidimensionales dan una imagen real de cómo entra cada variable al modelo.

Un modelo en el que x o y se transforma no se debe ver como un *modelo de regresión no lineal*. Por lo general nos referimos a un modelo de regresión como lineal cuando es *lineal en los parámetros*. En otras palabras, suponga que el aspecto de los datos u otra información científica sugiere que debemos realizar la regresión de y^* contra x^* , donde cada una es una transformación de las variables naturales x y y . Entonces el modelo de la forma

$$y_i^* = \alpha + \beta x_i^* + \varepsilon_i$$

es un modelo lineal, pues es lineal en los parámetros α y β . El material dado en las secciones 11.2 a 11.8 aún es válido con y_i^* y x_i^* que reemplazan y_i y x_i . Un ejemplo simple y útil es el modelo log-log

$$\log y_i = \alpha + \beta \log x_i + \varepsilon_i.$$

Aunque este modelo no es lineal en x y y , es lineal en los parámetros y por ello se trata como un modelo lineal. Por otro lado, un ejemplo de modelo realmente no lineal es

$$y_i = \beta_0 + \beta_1 x_i^{\beta_2} + \varepsilon_i,$$

donde se debe estimar el parámetro β_2 (así como también β_0 y β_1). El modelo no es lineal en β_2 .

Las transformaciones que pueden mejorar el ajuste y la capacidad de predicción del modelo son muy numerosas. Para una discusión completa de las transformaciones, se remite al lector a Myers (véase la bibliografía). Elegimos aquí indicar algunas de ellas y mostrar la apariencia de la gráfica que sirve como diagnóstico. Considere la tabla 11.6. Se dan varias funciones que describen relaciones entre y y

Tabla 11.6 Algunas transformaciones útiles para linealizar

Forma funcional que relaciona y con x	Transformación apropiada	Forma de la regresión lineal simple
Exponencial: $y = \alpha e^{\beta x}$	$y^* = \ln y$	Regresión de y^* contra x
Potencia: $y = \alpha x^{\beta}$	$y^* = \log y; \quad x^* = \log x$	Regresión de y^* contra x^*
Recíproca: $y = \alpha + \beta \left(\frac{1}{x}\right)$	$x^* = \frac{1}{x}$	Regresión de y contra x^*
Función hiperbólica: $y = \frac{x}{\alpha + \beta x}$	$y^* = \frac{1}{y}; \quad x^* = \frac{1}{x}$	Regresión de y^* contra x^*

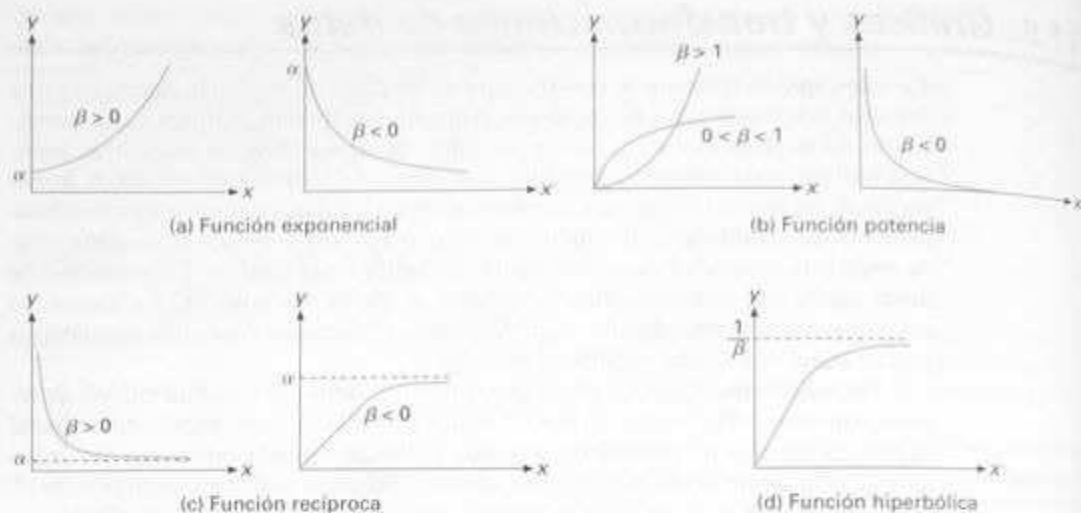


Figura 11.10 Diagramas que describen las funciones descritas en la tabla 11.5.

x que pueden producir una *regresión lineal* a través de la transformación que se indica. Además, para que quede completo se dan al lector las variables dependientes e independientes a utilizar en la *regresión lineal simple* que resulta. La figura 11.10 muestra las funciones que se listan en la tabla 11.6. Éstas sirven como guía para el analista al elegir una transformación de la observación de la gráfica de y contra x .

Lo anterior pretende ser una ayuda para el analista cuando es evidente que una transformación proporcionará una mejora. Sin embargo, antes de proporcionar un ejemplo, se deben señalar dos puntos importantes. El primero gira alrededor de la escritura formal del modelo cuando los datos se transforman. Con bastante frecuencia el analista no piensa acerca de esto. Simplemente lleva a cabo la transformación sin preocupación alguna acerca de la forma del modelo *antes* y *después* de la transformación. El modelo exponencial sirve como una buena ilustración. El modelo en las variables naturales (sin transformar) que produce un *modelo de error aditivo* en las variables transformadas está dado por

$$y_i = \alpha e^{\beta x_i} \cdot \varepsilon_i,$$

que es un *modelo de error multiplicativo*. Claramente, sacar logaritmos produce

$$\ln y_i = \ln \alpha + \beta x_i + \ln \varepsilon_i.$$

Como resultado, es en $\ln \varepsilon_i$ que se realizan las suposiciones básicas. El propósito de esta presentación simplemente es recordar al lector que no se debe ver a una transformación como una simple manipulación algebraica con un error agregado. A menudo un modelo en las variables transformadas que tiene una *estructura de error aditivo* propia es un resultado de un modelo en las variables naturales con una *estructura de error* de tipo diferente.

El segundo punto importante a tratar tiene que ver con la noción de mediciones de mejora. Medidas obvias de comparación son, por supuesto, R^2 y el cuadrado me-

dio del residuo, s^2 . (En el capítulo 12 se dan otras mediciones de comportamiento en comparaciones entre modelos en competencia). Ahora bien, si la respuesta y no se transforma, entonces es evidente que s^2 y R^2 se pueden utilizar para medir la utilidad de la transformación. Los residuos estarán en las mismas unidades para los modelos transformado y sin transformar. Pero cuando y se transforma, el criterio de comportamiento para el modelo transformado se debe basar en valores de los residuos en la métrica de la respuesta sin transformar. De esta manera las comparaciones que se realizan son apropiadas. El ejemplo que sigue proporciona una ilustración.

Ejemplo 11.9 La presión P de un gas que corresponde a varios volúmenes V se registra como sigue:

$V(\text{cm}^3)$	50	60	70	90	100
$P(\text{kg/cm}^2)$	64.7	51.3	40.5	25.9	7.8

La ley de los gases ideales está dada por la forma funcional $PV^\gamma = C$, donde γ y C son constantes. Estime las constantes C y γ .

SOLUCIÓN

Tomemos logaritmos naturales en ambos lados del modelo

$$P_i V_i^\gamma = C \cdot \varepsilon_i, \quad i = 1, 2, 3, 4, 5.$$

Como resultado, se puede escribir un modelo lineal

$$\ln P_i = \ln C - \gamma \ln V_i + \varepsilon_i^*, \quad i = 1, 2, 3, 4, 5,$$

donde $\varepsilon_i^* = \ln \varepsilon_i$. Los siguientes representan los resultados de la regresión lineal simple

$$\begin{aligned} \text{Intersección: } \widehat{\ln C} &= 14.7589, & \hat{C} &= 2,568,862.88 \\ \text{Pendiente: } \hat{\gamma} & & &= -2.65347221. \end{aligned}$$

La siguiente tabla representa la información que se toma del análisis de regresión.

P_i	V_i	$\ln P_i$	$\ln V_i$	$\widehat{\ln P_i}$	\hat{P}_i	$e_i = P_i - \hat{P}_i$
64.7	50	4.16976	3.91202	4.37853	79.7	-15.0
51.3	60	3.93769	4.09434	3.89474	49.1	2.2
40.5	70	3.70130	4.24850	3.48571	32.6	7.9
25.9	90	3.25424	4.49981	2.81885	16.8	9.1

Es ilustrativo graficar los datos y la ecuación de regresión. La figura 11.11 en la página 390 muestra una gráfica de los datos de presión y volumen sin transformar y la curva que representa la ecuación de regresión. ■

Gráficas de diagnóstico de residuos: detección gráfica de la violación de las suposiciones

Las gráficas de los datos crudos pueden ser en extremo útiles para determinar la naturaleza del modelo que se debe ajustar a los datos cuando hay una sola variable independiente. Intentamos ilustrar esto anteriormente. La detección de la forma

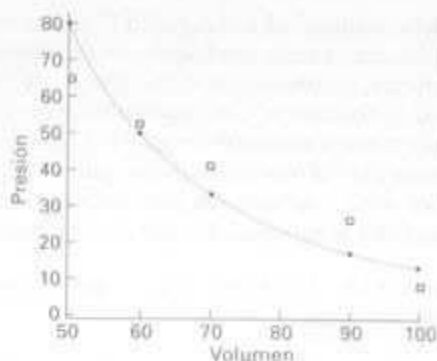


Figura 11.11 Datos de presión y volumen en una regresión ajustada.

apropiada del modelo no es, sin embargo, el único beneficio que se obtiene de la gráfica de diagnóstico. Como en gran parte del material que se asocia con la prueba de significancia en el capítulo 10, los métodos de graficación pueden ilustrar y detectar la violación de suposiciones. El lector debe recordar que mucho de lo que se ilustra en este capítulo requiere de las suposiciones que se hacen sobre los errores del modelo, E_i . De hecho, suponemos que las E_i son variables aleatorias independientes $N(0, \sigma)$. Ahora, por supuesto, las realizaciones de E_i , las e_i , no se observan. Sin embargo, las $e_i = y_i - \hat{y}_i$, los *residuos*, son el error en el ajuste de la línea de regresión y de esta manera sirven para simular las e_i . Así, el aspecto general de estos residuos a menudo puede resaltar las dificultades. Idealmente, por supuesto, las gráficas de los residuos son como se describe en la figura 11.12. Es decir, deben mostrar realmente *fluctuaciones aleatorias* alrededor de un valor cero.

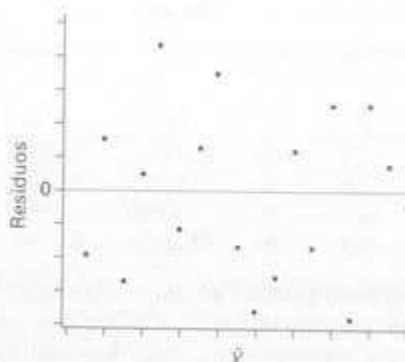


Figura 11.12 Gráfica ideal de los residuos.

Varianza no homogénea

La varianza homogénea es una suposición importante que se hace en el análisis de regresión. Las violaciones a menudo se pueden detectar a través de la apariencia de la gráfica del residuo. Aumentar la varianza del error con un incremento en la variable de regresión es una condición común en los datos científicos. Varianzas de

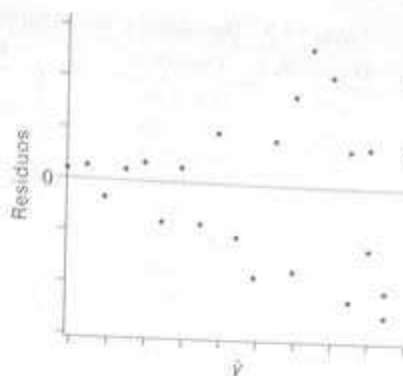


Figura 11.13 Gráfica de residuos que describen una varianza de error heterogénea.

error grandes producen grandes residuos y por ello una gráfica de residuos como la de la figura 11.13 es una señal de una varianza no homogénea. En el capítulo 12 aparece una presentación más amplia con respecto a estas gráficas de residuos e información respecto a diferentes tipos de residuos donde tratamos con la regresión lineal múltiple.

Graficación de probabilidad normal

La suposición de que los errores del modelo son normales se hace cuando el analista de datos trabaja con las pruebas de hipótesis o con una estimación de intervalos de confianza. De nuevo, la contraparte numérica de las ϵ_j , a saber, los residuos, son sujetos de la graficación de diagnóstico para detectar cualesquiera violaciones extremas. En el capítulo 8 se introducen las gráficas de cuantil-cuantil normal y se presentan con brevedad las gráficas de probabilidad normal. Estas gráficas sobre los residuos se ilustran en el estudio que se presenta en la siguiente sección.

11.10 Estudio de regresión lineal simple

En la fabricación de productos comerciales de madera es importante estimar la relación entre la densidad de un producto de madera y su dureza. Se considera un tipo relativamente nuevo de madera aglomerada que se puede formar con una facilidad considerablemente mayor que el producto comercial aceptado. Es necesario conocer a qué densidad se compara la dureza con el producto comercial bien conocido y bien documentado. El estudio lo realizó Terrance E. Conners, en *Investigation of Certain Mechanical Properties of a Wood-Foam Composite* (tesis de maestría, Departamento de Administración de Silvicultura y Fauna, Universidad de Massachusetts). Se produjeron treinta tableros de aglomerado con densidades de aproximadamente 8 a 26 libras por pie cúbico, y se midió la dureza en libras por pulgada cuadrada. La tabla 11.7 muestra los datos.

Es necesario para el analista de datos concentrarse en un ajuste apropiado para los datos y utilizar los métodos inferenciales que se presentan en este capítulo. La prueba de hipótesis sobre la pendiente de la regresión así como la estimación de

Tabla 11.7 Densidad y dureza para 30 tableros de aglomerado

Densidad, x	Dureza, y	Densidad, x	Dureza, y
9.50	14,814.00	15.40	25,312.00
8.40	17,502.00	15.00	26,222.00
9.80	14,007.00	14.50	22,148.00
11.00	19,443.00	14.80	26,751.00
8.30	7,573.00	13.60	18,036.00
9.90	14,191.00	25.60	96,305.00
8.60	9,714.00	23.40	104,170.00
6.40	8,076.00	24.40	72,594.00
7.00	5,304.00	23.30	49,512.00
8.20	10,728.00	19.50	32,207.00
17.40	43,243.00	21.20	48,218.00
15.00	25,319.00	22.80	70,453.00
15.20	28,028.00	21.70	47,661.00
16.40	41,792.00	19.80	38,138.00
16.70	49,499.00	21.30	53,045.00

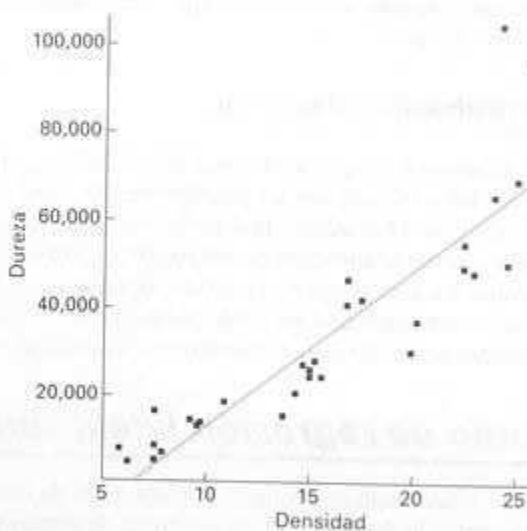


Figura 11.14 Gráfica de dispersión de los datos de densidad de la madera.

un intervalo de confianza o de predicción puede ser muy adecuado. Comenzamos con la demostración de una gráfica de dispersión simple de los datos crudos con una regresión lineal simple superpuesta. La figura 11.14 muestra esta gráfica.

La regresión lineal simple que se ajusta a los datos produce el modelo ajustado

$$\hat{y} = -25,433.739 + 3884.976x \quad (R^2 = 0.7975),$$

y se calculan los residuos. La figura 11.15 muestra los residuos graficados contra las mediciones de la densidad. Éste con dificultad es un conjunto ideal o bueno de residuos. No muestran una dispersión aleatoria alrededor de un valor cero. De hecho,

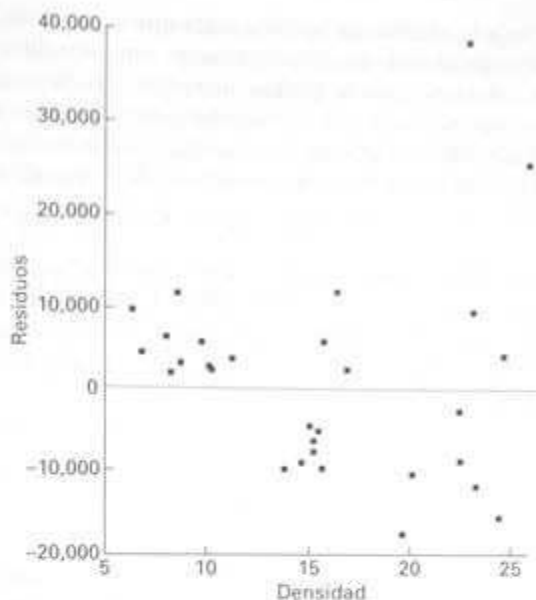


Figura 11.15 Gráfica de residuos para los datos de densidad de la madera.

los grupos de valores positivos y negativos pueden sugerir que se debe investigar una tendencia curvilínea en los datos.

Para obtener alguna idea con respecto a la suposición de error normal, se genera una gráfica de probabilidad normal de los residuos. Éste es el tipo de gráfica que se presenta en el capítulo 8 en la que el eje vertical representa la función de distribución empírica en una escala que produce una gráfica de línea recta cuando se grafica contra los mismos residuos. La figura 11.16 muestra la gráfica de la probabilidad normal de los residuos. Aquí $y_i - \bar{y}_i$ es el i -ésimo residuo más pequeño y

$f_i = \frac{i - 3/8}{n + 1/4}$ se grafica en el eje vertical. La gráfica de probabilidad normal no

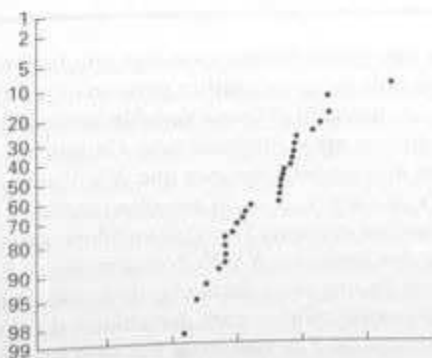


Figura 11.16 Gráfica de probabilidad normal de los residuos para los datos de densidad de la madera.

refleja la apariencia de línea recta que se desearía ver. Éste es otro síntoma de una falla, quizá una elección demasiado simplista del modelo de regresión.

Ambos tipos de gráficas de residuos y, en realidad, la gráfica de dispersión misma sugieren aquí que sería apropiado un modelo un poco más complicado. Un posible modelo es utilizar una transformación de logaritmo natural. En otras palabras, se podría elegir hacer una regresión de $\ln y$ contra x . Ésta produce la regresión

$$\ln y = 8.257 + 0.125x_i \quad (R^2 = 0.9016).$$

Para obtener una idea de si es más apropiado el modelo transformado, considere la figura 11.17 que revela una gráfica de los residuos en la dureza [es decir, y_i -antilog ($\ln y$)] contra la densidad]. La figura 11.17 parece estar más cercana a un patrón aleatorio alrededor de cero. Esto además del valor R^2 superior sugeriría que el modelo transformado es más apropiado.

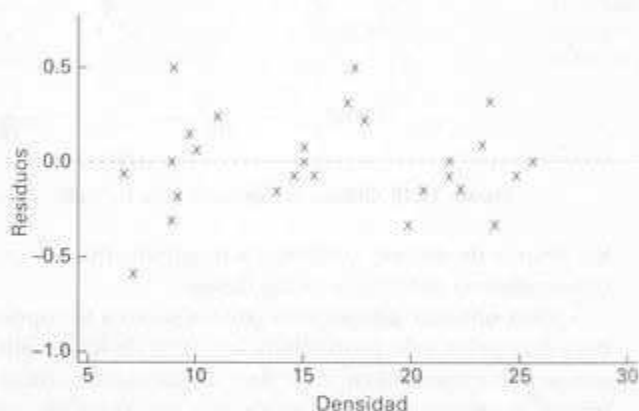


Figura 11.17 Gráfica de residuos con el uso de la transformación logarítmica para los datos de densidad de la madera.

11.11 Correlación

Hasta este punto hemos supuesto que la variable de regresión independiente x es una variable física o científica pero no una variable aleatoria. De hecho, en este contexto, x a menudo se llama **variable matemática**, que, en el proceso de muestreo, se mide con un error insignificante. En muchas aplicaciones de las técnicas de regresión es más realista suponer que X y Y son variables aleatorias y que las mediciones $\{(x_i, y_i); i = 1, 2, \dots, n\}$ son observaciones de una población que tiene la función de densidad conjunta $f(x, y)$. Consideraremos el problema de medir la relación entre las dos variables X y Y . Por ejemplo, si X y Y representan la longitud y circunferencia de una clase particular de hueso en el cuerpo adulto, podemos realizar un estudio antropológico para determinar si valores grandes de X se asocian con valores grandes de Y , y viceversa. Por otro lado, si X representa la edad de un automóvil usado y Y representa el valor en libras de reventa del automóvil, esperaríamos

que valores grandes de X correspondieran a valores pequeños de Y y que valores pequeños de X correspondieran a valores grandes de Y . El **análisis de correlación** intenta medir la fuerza de tales relaciones entre dos variables por medio de un solo número llamado **coeficiente de correlación**.

En teoría a menudo se supone que la distribución condicional $f(y|x)$ de Y , para valores fijos de X , es normal con media $\mu_{Y|x} = \alpha + \beta x$ y varianza $\sigma_{Y|x}^2 = \sigma^2$ y que X también se distribuye normalmente con media μ_x y varianza σ_x^2 . La densidad conjunta de X y Y es entonces

$$\begin{aligned} f(x, y) &= n(y|x; \alpha + \beta x, \sigma) n(x; \mu_x, \sigma_x) \\ &= \frac{1}{2\pi\sigma_x\sigma} \exp\left\{-\left(\frac{1}{2}\right)\left[\left(\frac{y - (\alpha + \beta x)}{\sigma}\right)^2 + \left(\frac{x - \mu_x}{\sigma_x}\right)^2\right]\right\} \end{aligned}$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$.

Escribamos la variable aleatoria Y en la forma

$$Y = \alpha + \beta X + E,$$

donde X es ahora una variable aleatoria independiente del error aleatorio E . Como la media del error aleatorio E es cero, se sigue que

$$\mu_Y = \alpha + \beta\mu_x \quad \text{y} \quad \sigma_Y^2 = \sigma^2 + \beta^2\sigma_x^2.$$

Al sustituir para α y σ^2 en la expresión anterior para $f(x, y)$, obtenemos la **distribución normal bivariada**

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_x\sigma_Y\sqrt{1-\rho^2}} \\ &\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x - \mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x - \mu_x}{\sigma_x}\right)\left(\frac{y - \mu_Y}{\sigma_Y}\right) + \left(\frac{y - \mu_Y}{\sigma_Y}\right)^2\right]\right\} \end{aligned}$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$, donde

$$\rho^2 = 1 - \frac{\sigma^2}{\sigma_Y^2} = \beta^2 \frac{\sigma_x^2}{\sigma_Y^2}.$$

La constante ρ (rho) se llama **coeficiente de correlación poblacional** y juega un papel importante en muchos problemas de análisis de datos de dos variables. Es importante para el lector comprender la interpretación física de este coeficiente de correlación y la distinción entre correlación y regresión. El término *regresión* aún tiene aquí un significado. De hecho, la línea recta dada por $\mu_{Y|x} = \alpha + \beta x$ todavía se llama línea de regresión como antes, y las estimaciones de α y β son idénticas a las que se dan en la sección 11.2. El valor de ρ es 0 cuando $\beta = 0$, que resulta cuando en esencia no hay una regresión lineal; es decir, la línea de regresión es horizontal y cualquier conocimiento de X no es de utilidad para predecir Y . Como debemos tener $\sigma_Y^2 \geq \sigma^2$, y $\rho^2 \leq 1$ por ello $-1 \leq \rho \leq 1$. Los valores de $\rho = \pm 1$ sólo ocurren cuando $\sigma^2 = 0$, en cuyo caso tenemos una relación lineal perfecta entre las dos variables. De esta manera un valor de ρ igual +1 implica una relación lineal perfecta

con una pendiente positiva; mientras que un valor de ρ igual a -1 resulta de una relación lineal perfecta con pendiente negativa. Se puede decir entonces que las estimaciones muestrales de ρ cercanas a la unidad en magnitud implican una buena correlación o una **asociación lineal** entre X y Y , mientras que valores cercanos a cero indican poca o ninguna correlación.

Para obtener una estimación muestral de ρ , recuerde de la sección 11.3 que la suma de cuadrados del error es

$$SSE = S_{yy} - bS_{xy}.$$

Al dividir ambos lados de esta ecuación entre S_{yy} y reemplazar S_{xy} por bS_{xx} , obtenemos la relación

$$b^2 \frac{S_{xx}}{S_{yy}} = 1 - \frac{SSE}{S_{yy}}.$$

El valor de $b^2 S_{xx}/S_{yy}$ es cero cuando $b = 0$, lo cual ocurre cuando los puntos de la muestra no presentan una relación lineal. Como $S_{yy} \geq SSE$, concluimos que $b^2 S_{xx}/S_{yy}$ debe estar entre 0 y 1. En consecuencia, $b \sqrt{S_{xx}/S_{yy}}$ debe ir de -1 a $+1$, en donde los valores negativos corresponden a líneas con pendientes negativas y los valores positivos a líneas con pendientes positivas. Un valor de -1 o $+1$ ocurrirá cuando $SSE = 0$, pero éste es el caso cuando todos los puntos muestrales yacen en una línea recta. Por ello una relación lineal perfecta aparece en los datos muestrales cuando $b \sqrt{S_{xx}/S_{yy}} = \pm 1$. Es claro que la cantidad $b \sqrt{S_{xx}/S_{yy}}$, que en adelante designaremos como r , se puede utilizar como una estimación del coeficiente de correlación poblacional ρ . Se acostumbra referirse a la estimación de r como el **coeficiente de correlación de producto-momento de Pearson** o simplemente **coeficiente de correlación muestral**.

Coefficiente de correlación

La medición ρ de la asociación lineal entre dos variables X y Y se estima mediante el **coeficiente de correlación muestral** r , donde

$$r = b \sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}.$$

Para valores de r entre -1 y $+1$ debemos ser cuidadosos en nuestra interpretación. Por ejemplo, los valores de r iguales a 0.3 y 0.6 sólo significan que tenemos dos correlaciones positivas, una algo más fuerte que la otra. Es incorrecto concluir que $r = 0.6$ indica una relación lineal el doble de buena que la que indica el valor $r = 0.3$. Por otro lado, si escribimos

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{SSR}{S_{yy}},$$

entonces r^2 , que por lo general se denomina **coeficiente de determinación muestral**, representa la proporción de la variación de S_{yy} explicada por la regresión de Y sobre x , a saber, SSR . Es decir, r^2 expresa la proporción de la variación total en los valores de la variable Y que dan razón o se pueden explicar mediante una relación

lineal con los valores de la variable aleatoria X . De esta manera una correlación de 0.6 significa que 0.36, o 36%, de la variación total de los valores de Y en nuestra muestra se explican mediante una relación lineal con los valores de X .

Ejemplo 11.10 Es importante que los investigadores científicos en el área de los productos forestales sean capaces de estudiar la correlación entre la anatomía y las propiedades mecánicas de los árboles. De acuerdo con el estudio *Quantitative Anatomical Characteristics of Plantation Grown Loblolly Pine (Pinus Taeda L.) and Cottonwood (Populus deltoides Bart. Ex Marsh.) and Their Relationship to Mechanical Properties* que llevó a cabo el Departamento de Silvicultura y Productos Forestales del Instituto Politécnico y Universidad Estatal de Virginia, un experimento en el que se seleccionaron aleatoriamente 29 pinos loblolly para investigación que produjeron los datos de la tabla 11.8 sobre la gravedad específica en gramos/cm³ y el módulo de ruptura en kilopascales (kPa). Calcule e interprete el coeficiente de correlación muestral.

Tabla 11.8

Gravedad específica, x (g/cm ³)	Módulo de ruptura, y (kPa)	Gravedad específica, x (g/cm ³)	Módulo de ruptura, y (kPa)
0.414	29,186	0.581	85,156
0.383	29,266	0.557	69,571
0.399	26,215	0.550	84,160
0.402	30,162	0.531	73,466
0.442	38,867	0.550	78,610
0.422	37,831	0.556	67,657
0.466	44,576	0.523	74,017
0.500	46,097	0.602	87,291
0.514	59,698	0.569	86,836
0.530	67,705	0.544	82,540
0.569	66,088	0.557	81,699
0.558	78,486	0.530	82,096
0.577	89,869	0.547	75,657
0.572	77,369	0.585	80,490
0.548	67,095		

SOLUCIÓN

De los datos encontramos que

$$S_{xx} = 0.11273, \quad S_{yy} = 11,807,324,786, \quad S_{xy} = 34,422.75972.$$

Por tanto,

$$r = \frac{34,422.75972}{\sqrt{(0.11273)(11,807,324,786)}} = 0.9435.$$

Un coeficiente de correlación de 0.9435 indica una buena relación lineal entre X y Y . Como $r^2 = 0.8902$, podemos decir que aproximadamente 89% de la variación en los valores de Y se explica por una relación lineal con X .

Una prueba de la hipótesis especial $\rho = 0$ contra una alternativa apropiada es equivalente a probar $\beta = 0$ para el modelo de regresión lineal simple y por tanto se aplican los procedimientos de la sección 11.7 que utilizan la distribución t con $n - 2$ grados de libertad o la distribución F con 1 y $n - 2$ grados de libertad. Sin embargo, si se desea evitar el procedimiento del análisis de varianza y calcular sólo el coeficiente de correlación muestral, se puede verificar (véase el ejercicio 3 de la página 400) que el valor t dado por

$$t = \frac{b}{s/\sqrt{S_{xx}}} = \frac{\sqrt{SSR}}{s} \quad \text{también se puede escribir como} \quad t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

que, como antes, es un valor de la estadística T que tiene una distribución t con $n - 2$ grados de libertad. ■

Ejemplo 11.11 Para los datos del ejemplo 11.10 pruebe la hipótesis de que no hay una asociación lineal entre las variables.

SOLUCIÓN

1. $H_0: \rho = 0$.
2. $H_1: \rho \neq 0$.
3. $\alpha = 0.05$.
4. Región crítica: $t < -2.052$ y $t > 2.052$.
5. Cálculos:

$$t = \frac{0.9435(\sqrt{27})}{\sqrt{1 - (0.9435)^2}} = 14.79, \quad P < 0.0001.$$

6. Decisión: rechazar la hipótesis de no asociación lineal.

Una prueba de la hipótesis más general $\rho = \rho_0$ contra una alternativa adecuada se lleva a cabo fácilmente a partir de la información muestral. Si X y Y siguen la distribución normal bivariada, la cantidad

$$\frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

es un valor de una variable aleatoria que sigue de manera aproximada la distribución normal con media $(1/2) \ln [(1 + \rho)/(1 - \rho)]$ y varianza $1/(n - 3)$. De esta forma el procedimiento de prueba es calcular

$$\begin{aligned} z &= \frac{\sqrt{n-3}}{2} \left[\ln \left(\frac{1+r}{1-r} \right) - \ln \left(\frac{1+\rho_0}{1-\rho_0} \right) \right] \\ &= \frac{\sqrt{n-3}}{2} \ln \left[\frac{(1+r)(1-\rho_0)}{(1-r)(1+\rho_0)} \right], \end{aligned}$$

y comparar con los puntos críticos de la distribución normal estándar. ■

Ejemplo 11.12 Para los datos del ejemplo 11.10 pruebe la hipótesis nula $\rho = 0,9$ contra la alternativa de que $\rho > 0,9$. Utilice un nivel de significancia de 0,05.

SOLUCIÓN

1. $H_0: \rho = 0,9$.
2. $H_1: \rho > 0,9$.
3. $\alpha = 0,05$.
4. Región crítica: $z > 1,645$.
5. Cálculos:

$$z = \frac{\sqrt{26}}{2} \ln \left[\frac{(1 + 0,9435)0,1}{(1 - 0,9435)1,9} \right] = 1,51,$$

$$P = 0,0655$$

6. Decisión: hay en realidad alguna evidencia de que el coeficiente de correlación excede 0,9. ■

Se debe señalar que en estudios de correlación, como en problemas de regresión lineal, los resultados que se obtienen sólo son tan buenos como el modelo que se supone. En las técnicas de correlación que aquí se estudian se supone una densidad normal bivariada para las variables X y Y , con el valor medio de Y en cada valor x linealmente relacionado con x . Para observar la conveniencia de la suposición de linealidad, a menudo es útil una graficación preliminar de los datos experimentales. Un valor del coeficiente de correlación muestral cercano a cero resultará de datos que muestren un efecto estrictamente aleatorio como en la figura 11.18(a), lo que implica poca o ninguna relación causal. Es importante recordar que el coeficiente de correlación entre dos variables es una medida de su relación lineal, y que un valor de $r = 0$ implica una falta de linealidad y no una falta de asociación. Por ello, si existe una fuerte relación cuadrática entre X y Y como se indica en la figura 11.18(b), podemos aún obtener una correlación cero que indique una relación no lineal.

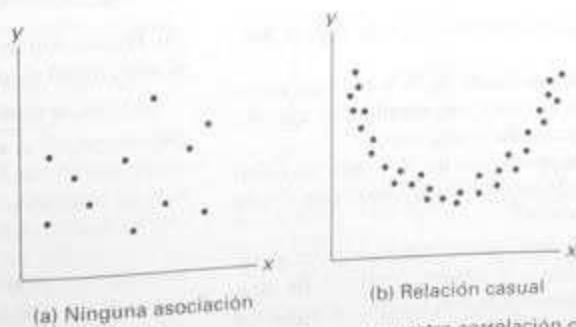


Figura 11.18 Diagrama de dispersión que muestra correlación cero.

Ejercicios

1. Calcule e interprete el coeficiente de correlación para las siguientes calificaciones de seis estudiantes seleccionados al azar:

Calificación en matemáticas	70	92	80	74	65	83
Calificación en inglés	74	84	63	87	78	90

2. Pruebe la hipótesis de que $\rho = 0$ en el ejercicio 1 contra la alternativa de que $\rho \neq 0$. Utilice un nivel de significancia de 0.05.

3. Muestre los pasos necesarios para convertir la ecuación $t = \frac{b}{s/\sqrt{S_{xx}}}$ a la forma equivalente

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

4. Los siguientes datos se obtienen en un estudio de la relación entre el peso y el tamaño del tórax de infantes al nacer:

Peso (kg)	Tamaño del tórax (cm)
2.75	29.5
2.15	26.3
4.41	32.2
5.52	36.5
3.21	27.2
4.32	27.7
2.31	28.3
4.30	30.3
3.71	28.7

(a) Calcule r .

(b) Pruebe la hipótesis nula $\rho = 0$ contra la alternativa $\rho > 0$ con un nivel de significancia de 0.01.

(c) ¿Qué porcentaje de la variación en los tamaños del tórax de los infantes se explica por la diferencia en peso?

5. Con referencia al ejercicio 1 de la página 363, suponga que x y y son variables aleatorias con una distribución normal bivariada:

(a) Calcule r .

(b) Pruebe la hipótesis de que $\rho = 0$ contra la alternativa de que $\rho \neq 0$ con un nivel de significancia de 0.05.

6. Con referencia al ejercicio 9 de la página 365, suponga una distribución normal bivariada para x y y :

(a) Calcule r .

(b) Pruebe la hipótesis nula $\rho = -0.5$ contra la alternativa $\rho < -0.5$ con un nivel de significancia de 0.025.

(c) Determine el porcentaje de variación en la cantidad de partículas eliminadas que deben sus cambios a la cantidad de lluvia diaria.

Ejercicios de repaso

1. Con referencia al ejercicio 6 de la página 364, construya

(a) un intervalo de confianza de 95% para las calificaciones promedio del curso de estudiantes que obtienen 35 en el examen de colocación;

(b) un intervalo de predicción de 95% para las calificaciones del curso de un estudiante que obtiene 35 en el examen de colocación.

2. Las cantidades de sólidos eliminados de un material particular cuando se le expone a periodos de secado de diferentes duraciones son los que se muestran en la tabla siguiente.

(a) Estime la línea de regresión lineal.

(b) Pruebe con un nivel de significancia de 0.05 si el modelo lineal es adecuado.

x (horas)	y (gramos)	
4.4	13.1	14.2
4.5	9.0	11.5
4.8	10.4	11.5
5.5	13.8	14.8
5.7	12.7	15.1
5.9	9.9	12.7
6.3	13.8	16.5
6.9	16.4	15.7
7.5	17.6	16.9
7.8	18.3	17.2

3. El Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia analizó datos de marmotas normales para el Departamento de Medicina Veterinaria. Las variables de interés fueron peso corporal en gramos y peso del corazón en gramos. También fue de interés desarrollar una ecuación de regresión lineal a fin de determinar si existe una relación lineal significativa entre el peso del corazón y el peso total del cuerpo. Utilice el peso del corazón como variable independiente y el peso del cuerpo como variable dependiente y ajuste una regresión lineal simple con el uso de los datos siguientes. Además, pruebe la hipótesis $H_0: \beta = 0$ contra $H_1: \beta \neq 0$. Extraiga conclusiones.

Peso corporal (gramos) Peso del corazón (gramos)

4050	11.2
2465	12.4
3120	10.5
5700	13.2
2595	9.8
3640	11.0
2050	10.8
4235	10.4
2935	12.2
4975	11.2
3690	10.8
2800	14.2
2775	12.2
2170	10.0
2370	12.3
2055	12.5
2025	11.8
2645	16.0
2675	13.8

4. Con referencia al ejercicio 7 de la página 364, construya

(a) un intervalo de confianza de 95% para las ventas semanales promedio cuando se gastan \$45 en publicidad;

(b) un intervalo de predicción de 95% para las ventas semanales cuando se gastan \$45 en publicidad.

5. Se diseñó un experimento para el Departamento de Ingeniería de Materiales del Instituto Politécnico y Universidad Estatal de Virginia para estudiar las propiedades de fragilidad del hidrógeno con base en mediciones de presión electrolytica de hidrógeno. La solución que se utilizó fue 0.1 N NaOH, y el material cierto tipo de acero inoxidable. Se controló la densidad de corriente de carga catódica y se varió en cuatro niveles. Se observó la presión efectiva del hidrógeno como la respuesta. Los datos son los siguientes.

Corrida	Densidad de corriente de carga, x (mA/cm ²)	Presión efectiva de hidrógeno, y (atm)
1	0.5	86.1
2	0.5	92.1
3	0.5	64.7
4	0.5	74.7
5	1.5	223.6
6	1.5	-202.1
7	1.5	132.9
8	2.5	413.5
9	2.5	231.5
10	2.5	466.7
11	2.5	365.3
12	3.5	493.7
13	3.5	382.3
14	3.5	447.2
15	3.5	563.8

(a) Realice una regresión lineal simple de y contra x .
 (b) Calcule la suma de cuadrados del error puro y haga una prueba de falta de ajuste.

(c) ¿La información de la parte (b) indica la necesidad de un modelo en x superior a una regresión de primer orden? Explique.

6. Los siguientes datos representan las calificaciones de química para una muestra aleatoria de 12 alumnos de primer grado en cierta universidad junto con sus calificaciones de una prueba de inteligencia que se les aplicó cuando aún eran alumnos de último año de preparatoria:

Estudiante	Calificación de la prueba, x	Calificación en química, y
1	65	85
2	50	74
3	55	76
4	65	90
5	55	85
6	70	87
7	65	94
8	70	98
9	55	81
10	70	91
11	50	76
12	55	74

(a) Calcule e interprete el coeficiente de correlación de la muestra.

(b) Establezca las suposiciones necesarias sobre las variables aleatorias.

(c) Pruebe la hipótesis de que $\rho = 0.5$ contra la alternativa de que $\rho > 0.5$. Utilice un valor P en las conclusiones.

7. Para el modelo de regresión lineal simple, pruebe que $E(s^2) = \sigma^2$.
8. La sección de negocios del *Washington Times* de marzo de 1997 lista 21 diferentes computadoras e impresoras y sus precios de venta. También se lista el pago promedio del comprador. En el siguiente impreso se muestran resultados parciales del análisis de regresión con el uso del software SAS.

```
Root MSE: 71.14129      R-square: 0.9672
Dep Mean: 894.04763     Adj R-Sq: 0.9655
C.V.: 7.95727
```

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEPT	1	53.119183	38.77463416	-1.370	0.1867
BUYER	1	1.052727	0.04448017	23.667	0.0001

PREDICT	BUYER	Dep Var PRICE	Predict Value	Std Err Predict	Lower 95%	Upper 95%	Lower 95%	Upper 95%
					Mean	Mean	Predict	Predict
IBM PS/2 486/66 40MB	325	375.0	365.3	26.176	340.5	450.0	239.5	555.5
EM ThinkPad 500	450	825.0	526.8	21.948	480.9	572.8	371.0	482.7
EM ThinkPad 755CX	1700	1850.0	1842.8	42.986	1752.8	1952.8	1658.8	2046.3
AST Pentium 80 540MB	900	875.0	899.3	15.525	862.8	927.8	742.9	1007.7
Dell Pentium 75 1GB	650	700.0	737.4	16.877	702.1	772.7	594.4	895.4
Gateway 486/75 320MB	700	750.0	790.0	16.135	756.1	823.8	637.3	942.7
Clone 586/133 1GB	500	600.0	632.1	19.065	592.2	672.0	478.0	766.3
Compaq Contura 4/75 120MB	450	600.0	526.8	21.948	480.9	572.8	371.0	482.7
Compaq Deskpro 490 1.2GB	805	850.0	895.3	15.525	862.8	927.8	742.9	1007.7
Micron 475 810MB	800	675.0	895.3	15.525	862.8	927.8	742.9	1007.7
Micron 9160 1.2GB	900	975.0	1000.6	16.164	966.7	1034.4	847.9	1153.9
Mac Quadra 840AV 500MB	450	575.0	526.8	21.948	480.9	572.8	371.0	482.7
Mac Performa 6116 700MB	700	775.0	790.0	16.135	756.1	823.8	637.3	942.7
PowerBook 540c 320MB	1400	1500.0	1528.9	30.921	1462.2	1591.7	1264.6	1697.3
PowerBook 5300 500MB	1350	1575.0	1474.3	29.618	1413.6	1535.0	1212.5	1651.2
Power Mac 7500/100 1GB	1150	1325.0	1263.8	22.023	1217.7	1309.9	1107.9	1419.5
DEC Versa 486 340MB	800	900.0	898.3	15.525	862.8	927.8	742.9	1007.7
Toshiba 1960CE 320MB	700	825.0	790.0	16.135	756.1	823.8	637.3	942.7
Toshiba 4800VCT 500MB	1000	1150.0	1105.8	17.919	1069.3	1143.4	952.1	1258.9
HP LaserJet III	350	475.0	421.6	25.289	368.6	474.5	263.5	579.4
Apple Laser Writer Pro 63	750	800.0	842.7	15.675	809.9	875.5	690.2	995.2

- (a) Explique la diferencia entre el intervalo de confianza sobre la media y el intervalo de predicción.
- (b) Explique por qué los errores estándar de la predicción varían de una observación a otra.
- (c) ¿Cuál observación tiene el error estándar de la media más bajo? ¿Por qué?

9. Considere los siguientes datos de vehículos tomados de *Consumer Reports*. El peso está en toneladas, el rendimiento en millas por galón, también se indica la razón de tracción. Se ajusta un modelo de regresión que relaciona el peso x con el rendimiento y . Un impreso parcial del SAS muestra algunos de los resultados del análisis de regresión y una gráfica de los residuos y el peso para cada vehículo.

OBS	MODEL	WT	MPG	DR_RATIO
1.	Buick Estate Wagon	4.360	18.9	2.73
2.	Ford Country Squire Wagon	4.054	15.5	2.56
3.	Chery Malibu Wagon	3.605	19.2	2.56
4.	Chrysler LeBaron Wagon	3.940	16.8	2.65
5.	Chevette	2.158	30.0	3.70
6.	Toyota Corona	3.860	27.5	3.05
7.	Datsun 510	2.300	27.2	3.54

[continuación]

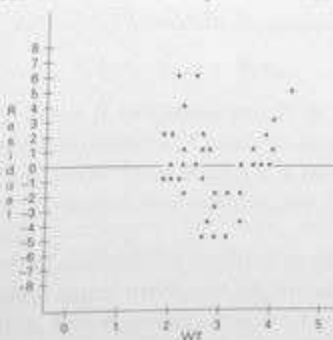
9	Dodge Omni			
9	Audi 5000	2.230	10.9	3.37
10	Volvo 240 GL	2.830	20.3	3.90
11	Saab 99 GL	3.140	17.0	3.56
12	Peugeot 694 SL	2.795	21.6	3.77
13	Buick Century Special	3.410	16.2	3.58
14	Mercury Zephyr	3.380	20.6	3.73
15	Dodge Aspen	3.070	20.8	3.58
16	AMC Concord D/L	3.620	18.6	2.71
17	Chevy Caprice Classic	3.410	18.1	2.73
18	Ford LTD	3.840	17.0	2.41
19	Mercury Grand Marquis	3.725	17.8	2.36
20	Dodge St Regis	3.855	16.5	2.26
21	Ford Mustang 4	3.830	18.2	2.45
22	Ford Mustang Ghia	2.585	26.5	3.08
23	Ford Mustang Ghia	2.910	21.9	3.08
24	Mazda GLC	1.975	34.1	3.73
24	Dodge Colt	1.915	35.1	2.87
25	AMC Spirit	2.670	27.4	3.08
26	VW Scirocco	1.990	31.5	3.78
27	Honda Accord LX	2.135	29.5	3.08
28	Buick Skylark	2.670	28.4	3.53
29	Chevy Citation	2.595	28.8	2.69
30	Datsun Omega	2.100	26.8	2.84
31	Pontiac Phoenix	2.554	33.5	2.69
32	Plymouth Horizon	2.380	34.2	3.37
33	Datsun 210	2.020	31.8	3.70
34	Fiat Strada	2.130	37.3	3.10
35	VW Dasher	2.190	30.5	2.70
36	Datsun 810	2.815	22.0	2.70
37	BMW 320i	2.600	21.5	3.84
38	VW Rabbit	1.925	31.9	3.70

Root MSE	1.85880	R-square	0.8155
Dep Mean	24.76053	Adj R-sq	0.8104
C.V.	11.51351		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	48.707495	1.95368170	24.931	0.0001
WT	1	-6.364600	0.66302032	-12.616	0.0001

plot of RESIDUAL*WT. Symbol used is '*'.



Note: 1 obs. hidden.

- (a) ¿Del análisis y de la gráfica del residuo parece que se puede encontrar un modelo mejorado con el uso de una transformación? Explique.
- (b) Ajuste el modelo mediante el reemplazo del peso por el logaritmo del peso. Comente los resultados.
- (c) Ajuste un modelo mediante el reemplazo de las mpg con mpg por 100 millas recorridas, como con frecuencia se reporta el rendimiento en otros países. ¿Cuál de los tres modelos es preferible? Explique.

Regresión lineal múltiple

12.1 Introducción

En la mayor parte de los problemas de investigación donde se aplica el análisis de regresión se necesita más de una variable independiente en el modelo de regresión. La complejidad de la mayor parte de los mecanismos científicos es tal que para ser capaces de predecir una respuesta importante se necesita un **modelo de regresión múltiple**. Cuando este modelo es lineal en los coeficientes se denomina **modelo de regresión lineal múltiple**. Para el caso de k variables independientes x_1, x_2, \dots, x_k , la media de $Y|x_1, x_2, \dots, x_k$ está dada por el modelo de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

y la respuesta estimada se obtiene de la ecuación de regresión de la muestra

$$\hat{y} = b_0 + b_1 x_1 + \dots + b_k x_k,$$

donde cada coeficiente de regresión β_i se estima por b_i de los datos de la muestra con el uso del método de mínimos cuadrados. Como en el caso de una sola variable independiente, el modelo de regresión lineal múltiple a menudo puede ser una representación adecuada de una estructura más complicada dentro de ciertos rangos de las variables independientes.

Técnicas de mínimos cuadrados similares también se pueden aplicar al estimar los coeficientes cuando el modelo lineal involucra, digamos, potencias y productos de las variables independientes. Por ejemplo, cuando $k = 1$, el experimentador puede pensar que las medias $\mu_{Y|x}$ no caen en una línea recta pero que se describen de forma más apropiada con el **modelo de regresión polinomial**

$$\mu_{y|x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_r x^r,$$

y la respuesta estimada se obtiene de la ecuación de regresión polinomial

$$\hat{y} = b_0 + b_1 x + b_2 x^2 + \cdots + b_r x^r.$$

En ocasiones surge confusión cuando hablamos de un modelo polinomial como de un modelo lineal. Sin embargo, los estadísticos por lo general se refieren a un modelo lineal como uno en el cual los parámetros ocurren linealmente, sin importar cómo entran las variables independientes al modelo. Un ejemplo de un modelo no lineal es la **relación exponencial**

$$\mu_{y|x} = \alpha\beta^x,$$

que se estima con la ecuación de regresión

$$\hat{y} = ab^x.$$

Existen muchos fenómenos en la ciencia y en la ingeniería que son inherentemente no lineales por naturaleza y, cuando se conoce la estructura real, desde luego se debe hacer un intento para ajustar el modelo presente. La literatura sobre estimación por mínimos cuadrados de modelos no lineales es voluminosa y, excepto por el ejercicio 10 de la página 365, no intentamos cubrir el tema en este libro. El estudiante que quiera una buena explicación de algunos aspectos de este tema debe consultar *Classical and Modern Regression with Applications* de Myers (véase la bibliografía).

12.2 Estimación de los coeficientes

En esta sección obtenemos los estimadores de mínimos cuadrados de los parámetros $\beta_0, \beta_1, \dots, \beta_k$ mediante el ajuste del modelo de regresión lineal múltiple

$$\mu_{y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

a los puntos de datos

$$\{(x_{1i}, x_{2i}, \dots, x_{ki}, y_i), \quad i = 1, 2, \dots, n \text{ y } n > k\},$$

donde y_i es la respuesta observada para los valores $x_{1i}, x_{2i}, \dots, x_{ki}$ de las k variables independientes x_1, x_2, \dots, x_k . Cada observación $(x_{1i}, x_{2i}, \dots, x_{ki}, y_i)$ satisface la ecuación

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + e_i$$

o

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \cdots + b_k x_{ki} + e_i,$$

donde e_i y e_i son el error aleatorio y residual, respectivamente, asociados con la respuesta y_i . Al utilizar el concepto de mínimos cuadrados para llegar a las estimaciones b_0, b_1, \dots, b_k , minimizamos la expresión

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{1i} - b_2 x_{2i} - \dots - b_k x_{ki})^2.$$

Al diferenciar SSE a su vez con respecto a $b_0, b_1, b_2, \dots, b_k$, e igualar a cero, generamos un conjunto de $k + 1$ ecuaciones normales

$$\begin{aligned} nb_0 + b_1 \sum_{i=1}^n x_{1i} + b_2 \sum_{i=1}^n x_{2i} + \dots + b_k \sum_{i=1}^n x_{ki} &= \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_{1i} + b_1 \sum_{i=1}^n x_{1i}^2 + b_2 \sum_{i=1}^n x_{1i} x_{2i} + \dots + b_k \sum_{i=1}^n x_{1i} x_{ki} &= \sum_{i=1}^n x_{1i} y_i \\ \vdots & \\ b_0 \sum_{i=1}^n x_{ki} + b_1 \sum_{i=1}^n x_{ki} x_{1i} + b_2 \sum_{i=1}^n x_{ki} x_{2i} + \dots + b_k \sum_{i=1}^n x_{ki}^2 &= \sum_{i=1}^n x_{ki} y_i. \end{aligned}$$

Estas ecuaciones se pueden resolver para $b_0, b_1, b_2, \dots, b_k$ mediante cualquier método apropiado para resolver sistemas de ecuaciones lineales.

Ejemplo 12.1 Se realizó un estudio sobre un camión de reparto ligero a diesel para ver si la humedad, temperatura del aire y presión barométrica influyen en la emisión de óxido nítrico (en ppm). Las mediciones de las emisiones se tomaron en diferentes momentos, con condiciones experimentales variantes. Los datos son los siguientes:

Óxido nítrico, y	Humedad x_1	Temperatura, x_2	Presión, x_3	Óxido nítrico, y	Humedad, x_1	Temperatura, x_2	Presión, x_3
0.90	72.4	76.3	29.18	1.07	23.2	76.8	29.38
0.91	41.6	70.3	29.35	0.94	47.4	86.6	29.35
0.96	34.3	77.1	29.24	1.10	31.5	76.9	29.63
0.89	35.1	68.0	29.27	1.10	10.6	86.3	29.56
1.00	10.7	79.0	29.78	1.10	11.2	86.0	29.48
1.10	12.9	67.4	29.39	0.91	73.3	76.3	29.40
1.15	8.3	66.8	29.69	0.87	75.4	77.9	29.28
1.03	20.1	76.9	29.48	0.78	96.6	78.7	29.29
0.77	72.2	77.7	29.09	0.82	107.4	86.8	29.03
1.07	24.0	67.7	29.60	0.95	54.9	70.9	29.37

Fuente: Charles T. Hare, "Light-Duty Diesel Emission Correction Factors for Ambient Conditions", EPA-600/2-77-116, Agencia de Protección Ambiental de Estados Unidos.

El modelo es

$$\mu_{y|x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

Ajuste este modelo de regresión lineal múltiple a los datos dados y después estime la cantidad de óxido nítrico para las condiciones donde la humedad es 50%, la temperatura 76°F y la presión barométrica 29.30.

SOLUCIÓN

Para las ecuaciones normales encontramos que

$$20b_0 + 863.1000b_1 + 1530.4000b_2 + 587.8400b_3 = 19.4200,$$

$$863.1000b_0 + 54,876.8900b_1 + 67,000.0900b_2 + 25,283.3950b_3 = 779.4770,$$

$$1530.4000b_0 + 67,000.0900b_1 + 117,912.3200b_2 + 44,976.8670b_3 = 1483.4370,$$

$$587.8400b_0 + 25,283.3950b_1 + 44,976.8670b_2 + 17,278.5086b_3 = 571.1219.$$

La solución de este conjunto de ecuaciones da las estimaciones únicas

$$b_0 = -3.507778, \quad b_1 = -0.002625, \quad b_2 = 0.000799, \quad b_3 = 0.154155.$$

Por tanto, la ecuación de regresión es

$$\hat{y} = -3.507778 - 0.002625x_1 + 0.000799x_2 + 0.154155x_3.$$

Para 50% de humedad, una temperatura de 76 °F y una presión barométrica de 29.30, la cantidad estimada de óxido nítrico es

$$\begin{aligned} \hat{y} &= -3.507778 - 0.002625(50.0) + 0.000799(76.0) + 0.154155(29.30) \\ &= 0.9384 \quad \blacksquare \end{aligned}$$

Regresión polinomial

Suponga ahora que deseamos ajustar la ecuación polinomial

$$\mu_{Y|x} = \beta_0 + \beta_1x + \beta_2x^2 + \cdots + \beta_r x^r$$

a los n pares de observaciones $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Cada observación, y_i , satisface la ecuación

$$y_i = \beta_0 + \beta_1x_i + \beta_2x_i^2 + \cdots + \beta_r x_i^r + \varepsilon_i$$

o

$$y_i = b_0 + b_1x_i + b_2x_i^2 + \cdots + b_r x_i^r + \varepsilon_i,$$

donde r es el grado del polinomio, y ε_i y e_i son de nuevo el error aleatorio y residual asociados con la respuesta y_i . Aquí, el número de pares, n , debe ser al menos tan grande como $r + 1$, el número de parámetros a estimar. Nótese que el modelo

12.3 Modelo de regresión lineal con el uso de matrices

Al ajustar un modelo de regresión lineal múltiple, en particular cuando el número de variables pasa de dos, el conocimiento de la teoría matricial puede facilitar las manipulaciones matemáticas de forma considerable. Suponga que el experimentador tiene k variables independientes x_1, x_2, \dots, x_k y n observaciones y_1, y_2, \dots, y_n , cada una de las cuales se pueden expresar por la ecuación

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i.$$

Este modelo en esencia representa n ecuaciones que describen cómo se generan los valores de respuesta en el proceso científico. Con el uso de la notación matricial, podemos escribir la ecuación

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

donde

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Entonces la solución de mínimos cuadrados para la estimación de $\boldsymbol{\beta}$ que se ilustra en la sección 12.2 implica encontrar \mathbf{b} para la que

$$SSE = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})$$

se minimiza. Este proceso de minimización implica resolver para \mathbf{b} en la ecuación

$$\frac{\partial}{\partial \mathbf{b}} (SSE) = \mathbf{0}.$$

No presentaremos los detalles relacionados con las soluciones de las ecuaciones anteriores. El resultado se reduce a la solución de \mathbf{b} en

$$(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}.$$

Nótese la naturaleza de la matriz \mathbf{X} . Aparte del elemento inicial, el i -ésimo renglón representa los valores x que dan lugar a la respuesta y_i . Al escribir

$$A = X'X = \begin{bmatrix} n & \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \dots & \sum_{i=1}^n x_{ki} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}x_{2i} & \dots & \sum_{i=1}^n x_{1i}x_{ki} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{ki} & \sum_{i=1}^n x_{ki}x_{1i} & \sum_{i=1}^n x_{ki}x_{2i} & \dots & \sum_{i=1}^n x_{ki}^2 \end{bmatrix} \quad y \quad g = X'y = \begin{bmatrix} g_0 = \sum_{i=1}^n y_i \\ g_1 = \sum_{i=1}^n x_{1i}y_i \\ \vdots \\ g_k = \sum_{i=1}^n x_{ki}y_i \end{bmatrix}$$

las ecuaciones normales se pueden escribir en la forma matricial

$$AB = g.$$

Si la matriz A es no singular, podemos escribir la solución para el coeficiente de regresión como

$$b = A^{-1}g = (X'X)^{-1}X'y.$$

De esta forma se puede obtener la ecuación de predicción o la ecuación de regresión al resolver un conjunto de $k + 1$ ecuaciones con un número igual de incógnitas. Esto implica la inversión de la matriz $X'X$ de $k + 1$ por $k + 1$. Las técnicas para invertir esta matriz se explican en la mayoría de los libros de texto sobre determinantes y matrices elementales. Por supuesto, se dispone de muchos paquetes de computadora de alta velocidad para problemas de regresión múltiple, paquetes que no sólo imprimen estimaciones de los coeficientes de regresión, sino que también proporcionan otra información relevante para hacer inferencias respecto a la ecuación de regresión.

Ejemplo 12.3 Se midió el porcentaje de sobrevivencia de cierto tipo de semen animal, después del almacenamiento, en varias combinaciones de concentraciones de tres materiales que se utilizan para aumentar su oportunidad de sobrevivencia. Los datos son los siguientes:

y (% sobrevivencia)	x_1 (peso %)	x_2 (peso %)	x_3 (peso %)
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

Estíme el modelo de regresión lineal múltiple para los datos dados.

SOLUCIÓN

Las ecuaciones de estimación de mínimos cuadrados, $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}$, son

$$\begin{bmatrix} 13 & 59.43 & 81.82 & 115.40 \\ 59.43 & 394.7255 & 360.6621 & 522.0780 \\ 81.82 & 360.6621 & 576.7264 & 728.3100 \\ 115.40 & 522.0780 & 728.3100 & 1035.9600 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 377.5 \\ 1877.567 \\ 2246.661 \\ 3337.780 \end{bmatrix}$$

De los resultados de una computadora obtenemos los elementos de la matriz inversa

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 8.0648 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix}$$

y después, con el uso de la relación $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, los coeficientes estimados de regresión son

$$b_0 = 39.1574, \quad b_1 = 1.0161, \quad b_2 = -1.8616, \quad b_3 = -0.3433.$$

De aquí nuestra ecuación de regresión estimada es

$$\hat{y} = 39.1574 + 1.0161x_1 - 1.8616x_2 - 0.3433x_3.$$

Para el caso de una sola variable independiente, el *grado* del polinomio de mejor ajuste a menudo se puede determinar al graficar un diagrama de dispersión de los datos que se obtienen de un experimento que da n pares de observaciones de la forma $\{(x_i, y_i); i = 1, 2, \dots, n\}$. El conjunto de ecuaciones normales que se derivan en la sección 12.2 para estimar los coeficientes del polinomio general de grado r toma la forma

$$\begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \dots & \sum_{i=1}^n x_i^r \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \dots & \sum_{i=1}^n x_i^{r+1} \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 & \dots & \sum_{i=1}^n x_i^{r+2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \sum_{i=1}^n x_i^r & \sum_{i=1}^n x_i^{r+1} & \sum_{i=1}^n x_i^{r+2} & \dots & \sum_{i=1}^n x_i^{2r} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_r \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \vdots \\ \sum_{i=1}^n x_i^r y_i \end{bmatrix}$$

Al resolver estas $r+1$ ecuaciones, obtenemos las estimaciones b_0, b_1, \dots, b_r , y por ello generamos la ecuación de predicción de regresión polinomial

$$\hat{y} = b_0 + b_1x + b_2x^2 + \dots + b_r x^r.$$

El procedimiento para ajustar un modelo de regresión polinomial se puede generalizar al caso de más de una variable independiente. De hecho, el estudiante de

análisis de regresión debe, en esta etapa, tener la facilidad para ajustar cualquier modelo lineal en, digamos, k variables independientes. Suponga, por ejemplo, que tenemos una respuesta Y con $k = 2$ variables independientes y se postula un modelo cuadrático del tipo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \varepsilon_i,$$

donde y_i , $i = 1, 2, \dots, n$, es la respuesta para la combinación (x_{1i}, x_{2i}) de las variables independientes en el experimento. En esta situación n debe ser al menos 6, pues hay seis parámetros a estimar mediante el procedimiento de mínimos cuadrados. Además, como el modelo contiene términos cuadráticos en ambas variables, se deben usar al menos tres niveles de cada variable. El lector debe verificar con facilidad que las ecuaciones normales de mínimos cuadrados $(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{y}$ están dadas por

$$\begin{bmatrix} n & \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{2i}^2 & \sum_{i=1}^n x_{1i} x_{2i} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i} x_{2i} & \sum_{i=1}^n x_{1i}^3 & \sum_{i=1}^n x_{1i} x_{2i}^2 & \sum_{i=1}^n x_{1i}^2 x_{2i} \\ \sum_{i=1}^n x_{2i} & \sum_{i=1}^n x_{1i} x_{2i} & \sum_{i=1}^n x_{2i}^2 & \sum_{i=1}^n x_{1i}^2 x_{2i} & \sum_{i=1}^n x_{2i}^3 & \sum_{i=1}^n x_{1i} x_{2i}^2 \\ \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}^3 & \sum_{i=1}^n x_{1i}^2 x_{2i} & \sum_{i=1}^n x_{1i}^4 & \sum_{i=1}^n x_{1i}^2 x_{2i}^2 & \sum_{i=1}^n x_{1i}^3 x_{2i} \\ \sum_{i=1}^n x_{2i}^2 & \sum_{i=1}^n x_{1i} x_{2i}^2 & \sum_{i=1}^n x_{2i}^3 & \sum_{i=1}^n x_{1i}^2 x_{2i}^2 & \sum_{i=1}^n x_{2i}^4 & \sum_{i=1}^n x_{1i} x_{2i}^3 \\ \sum_{i=1}^n x_{1i} x_{2i} & \sum_{i=1}^n x_{1i}^2 x_{2i} & \sum_{i=1}^n x_{1i} x_{2i}^2 & \sum_{i=1}^n x_{1i}^3 x_{2i} & \sum_{i=1}^n x_{1i} x_{2i}^3 & \sum_{i=1}^n x_{1i}^2 x_{2i}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_{11} \\ b_{22} \\ b_{12} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{1i} y_i \\ \sum_{i=1}^n x_{2i} y_i \\ \sum_{i=1}^n x_{1i}^2 y_i \\ \sum_{i=1}^n x_{2i}^2 y_i \\ \sum_{i=1}^n x_{1i} x_{2i} y_i \end{bmatrix}$$

Ejemplo 12.4 Los siguientes datos representan el porcentaje de impurezas que ocurren a varias temperaturas y tiempos de esterilización durante una reacción asociada con la fabricación de cierta bebida.

Tiempo de esterilización, x_2 (min)	Temperatura, x_1 ($^{\circ}\text{C}$)		
	75	100	125
15	14.05	10.55	7.55
	14.93	9.48	6.59
20	16.56	13.63	9.23
	15.83	11.75	8.78
25	22.41	18.55	15.93
	21.66	17.98	16.44

Estimar los coeficientes de regresión en el modelo

$$\mu_{Y|X_1, X_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2.$$

SOLUCIÓN

$$b_0 = 56.4668, \quad b_{11} = 0.00081,$$

$$b_1 = -0.36235, \quad b_{22} = 0.08171,$$

$$b_2 = -2.75299, \quad b_{12} = 0.00314,$$

y nuestra ecuación de regresión estimada es

$$\hat{y} = 56.4668 - 0.36235x_1 - 2.75299x_2 + 0.00081x_1^2 + 0.08171x_2^2 + 0.00314x_1x_2.$$

Muchos de los principios y procedimientos asociados con la estimación de funciones de regresión polinomial caen en la categoría de la **metodología de respuesta superficial**, un conjunto de técnicas que los científicos e ingenieros han utilizado con bastante éxito en muchos campos. Problemas como la selección de un diseño experimental apropiado, en particular para casos donde hay un número grande de variables en el modelo, y la elección de las condiciones "óptimas" de operación sobre x_1, x_2, \dots, x_k a menudo se aproximan a través del uso de estos métodos. Para una exposición más amplia se remite al lector a *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* de Myers y Montgomery (véase la bibliografía). ■

Ejercicios

1. Suponga en el ejercicio 6 de la página 401 que también se nos da el número de periodos de clases perdidos por los 12 estudiantes que toman el curso de química. Los datos completos se registran a continuación:

Estudiante	Calificación de química, y	Calificación del examen, x_1	Clases perdidas, x_2
1	85	65	1
2	74	50	7
3	76	55	5
4	90	65	2
5	85	55	6
6	87	70	3
7	94	65	2
8	98	70	5
9	81	55	4
10	91	70	3
11	76	50	1
12	74	55	4

(a) Ajuste una ecuación de regresión lineal múltiple de la forma $\mu_{y|x_1, x_2} = \beta_0 + \beta_1x_1 + \beta_2x_2$.

(b) Estime la calificación de química para un estudiante que tiene una calificación en un examen de inteligencia de 60 y pierde cuatro clases.

2. En *Applied Spectroscopy* aparece un estudio sobre las propiedades espectrales de la reflectancia infrarroja de un líquido viscoso que se utiliza en la

industria de la electrónica como lubricante. El diseño experimental consiste en el efecto de la frecuencia de banda x_1 y el espesor de la película x_2 sobre la densidad óptica y mediante el uso de un espectrómetro infrarrojo de Perkin-Elmer modelo 621. [Fuente: I. Pachansky, C.D. England y R. Wattman, "Infrared spectroscopic studies of poly (perfluoropropyleneoxide). A classical dispersion analysis for the refractive index", *Applied Spectroscopy*, Vol. 40, Núm. 1, enero de 1986, p. 9, tabla 1.]

y	x_1	x_2
0.231	740	1.10
0.107	740	0.62
0.053	740	0.31
0.129	805	1.10
0.069	805	0.62
0.030	805	0.31
1.005	980	1.10
0.559	980	0.62
0.321	980	0.31
2.948	1,235	1.10
1.633	1,235	0.62
0.934	1,235	0.31

Estime la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2} = \beta_0 + \beta_1x_1 + \beta_2x_2.$$

3. Se realiza un conjunto de corridas experimentales para determinar una forma de predecir el tiempo

de cocción y a varios niveles de ancho del horno x_1 , y temperatura de combustión x_2 . Los datos codificados se registran a continuación:

y	x_1	x_2
6.40	1.32	1.15
15.05	2.69	3.40
18.75	3.56	4.10
30.25	4.41	8.75
44.85	5.35	14.82
48.94	6.20	15.15
51.55	7.12	15.32
61.50	8.87	18.18
100.44	9.80	35.19
111.42	10.65	40.40

Estime la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$

4. Se lleva a cabo un experimento para determinar si el peso de un animal se puede predecir después de un tiempo dado sobre la base del peso inicial del animal y la cantidad de alimento que consume. Se registran los datos siguientes, en kilogramos:

Peso final, y	Peso inicial, x_1	Alimento consumido, x_2
95	42	272
77	33	226
80	33	259
100	45	292
97	39	311
70	36	183
50	32	173
80	41	236
92	40	230
84	38	235

(a) Ajuste una ecuación de regresión múltiple de la forma $\mu_{y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.

(b) Pronostique el peso final de un animal que tiene un peso inicial de 35 kilogramos que se alimenta con 250 kilogramos de comida.

5. (a) Ajuste una ecuación de regresión múltiple de la forma $\mu_{y|x} = \beta_0 + \beta_1 x + \beta_2 x^2$ a los datos del ejemplo 11.8.

(b) Estime el rendimiento de la reacción química para una temperatura de 225°C.

6. Se lleva a cabo un experimento sobre un nuevo modelo de una marca particular de automóvil para determinar la distancia de frenado a varias velocidades. Se registran los siguientes datos.

Velocidad, v (km/hr)	35	50	65	80	95	110
Distancia de frenado, d (m)	16	26	41	62	88	119

(a) Ajuste una curva de regresión múltiple de la forma $\mu_{d|v} = \beta_0 + \beta_1 v + \beta_2 v^2$.

(b) Estime la distancia de frenado cuando el auto se desplaza a 70 kilómetros por hora.

7. Se lleva a cabo un experimento a fin de determinar si el flujo sanguíneo cerebral en seres humanos se puede predecir a partir de la tensión de oxígeno arterial (milímetros de mercurio). Se utilizan quince pacientes en el estudio y se observan los datos siguientes:

Flujo sanguíneo, y	Tensión de oxígeno arterial, x
84.33	603.40
87.80	582.50
82.20	556.20
78.21	594.60
78.44	558.90
80.01	575.20
83.53	580.10
79.46	451.20
75.22	404.00
76.58	484.00
77.90	452.40
78.80	448.40
80.67	334.80
86.60	320.30
78.20	350.30

Estime la ecuación de regresión cuadrática

$$\mu_{y|x} = \beta_0 + \beta_1 x + \beta_2 x^2.$$

8. El siguiente es un conjunto de datos experimentales de la resistencia a la compresión de una aleación particular a diversos valores de la concentración de algún aditivo:

Concentración, x	Resistencia a la compresión, y		
10.0	25.2	27.3	28.7
15.0	29.8	31.1	27.8
20.0	31.2	32.6	29.7
25.0	31.7	30.1	32.3
30.0	29.4	30.8	32.8

(a) Estime la ecuación de regresión cuadrática

$$\mu_{y|x} = \beta_0 + \beta_1 x + \beta_2 x^2.$$

(b) Pruebe la falta de ajuste del modelo.

9. Se considera que la energía eléctrica que consume una planta química cada mes está relacionada con la temperatura ambiente promedio x_1 , el número de días al mes x_2 , la pureza promedio del producto x_3 y las toneladas de producto fabricadas x_4 . Se dispone de los datos históricos del año pasado y se presentan en la siguiente tabla.

y	x_1	x_2	x_3	x_4
240	25	24	91	100
236	31	21	90	95
290	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

(a) Ajuste un modelo de regresión lineal múltiple con el uso del conjunto anterior de datos.

(b) Pronostique el consumo de energía para un mes en el que $x_1 = 75^\circ\text{F}$, $x_2 = 24$ días, y $x_3 = 98$ toneladas.

10. Dados los datos

x	0	1	2	3	4	5	6
y	1	4	5	3	2	3	4

(a) Ajuste el modelo cúbico $\mu_{y|x} = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3$.

(b) Pronostique Y cuando $x = 2$.

11. El departamento de personal de cierta empresa industrial utiliza 12 sujetos en un estudio para determinar la relación entre la clasificación de rendimiento en el trabajo (y) y las calificaciones de cuatro pruebas. Los datos son los siguientes:

y	x_1	x_2	x_3	x_4
11.2	56.5	71.0	38.5	43.0
14.5	59.5	72.5	38.2	44.8
17.2	69.2	76.0	42.5	49.0
17.8	74.5	79.5	43.4	56.3
19.3	81.2	84.0	47.5	60.2
24.5	88.0	86.2	47.4	62.0
21.2	78.2	80.5	44.5	58.1
16.9	69.0	72.0	41.8	48.1
14.8	58.1	68.0	42.1	46.0
20.0	80.5	85.0	48.1	60.3
13.2	58.3	71.0	37.5	47.1
22.5	84.0	87.2	51.0	65.2

Estime los coeficientes de regresión en el modelo

$$\mu_{y|x_1, x_2, x_3, x_4} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4$$

12. Los siguientes datos reflejan la información que se toma de 17 hospitales navales de Estados Unidos en varios lugares alrededor del mundo. Los regresores son variables de carga de trabajo, es decir, asuntos que tienen como resultado la necesidad de personal en las instalaciones de un hospital. La siguiente es una breve descripción de las variables:

- y = horas laborables mensuales,
- x_1 = carga diaria de pacientes promedio,
- x_2 = exposiciones mensuales de rayos X,
- x_3 = días-cama ocupadas mensualmente,
- x_4 = población derechohabiente en el área/1000,
- x_5 = duración promedio de la estancia del paciente, en días.

Lugar	x_1	x_2	x_3	x_4	x_5	y
1	15.57	2,463	472.92	18.0	4.45	566.51
2	44.92	2,048	1,339.75	9.5	6.92	696.82
3	20.42	3,940	620.25	12.8	4.28	1,033.15
4	18.74	6,505	568.33	36.7	3.90	1,603.62
5	49.20	5,723	1,497.60	35.7	5.50	1,611.37
6	44.92	11,520	1,365.83	24.0	4.60	1,613.27
7	55.48	5,779	1,687.00	43.3	5.62	1,854.17
8	59.28	5,969	1,639.92	46.7	5.15	2,160.55
9	94.39	8,461	2,872.33	78.7	6.18	2,305.58
10	128.02	20,106	3,655.08	180.5	6.15	3,503.93
11	96.00	13,313	2,912.00	60.9	5.88	3,571.89
12	131.42	10,771	3,921.00	103.7	4.88	3,741.40
13	127.21	15,543	3,865.67	126.8	5.50	4,026.32
14	252.90	36,194	7,684.10	157.7	7.00	10,343.81
15	409.20	34,703	12,446.33	169.4	10.78	11,732.17
16	463.70	39,204	14,098.40	331.4	7.05	15,414.94
17	510.22	86,533	15,524.00	371.6	6.35	18,854.45

El objetivo aquí es producir una ecuación empírica que estime (o pronostique) las necesidades personales de hospitales navales. Estime la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5$$

13. Se llevó a cabo un experimento para estudiar el tamaño de los calamares que comen los tiburones y atunes. Las variables de regresión son las características del pico o boca del calamar. Las variables regresoras y la respuesta que se consideran en el estudio son

- x_1 = longitud rostral, en pulgadas,
- x_2 = longitud de aleta, en pulgadas.

- x_3 = longitud rostral a muesca, en pulgadas,
- x_4 = longitud de muesca a aleta, en pulgadas,
- x_5 = ancho, en pulgadas,
- y = peso, en libras.

x_1	x_2	x_3	x_4	x_5	y
1.31	1.07	0.44	0.75	0.35	1.95
1.55	1.49	0.53	0.90	0.47	2.90
0.99	0.84	0.34	0.57	0.32	0.72
0.99	0.83	0.34	0.54	0.27	0.81
1.05	0.90	0.36	0.64	0.30	1.09
1.09	0.93	0.42	0.61	0.31	1.22
1.08	0.90	0.40	0.51	0.31	1.02
1.27	1.08	0.44	0.77	0.34	1.93
0.99	0.85	0.36	0.56	0.29	0.64
1.34	1.13	0.45	0.77	0.37	2.08
1.30	1.10	0.45	0.76	0.38	1.98
1.33	1.10	0.48	0.77	0.38	1.90
1.86	1.47	0.60	1.01	0.65	8.56
1.58	1.34	0.52	0.95	0.50	4.49
1.97	1.59	0.67	1.20	0.59	8.49
1.80	1.56	0.66	1.02	0.59	6.17
1.75	1.58	0.63	1.09	0.59	7.54
1.72	1.43	0.64	1.02	0.63	6.36
1.68	1.57	0.72	0.96	0.68	7.63
1.75	1.59	0.68	1.08	0.62	7.78
2.19	1.86	0.75	1.24	0.72	10.15
1.73	1.67	0.64	1.14	0.55	6.88

Estime la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2, x_3, x_4, x_5} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5$$

14. Veintitrés estudiantes toman parte en un programa de evaluación diseñado para medir la efectividad del maestro y determinar qué factores son importantes. Toman parte once instructoras. La medida de la respuesta es una evaluación cuantitativa que se realiza sobre el maestro que coopera. Las variables regresoras son las calificaciones de cuatro pruebas estandarizadas que se aplican a cada instructor. Los datos son:

y	x_1	x_2	x_3	x_4
410	69	125	59.00	55.66
569	57	131	31.75	63.97
425	77	141	80.50	45.32
344	81	122	75.00	46.67
324	0	141	49.00	41.21
505	53	152	49.35	43.83
235	77	141	60.75	41.61
501	76	132	41.25	64.57
400	65	157	50.75	42.41
584	97	166	32.25	57.95
434	76	141	54.50	57.90

Estime la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

15. Se lleva a cabo un estudio sobre el desgaste de un cojinete y y su relación con x_1 = viscosidad del aceite y x_2 = carga. Se obtienen los siguientes datos. [(De *Response Surface Methodology*, Myers y Montgomery (1995).]

y	x_1	x_2
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

(a) Estime los parámetros desconocidos de la ecuación de regresión lineal múltiple

$$\mu_{y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

(b) Pronostique el uso cuando la viscosidad del aceite es 20 y la carga es 1200.

16. Un ingeniero de una compañía de semiconductores desea modelar la relación entre la ganancia del dispositivo o $hFE(y)$ y tres parámetros: emisor-RS (x_1), base-RS (x_2) y emisor-base-RS (x_3). Los datos se muestran a continuación:

x_1 , Emisor-RS	x_2 , Base-RS	x_3 , E-B-RS	y , hFE-1M-5V
14,620	226.00	7.000	128.40
15,630	220.00	3,375	52.62
14,620	217.40	6,375	113.90
15,000	220.00	6,000	98.01
14,500	226.50	7,625	139.90
15,250	224.10	6,000	102.60
16,120	220.50	3,375	48.14
15,130	223.50	6,125	109.60
15,500	217.60	5,000	82.68
15,130	228.50	6,625	112.60
15,500	230.20	5,750	97.52
16,120	226.50	3,750	59.06
15,130	226.60	6,125	111.80
15,630	225.60	5,375	89.09
15,380	234.00	8,875	171.90
15,500	230.00	4,000	66.80
14,250	224.30	8,000	157.10
14,500	240.50	10,870	208.40
14,620	223.70	7,375	133.40

- (a) Ajuste una regresión lineal múltiple a los datos.
- (b) Pronostique hFE cuando $x_1 = 14$, $x_2 = 220$ y $x_3 = 5$. [Datos de Myers y Montgomery (1995).]

12.4 Propiedades de los estimadores de mínimos cuadrados

Las medias y varianzas de los estimadores B_0, B_1, \dots, B_k se obtienen fácilmente bajo ciertas suposiciones sobre los errores aleatorios E_1, E_2, \dots, E_n que son idénticas a las que se hacen en el caso de la regresión lineal simple. Suponga que consideramos que estos errores son independientes, cada uno con media cero y varianza σ^2 . Se puede mostrar entonces que B_0, B_1, \dots, B_k son, respectivamente, estimadores insesgados de los coeficientes de regresión $\beta_0, \beta_1, \dots, \beta_k$. Además, las varianzas de las B se obtienen a través de los elementos del inverso de la matriz \mathbf{A} . Se notará que los elementos fuera de la diagonal de $\mathbf{A} = \mathbf{X}'\mathbf{X}$ representan sumas de productos de elementos en las columnas de \mathbf{X} , mientras que los elementos de la diagonal de \mathbf{A} representan sumas de cuadrados de los elementos en las columnas de \mathbf{X} . La matriz inversa, \mathbf{A}^{-1} , además del multiplicador σ^2 , representa la **matriz de varianza-covarianza** de los coeficientes de regresión estimados. Es decir, los elementos de la matriz $\mathbf{A}^{-1}\sigma^2$ representan las varianzas de B_0, B_1, \dots, B_k en la diagonal principal y las covarianzas fuera de la diagonal. Por ejemplo, en un problema de regresión lineal múltiple con $k = 2$ podemos escribir

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} c_{00} & c_{01} & c_{02} \\ c_{10} & c_{11} & c_{12} \\ c_{20} & c_{21} & c_{22} \end{bmatrix}$$

con los elementos debajo de la diagonal principal determinados por la simetría de la matriz. Entonces podemos escribir

$$\begin{aligned} \sigma_{B_i}^2 &= c_{ii}\sigma^2, & i &= 0, 1, 2 \\ \sigma_{B_i, B_j} &= \text{Cov}(B_i, B_j) = c_{ij}\sigma^2, & i &\neq j. \end{aligned}$$

Por supuesto, las estimaciones de las varianzas y por ello los errores estándar de estos estimadores se obtienen al reemplazar σ^2 con la estimación apropiada que se obtiene de los datos experimentales. Una estimación insesgada de σ^2 de nuevo se define en términos de la suma de cuadrados del error, que se calcula con la fórmula que se establece en el teorema 12.1, en el cual hacemos la suposición sobre las E_i que se describieron antes.

Teorema 12.1

Para la ecuación de regresión lineal

$$y = X\beta + \varepsilon,$$

una estimación insesgada de σ^2 está dada por el cuadrado medio del error o del residuo

$$s^2 = \frac{SSE}{n - k - 1}, \quad \text{donde} \quad SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Se puede ver que el teorema 12.1 representa una generalización del teorema 11.1 para el caso de la regresión lineal simple. Se deja la demostración al lector. Como en el caso de la regresión lineal simple, la estimación s^2 es una medida de la variación

en los errores de predicción o residuos. En las secciones 12.9 y 12.10 se presentan otras inferencias importantes con respecto a la ecuación de regresión ajustada, que se basa en los valores de los residuos individuales $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$.

La suma de cuadrados de los errores y la regresión adoptan la misma forma y juegan el mismo papel que en el caso de la regresión lineal simple. De hecho, la identidad de suma de cuadrados

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

continúa siendo válida y retenemos nuestra notación anterior, a saber,

$$SST = SSR + SSE$$

con

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 = \text{suma total de cuadrados}$$

y

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \text{suma de regresión de cuadrados.}$$

Hay k grados de libertad que se asocian con SSR y, como siempre, SST tiene $n - 1$ grados de libertad. Por tanto, después de la resta, SSE tiene $n - k - 1$ grados de libertad. De esta manera nuestra estimación de σ^2 de nuevo está dada por la suma de cuadrados del error dividida entre sus grados de libertad. Estas tres sumas de cuadrados aparecerán en los impresos de la mayor parte de los paquetes de regresión múltiple para computadora.

Análisis de varianza en la regresión múltiple

La división de la suma total de cuadrados en sus componentes, la regresión y suma de cuadrados del error juega un papel importante. Se puede llevar a cabo un **análisis de varianza** para aclarar la calidad de la ecuación de regresión. Una hipótesis útil que determina si una cantidad significativa de la variación se explica mediante el modelo es

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0.$$

El análisis de varianza incluye una prueba F vía una tabla dada como a continuación:

Fuente	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Regresión	SS_R	k	$MS_R = \frac{SS_R}{k}$	$f = \frac{MS_R}{MS_E}$
Error	SS_E	$n - (k + 1)$	$MS_E = \frac{SS_E}{n - (k + 1)}$	
Total	SS_T	$n - 1$		

La prueba que se incluye es una **prueba de la cola superior**. El rechazo de H_0 implica que la **ecuación de regresión es distinta de una constante**. Es decir, al menos una variable regresora es importante. En las secciones siguientes aparece una discusión adicional del uso del análisis de varianza.

Una utilidad adicional del cuadrado medio del error (o cuadrado medio del residuo) yace en su utilización en la prueba de hipótesis y en la estimación del intervalo de confianza, que se presenta en la sección 12.5. Además, el cuadrado medio del error juega un papel importante en situaciones donde el científico busca el mejor de un conjunto de modelos en competencia. Muchos criterios de construcción de modelos incluyen la estadística s^2 . Los criterios para la comparación de modelos en competencia se presentan en la sección 12.10.

12.5 Inferencias en la regresión lineal múltiple

Una de las inferencias más útiles que se pueden hacer con respecto a la calidad de la respuesta pronosticada \hat{y}_0 que corresponde a los valores $x_{10}, x_{20}, \dots, x_{k0}$ es el intervalo de confianza sobre la respuesta media $\mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}$. Nos interesa construir un intervalo de confianza sobre la respuesta media para el conjunto de condiciones dado por

$$\mathbf{x}'_0 = [1, x_{10}, x_{20}, \dots, x_{k0}]$$

Aumentamos las condiciones sobre las x por el número 1 a fin de facilitar el uso de la notación matricial. Como en el caso $k = 1$ si hacemos la suposición adicional de que los errores son independientes y se distribuyen de forma normal, entonces las B_j son normales, con media, varianzas y covarianzas como se indica en la sección 12.4, y por ello

$$\hat{Y} = B_0 + \sum_{j=1}^k B_j x_{j0}$$

también está normalmente distribuida y es, de hecho, un estimador insesgado para la **respuesta media** sobre el que intentamos unir los intervalos de confianza. La varianza de \hat{Y}_0 escrita en notación matricial simplemente como función de σ^2 , $(\mathbf{X}'\mathbf{X})^{-1}$, y el vector de condición \mathbf{x}'_0 , es

$$\sigma_{\hat{Y}_0}^2 = \sigma^2 \mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0$$

Si esta expresión se expande para un caso dado, digamos $k = 2$, se ve fácilmente que explica de manera apropiada las varianzas y covarianzas de las B_j . Después de reemplazar σ^2 por s^2 como se da en el teorema 12.1, el intervalo de confianza de $100(1 - \alpha)\%$ sobre $\mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}$ se puede construir a partir de la estadística

$$T = \frac{\hat{Y}_0 - \mu_{Y|x_{10}, x_{20}, \dots, x_{k0}}}{S \sqrt{\mathbf{x}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0}}$$

que tiene una distribución t con $n - k - 1$ grados de libertad.

Intervalo de confianza para
 $\mu_{Y|x_1, x_2, \dots, x_k}$

Un intervalo de confianza de $(1 - \alpha)100\%$ para la **respuesta media** $\mu_{Y|x_1, x_2, \dots, x_k}$ es

$$\hat{y}_0 - t_{\alpha/2, s} \sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0} < \mu_{Y|x_1, x_2, \dots, x_k} < \hat{y}_0 + t_{\alpha/2, s} \sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0},$$

donde $t_{\alpha/2, s}$ es un valor de la distribución t con $n - k - 1$ grados de libertad.

La cantidad $s \sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}$ a menudo se llama **error estándar de predicción** y por lo general aparece en el impreso de muchos paquetes de regresión para computadora.

Ejemplo 12.5 Con el uso de los datos del ejemplo 12.3, construya un intervalo de confianza de 95% para la respuesta media cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$.

SOLUCIÓN

De la ecuación de regresión del ejemplo 12.3, el porcentaje estimado de sobrevivencia cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$ es

$$\hat{y} = 39.1574 + (1.0161)(3) - (1.8616)(8) - (0.3433)(9) = 24.2232.$$

A continuación encontramos que

$$\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0 = [1, 3, 8, 9] \begin{bmatrix} 8.064 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 8 \\ 9 \end{bmatrix}$$

$$= 0.1267.$$

Con el uso del cuadrado medio del error, $s^2 = 4.298$ o $s = 2.073$, y de la tabla A.4, vemos que $t_{0.025, 9} = 2.262$ para 9 grados de libertad. Por tanto, un intervalo de confianza de 95% para el porcentaje medio de sobrevivencia para $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$ está dado por

$$24.2232 - (2.262)(2.073) \sqrt{0.1267} < \mu_{Y|3, 8, 9} < 24.2232 + (2.262)(2.073) \sqrt{0.1267}$$

o simplemente

$$22.5541 < \mu_{Y|3, 8, 9} < 25.8923.$$

Como en el caso de la regresión lineal simple, necesitamos hacer una clara distinción entre el intervalo de confianza de la respuesta media y el intervalo de predicción sobre una *respuesta observada*. Esta última proporciona un límite dentro del cual podemos decir con un grado de certeza preestablecido que caerá una nueva respuesta observada.

Un intervalo de predicción para una sola respuesta pronosticada y_0 se establece de nuevo al considerar las diferencias $\hat{y}_0 - y_0$ de la variable aleatoria $\hat{Y}_0 - Y_0$. Se puede mostrar que la distribución muestral es normal con media

$$\mu_{\hat{Y}_0 - Y_0} = 0$$

y varianza

$$\sigma_{\hat{Y}_0 - Y_0}^2 = \sigma^2 [1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0].$$

De esta manera el intervalo de predicción de $(1 - \alpha)100\%$ para un solo valor de predicción y_0 se puede construir a partir de la estadística

$$T = \frac{\hat{Y}_0 - Y_0}{S \sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}}$$

que tiene una distribución t con $n - k - 1$ grados de libertad. ■

Intervalo de predicción para y_0

Un intervalo de predicción de $(1 - \alpha)100\%$ para una sola respuesta y_0 está dado por

$$\hat{y}_0 - t_{\alpha/2} s \sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0} < y_0 < \hat{y}_0 + t_{\alpha/2} s \sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}$$

donde $t_{\alpha/2}$ es un valor de la distribución t con $n - k - 1$ grados de libertad.

Ejemplo 12.6 Con el uso de los datos del ejemplo 12.3, construya un intervalo de predicción de 95% para una respuesta individual del porcentaje de sobrevivencia cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$.

SOLUCIÓN

Con referencia a los resultados del ejemplo 12.5, encontramos que el intervalo de predicción de 95% para la respuesta y_0 cuando $x_1 = 3\%$, $x_2 = 8\%$, y $x_3 = 9\%$ es

$$24.2232 - (2.262)(2.073) \sqrt{1.1267} < y_0 < 24.2232 + (2.262)(2.073) \sqrt{1.1267}$$

que se reduce a $19.2459 < y_0 < 29.2005$. Nótese, como se espera, que el intervalo de predicción es considerablemente menos estrecho que el intervalo de confianza para el porcentaje de sobrevivencia media en el ejemplo 12.5. ■

Un conocimiento de las distribuciones de los estimadores de los coeficientes individuales permite al experimentador construir intervalos de confianza para los coeficientes y probar hipótesis acerca de ellos. Recuerdese de la sección 12.4 que las B_j ($j = 0, 1, 2, \dots, k$) se distribuyen de forma normal con media β_j y varianza $c_j \sigma^2$. De esta manera podemos utilizar la estadística

$$T = \frac{B_j - \beta_j}{S \sqrt{c_j}}$$

con $n - k - 1$ grados de libertad para probar las hipótesis y construir intervalos de confianza sobre β_j . Por ejemplo, si deseamos probar

$$H_0: B_j = \beta_{j0}$$

$$H_1: B_j \neq \beta_{j0}$$

calculamos la estadística

$$t = \frac{b_j - \beta_{j0}}{s \sqrt{c_j}}$$

y no rechazamos H_0 si $-t_{\alpha/2} < t < t_{\alpha/2}$, donde $t_{\alpha/2}$ tiene $n - k - 1$ grados de libertad.

Ejemplo 12.7 Para el modelo del ejemplo 12.3, pruebe la hipótesis de que $\beta_2 = -2.5$ en el nivel de significancia 0.05 contra la alternativa de que $\beta_2 > -2.5$.

SOLUCIÓN

$$H_0: \beta_2 = -2.5,$$

$$H_1: \beta_2 > -2.5.$$

Cálculos:

$$t = \frac{b_2 - \beta_{20}}{s \sqrt{c_{22}}} = \frac{-1.8616 + 2.5}{2.073 \sqrt{0.0166}} = 2.391,$$

$$P = P(T > 2.391) = 0.04.$$

Decisión: rechazar H_0 y concluir que $\beta_2 > -2.5$. ■

Impreso de computadora comentado para los datos del ejemplo 12.3

La figura 12.1 muestra un impreso de computadora comentado para el ajuste de una regresión lineal para los datos del ejemplo 12.3. El paquete que se utiliza es el SAS.

Nótese las estimaciones del parámetro del modelo, los errores estándar y la estadística t . Los errores estándar se calculan a partir de las raíces cuadradas de los elementos de la diagonal de $(X'X)^{-1}s^2$. En esta ilustración la variable x_3 es insignificante ante x_1 y x_2 . Los términos CLM y CLI son intervalos de confianza sobre la

Analysis of Variance							
Source	DF	Sum of Squares	Mean Square	F Value	Prob > F		
Model	3	399.45437	133.15146	30.984	0.0001		
Error	9	36.67640	4.09738				
C Total	12	436.13077					
Root MSE		2.0730	R-square	0.9117			
Dep Mean		29.03646					
C.V.		7.13665					
Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	T for HO: Parameter = 0	Prob > T		
INTERCEPT	1	39.157350	5.66705963	6.911	0.0001		
X1	1	1.016100	0.190089520	5.323	0.0005		
X2	1	-1.661649	0.26732550	-6.964	0.0001		
X3	1	-0.343260	0.61705204	-0.556	0.5916		
OBS	Y	PREDICTED	RESIDUAL	CLM_L	CLM_U	CLI_L	CLI_U
1	25.9	27.3914	-1.85241	24.1500	30.5928	21.6733	31.0299
2	31.2	32.2623	-1.06212	30.4875	34.0371	27.2462	37.2744
3	25.9	27.3495	-1.44955	24.2757	30.4234	21.7424	32.9567
4	36.4	36.3096	0.09042	35.4098	41.2093	32.7960	43.8232
5	18.4	15.5447	0.65527	11.9729	19.1165	9.16499	21.4396
6	26.7	26.1081	0.59193	23.7649	28.4512	20.8657	31.3504
7	26.4	28.2532	-1.85316	26.4222	30.0641	23.2189	33.2874
8	25.9	26.2219	-0.32155	24.0204	28.4233	21.0413	31.4024
9	32.0	32.0882	-0.08838	30.3174	33.6589	27.0255	37.1009
10	25.2	26.0674	-0.86830	24.5824	27.6329	21.1224	31.0115
11	39.7	37.2524	2.44764	34.2957	40.2090	31.7066	42.7961
12	35.7	32.4879	3.21208	29.1743	35.6016	26.7458	38.2300
13	26.5	26.2032	-1.70324	25.9771	30.4894	23.0121	33.3943

Figura 12.1 Impreso del SAS para los datos del ejemplo 12.3.

respuesta media y los límites de predicción sobre una observación individual, respectivamente. La prueba f en el análisis de varianza indica que se explica una cantidad insignificante de variabilidad. Como ejemplo de la interpretación de CLM y CLI, considérese la observación 10. Con una observación de 25.2 y un valor pronosticado de 26.067, tenemos una confianza de 95% de que la respuesta media esté entre 24.502 y 27.633, y que una nueva observación caerá entre 21.124 y 31.011 con probabilidad 0.95. El valor R^2 de 0.9117 implica que el modelo explica 91.17% de la variabilidad en la respuesta. Más discusión acerca de R^2 aparece en la sección 12.5.

Más sobre el análisis de varianza en la regresión múltiple

En la sección 12.4 discutimos brevemente la división de la suma total de cuadrados $\sum_{i=1}^n (y_i - \bar{y})^2$ en sus dos componentes, las sumas de cuadrados de regresión y del error (que se ilustran en la figura 12.1). El análisis de varianza conduce a una prueba de

$$H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_k = 0.$$

El rechazo de la hipótesis nula tiene una interpretación importante para el científico o el ingeniero. (Para los que se interesan en un tratamiento adicional del tema con el uso de matrices, es útil discutir el desarrollo de estas sumas de cuadrados que se utilizan en el ANOVA.)

Primero recuerde la definición de \mathbf{y} , \mathbf{X} y $\boldsymbol{\beta}$ en la sección 12.3, así como \mathbf{b} , el vector de estimadores de mínimos cuadrados dado por

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

Una división de la **suma de cuadrados sin corrección**

$$\mathbf{y}'\mathbf{y} = \sum_{i=1}^n y_i^2$$

en dos componentes está dada por

$$\mathbf{y}'\mathbf{y} = \mathbf{b}'\mathbf{X}'\mathbf{y} + (\mathbf{y}'\mathbf{y} - \mathbf{b}'\mathbf{X}'\mathbf{y})$$

$$= \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + [\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}].$$

El segundo término (entre corchetes) en el lado derecho simplemente es la suma de cuadrados del error $\sum_{i=1}^n (y_i - \hat{y}_i)^2$. El lector debe ver que una expresión alternativa para la suma de cuadrados del error es

$$SSE = \mathbf{y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}.$$

El término $y'X(X'X)^{-1}X'y$ se llama **suma de cuadrados de regresión**. Sin embargo, no es la expresión $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ que se usa para probar la "importancia" de los términos b_1, b_2, \dots, b_k sino, más bien,

$$y'X(X'X)^{-1}X'y = \sum_{i=1}^n \hat{y}_i^2,$$

que es una suma de cuadrados de regresión sin corrección para la media. Como tal sólo se utilizaría para probar si "la ecuación de regresión difiere de manera significativa de cero". Es decir,

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0.$$

En general, esto no es tan importante como probar

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

pues ésta establece que la respuesta media es una constante, no necesariamente cero.

Grados de libertad

De esta forma la división de sumas de cuadrados y grados de libertad se reducen a

Fuente	Suma de cuadrados	g.l.
Regresión	$\sum_{i=1}^n \hat{y}_i^2 = y'X(X'X)^{-1}X'y$	$k + 1$
Error	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = y'[I - X(X'X)^{-1}X']y$	$n - (k + 1)$
Total	$\sum_{i=1}^n y_i^2 = y'y$	n

Hipótesis de interés

Ahora bien, por supuesto, la hipótesis de interés para una ANOVA debe eliminar el papel de la intersección por lo que se describe arriba. Hablando estrictamente, si $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, entonces la línea de regresión estimada es simplemente $\hat{y}_i = \bar{y}$. Como resultado en realidad se busca evidencia de que la ecuación de regresión "varíe desde una constante". Así, las sumas de cuadrados total y de regresión se deben "corregir por la media". Como resultado tenemos

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

que se presenta en la sección 12.3.

En notación matricial ésta es simplemente:

$$y'[I - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']y = y'[X(X'X)^{-1}X' - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']y + [y'y - y'X(X'X)^{-1}X'y].$$

En esta expresión $\mathbf{1}$ es sólo un vector de n unos. Como resultado simplemente restamos

$$\mathbf{y}'\mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{y} = \frac{\left(\sum_{i=1}^n y_i\right)^2}{n}$$

de $\mathbf{y}'\mathbf{y}$ y de $\mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, es decir, corregimos las sumas de cuadrados total y de regresión por la media.

Finalmente, la división apropiada de sumas de cuadrados con grados de libertad es la siguiente:

Fuente	Suma de cuadrados	g.l.
Regresión	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \mathbf{y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$	k
Error	$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \mathbf{y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y}$	$n - (k + 1)$
Total	$\sum_{i=1}^n (y_i - \bar{y})^2 = \mathbf{y}'[\mathbf{I} - \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$	$n - 1$

Ésta es la tabla de ANOVA que aparece en el impreso de computadora de la figura 12.1. La expresión $\mathbf{y}'[\mathbf{I}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}']\mathbf{y}$ a menudo se llama **suma de cuadrados de regresión asociada con la media**, y se le asigna 1 grado de libertad.

Ejercicios

- Estime σ^2 para los datos del ejercicio 2 de la página 414.
- Estime σ^2 para los datos del ejercicio 3 de la página 414.
- Estime σ^2 para los datos del ejercicio 9 de la página 416.
- Obtenga estimaciones de las varianzas y covarianzas de los estimadores B_1 y B_2 del ejercicio 2 de la página 414.
- Con referencia al ejercicio 9 de la página 416, encuentre la estimación de
 - $\sigma_{B_1}^2$.
 - $\text{Cov}(B_1, B_2)$.
- Con el uso de los datos del ejercicio 2 de la página 414 y la estimación de σ^2 del ejercicio 1, calcule intervalos de confianza del 95% para la respuesta pronosticada y la respuesta media cuando $x_1 = 900$ y $x_2 = 1.00$.
- Para el ejercicio 8 de la página 415, construya un intervalo de confianza de 90% para la resistencia media a la compresión cuando la concentración es $x = 19.5$ y se utiliza un modelo cuadrático.
- Con el uso de los datos del ejercicio 9 de la página 416 y la estimación de σ^2 del ejercicio 3 anterior,

calcule intervalos de confianza de 95% para la respuesta pronosticada y la respuesta media cuando $x_1 = 8.2$, $x_2 = 6.0$, $x_3 = 10.3$, y $x_4 = 5.8$.

- Para el modelo del ejercicio 7 en la página 415, pruebe la hipótesis de que $\beta_2 = 0$ en el nivel de significancia de 0.05 contra la alternativa de que $\beta_2 \neq 0$.
- Para el modelo del ejercicio 2 en la página 414, pruebe la hipótesis de que $\beta_1 = 0$ en el nivel de significancia de 0.05 contra la alternativa de que $\beta_1 \neq 0$.
- Para el modelo del ejercicio 3 de la página 414, pruebe la hipótesis de que $\beta_1 = 2$ contra la alternativa de que $\beta_1 \neq 2$. Utilice un valor P en su conclusión.
- Considere los siguientes datos que se listan en el ejercicio 15 de la página 417.

y (desgaste)	x_1 (viscosidad del aceite)	x_2 (carga)
193	1.6	853
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

- (a) Estime σ^2 con el uso de la regresión múltiple de y sobre x_1 y x_2 .
- (b) Calcule valores pronosticados, un intervalo de confianza de 95% para el desgaste medio y un intervalo de predicción de 95% para el desgaste observado si $x_1 = 20$ y $x_2 = 1000$.

13. Con el uso de los datos del ejercicio 12 anterior, pruebe.

$$(a) H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

y

$$(b) H_0: \beta_2 = 0$$

$$H_1: \beta_2 \neq 0$$

(c) ¿Tiene alguna razón para creer que el modelo del ejercicio 12 se debe cambiar? ¿Por qué sí o por qué no?

12.6 Elección de un modelo de ajuste a través de la prueba de hipótesis

En muchas situaciones de regresión, los coeficientes individuales son de importancia para el experimentador. Por ejemplo, en una aplicación de economía, β_1, β_2, \dots , pueden tener algún significado particular, y por ello los intervalos de confianza y pruebas de hipótesis sobre estos parámetros son de interés para el economista. Sin embargo, considere una situación química industrial en la que el modelo postulado supone que el rendimiento de la reacción depende linealmente de la temperatura de reacción y concentración de cierto catalizador. Probablemente se sabe que éste no es el modelo correcto sino una aproximación adecuada, por lo que el interés probablemente no está en los parámetros individuales sino más bien en la capacidad de toda la función para predecir la respuesta correcta en el rango de las variables que se consideran. Por tanto, en esta situación, se podría poner más énfasis en σ_y^2 , en los intervalos de confianza sobre la respuesta media, etcétera, y probablemente quitar el énfasis a las inferencias sobre los parámetros individuales.

El experimentador que utiliza el análisis de regresión también se interesa en la supresión de variables cuando la situación dicta que, además de llegar a una ecuación de predicción más fácil de trabajar, debe encontrar la "mejor regresión" que incluya sólo variables que sean pronosticadores útiles. Se dispone de una cantidad de programas de computadora que llegan de manera secuencial a la así llamada mejor ecuación de regresión que depende de ciertos criterios. Discutimos esto con más detalle en la sección 12.8.

Un criterio que por lo general se utiliza para ilustrar lo adecuado de un modelo de regresión ajustado es el **coeficiente de determinación múltiple**:

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Esta cantidad sólo indica qué proporción de la variación total en la respuesta Y se explica con el modelo ajustado. A menudo un experimentador reportará $R^2 \times 100\%$ e interpretará el resultado como una variación porcentual explicada por el modelo que se postula. La raíz cuadrada de R^2 se llama **coeficiente de correlación múltiple** entre Y y el conjunto x_1, x_2, \dots, x_k . En el ejemplo 12.3, se encuentra que

el valor de R^2 que indica la proporción de variación explicada por las tres variables independientes x_1 , x_2 , y x_3 es

$$R^2 = \frac{SSR}{SST} = \frac{399.45}{438.13} = 0.9117,$$

que significa que 91.17% de la variación en el porcentaje de sobrevivencia se explica con el modelo de regresión lineal.

La suma de cuadrados de regresión se puede utilizar para dar alguna indicación con respecto a si el modelo es o no una explicación adecuada de la situación real. Se puede probar la hipótesis H_0 de que la regresión no es significativa al formar simplemente la razón

$$f = \frac{SSR/k}{SSE/(n - k - 1)} = \frac{SSR/k}{s^2}$$

y rechazar H_0 en el nivel de significancia α cuando $f > f_{\alpha}(k, n - k - 1)$. Para los datos del ejemplo 12.3 obtenemos

$$f = \frac{399.45/3}{4.298} = 30.98.$$

Del impreso de la figura 12.1 el valor P es menor que 0.0001. Esto no se debe interpretar mal. Aunque indica que la regresión que explica el modelo es significativa, no excluye la posibilidad de que

1. El modelo de regresión lineal en este conjunto de x no es el único modelo que se puede utilizar para explicar los datos; en realidad, puede haber otros modelos con transformaciones en las x que pueden dar un valor más grande de la estadística F .
2. El modelo podría ser más efectivo con la inclusión de otras variables además de x_1 , x_2 , y x_3 o quizá con la supresión de una o más de las variables en el modelo, digamos x_3 , que muestra $P = 0.5916$.

Pruebas sobre subconjuntos y coeficientes individuales

La inclusión de cualquier variable única en un sistema de regresión *aumentará la suma de cuadrados de regresión* y por ello *reducirá la suma de cuadrados del error*. En consecuencia, debemos decidir si el aumento en la regresión es suficiente para garantizar su uso en el modelo. Como podríamos esperar, el uso de variables sin importancia puede reducir la efectividad de la ecuación de predicción al aumentar la varianza de la respuesta estimada. Proseguiremos este punto con más detalle mediante la consideración de la importancia de x_3 en el ejemplo 12.3. Inicialmente, podemos probar

$$H_0: \beta_3 = 0,$$

$$H_1: \beta_3 \neq 0$$

con el uso de la distribución t con 9 grados de libertad. Tenemos

$$t = \frac{b_3 - 0}{S \sqrt{c_{33}}} = \frac{-0.3433}{2.073 \sqrt{0.0886}} = -0.556,$$

que indica que β_3 no difiere de forma significativa de cero, y por ello se puede justificar muy bien la eliminación de x_3 del modelo. Suponga que consideramos la regresión de Y sobre el conjunto (x_1, x_2) , la ecuación normal de mínimos cuadrados se reduce ahora a

$$\begin{bmatrix} 13 & 59.43 & 81.82 \\ 59.43 & 394.7255 & 360.6621 \\ 81.82 & 360.6621 & 576.7264 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 377.75 \\ 1877.5670 \\ 2246.6610 \end{bmatrix}.$$

Los coeficientes de regresión estimados para este modelo reducido son

$$b_0 = 36.094, \quad b_1 = 1.031, \quad b_2 = -1.870,$$

y la suma de cuadrados de regresión que resulta con 2 grados de libertad es:

$$R(\beta_1, \beta_2) = 398.12.$$

Aquí utilizamos la notación $R(\beta_1, \beta_2)$ para indicar la suma de cuadrados de regresión del modelo restringido y no se debe confundir con SSR , la suma de cuadrados de regresión del modelo original con 3 grados de libertad. La nueva suma de cuadrados del error es entonces

$$SST - R(\beta_1, \beta_2) = 438.13 - 398.12 = 40.01,$$

y el cuadrado medio del error con 10 grados de libertad que resulta se convierte en

$$s^2 = \frac{40.01}{10} = 4.001.$$

La cantidad de variación en la respuesta, el porcentaje de sobrevivencia, que se atribuye a x_3 , el porcentaje por peso de la tercera adición, en presencia de las variables x_1 y x_2 , es

$$R(\beta_3 | \beta_1, \beta_2) = SSR - R(\beta_1, \beta_2) = 399.45 - 398.12 = 1.33,$$

que representa una pequeña parte de toda la variación de la regresión. Esta cantidad de regresión adicional es estadísticamente insignificante, como lo indica nuestra prueba anterior sobre β_3 . Una prueba equivalente que incluye la formación de la razón

$$f = \frac{R(\beta_3 | \beta_1, \beta_2)}{s^2} = \frac{1.33}{4.298} = 0.309,$$

que es un valor de la distribución F con 1 y 9 grados de libertad. Recuerde que la relación básica entre la distribución t con v grados de libertad y la distribución F con 1 y v grados de libertad es

$$t_{\alpha/2}^2 = f_{\alpha}(1, v).$$

y notamos que el valor f de 0.309 es en realidad el cuadrado del valor t de 0.556,

Podemos proporcionar un apoyo adicional para eliminar x_3 del modelo al considerar $\hat{\sigma}_{\hat{Y}_0}^2$ en la ecuación de regresión completa y en la reducida. Notamos, primero, que el cuadrado medio del error, s^2 , se reduce de 4,298 a 4,001 al eliminar x_3 del modelo. En el caso del modelo completo, una estimación de la varianza de la respuesta estimada \hat{Y}_0 , cuando $x_1 = 4.00$, $x_2 = 5.5$, y $x_3 = 8.90$, es

$$\begin{aligned}\hat{\sigma}_{\hat{Y}_0}^2 &= s^2 \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0 \\ &= 4.298 [1, 4.00, 5.5, 8.90] \\ &\quad \times \begin{bmatrix} 8.064 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & -0.0021 & 0.0886 \end{bmatrix} \begin{bmatrix} 1 \\ 4.00 \\ 5.5 \\ 8.90 \end{bmatrix} \\ &= 0.3936.\end{aligned}$$

Para el modelo reducido nuestro punto de interés se vuelve $\mathbf{x}_0' = [1, 4.00, 5.5]$; nuestra matriz $\mathbf{X}'\mathbf{X}$ se reduce a una matriz 3×3 con inversa

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 1.0114 & -0.0494 & -0.1126 \\ -0.0494 & 0.0083 & 0.0018 \\ -0.1126 & 0.0018 & 0.0166 \end{bmatrix}$$

y la varianza de la respuesta estimada se estima ahora como

$$\begin{aligned}\hat{\sigma}_{\hat{Y}_0}^2 &= 4.001 [1, 4.00, 5.5] \begin{bmatrix} 1.0114 & -0.0494 & -0.1126 \\ -0.0494 & 0.0083 & 0.0018 \\ -0.1126 & 0.0018 & 0.0166 \end{bmatrix} \begin{bmatrix} 1 \\ 4.00 \\ 5.5 \end{bmatrix} \\ &= 0.3668,\end{aligned}$$

que representa una reducción sobre la que se encuentra para el modelo completo.

Para generalizar los conceptos anteriores, se puede evaluar el trabajo de una variable independiente x_i en el modelo general de regresión lineal múltiple

$$\mu_{Y|x_1, x_2, \dots, x_k} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

al observar la cantidad de regresión que se atribuye a x_i , además de la que se atribuye a las otras variables; es decir, la regresión sobre x_i ajustada para las otras variables. Esto se calcula mediante la resta de la suma de cuadrados de regresión para un modelo con x_i que se elimina de SSR. Por ejemplo, decimos que x_1 se evalúa al calcular

$$R(\beta_1 | \beta_2, \beta_3, \dots, \beta_k) = SSR - R(\beta_2, \beta_3, \dots, \beta_k),$$

donde $R(\beta_2, \beta_3, \dots, \beta_k)$ es la suma de cuadrados de regresión con $\beta_1 x_1$ eliminada del modelo. Para probar la hipótesis

$$H_0: \beta_1 = 0,$$

$$H_1: \beta_1 \neq 0$$

calcule

$$f = \frac{R(\beta_1 | \beta_2, \beta_3, \dots, \beta_k)}{s^2}$$

y compararla con $f_\alpha(1, n - k - 1)$.

De manera similar podemos probar la significancia de un conjunto de las variables. Por ejemplo, para investigar de manera simultánea la importancia de incluir x_1 y x_2 en el modelo, probamos la hipótesis

$$H_0: \beta_1 = \beta_2 = 0,$$

$$H_1: \beta_1 \text{ y } \beta_2 \text{ no son ambas cero,}$$

mediante el cálculo de

$$f = \frac{[R(\beta_1, \beta_2 | \beta_3, \beta_4, \dots, \beta_k)]/2}{s^2} = \frac{[SSR - R(\beta_3, \beta_4, \dots, \beta_k)]/2}{s^2}$$

y la comparación de ésta con $f_\alpha(2, n - k - 1)$. El número de grados de libertad asociados con el numerador, en este caso 2, es igual al número de variables en el conjunto.

12.7 Caso especial de ortogonalidad

Antes de nuestro desarrollo original del problema general de regresión lineal se realizó la suposición de que las variables independientes se miden sin error y por lo general están controladas por el experimentador. Muy a menudo éstas ocurren como resultado de un *diseño experimental* muy elaborado. De hecho, se puede aumentar la efectividad de la ecuación de predicción que resulta con el uso de un plan experimental adecuado.

Suponga que una vez más consideramos la matriz \mathbf{X} como se define en la sección 12.3. La podemos reescribir como

$$\mathbf{X} = [\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k],$$

donde $\mathbf{1}$ representa una columna de unos y \mathbf{x}_j es un vector columnas que representa los niveles de x_j . Si

$$\mathbf{x}_p' \mathbf{x}_q = 0, \quad p \neq q,$$

se dice que las variables x_p y x_q son *ortogonales* entre ellas. Hay ciertas ventajas obvias para tener una situación completamente ortogonal por la que $\mathbf{x}_p' \mathbf{x}_q = 0$ para todas las posibles p y q , $p \neq q$, y, además,

$$\sum_{i=1}^n x_{ij} = 0, \quad j = 1, 2, \dots, k.$$

La $X'X$ que resulta es una matriz diagonal y las ecuaciones normales de la sección 10.3 se reducen a

$$\begin{aligned} nb_0 &= \sum_{i=1}^n y_i, \\ b_1 \sum_{i=1}^n x_{1i}^2 &= \sum_{i=1}^n x_{1i} y_i, \\ &\vdots \\ b_k \sum_{i=1}^n x_{ki}^2 &= \sum_{i=1}^n x_{ki} y_i. \end{aligned}$$

Una ventaja importante es que con facilidad se es capaz de dividir SSR en **componentes de un solo grado de libertad**, cada uno de los cuales corresponde a la cantidad de variación en Y explicada por una variable controlada dada. En la situación ortogonal podemos escribir

$$\begin{aligned} SSR &= \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^n (b_0 + b_1 x_{1i} + \dots + b_k x_{ki} - b_0)^2 \\ &= b_1^2 \sum_{i=1}^n x_{1i}^2 + b_2^2 \sum_{i=1}^n x_{2i}^2 + \dots + b_k^2 \sum_{i=1}^n x_{ki}^2 \\ &= R(\beta_1) + R(\beta_2) + \dots + R(\beta_k). \end{aligned}$$

La cantidad $R(\beta_i)$ es la cantidad de la suma de cuadrados de regresión que se asocia con un modelo que incluye una sola variable independiente x_i .

Para probar de manera simultánea la significancia de un conjunto de m variables en una situación ortogonal, la suma de cuadrados de regresión se convierte en

$$R(\beta_1, \beta_2, \dots, \beta_m | \beta_{m+1}, \beta_{m+2}, \dots, \beta_k) = R(\beta_1) + R(\beta_2) + \dots + R(\beta_m),$$

y de esta forma tenemos la simplificación adicional

$$R(\beta_1 | \beta_2, \beta_3, \dots, \beta_k) = R(\beta_1)$$

cuando evaluamos una sola variable independiente. Por tanto, la contribución de una variable dada o conjunto de variables se encuentra esencialmente al *ignorar* las otras variables en el modelo. Las evaluaciones independientes del mérito de las variables individuales se llevan a cabo con el uso de las técnicas de análisis de varianza que se dan en la tabla 12.1. La variación total en la respuesta se divide en componentes de un solo grado de libertad más el término de error con $n - k - 1$ grados de libertad. Cada valor f calculado se utiliza para probar una de las hipótesis

$$\left. \begin{aligned} H_0: & \beta_i = 0 \\ H_1: & \beta_i \neq 0 \end{aligned} \right\} \quad i = 1, 2, \dots, k,$$

al compararlás con el punto crítico $f_{\alpha}(1, n - k - 1)$.

Ejemplo 12.8 Suponga que un científico recopila datos experimentales del radio de un grano de propulsor Y como función de la temperatura del polvo x_1 , la tasa de

Tabla 12.1 Análisis de varianza para variables ortogonales

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f Calculada
β_1	$R(\beta_1) = b_1^2 \sum_{i=1}^n x_{1i}^2$	1	$R(\beta_1)$	$\frac{R(\beta_1)}{s^2}$
β_2	$R(\beta_2) = b_2^2 \sum_{i=1}^n x_{2i}^2$	1	$R(\beta_2)$	$\frac{R(\beta_2)}{s^2}$
\vdots	\vdots	\vdots	\vdots	\vdots
β_k	$R(\beta_k) = b_k^2 \sum_{i=1}^n x_{ki}^2$	1	$R(\beta_k)$	$\frac{R(\beta_k)}{s^2}$
Error	SSE	$n - k - 1$	$s^2 = \frac{SSE}{n - k - 1}$	
Total	$SST = S_{yy}$	$n - 1$		

expulsión x_2 , y la temperatura de la matriz x_3 . Ajuste un modelo de regresión lineal para predecir el radio del grano y determine la efectividad de cada variable en el modelo. Los datos son los siguientes:

Radio del grano	Temperatura del polvo	Tasa de expulsión	Temperatura de la matriz
82	150 (-1)	12 (-1)	220 (-1)
93	190 (1)	12 (-1)	220 (-1)
114	150 (-1)	24 (1)	220 (-1)
124	150 (-1)	12 (-1)	250 (1)
111	190 (1)	24 (1)	220 (-1)
129	190 (1)	12 (-1)	250 (1)
157	150 (-1)	24 (1)	250 (1)
164	190 (1)	24 (1)	250 (1)

SOLUCIÓN

Nótese que cada variable se controla en dos niveles, y que el experimento representa cada una de las ocho combinaciones posibles. Los datos de las variables independientes se codifican por conveniencia mediante las siguientes fórmulas:

$$x_1 = \frac{\text{temperatura del polvo} - 170}{20}$$

$$x_2 = \frac{\text{tasa de expulsión} - 18}{6}$$

$$x_3 = \frac{\text{temperatura de la matriz} - 235}{15}$$

Los niveles que resultan de x_1 , x_2 , y x_3 toman los valores -1 y $+1$ como se indica en la tabla de datos. Este *diseño experimental* particular proporciona la ortogonalidad que aquí ilustramos. Un tratamiento más completo de este tipo de disposición experimental se presenta en el capítulo 15. La matriz \mathbf{X} es

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

y las condiciones de ortogonalidad se verifican con facilidad. Se pueden calcular ahora los coeficientes

$$b_0 = \frac{\sum_{i=1}^8 y_i}{8} = 121.75, \quad b_1 = \frac{\sum_{i=1}^8 x_{1i} y_i}{8} = \frac{20}{8} = 2.5,$$

$$b_2 = \frac{\sum_{i=1}^8 x_{2i} y_i}{\sum_{i=1}^8 x_{2i}^2} = \frac{118}{8} = 14.75, \quad b_3 = \frac{\sum_{i=1}^8 x_{3i} y_i}{\sum_{i=1}^8 x_{3i}^2} = \frac{174}{8} = 21.75,$$

por lo que en términos de las variables *codificadas*, la ecuación de predicción es

$$\hat{y} = 121.75 + 2.5x_1 + 14.75x_2 + 21.75x_3.$$

El análisis de varianza de la tabla 12.2 muestra contribuciones independientes a *SSR* para cada variable. Los resultados, cuando se comparan con el punto crítico $f_{0.05}(1, 4)$ de 7.71, indican que x_1 no contribuye de manera significativa en el nivel 0.05, mientras que las variables x_2 y x_3 son significativas. En este ejemplo la estimación para σ^2 es 23.1250. Como para el caso de una sola variable independiente, se señala que esta estimación no sólo contiene la variación del error experimental a menos que el modelo que se postula sea correcto. De otra manera, la estimación está "contaminada" por la falta de ajuste además del error puro, y la falta de ajuste se puede separar sólo si se obtienen múltiples observaciones experimentales en las diversas combinaciones (x_1, x_2, x_3) . ■

Tabla 12.2 Análisis de varianza para los datos del radio del grano.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
β_1	$(2.5)^2(8) = 50$	1	50	2.16
β_2	$(14.75)^2(8) = 1740.50$	1	1740.50	75.26
β_3	$(21.75)^2(8) = 3784.50$	1	3784.50	163.65
Error	92.5	4	23.1250	
Total	5667.50	7		

Ejercicios

1. Calcule e interprete el coeficiente de determinación múltiple para las variables del ejercicio 3 de la página 414.

2. Pruebe si la regresión que se explica con el modelo del ejercicio 3 de la página 414 es significativa en el nivel de significancia de 0.01.

3. Pruebe si la regresión que explica el modelo del ejercicio 9 de la página 416 es significativa en el nivel de significancia de 0.01.

4. Para el modelo del ejercicio 9 de la página 416, pruebe la hipótesis:

$$H_0: \beta_1 = \beta_2 = 0,$$

$$H_1: \beta_1 \text{ y } \beta_2 \text{ no son ambos cero.}$$

5. Repita el ejercicio 10 de la página 426 con el uso de una estadística F .

6. Se lleva a cabo un experimento para ajustar una ecuación de regresión múltiple que relacione la producción y con la temperatura x_1 , el tiempo de reacción x_2 y la concentración de uno de los reactivos x_3 . Se eligen dos niveles de cada variable y las mediciones correspondientes a las variables independientes codificadas se registran a continuación:

y	x_1	x_2	x_3
7.6	-1	-1	-1
8.4	1	-1	-1
9.2	-1	1	-1
10.3	-1	-1	1
9.8	1	1	-1
11.1	1	-1	1
10.2	-1	1	1
12.6	1	1	1

(a) Utilice las variables codificadas y estime la ecuación de regresión lineal múltiple

$$\mu_{Y|X_1, X_2, X_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

(b) Divida SSR , la suma de cuadrados de regresión, en tres componentes de un solo grado de libertad atribuibles a x_1 , x_2 y x_3 , respectivamente. Muestre una tabla de análisis de varianzas e indique las pruebas de significancia en cada variable.

7. Considere los datos de energía eléctrica del ejercicio 9 de la página 416. Pruebe $H_0: \beta_1 = \beta_2 = 0$ haga uso de $R(\beta_1, \beta_2 | \beta_0, \beta_3)$. Dé un valor P y extraiga conclusiones.

8. Considere los datos para el ejercicio 6 anterior. Calcule la tabla

$$R(\beta_1 | \beta_0), \quad R(\beta_1 | \beta_0, \beta_2, \beta_3),$$

$$R(\beta_2 | \beta_1, \beta_0), \quad R(\beta_2 | \beta_0, \beta_1, \beta_3),$$

$$R(\beta_3 | \beta_0, \beta_1, \beta_2).$$

Haga comentarios.

9. Considere los datos del ejercicio 9 de la página 402. Ajuste un modelo de regresión con el uso del peso y la razón de tracción como variables explicatorias. Compare este modelo con el modelo SLR que usa sólo el peso.

12.8 Métodos secuenciales para la selección del modelo

En ocasiones las pruebas de significancia que se señalan en la sección 12.6 son bastante adecuadas para determinar qué variables se deben utilizar en el modelo final de regresión. Estas pruebas son en realidad efectivas si el experimento se puede planear y las variables son ortogonales entre ellas. Incluso si las variables no son ortogonales, las pruebas t individuales pueden ser de alguna utilidad en muchos problemas donde el número de variables bajo investigación es pequeño. Sin embargo, hay muchos problemas donde es necesario utilizar técnicas más elaboradas para depurar variables, en particular cuando el experimento exhibe una desviación considerable de la ortogonalidad. Mediciones útiles de **multicolinealidad** (dependencia lineal) entre las variables independientes las proporcionan los coeficientes de correlación muestral $r_{x_i x_j}$. Como sólo nos interesa la dependencia lineal entre las variables independientes, no habrá ninguna confusión si eliminamos las x de nuestra notación y simplemente escribimos $r_{x_i x_j} = r_{ij}$, donde

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii} S_{jj}}}$$

Nótese que las r_{ij} no dan estimaciones reales de coeficientes de correlación poblacionales en sentido estricto, pues las x en realidad no son variables aleatorias en el contexto que aquí se discute. De esta manera el término "correlación", aunque estándar, quizá es inapropiado.

Cuando uno o más de estos coeficientes de correlación muestral se desvían considerablemente de cero, puede ser bastante difícil encontrar el subconjunto de variables más efectivo para su inclusión en nuestra ecuación de predicción. De hecho, para algunos problemas la multicolinealidad será tan extrema que no se puede encontrar un pronosticador adecuado a menos que se investiguen todos los posibles subconjuntos de las variables. En la bibliografía se citan presentaciones informativas de la selección de modelos en la regresión de Hocking. Los procedimientos para la detección de la multicolinealidad se presentan en el texto de Myers, también citado.

El usuario de la regresión lineal múltiple intenta llevar a cabo uno de los tres objetivos:

1. Obtener estimaciones de coeficientes individuales en un modelo completo.
2. Depurar variables para determinar cuáles tienen un efecto significativo sobre la respuesta.
3. Llegar a la ecuación de predicción más efectiva.

En (1) se sabe de antemano que todas las variables se deben incluir en el modelo. En (2) la predicción es secundaria, mientras que en (3) los coeficientes de regresión individuales no son tan importantes como la calidad de la respuesta estimada \hat{y} . Para cada una de las situaciones anteriores, la multicolinealidad en el experimento puede tener un efecto profundo en el éxito de la regresión.

En esta sección se presentan algunos procedimientos secuenciales estándar para seleccionar variables. Éstos se basan en la noción de que una sola variable o una colección de variables no debe aparecer en la ecuación de estimación a menos que tengan como resultado un aumento significativo en la suma de cuadrados de regresión o, de manera equivalente, un aumento significativo en R^2 , el coeficiente de determinación múltiple.

Ejemplo 12.9 Considere los datos de la tabla 12.3 donde se tomaron mediciones de nueve niños. El propósito del experimento fue llegar a una ecuación de estimación adecuada que relacionara la estatura de un niño con todas o con un subconjunto de las variables independientes. Los coeficientes de correlación muestral, que indican la dependencia lineal entre las variables independientes, se muestran en la matriz simétrica

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1.0000 & 0.9523 & 0.5340 & 0.3900 \\ 0.9523 & 1.0000 & 0.2626 & 0.1549 \\ 0.5340 & 0.2626 & 1.0000 & 0.7847 \\ 0.3900 & 0.1549 & 0.7847 & 1.0000 \end{bmatrix}$$

Nótese que parece haber una cantidad apreciable de multicolinealidad. Con el uso de la técnica de mínimos cuadrados que se esbozó en la sección 10.3, se ajusta la ecuación de regresión estimada con el uso del modelo completo y es

$$\hat{y} = 7.1475 + 0.1000x_1 + 0.7264x_2 + 3.0758x_3 + 0.0300x_4$$

Tabla 12.3 Datos relacionados con la estatura de los niños*

Estatura del niño, y (cm)	Edad, x_1 (días)	Estatura al nacer, x_2 (cm)	Peso al nacer, x_3 (kg)	Talla del tórax al nacer, x_4 (cm)
57.5	78	48.2	2.75	29.5
52.8	69	45.5	2.15	26.3
61.3	77	46.3	4.41	32.2
67.0	88	49.0	5.52	36.5
53.5	67	43.0	3.21	27.2
62.7	80	48.0	4.32	27.7
56.2	74	48.0	2.31	28.3
68.5	94	53.0	4.30	30.3
69.2	102	58.0	3.71	28.7

*Datos analizados por el Centro de Consulta Estadística, Instituto Politécnico y Universidad Estatal de Virginia, Blacksburg, Virginia.

El valor de s^2 con 4 grados de libertad es 0.7414, y el valor para el coeficiente de determinación para este modelo es 0.9907. En la tabla 12.4 se presentan la suma de cuadrados de la regresión que miden la variación que se atribuye a cada variable individual en presencia de las otras y los valores t correspondientes

Tabla 12.4 Valores t para los datos de la regresión de la tabla 12.3

Variable x_1	Variable x_2	Variable x_3	Variable x_4
$R(\beta_1 \beta_2, \beta_3, \beta_4)$	$R(\beta_2 \beta_1, \beta_3, \beta_4)$	$R(\beta_3 \beta_1, \beta_2, \beta_4)$	$R(\beta_4 \beta_1, \beta_2, \beta_3)$
= 0.0644	= 0.6334	= 6.2523	= 0.0241
$t = 0.2947$	$t = 0.9243$	$t = 2.9040$	$t = -0.1805$

Una región crítica con dos colas y 4 grados de libertad en el nivel de significancia 0.05 está dada por $|t| > 2.776$. De los cuatro valores t calculados, sólo la variable x_3 parece ser significativa. Sin embargo, recuerde que aunque la estadística t que se describe en la sección 12.6 mide el mérito de una variable ajustada para todas las otras variables; no detecta la importancia potencial de una variable en combinación con un subconjunto de las variables. Por ejemplo, considere el modelo sólo con las variables x_2 y x_3 en la ecuación. El análisis de datos da la función de regresión

$$\hat{y} = 2.1833 + 0.9576x_2 + 3.3253x_3,$$

con $R^2 = 0.9905$, en realidad sin una reducción considerable de $R^2 = 0.9907$ para el modelo completo. Sin embargo, a menos que se observen las características de comportamiento de esta combinación particular, no se estaría seguro de su potencial predictivo. Esto, por supuesto, da apoyo a una metodología que observa *todas las posibles regresiones* o un procedimiento secuencial sistemático diseñado para probar varios subconjuntos. ■

Regresión por pasos

Un procedimiento estándar para buscar el "subconjunto óptimo" de variables en ausencia de ortogonalidad es una técnica que se llama **regresión por pasos**. Se ba-

sa en el procedimiento de introducir en forma secuencial las variables al modelo una por una. La descripción de la rutina paso por paso se comprenderá mejor si primero se describen los métodos de **selección hacia adelante** y **eliminación hacia atrás**.

La **selección hacia adelante** se basa en la noción de que las variables se deben introducir una por una hasta que se encuentra una ecuación de regresión satisfactoria. El procedimiento es el siguiente:

PASO 1. Elegir la variable que dé la suma de cuadrados de la regresión más grande cuando se lleva a cabo una regresión lineal simple con y o, de manera equivalente, la que dé el valor más grande de R^2 . Llamaremos a esta variable inicial x_1 .

PASO 2. Elegir la variable que cuando se introduce en el modelo da el mayor incremento en R^2 , en presencia de x_1 , sobre la R^2 que se encuentra en el paso 1. Ésta, por supuesto, es la variable x_2 , para la que

$$R(\beta_2 | \beta_1) = R(\beta_1, \beta_2) - R(\beta_1)$$

es la más grande. Llamemos a esta variable x_2 . El modelo de regresión con x_1 y x_2 se ajusta entonces y se observa R^2 .

PASO 3. Elegir la variable x_3 que da el valor más grande de

$$R(\beta_3 | \beta_1, \beta_2) = R(\beta_1, \beta_2, \beta_3) - R(\beta_1, \beta_2)$$

que de nuevo tiene como resultado el incremento más grande de R^2 sobre el que se da en el paso 2. Llamamos a esta variable x_3 , y tenemos ahora un modelo de regresión que incluye x_1 , x_2 , y x_3 .

Este proceso se continúa hasta que la variable introducida más reciente deje de inducir un aumento significativo en la regresión que se expone. Tal incremento se puede determinar en cada paso con el uso de la prueba F o prueba t apropiada. Por ejemplo, en el paso 2 el valor

$$f = \frac{R(\beta_2 | \beta_1)}{s^2}$$

se puede determinar para probar lo apropiado de x_2 en el modelo. Aquí el valor de s^2 es el cuadrado medio del error para el modelo que contiene las variables x_1 y x_2 . De manera similar, en el paso 3 la razón

$$f = \frac{R(\beta_3 | \beta_1, \beta_2)}{s^2}$$

prueba la conveniencia de x_3 en el modelo. Ahora, sin embargo, el valor para s^2 es el cuadrado medio del error para el modelo que contiene las tres variables x_1 , x_2 , y x_3 . Si $f < f_{\alpha}(1, n - 3)$ en el paso 2, para un nivel de significancia preestablecido, x_2 no se incluye y el proceso termina, con lo que resulta una ecuación lineal simple que relaciona y y x_1 . Sin embargo, si $f > f_{\alpha}(1, n - 3)$ procedemos al paso 3. De nuevo, si $f < f_{\alpha}(1, n - 4)$ en el paso 3, x_3 no se incluye y se detiene el proceso con la ecuación de regresión apropiada que contiene las variables x_1 y x_2 .

La **eliminación hacia atrás** involucra los mismos conceptos que la selección hacia adelante excepto que se comienza con todas las variables en el modelo. Suponga, por ejemplo, que hay cinco variables bajo consideración. Los pasos son los siguientes:

PASO 1. Ajustar una ecuación de regresión con las cinco variables incluidas en el modelo. Elegir la variable que da el menor valor de la suma de cuadrados de la regresión **ajustada por las otras**. Suponga que esta variable es x_2 . Eliminar x_2 del modelo si

$$f = \frac{R(\beta_2 | \beta_1, \beta_3, \beta_4, \beta_5)}{s^2}$$

es significativa.

PASO 2. Ajustar una ecuación de regresión con el uso de las variables restantes x_1 , x_3 , x_4 , y x_5 , y repetir el paso 1. Suponga que esta vez se elige la variable x_5 . Una vez más si

$$f = \frac{R(\beta_5 | \beta_1, \beta_3, \beta_4)}{s^2}$$

no es significativa, se elimina la variable x_5 del modelo. En cada paso la s^2 que se utiliza en la prueba F es el cuadrado medio del error para el modelo de regresión en esta etapa.

Este proceso se repite hasta algún paso en el que la variable con la suma de cuadrados de la regresión ajustada más pequeña tiene como resultado un valor f significativo para algún nivel de significancia predeterminado.

La **regresión por pasos** se lleva a cabo con una modificación ligera pero importante del procedimiento de selección hacia adelante. La modificación implica la prueba adicional en cada etapa para asegurar la efectividad continua de las variables insertadas en el modelo en una etapa anterior. Esto representa una mejora sobre la selección hacia adelante, pues es muy posible que una variable que entró a la ecuación de regresión en una etapa anterior se considere sin importancia o redundante debido a las relaciones que existen entre ésta y otras variables que entran en etapas posteriores. Por tanto, en una etapa en la que se introduce una nueva variable a la ecuación de regresión a través de un incremento significativo en R^2 determinado por la prueba F , todas las variables ya en el modelo se sujetan a pruebas F (o, de manera equivalente, a pruebas t) a la luz de esta nueva variable, y se eliminan si no muestran un valor f significativo. El procedimiento continúa hasta que se alcanza una etapa donde no se pueden insertar o eliminar variables adicionales. Ilustramos el procedimiento por pasos mediante el siguiente ejemplo.

Ejemplo 12.10 Con el uso de las técnicas de la regresión por pasos, encuentre un modelo de regresión lineal apropiado para predecir la estatura de niños con los datos de la tabla 12.3.

SOLUCIÓN

PASO 1. Al considerar cada variable por separado, se ajustan cuatro ecuaciones de regresión lineal simple. Se calculan las siguientes sumas de cuadrados de regresión pertinentes:

$$R(\beta_1) = 288.1468, \quad R(\beta_2) = 215.3013,$$

$$R(\beta_3) = 186.1065, \quad R(\beta_4) = 100.8594.$$

La variable x_1 claramente da la suma de cuadrados de regresión más grande. El cuadrado medio del error para la ecuación que involucra sólo a x_1 es $s^2 = 4.7276$, y como

$$f = \frac{R(\beta_1)}{s^2} = \frac{288.1468}{4.7276} = 60.9500,$$

que excede a $f_{0.05}(1, 7) = 5.59$, la variable x_1 se introduce al modelo.

PASO 2. En esta etapa se ajustan tres ecuaciones de regresión, las cuales contienen todas las x_i . Los resultados importantes para las combinaciones (x_1, x_2) , (x_1, x_3) y (x_1, x_4) son

$$R(\beta_2|\beta_1) = 23.8703, \quad R(\beta_3|\beta_1) = 29.3086, \quad R(\beta_4|\beta_1) = 13.8178.$$

La variable x_3 demuestra la suma de cuadrados de regresión más grande en presencia de x_1 . La regresión que incluye a x_1 y x_3 da un valor nuevo de $s^2 = 0.6307$, y como

$$f = \frac{R(\beta_3|\beta_1)}{s^2} = \frac{29.3086}{0.6307} = 46.47,$$

que excede a $f_{0.05}(1, 6) = 5.99$, la variable x_3 se incluye junto con x_1 en el modelo. Ahora debemos sujetar a x_1 en presencia de x_3 a una prueba de significancia. Encontramos que $R(\beta_2|\beta_3) = 131.349$, y de aquí

$$f = \frac{R(\beta_2|\beta_3)}{s^2} = \frac{131.349}{0.6307} = 208.26,$$

que es altamente significativa. Por tanto, x_1 se retiene junto con x_3 .

PASO 3. Con x_1 y x_3 ya en el modelo, requerimos ahora $R(\beta_2|\beta_1, \beta_3)$ y $R(\beta_4|\beta_1, \beta_3)$ a fin de determinar cuál, si hay alguna, de las dos variables restantes se introduce en esta etapa. Del análisis de regresión con el uso de x_2 junto con x_1 y x_3 , encontramos que $R(\beta_2|\beta_1, \beta_3) = 0.7948$, y cuando x_4 se utiliza junto con x_1 y x_3 , obtenemos $R(\beta_4|\beta_1, \beta_3) = 0.1855$. El valor de s^2 es 0.5979 para la combinación (x_1, x_2, x_3) y 0.7198 para la combinación (x_1, x_2, x_4) . Como ningún valor f es significativo en el nivel $\alpha = 0.05$, el modelo final de regresión incluye sólo las variables x_1 y x_3 . Se encuentra que la ecuación de estimación es

$$\hat{y} = 20.1084 + 0.4136x_1 + 2.0253x_3,$$

y el coeficiente de determinación para este modelo es $R^2 = 0.9882$.

Aunque (x_1, x_3) es la combinación que se elige mediante la regresión por pasos, no necesariamente es la combinación de dos variables la que da el mayor valor de R^2 . De hecho, ya observamos que la combinación (x_2, x_3) da $R^2 = 0.9905$. Por supuesto, el procedimiento por pasos nunca observa realmente esta combinación. Se podría elaborar un argumento racional de que en realidad hay una diferencia insignificante en el comportamiento entre estas dos ecuaciones de estimación, al menos en términos de la variación porcentual explicada. Es de interés observar, sin embargo, que el procedimiento de eliminación hacia atrás da la combinación (x_2, x_3) en la ecuación final (véase el ejercicio 3 en la página 453).

La función principal de cada uno de los procedimientos que se explican en esta sección es exponer las variables a una metodología sistemática diseñada para asegurar la inclusión eventual de las mejores combinaciones de las variables. Obviamente, no hay seguridad de que esto suceda en todos los problemas y, por supuesto, es posible que la multicolinealidad sea tan extensa que no se tenga más alternativa que recurrir a procedimientos de estimación distintos al de mínimos cuadrados. Estos procedimientos de estimación se discuten en Myers, listado en la bibliografía.

Los procedimientos secuenciales que se presentan aquí representan tres de los muchos métodos que se proponen en la literatura y que aparecen en varios paquetes disponibles de regresión por computadora. Estos métodos están diseñados para ser eficientes en el cálculo pero, por supuesto, no dan resultados para todos los posibles subconjuntos de las variables. Como resultado, los procedimientos son más efectivos en los conjuntos de datos que incluyen un **número grande de variables**. En problemas de regresión que incluyen un número relativamente pequeño de variables, en realidad cuando $k \leq 15$, los paquetes de regresión por computadora modernos permiten el cálculo y el resumen de la información cuantitativa sobre todos los modelos para cada posible subconjunto de las variables. En la sección 12.10 se proporcionan ejemplos. ■

12.9 Estudio de residuos y violación de suposiciones

Se sugirió antes en este capítulo que los residuos, o errores en el ajuste de regresión, a menudo llevan información que puede ser muy informativa para el analista de datos. Los $e_i = y_i - \hat{y}_i$, $i = 1, 2, \dots, n$, que son la contraparte numérica de los ε_i , los errores del modelo, con frecuencia dan luz sobre la posible violación de las suposiciones o sobre la presencia de puntos de datos "sospechosos". Suponga que denotamos con el vector \mathbf{x}_i los valores de las variables regresoras que corresponden al i -ésimo punto de datos, que se complementa con un 1 en la posición inicial. Es decir,

$$\mathbf{x}_i' = [1, x_{i1}, x_{i2}, \dots, x_{ik}]$$

Considere la cantidad

$$h_{ii} = \mathbf{x}_i' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_i, \quad i = 1, 2, \dots, n.$$

El lector debe reconocer que h_{ii} se utiliza en el cálculo de los intervalos de confianza sobre la respuesta media en la sección 12.5. Además de σ^2 , h_{ii} representa la varianza del valor ajustado \hat{Y}_i . Los valores h_{ii} son los elementos diagonales de la **matriz HAT**

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

que juega un papel importante en cualquier estudio de residuos y en otros aspectos modernos del análisis de regresión (véase la referencia a Myers que se lista en la bibliografía). El término *matriz HAT* se deriva del hecho que \mathbf{H} genera las "y con sombrero" o los valores ajustados cuando se multiplica por el vector \mathbf{y} de respuestas observadas. Es decir,

$$\hat{\mathbf{y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y}$$

dónde $\hat{\mathbf{y}}$ es el vector cuyo i -ésimo elemento es \hat{y}_i .

Si hacemos las suposiciones acostumbradas que las E_i son independientes y se distribuyen normalmente con media cero y varianzas σ^2 , las propiedades estadísticas de los residuos se caracterizan con facilidad. Entonces

$$E(E_i) = E(Y_i - \hat{Y}_i) = 0, \quad \text{y} \quad \sigma_{E_i}^2 = (1 - h_{ii})\sigma^2$$

para $i = 1, 2, \dots, n$. (Véase la referencia de Myers para más detalles). Se puede mostrar que los valores diagonales HAT se limitan de acuerdo con la desigualdad

$$\frac{1}{n} \leq h_{ii} \leq 1.$$

Además, $\sum_{i=1}^n h_{ii} = k + 1$, el número de parámetros de la regresión. Como resultado, cualquier punto de datos cuyo elemento diagonal HAT es grande, es decir, muy por arriba del valor promedio de $(k + 1)/n$, está en una posición en el conjunto de datos donde la varianzas de \hat{Y}_i es relativamente grande, y la varianzas de un residuo es relativamente pequeña. Como resultado, el analista de datos puede obtener alguna idea de qué tan grande se puede hacer un residuo antes que su desviación de cero se pueda atribuir a algo diferente del azar. Muchos de los paquetes de regresión por computadora comerciales producen el conjunto de **residuos studentizados**.

Residuos studentizados

$$r_i = \frac{e_i}{s \sqrt{1 - h_{ii}}}, \quad i = 1, 2, \dots, n.$$

Aquí cada residuo se divide entre una estimación de su desviación estándar, lo que genera una estadística tipo t diseñada para dar al analista una cantidad sin escala que proporcione información con respecto al *tamaño* de los residuos. Además, los paquetes estándar de cómputo a menudo proporcionan valores de otro conjunto de residuos tipo Student, que se llaman **valores R de Student**.

Valores R de Student

$$t_i = \frac{e_i}{s_{-i} \sqrt{1 - h_{ii}}}, \quad i = 1, 2, \dots, n,$$

donde s_{-i} es una estimación de la desviación estándar del error, que se calcula con el i -ésimo punto de datos eliminado.

Existen tres tipos de violaciones de las suposiciones que se detectan con facilidad mediante el uso de residuos o *gráficas de residuos*. Mientras que las gráficas de los residuos crudos, e_i , pueden ser útiles, con frecuencia es más informativo graficar los residuos studentizados. Las tres violaciones son las siguientes:

1. Presencia de datos alejados.
2. Varianza del error heterogénea.
3. Modelo mal especificado.

En el caso 1, elegimos definir un *dato alejado* como un punto de los datos donde hay una desviación de la suposición acostumbrada $E(E_i) = 0$ para un valor específico de i . Si hay una razón para creer que un punto de datos específico es un

dato alejado que ejerce una gran influencia sobre el modelo ajustado, r_i o t_i pueden ser informativos. Se puede esperar que los valores R de Student sean más sensibles a los datos alejados que los valores r_i .

De hecho, bajo la condición $E(E_i) = 0$, t_i es un valor de una variable aleatoria que sigue una distribución t con

$$n - 1 - (k + 1) = n - k - 2 \text{ grados de libertad.}$$

De esta manera se puede utilizar una prueba t bilateral para proporcionar información para detectar si el i -ésimo punto es un dato alejado o no.

Aunque la estadística R de Student t_i produce una prueba t exacta para la detección de un dato alejado en una posición específica de los datos, la distribución t no se aplicaría para la prueba simultánea de datos alejados en todos los lugares. Como resultado, los residuos studentizados o valores R de Student se deben utilizar estrictamente como herramientas de diagnóstico *sin* prueba de hipótesis formal como mecanismo. La implicación es que estas estadísticas resaltan puntos de datos donde el error de ajuste es más grande que lo esperado debido al azar. Valores R de Student grandes en magnitud sugieren la necesidad de "verificar" los datos con cualesquiera recursos posibles. La práctica de eliminar observaciones de los conjuntos de datos de regresión no se debe realizar indiscriminadamente. (Para mayor información con respecto al uso de diagnósticos de datos alejados, véase Myers, listado en la bibliografía).

Ejemplo 12.11 En un experimento biológico llevado a cabo en el Instituto Politécnico y Universidad Estatal de Virginia por el Departamento de Entomología, se realizaron n corridas experimentales con dos métodos diferentes para capturar saltamontes. Los métodos son: dejar caer la red y arrastrarla. Se registró el número promedio de saltamontes atrapados en un conjunto de cuadrantes de campo en una fecha dada para cada uno de los dos métodos. También se registró una variable regresora adicional, la altura promedio de las plantas en los cuadrantes. Los datos experimentales son los siguientes:

Observación	Dejar caer la red, y	Arrastrar la red, x_1	Altura de las plantas, x_2 (cm)
1	18.000	4.15476	52.705
2	8.8750	2.02381	42.069
3	2.0000	0.15909	34.766
4	20.0000	2.32812	27.622
5	2.3750	0.25521	45.879
6	2.7500	0.57292	97.472
7	3.3333	0.70139	102.062
8	1.0000	0.13542	97.790
9	1.3333	0.12121	88.265
10	1.7500	0.10937	58.737
11	4.1250	0.56250	42.386
12	12.8750	2.45312	31.274
13	5.3750	0.45312	31.750
14	28.0000	6.68750	35.401
15	4.7500	0.86979	64.516
16	1.7500	0.14583	25.241
17	0.1333	0.01562	36.354

El objetivo es ser capaz de estimar los saltamontes capturados sólo con el método de arrastre de la red, que es menos costoso. Había preocupación por la validez del cuarto punto de datos. La captura observada que se reporta con el uso del método de dejar caer la red parecía inusualmente alta dadas las otras condiciones y, en realidad, se consideró que la cifra podría ser errónea. Ajuste un modelo del tipo

$$\mu_{Y|X_1, X_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

a los 17 puntos de datos y estudie los residuos para determinar si el punto de datos 4 es un dato alejado.

SOLUCIÓN

Un paquete de computadora generó el modelo de regresión ajustado

$$\hat{y} = 3.6870 + 4.1050x_1 - 0.0367x_2$$

junto con las estadísticas $R^2 = 0.9244$ y $s^2 = 5.580$. Los residuos y otra información de diagnóstico también se generaron y se registraron como sigue:

Obs.	y_i	\hat{y}_i	$y_i - \hat{y}_i$	h_{ii}	$s\sqrt{1 - h_{ii}}$	r_i	t_i
1	18.000	18.809	-0.809	0.2291	2.074	-0.390	-0.3780
2	8.875	10.452	-1.577	0.0766	2.270	-0.695	-0.6812
3	2.000	3.065	-1.065	0.1364	2.195	-0.485	-0.4715
4	20.000	12.231	7.769	0.1256	2.209	3.517	9.9315
5	2.375	3.052	-0.677	0.0931	2.250	-0.301	-0.2909
6	2.750	2.464	0.286	0.2276	2.076	0.138	0.1329
7	3.333	2.823	0.510	0.2669	2.023	0.252	0.2437
8	1.000	0.656	0.344	0.2318	2.071	0.166	0.1601
9	1.333	0.947	0.386	0.1691	2.153	0.179	0.1729
10	1.750	1.982	-0.232	0.0852	2.260	-0.103	-0.0989
11	4.125	4.442	-0.317	0.0884	2.255	-0.140	-0.1353
12	12.875	12.610	0.265	0.1152	2.222	0.119	0.1149
13	5.375	4.383	0.992	0.1339	2.199	0.451	0.4382
14	28.000	29.841	-1.841	0.6233	1.450	-1.270	-1.3005
15	4.750	4.891	-0.141	0.0699	2.278	-0.062	-0.0598
16	1.750	3.360	-1.610	0.1891	2.127	-0.757	-0.7447
17	0.133	2.418	-2.285	0.1386	2.193	-1.042	-1.0454

Como se esperaba, el residuo en la cuarta posición parece ser inusualmente alto, a saber, 7.769. El punto vital aquí es si este residuo es más grande o no que el que se esperaría por el azar. El error estándar del residuo para el punto 4 es 2.209. Se encuentra que el valor R de Student t_4 es 9.9315. Al ver esto como un valor de una variable aleatoria que tiene una distribución t con 13 grados de libertad, en realidad se concluiría que el residuo de la cuarta observación se estima algo más grande que 0 y que el error de medición que se sospecha está apoyado por el estudio de los residuos. Nótese que ningún otro residuo tiene como resultado un valor R de Student que produzca alguna causa de alarma. ■

Graficación de residuos

En el capítulo 11 presentamos, con algún detalle, la utilidad de graficar residuos en el análisis de regresión. La violación de las suposiciones del modelo a menudo se pueden detectar a través de estas gráficas. En la probabilidad normal de regresión múltiple la graficación de residuos o gráficas de residuos contra \hat{y} puede ser útil. Sin embargo, a menudo se prefiere graficar residuos studentizados.

Tenga en cuenta que la preferencia de los residuos studentizados sobre los residuos ordinarios con propósitos de graficación se deriva del hecho que como la varianza del i -ésimo residuo depende de la i -ésima diagonal HAT, las varianzas de los residuos serán diferentes si hay una dispersión en las diagonales HAT. De esta manera la apariencia de una gráfica de residuos puede describir heterogeneidad debido a que los residuos mismos no se comportan, en general, en forma ideal. El propósito de utilizar residuos studentizados es proporcionar una *estandarización*. Claramente, si se conociera σ , entonces bajo condiciones ideales (es decir, un modelo correcto y varianza homogénea), tenemos

$$E\left(\frac{E_i}{\sigma\sqrt{1-h_{ii}}}\right) = 0 \quad \text{y} \quad \text{Var}\left(\frac{E_i}{\sigma\sqrt{1-h_{ii}}}\right) = 1.$$

Por lo que los residuos studentizados producen un conjunto de estadísticas que se comportan en forma estándar bajo condiciones ideales.

La figura 12.2 de la página 446 muestra una gráfica de los valores **R de Student** para los datos de los saltamontes del ejemplo 12.11. Nótese cómo los valores para la observación 4 se apartan del resto. La gráfica R de Student se generó con el software SAS. La gráfica muestra el residuo contra los valores \hat{y} .

Verificación de la normalidad

El lector debe recordar la importancia de la verificación de la normalidad a través del uso de la graficación de la probabilidad normal como se presentó en el capítulo 11. La misma recomendación es válida para el caso de la regresión lineal múltiple. Las gráficas de probabilidad normal se pueden generar con el uso del software de regresión estándar. De nuevo, sin embargo, pueden ser más efectivas cuando no se utilizan residuos comunes sino, más bien, residuos studentizados o valores R de Student.

12.10 Validación cruzada, C_p , y otros criterios para la selección del modelo

Para muchos problemas de regresión el experimentador debe elegir entre varios modelos alternativos o formas de modelo que se desarrollan del mismo conjunto de datos. Muy a menudo, de hecho, se requiere el modelo que mejor predice o estima la respuesta media. El experimentador debe tomar en cuenta los tamaños relativos de los valores s^2 para los modelos candidatos y ciertamente la naturaleza general de los intervalos de confianza sobre la respuesta media. También se debe considerar qué tan bien predice el modelo valores de respuesta que **no se utilizaron en la cons-**

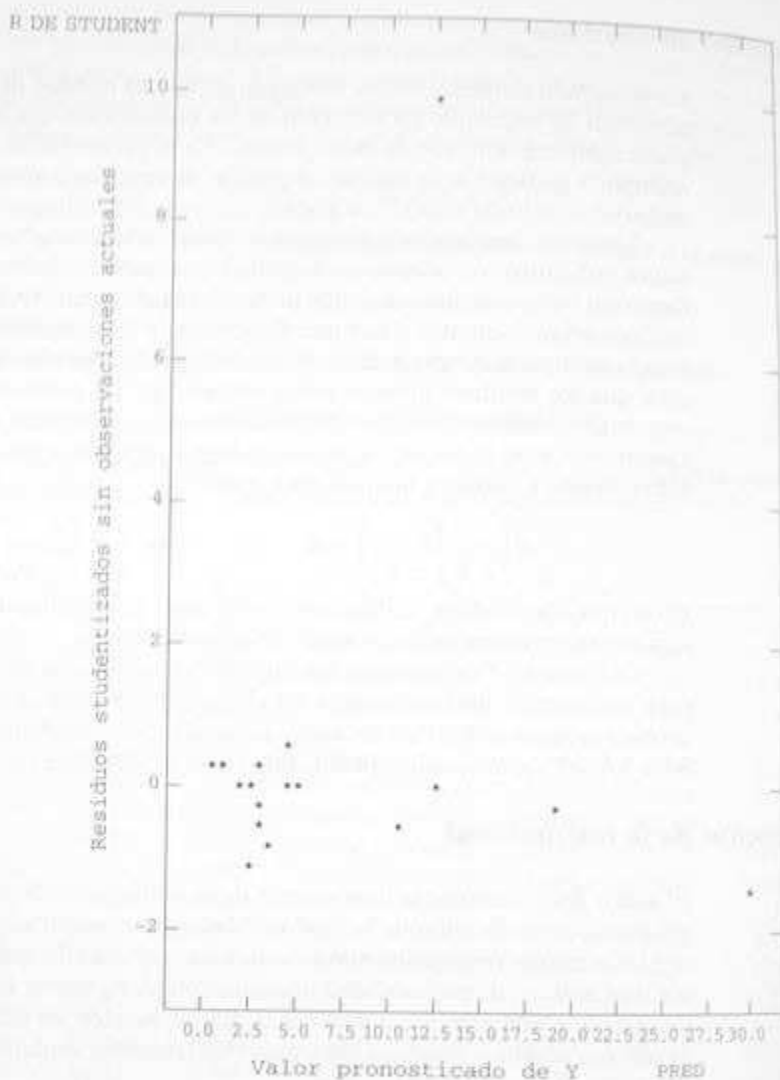


Figura 12.2 Valores R de Student graficados contra valores pronosticados para los datos de los saltamontes del ejemplo 12.11.

trucción de los modelos candidatos. Los modelos se deben sujetar a la **validación cruzada**. Lo que se requiere, entonces, son los errores de validación cruzada en lugar de los errores de ajuste. Tales errores en la predicción son los **residuos PRESS**

$$\hat{\delta}_i = y_i - \hat{y}_{i(-i)}, \quad i = 1, 2, \dots, n,$$

donde $\hat{y}_{i(-i)}$ es la predicción del i -ésimo punto de datos mediante un modelo que no hace uso del i -ésimo punto en el cálculo de los coeficientes. Estos residuos PRESS se calculan a partir de la fórmula

$$\delta_i = \frac{e_i}{1 - h_{ii}}, \quad i = 1, 2, \dots, n.$$

(La derivación se puede encontrar en el texto sobre regresión de Myers).

Uso de la estadística PRESS

La motivación de los residuos PRESS y la utilidad de éstos es muy simple de entender. El propósito de extraer o *hacer a un lado* los datos de puntos uno por uno es para permitir el uso de metodologías separadas para ajustar y evaluar un modelo específico. Para la evaluación de un modelo “-i” indica que el residuo PRESS da un error de predicción donde la observación que se pronostica es independiente del ajuste del modelo.

Los criterios que hacen uso de los residuos PRESS están dados por

$$\sum_{i=1}^n |\delta_i| \quad \text{y} \quad \text{PRESS} = \sum_{i=1}^n \delta_i^2.$$

El término PRESS es la sigla de **suma de cuadrados de predicción**. Sugerimos que se utilicen ambos criterios. Es posible que PRESS sea dominado por uno o sólo algunos residuos PRESS grandes. Claramente, el criterio $\sum_{i=1}^n |\delta_i|$ es menos sensible a un pequeño número de valores grandes.

Además para la estadística PRESS misma, el analista puede simplemente calcular una estadística “tipo R^2 ” que refleje el comportamiento de la predicción. La estadística a menudo se llama R_{pred}^2 y está dada como sigue:

Dado un modelo ajustado con un valor específico para PRESS, R_{pred}^2 está dado por

$$R_{\text{pred}}^2 = 1 - \frac{\text{PRESS}}{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

En el siguiente ejemplo se proporciona un ejemplo de “estudio de caso” en el que muchos modelos candidatos se ajustan a un conjunto de datos y se elige el mejor modelo. Los procedimientos secuenciales que se describen en la sección 12.8 no se utilizan. En su lugar, se ilustra el papel de los residuos PRESS y otros valores estadísticos al seleccionar la mejor ecuación de regresión.

Ejemplo 12.12. Estudio de caso La tensión de la pierna es un ingrediente necesario para un pateador exitoso en el fútbol americano. Una medida de la calidad de una buena patada es el “tiempo de suspensión”. Éste es el tiempo que el balón permanece en el aire antes de que sea atrapado por el regresador de la patada. Para determinar cuáles factores de tensión de la pierna influyen en el tiempo de suspensión y desarrollar un modelo empírico para predecir esta respuesta, el Departamento de Salud, Educación Física y Recreación del Instituto Politécnico y Universidad Estatal de Virginia llevó a cabo el estudio *The Relationship Between Selected Physical Performance Variables and Football Punting Ability*. Se eligieron 13 pateadores para el

experimento y cada uno pateó 10 veces un balón. El tiempo promedio de suspensión, junto con las medidas de tensión que se usaron en el análisis, se registraron como sigue:

Pateador	Tiempo de suspensión,					POTENCIA, x_5
	y (s)	RIS, x_1	LIS, x_2	RHF, x_3	LHF, x_4	
1	4.75	170	170	106	106	240.57
2	4.07	140	130	92	93	195.49
3	4.04	180	170	93	78	152.99
4	4.18	160	160	103	93	197.09
5	4.35	170	150	104	93	266.56
6	4.16	150	150	101	87	260.56
7	4.43	170	180	108	106	219.25
8	3.20	110	110	86	92	132.68
9	3.02	120	110	90	86	130.24
10	3.64	130	120	85	80	205.88
11	3.68	120	140	89	83	153.92
12	3.60	140	130	92	94	154.64
13	3.85	160	150	95	95	240.57

Cada variable regresora se define como sigue:

1. RIS, tensión de la pierna derecha (libras).
2. LIS, tensión de la pierna izquierda (libras).
3. RHF, flexibilidad del músculo del tendón de la corva derecha (grados).
4. LHF, flexibilidad del músculo del tendón de la corva izquierda (grados).
5. POTENCIA, tensión total de la pierna (libras pie).

Determine el modelo más apropiado para predecir el tiempo de suspensión.

SOLUCIÓN

En la búsqueda del "mejor" de los modelos candidatos para predecir el tiempo de suspensión, se obtuvo la información de la tabla 12.5 a partir de un paquete de regresión por computadora. Los modelos se clasifican en orden ascendente de los valores de la estadística PRESS. Esta tabla proporciona suficiente información sobre todos los modelos posibles para permitir al usuario eliminar de la consideración casi todos los modelos. El modelo x_2, x_5 (LIS y POTENCIA) parece ser superior para predecir el tiempo de suspensión del pateador. Nótese también que todos los modelos

con valores PRESS bajos, s^2 bajos, $\sum_{i=1}^n |\delta_i|$, bajos y R^2 altos contienen estas dos variables.

A fin de obtener alguna idea de los residuos de la regresión ajustada

$$\hat{y}_i = b_0 + b_2 x_{2i} + b_5 x_{5i},$$

se generan los residuos comunes y los residuos PRESS. El modelo de predicción real (véase el ejercicio 1 de la página 453) está dado por

$$\hat{y} = 1.10765 + 0.01370x_2 + 0.00429x_5.$$

Tabla 12.5 Comparación de diferentes modelos de regresión

Modelo	s^2	$\sum \delta_i $	PRESS	R^2
x_2x_3	0.036907	1.93583	0.54683	0.871300
$x_1x_2x_3$	0.041001	2.06489	0.58998	0.871321
$x_2x_4x_5$	0.037708	2.18797	0.59915	0.881658
$x_2x_3x_5$	0.039636	2.09553	0.66182	0.875606
$x_1x_2x_4x_5$	0.042265	2.42194	0.67840	0.882093
$x_1x_2x_3x_5$	0.044578	2.26283	0.70958	0.875642
$x_2x_3x_4x_5$	0.042421	2.55789	0.86236	0.881658
$x_1x_3x_5$	0.053664	2.65276	0.87325	0.831580
$x_1x_4x_5$	0.056279	2.75390	0.89551	0.823375
x_1x_5	0.059621	2.99434	0.97483	0.792094
x_2x_3	0.056153	2.95310	0.98815	0.804187
x_1x_3	0.059400	3.01436	0.99697	0.792864
$x_1x_2x_3x_4x_5$	0.048302	2.87302	1.00920	0.882096
x_2	0.066894	3.22319	1.04564	0.743404
x_1x_3	0.065678	3.09474	1.05708	0.770971
x_1x_2	0.068402	3.09047	1.09726	0.761474
x_3	0.074518	3.06754	1.13555	0.714161
$x_1x_3x_4$	0.065414	3.36304	1.15043	0.794705
$x_2x_3x_4$	0.062082	3.32392	1.17491	0.805163
x_2x_4	0.063744	3.59101	1.18531	0.777716
$x_1x_2x_3$	0.059670	3.41287	1.26558	0.812730
x_3x_4	0.080605	3.28004	1.28314	0.718921
x_1x_4	0.069965	3.64415	1.30194	0.756023
x_1	0.080208	3.31562	1.30275	0.692334
$x_1x_3x_4x_5$	0.059169	3.37362	1.36867	0.834936
$x_1x_2x_4$	0.064143	3.89402	1.39834	0.798692
$x_3x_4x_5$	0.072505	3.49695	1.42036	0.772450
$x_1x_2x_3x_4$	0.066088	3.95854	1.52344	0.815633
x_5	0.111779	4.17839	1.72511	0.571234
x_4x_5	0.105648	4.12729	1.87734	0.631593
x_4	0.186708	4.88870	2.82207	0.283819

Los residuos, valores diagonales HAT y valores PRESS se listan en la tabla 12.6. Nótese el ajuste relativamente bueno del modelo de regresión de dos variables a los datos. Los residuos PRESS reflejan la capacidad de la ecuación de regresión para predecir el tiempo de suspensión si se realizan predicciones independientes. Por ejemplo, para el pateador número 4, el tiempo de suspensión de 4.180 encontraría un error de predicción de 0.039 si se utilizara el modelo construido con los restantes 12 pateadores. Para este modelo, el error promedio de predicción o error de validación cruzada es

$$\frac{\sum_{i=1}^n |\delta_i|}{13} = 0.1489 \text{ segundo,}$$

que es bastante pequeño comparado con el tiempo promedio de suspensión para los 13 pateadores.

Tabla 12.6 PRESS Residuals

Pateador	y_i	\hat{y}_i	$\epsilon_i = y_i - \hat{y}_i$	h_{ii}	δ_i
1	4.750	4.470	0.280	0.198	0.349
2	4.070	3.728	0.342	0.118	0.388
3	4.040	4.094	-0.054	0.444	-0.097
4	4.180	4.146	0.034	0.132	0.039
5	4.350	4.307	0.043	0.286	0.060
6	4.160	4.281	-0.121	0.250	-0.161
7	4.430	4.515	-0.085	0.298	-0.121
8	3.200	3.184	0.016	0.294	0.023
9	3.020	3.174	-0.154	0.301	-0.220
10	3.640	3.636	0.004	0.231	0.005
11	3.680	3.687	-0.007	0.152	-0.008
12	3.600	3.553	0.047	0.142	0.055
13	3.850	4.196	-0.346	0.154	-0.409

Indicamos en la sección 12.8 que el uso de todas las posibles regresiones de subconjuntos a menudo es aconsejable cuando se busca el mejor modelo. Los métodos secuenciales que se presentan en la sección 12.8 se deben utilizar sólo cuando una restricción en el tiempo de cálculo prohíbe el cálculo de todas las regresiones posibles. La mayor parte de los paquetes de software estadístico comerciales contiene una rutina de *todas las posibles regresiones*. Estos algoritmos calculan varios criterios para todos los subconjuntos de los términos del modelo. Obviamente, criterios como R^2 , s^2 y PRESS son razonables para elegir entre los subconjuntos candidatos. Otra estadística muy popular y útil, en particular para áreas en las ciencias físicas y en la ingeniería, es la estadística C_p , que se describe a continuación. ■

Estadística C_p

Muy a menudo la elección del modelo más adecuado implica muchas consideraciones. Obviamente, el número de términos del modelo es importante. El asunto de la parquedad es una consideración que no se puede ignorar. Por otro lado, el analista no puede estar complacido con un modelo que es demasiado simple, al punto donde hay una seria subespecificación. Una sola estadística que representa un buen compromiso a este respecto es la estadística C_p . (Véase la referencia a Mallows en la bibliografía.)

La estadística C_p recurre muy bien al sentido común y se desarrolla a partir de consideraciones del compromiso apropiado entre el sesgo excesivo en el que se incurre cuando se subajusta (se seleccionan pocos términos del modelo) y la varianza excesiva de predicción que se produce cuando se ajusta en forma excesiva (se tienen redundancias en el modelo). La estadística C_p es una función simple del número total de parámetros en el modelo candidato y del cuadrado medio del error s^2 .

No presentaremos todo el desarrollo de la estadística C_p . (Para los detalles se remite al lector al texto de Myers en la bibliografía). La C_p para un modelo de un subconjunto particular es una estimación de lo siguiente:

$$\Gamma_{(p)} = \frac{\sum_{i=1}^n \text{Var}(\hat{y}_i)}{\sigma^2} + \frac{\sum_{i=1}^n (\text{Bias } \hat{y}_i)^2}{\sigma^2}$$

Resulta que bajo las suposiciones de mínimos cuadrados estándar que se indican con anterioridad en este capítulo, y al suponer que el modelo "real" es el modelo que contiene todas las variables candidatas,

$$\frac{\sum_{i=1}^n \text{Var}(\hat{y}_i)}{\sigma^2} = p \quad (\text{número de parámetros en el modelo candidato})$$

(véase el ejercicio de revisión 4) y una estimación insesgada de

$$\frac{\sum_{i=1}^n (\text{Bias } \hat{y}_i)^2}{\sigma^2} \quad \text{está dada por} \quad \frac{\sum_{i=1}^n (\widehat{\text{Bias } \hat{y}_i})^2}{\sigma^2} = \frac{(s^2 - \sigma^2)(n - p)}{\sigma^2}$$

En lo anterior, s^2 es el cuadrado medio del error para el modelo candidato, y σ^2 es la varianza del error poblacional. De esta forma si suponemos que se dispone de alguna estimación $\hat{\sigma}^2$ para σ^2 , C_p está dada por

Definición 12.1

Estadística C_p :

$$C_p = p + \frac{(s^2 - \hat{\sigma}^2)(n - p)}{\hat{\sigma}^2}$$

donde p es el número de parámetros del modelo, s^2 es el cuadrado medio del error para el modelo candidato y $\hat{\sigma}^2$ es una estimación de σ^2 . Obviamente, los científicos deben adoptar modelos con valores pequeños de C_p .

El lector debe notar primero que, a diferencia de la estadística PRESS, C_p es libre de escala. Además, se puede obtener alguna idea con respecto a lo adecuado de un modelo candidato al observar su valor C_p . Por ejemplo, $C_p > p$ indica un modelo sesgado debido a que es un modelo subajustado, mientras que $C_p = p$ indica un modelo razonable.

A menudo hay confusión con respecto a de dónde proviene $\hat{\sigma}^2$ en la fórmula para C_p . Obviamente el científico o el ingeniero no tienen acceso a la cantidad poblacional σ^2 . En aplicaciones donde se dispone de corridas repetidas, digamos en una situación de diseño experimental, se dispone de una estimación de σ^2 independiente del modelo (véanse los capítulos 11 y 15). Sin embargo, muchos de los paquetes de software utilizan como $\hat{\sigma}^2$ al cuadrado medio del error del modelo más

completo. Obviamente, si ésta no es una buena estimación, la parte sesgada de la estadística C_p puede ser negativa. Entonces C_p puede ser menor que p .

Ejemplo 12.13 Considere el siguiente conjunto de datos, en el que un fabricante de asfalto se interesa en la relación entre las ventas de un año particular y los factores que influyen en las ventas. (Los datos se tomaron de Neter, Wassermann y Kutner; véase la bibliografía).

Distrito	Cuentas promocionales, x_1	Cuentas activas, x_2	Marcas de la competencia, x_3	Potencial, x_4	Ventas, y (miles)
1	5.5	31	10	8	\$ 79.3
2	2.5	55	8	6	200.1
3	8.0	67	12	9	163.2
4	3.0	50	7	16	200.1
5	3.0	38	8	15	146.0
6	2.9	71	12	17	177.7
7	8.0	30	12	8	30.9
8	9.0	56	5	10	291.9
9	4.0	42	8	4	160.0
10	6.5	73	5	16	339.4
11	5.5	60	11	7	159.6
12	5.0	44	12	12	86.3
13	6.0	50	6	6	237.5
14	5.0	39	10	4	107.2
15	3.5	55	10	4	155.0

De los subconjuntos de modelos posibles, tres son de particular interés. Estos tres contienen las combinaciones de variables (x_2, x_3) , (x_1, x_2, x_3) , y (x_1, x_2, x_3, x_4) . La siguiente representa información pertinente para comparar los tres modelos. Incluimos la estadística PRESS para los tres modelos para complementar la toma de decisiones.

Modelo	R^2	R^2_{pred}	s^2	PRESS	C_p
x_2, x_3	0.9940	0.9913	44.5552	782.1896	11.4013
x_1, x_2, x_3	0.9970	0.9928	24.7956	643.3578	3.4075
x_1, x_2, x_3, x_4	0.9971	0.9917	26.2073	741.7557	5.0

Parece claro a partir de la información en la tabla que el modelo x_1, x_2, x_3 es preferible a los otros dos. Nótese que, para el modelo completo, $C_p = 5.0$. Esto ocurre debido que la *parte sesgada* es cero y $\hat{\sigma}^2 = 26.2073$ es el cuadrado medio del error del modelo completo. ■

La figura 12.3 es un impreso comentado del SAS PROG REG que muestra información para todas las posibles regresiones. Aquí somos capaces de mostrar comparaciones de otros modelos con (x_1, x_2, x_3) . Nótese que (x_1, x_2, x_3) parece ser bastante bueno cuando se compara con todos los modelos.

N = 15 Regression Models for Dependent Variable: Y

C(p)	R-square	In	Adjusted R-square	MSE	Variables in Model
3.40747	0.99695410	3	0.99612340	24.7956	X1 X2 X3
5.00000	0.99707335	4	0.99590269	26.2073	X1 X2 X3 X4
11.40128	0.99402927	2	0.99303414	44.5552	X2 X3
13.37703	0.99403636	3	0.99240992	48.5479	X2 X3 X4
1054	0.68958715	0	0.60492910	2526.9614	X1 X3 X4
1083	0.68050659	2	0.62725768	2384.1429	X3 X4
1215	0.64168582	2	0.58196679	2573.8335	X1 X3
1228	0.63725372	1	0.60935016	2488.6833	X3
1654	0.51395107	3	0.38139228	3956.7527	X1 X2 X4
1669	0.50899678	2	0.42716291	3663.9936	X1 X2
1685	0.50421880	2	0.42158860	3699.6481	X2 X4
1694	0.50101498	1	0.46263152	3437.1285	X2
3015	0.11508695	2	-.03239856	6603.4511	X1 X4
3089	0.09284186	1	0.02306046	6248.7228	X4
3365	0.01199786	1	-.06400231	6805.5957	X1

Figura 12.3 Todos los posibles subconjuntos de los datos de ventas con el SAS.

Ejercicios

1. Considere los datos de "tiempo de suspensión" dados en el ejemplo 12.12; utilice sólo las variables x_2 y x_3 .

(a) Verifique la ecuación de regresión que se muestra en la página 448.

(b) Pronostique el tiempo de suspensión para un pateador con $LLS = 180$ libras y $POTENCIA = 260$ libras-pie.

(c) Construya un intervalo de confianza de 95% para el tiempo de suspensión medio de un pateador con $LLS = 180$ libras y $POTENCIA = 260$ libras-pie.

2. Para los datos del ejercicio 11 de la página 416, utilice las técnicas de

(a) *selección hacia adelante* con un nivel de significancia de 0.05 para elegir un modelo de regresión lineal;

(b) *eliminación hacia atrás* con un nivel de significancia de 0.05 para elegir un modelo de regresión lineal;

(c) *regresión por pasos* con un nivel de significancia de 0.05 para elegir un modelo de regresión lineal.

3. Utilice la técnica de *eliminación hacia atrás* con $\alpha = 0.05$ para elegir una ecuación de predicción para los datos de la tabla 12.3.

4. Para los datos del pateador del ejemplo 12.12, también se registra una respuesta adicional, "distancia de pateo".

Los siguientes son valores de distancia promedio para cada uno de los 13 pateadores:

Pateador	Distancia, y (pie)
1	162.50
2	144.00
3	147.50
4	163.50
5	192.00
6	171.75
7	162.00
8	104.93
9	105.67
10	117.59
11	140.25
12	150.17
13	165.16

(a) Con el uso de los datos de distancia en lugar de los tiempos de suspensión, estime un modelo de regresión lineal múltiple del tipo:

$$\mu_Y | x_1, x_2, x_3, x_4, x_5 \\ = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5$$

para predecir la distancia de pateo.

(b) Utilice la regresión por pasos con un nivel de significancia de 0.10 para seleccionar una combinación de variables.

(c) Genere valores para s^2 , R^2 , PRESS y $\sum_{i=1}^n |\delta_i|$ para todo el conjunto de 31 modelos. Utilice esta información para determinar la mejor combinación de variables para predecir la distancia de pateo.

5. El siguiente es un conjunto de datos para y , la cantidad de dinero (miles de dólares) de contribución para la asociación de alumnos en el tecnológico de Virginia por la generación de 1960, y x , el número de años después de la graduación:

y	x	y	x
812.52	1	2755.00	11
822.50	2	4390.50	12
1211.50	3	5581.50	13
1348.00	4	5548.00	14
1301.00	8	6086.00	15
2567.50	9	5764.00	16
2526.50	10	8903.00	17

(a) Ajuste un modelo de regresión del tipo

$$\mu_{y|x} = \beta_0 + \beta_1 x.$$

(b) Ajuste un modelo cuadrático del tipo

$$\mu_{y|x} = \beta_0 + \beta_1 x + \beta_{11} x^2.$$

(c) Determine cuál de los modelos en (a) o (b) es preferible. Utilice s^2 , R^2 y los residuos PRESS para apoyar su decisión.

6. Para el modelo del ejercicio 4(a), pruebe la hipótesis

$$H_0: \beta_4 = 0,$$

$$H_1: \beta_4 \neq 0.$$

Utilice un valor P en su conclusión.

7. Para el modelo cuadrático del ejercicio 5(b), proporcione estimaciones de las varianzas y covarianzas de las estimaciones de β_1 y β_{11} .

8. Un cliente del departamento de ingeniería mecánica acude al Centro de Consulta del Instituto Politécnico y Universidad Estatal de Virginia para ayudar a analizar un experimento sobre turbinas de gas. Se mide el voltaje de salida de las turbinas para varias combinaciones de la velocidad de las aspas y longitudes del sensor de medición de voltaje. Los datos son los siguientes:

y (voltios)	Velocidad, x_1 (pulgada/s)	Longitud, x_2 (pulgada)
1.95	6336	0.000
2.50	7099	0.000
2.93	8026	0.000
1.69	6230	0.000
1.23	5369	0.000
3.13	8343	0.000
1.55	6522	0.006
1.94	7310	0.006
2.18	7974	0.006
2.70	8501	0.006
1.32	6646	0.012
1.60	7384	0.012
1.89	8000	0.012
2.15	8545	0.012
1.09	6755	0.018
1.26	7362	0.018
1.57	7934	0.018
1.92	8554	0.018

(a) Ajuste una regresión lineal múltiple a los datos.
(b) Calcule pruebas t sobre los coeficientes. Proporcione valores P .
(c) Comente la calidad del modelo ajustado.

9. La blancura del rayón es un factor importante para los científicos que estudian la calidad de la tela. La blancura es afectada por la calidad de la pulpa y otras variables del proceso. Algunas de las variables incluyen: temperatura del baño ácido, $^{\circ}\text{C}$ (x_1); concentración en cascada del ácido, % (x_2); temperatura del agua, $^{\circ}\text{C}$ (x_3); concentración de sulfuro, % (x_4); cantidad de blanqueador de cloro, lb/min (x_5); temperatura de acabado de la manta, $^{\circ}\text{C}$ (x_6). A continuación hay un conjunto de datos tomados de una muestra de rayón. La respuesta, y , es la medición de la blancura.

y	x_1	x_2	x_3	x_4	x_5	x_6
88.7	43	0.211	85	0.243	0.606	48
89.3	42	0.604	89	0.237	0.600	55
75.5	47	0.450	87	0.198	0.527	61
92.1	46	0.641	90	0.194	0.500	65
83.4	52	0.370	93	0.198	0.485	54
44.8	50	0.526	85	0.221	0.533	60
50.9	43	0.486	83	0.203	0.510	57
78.0	49	0.504	93	0.279	0.489	49
86.8	51	0.609	90	0.220	0.462	64
47.3	51	0.702	86	0.198	0.478	63
53.7	48	0.397	92	0.231	0.411	61
92.0	46	0.488	88	0.211	0.387	88
87.9	43	0.525	85	0.199	0.437	63
90.3	45	0.486	84	0.189	0.499	58
94.2	53	0.527	87	0.245	0.530	65
89.5	47	0.601	95	0.208	0.500	67

Utilice los criterios MSE, C_p y PRESS para obtener el mejor modelo entre todos los modelos de subconjuntos.

10. En un esfuerzo para modelar la compensación ejecutiva para el año 1979, se seleccionaron 33 empresas y se recopilaron datos sobre compensación, ventas, utilidades y empleo. Los datos siguientes se recopilaron para el año 1979. Considere el modelo.

$$y_i = \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + \beta_3 \ln x_{3i} + e_i \quad (i = 1, 2, \dots, 33).$$

Ajuste la regresión con el modelo anterior.

Empresa	Compensación, y (miles)	Ventas, x_1 (millones)	Utilidades, x_2 (millones)	Empleo x_3
1	\$450	\$4,600.6	\$128.1	48,000
2	387	9,255.4	783.9	55,900
3	368	1,526.2	136.0	13,783
4	277	1,683.2	179.0	27,765
5	676	2,752.8	231.5	34,000
6	454	2,205.8	329.5	26,500
7	507	2,384.6	381.8	30,800
8	496	2,746.0	237.9	41,000
9	487	1,434.0	222.3	25,900
10	383	470.6	63.7	8,600
11	311	1,508.0	149.5	21,075
12	271	464.4	30.0	6,874
13	524	9,329.3	577.3	39,000
14	498	2,377.5	250.7	34,300
15	343	1,174.3	82.6	19,405
16	354	409.3	61.5	3,586
17	324	724.7	90.8	3,905
18	225	578.9	63.3	4,139
19	254	966.8	42.8	6,255
20	208	591.0	48.5	10,605
21	518	4,933.1	310.6	65,392
22	406	7,613.2	491.6	89,400
23	332	3,457.4	228.0	55,200
24	340	545.3	54.6	7,800
25	698	22,862.8	3,011.3	337,119
26	306	2,361.0	203.0	52,000
27	613	2,614.1	201.0	50,500
28	302	1,013.2	121.3	18,625
29	540	4,560.3	194.6	97,937
30	293	855.7	63.4	12,300
31	528	4,211.6	352.1	71,800
32	456	5,440.4	655.2	87,700
33	417	1,229.9	97.5	14,600

11. La resistencia a la ruptura de un fusible es una característica importante. La tabla siguiente proporciona información importante sobre la resistencia a la ruptura y , altura del dado x_1 , ancho del poste x_2 , altura del rizo x_3 , longitud del alambre x_4 , ancho de la unión con el dado x_5 y ancho de la unión con el poste x_6 . [Datos de Myers y Montgomery (1995).]

y	x_1	x_2	x_3	x_4	x_5	x_6
8.0	5.2	19.6	29.6	94.9	2.1	2.3
8.3	5.2	19.8	32.4	89.7	2.1	1.8
8.5	5.8	19.6	31.0	96.2	2.0	2.0
8.8	6.4	19.4	32.4	95.6	2.2	2.1
9.0	5.8	18.6	28.6	86.5	2.0	1.8
9.3	5.2	18.8	30.6	84.5	2.1	2.1
9.3	5.6	20.4	32.4	88.8	2.2	1.9
9.5	6.0	19.0	32.6	85.7	2.3	1.9
9.8	5.2	20.8	32.2	93.6	2.3	2.1
10.0	5.8	19.9	31.8	86.0	2.1	1.8
10.3	6.4	18.0	32.6	87.1	2.0	1.6
10.5	6.0	20.6	33.4	93.1	2.1	2.1
10.8	6.2	20.2	31.8	83.4	2.2	2.1
11.0	6.2	20.2	32.4	94.5	2.1	1.9
11.3	6.2	19.2	31.4	83.4	1.9	1.8
11.5	5.6	17.0	33.2	85.2	2.1	2.1
11.8	6.0	19.8	35.4	84.1	2.0	1.8
12.3	5.8	18.8	34.0	86.9	2.1	1.8
12.5	5.6	18.6	34.2	83.0	1.9	2.0

(a) Ajuste un modelo de regresión, utilice todas las variables independientes.

(b) Utilice la regresión por pasos con un nivel de significancia de ingreso de 0.25 y un nivel de significancia de eliminación de 0.05. Proporcione su modelo final.

(c) Utilice todos los modelos de regresión posibles y calcule al menos R^2 , C_p , s^2 , R^2 ajustada para todos los modelos.

(d) Proporcione el modelo final.

(e) Para su modelo de la parte (d), grafique los residuos studentizados (o R de Student) y coméntelo.

12. Para el ejercicio 11 anterior, pruebe $H_0: \beta_1 = \beta_6 = 0$. Proporcione valores P y comente.

13. En otra parte del texto (ejercicio 12, página 426) tenemos los siguientes datos con respecto al desgaste de un cojinete.

y (desgaste)	x_1 (viscosidad del aceite)	x_2 (carga)
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

(a) Se puede considerar que el siguiente modelo describe estos datos:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{12} x_{1i} x_{2i} + e_i, \quad i = 1, 2, \dots, 6.$$

La $x_1 x_2$ es un término de "interacción". Ajuste este modelo y estime el parámetro.

(b) Utilice los modelos (x_1) , (x_1, x_2) , (x_2) , $(x_1, x_2, x_1 x_2)$ y calcule PRESS, C_p y s^2 para determinar el "mejor" modelo.

Ejercicios de repaso

1. En el departamento de pesca y fauna del Instituto Politécnico y Universidad Estatal de Virginia se lleva a cabo un experimento para estudiar el efecto de las características de la corriente sobre la biomasa de peces. Las variables regresoras son las siguientes: profundidad promedio (de 50 celdas) (x_1); área del refugio dentro de la corriente (es decir, riberas socavadas, troncos, cantos rodados, etcétera) (x_2); porcentaje de cubierta en bóveda (promedio de 12) (x_3); área ≥ 25 centímetros de profundidad (x_4). La respuesta es y , la biomasa de peces. Los datos son los siguientes:

Observación	y	x_1	x_2	x_3	x_4
1	100	14.3	15.0	12.2	48.0
2	388	19.1	29.4	26.0	152.2
3	755	54.6	58.0	24.2	469.7
4	1288	28.8	42.6	26.1	485.9
5	230	16.1	15.9	31.6	87.6
6	0	10.0	56.4	23.3	6.9
7	551	28.5	95.1	13.0	192.9
8	345	13.8	60.6	7.5	105.8
9	0	10.7	35.2	40.3	0.0
10	348	25.9	52.0	40.3	116.6

- (a) Ajuste una regresión lineal múltiple que incluya las cuatro variables de regresión.
 (b) Utilice C_p , R^2 y s^2 para determinar el mejor subconjunto de variables. Calcule estas estadísticas para todos los subconjuntos posibles.
 (c) Compare lo apropiado de los modelos de las partes (a) y (b) para predecir la biomasa de peces.

2. En un experimento de ingeniería química que trata de la transferencia de calor en un lecho fluidizado superficial se recopilan datos para las siguientes cuatro variables regresoras: tasa de flujo del gas fluidizante, lb/h (x_1); tasa de flujo de gas flotante, lb/h (x_2); abertura de la boquilla de entrada del gas flotante, milímetros (x_3); temperatura de entrada del gas flotante, °F (x_4). Las respuestas medidas son: eficiencia de la transferencia de calor (y_1); eficiencia térmica (y_2). Los datos son los siguientes:

Observación	y_1	y_2	x_1	x_2	x_3	x_4
1	41.852	38.75	69.69	170.83	45	219.74
2	155.329	51.87	113.46	230.06	25	181.22
3	99.628	53.79	113.54	228.19	65	179.06
4	49.409	53.84	118.75	117.73	65	281.30
5	72.958	49.17	119.72	117.69	25	282.20
6	107.702	47.61	168.38	173.46	45	216.14
7	97.239	64.19	169.85	169.85	45	223.88
8	105.856	52.73	169.85	170.86	45	222.80
9	99.348	51.00	170.89	173.92	80	218.84
10	111.907	47.37	171.31	173.34	25	218.12
11	100.008	43.18	171.43	171.43	45	219.20
12	175.380	71.23	171.59	263.49	45	168.62
13	117.800	49.30	171.63	171.63	45	217.58
14	217.409	50.87	171.93	170.91	10	219.92
15	41.725	54.44	173.92	71.73	45	296.60
16	151.139	47.93	221.44	217.39	65	189.14
17	220.630	42.91	222.74	221.73	25	186.08
18	131.666	66.60	228.90	114.40	25	285.80
19	80.537	64.94	231.19	113.52	65	286.34
20	152.966	43.18	236.84	167.77	45	221.72

Considere el modelo para predecir la respuesta del coeficiente de transferencia de calor

$$y_{1i} = \beta_0 + \sum_{j=1}^4 \beta_j x_{ij} + \sum_{j=1}^4 \beta_{1j} x_{ij}^2 + \sum_{j=1}^4 \beta_{2j} x_{ij} x_{0j} + \epsilon_i \quad (i = 1, 2, \dots, 20)$$

- (a) Calcule PRESS y $\sum_{i=1}^n |y_i - \hat{y}_{i-1}|$ para el ajuste de regresión de mínimos cuadrados al modelo anterior.
 (b) Ajuste un modelo de segundo orden con x_4 eliminada por completo (es decir, eliminar todos los términos que incluyen a x_4). Calcule el criterio de predicción para el modelo reducido. Comente lo apropiado de x_4 para la predicción del coeficiente de transferencia de calor.
 (c) Repita las partes (a) y (b) para la eficiencia térmica.

3. En el ejercicio de la fisiología, una medición objetiva del ajuste aeróbico es el consumo de oxígeno en

volumen por unidad de peso corporal por unidad de tiempo. Se utilizan 31 individuos en un experimento a fin de poder modelar el consumo de oxígeno (edad en años (x_1); peso en kilogramos (x_2); tiempo para correr $1\frac{1}{2}$ millas (x_3); ritmo del pulso en descanso (x_4); ritmo del pulso al final de la carrera (x_5); ritmo máximo del pulso durante la carrera (x_6).

Individuo	y	x_1	x_2	x_3	x_4	x_5	x_6
1	44.609	44	89.47	11.37	62	178	182
2	45.313	40	75.07	10.07	62	185	185
3	54.297	44	85.84	8.65	45	156	168
4	59.571	42	68.15	8.17	40	166	172
5	49.874	38	89.02	9.22	55	178	180
6	44.811	47	77.45	11.63	58	176	176
7	45.681	40	75.98	11.95	70	176	180
8	49.091	43	81.19	10.85	64	162	170
9	39.442	44	81.42	13.08	63	174	176
10	60.055	38	81.87	8.63	48	170	186
11	50.541	44	73.03	10.13	45	168	168
12	37.388	45	87.66	14.03	56	186	192
13	44.754	45	66.45	11.12	51	176	176
14	47.273	47	79.15	10.60	47	162	164
15	51.855	54	83.12	10.33	50	166	170
16	49.156	49	81.42	8.95	44	180	185
17	40.836	51	69.63	10.95	57	168	172
18	46.672	51	77.91	10.00	48	162	168
19	46.774	48	91.63	10.25	48	162	164
20	50.388	49	73.37	10.08	76	168	168
21	39.407	57	73.37	12.63	58	174	176
22	46.080	54	79.38	11.17	62	156	165
23	45.441	52	76.32	9.63	48	164	166
24	54.625	50	70.87	8.92	48	146	155
25	45.118	51	67.25	11.08	48	172	172
26	39.203	54	91.63	12.88	44	168	172
27	45.790	51	73.71	10.47	59	186	188
28	50.545	57	59.08	9.93	49	148	155
29	48.673	49	76.32	9.40	56	186	188
30	47.920	48	61.24	11.50	52	170	176
31	47.467	52	82.78	10.50	53	170	172

(a) Realice una regresión por pasos con nivel de significancia de 0.25. Muestre el modelo final.

(b) Construya todos los subconjuntos posibles con el uso de s^2 , C_p , R^2 y R^2 ajustada. Tome una decisión y muestre el modelo final.

4. Muestre que, en un conjunto de datos de regresión lineal múltiple,

$$\sum_{i=1}^n h_{ii} = p.$$

5. Se lleva a cabo un pequeño experimento para ajustar una ecuación de regresión lineal múltiple que relaciona la producción y con la temperatura x_1 , tiempo de reacción x_2 y concentración de uno de los reactivos x_3 . Se eligen dos niveles de cada variable y las mediciones correspondientes a las variables independientes codificadas se registran como sigue:

y	x_1	x_2	x_3
7.6	-1	-1	-1
5.5	1	-1	-1
9.2	-1	1	-1
10.3	-1	-1	1
11.6	1	1	-1
11.1	1	-1	1
10.2	-1	1	1
14.0	1	1	1

(a) Con el uso de las variables codificadas, estime la ecuación de regresión lineal múltiple

$$\mu_{Y|X_1, X_2, X_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

(b) Divida SSR , la suma de cuadrados de la regresión, en tres componentes de un solo grado de libertad atribuibles a x_1 , x_2 y x_3 , respectivamente. Muestre una tabla de análisis de varianza que indique las pruebas de significancia sobre cada variable. Comente los resultados.

6. Considere los datos del ejercicio 5 anterior. Suponga que es de interés agregar algunos términos de "interacción". A saber, considere el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i} + \beta_{123} x_{1i} x_{2i} x_{3i} + \epsilon_i.$$

(a) ¿Aún tenemos ortogonalidad? Comente.

(b) ¿Con el modelo ajustado de la parte (a) puede encontrar intervalos de predicción e intervalos de confianza sobre la respuesta media? ¿Por qué sí o por qué no?

(c) Considere un modelo con $\beta_{123} x_1 x_2 x_3$ eliminada. Para determinar si se necesitan las interacciones (como un todo), pruebe

$$H_0: \beta_{12} = \beta_{13} = \beta_{23} = 0.$$

Proporcione un valor P y conclusiones.

7. Un artículo en *Journal of Pharmaceutical Sciences* (vol. 80, 1991) presenta los datos sobre la solubilidad en fracción molar de un soluto a una temperatura constante. También se miden la dispersión x_1 y los parámetros de solubilidad de unión bipolar y de hidrógeno x_2 y x_3 . Una parte de los datos se presenta en la tabla contigua.

En lo anterior, y es el logaritmo negativo de la fracción molar. Ajuste el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + e_i \quad (i = 1, 2, \dots, 20)$$

(a) Pruebe $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ con el uso de un ANOVA.

(b) Grafique los residuos studentizados contra x_1 , x_2 y x_3 (tres gráficas). Comente.

(c) Considere dos modelos adicionales que compiten con el modelo anterior:

Modelo 2 Agregar x_1^2, x_2^2, x_3^2

Modelo 3 Agregar $x_1^2, x_2^2, x_3^2, x_1 x_2, x_1 x_3, x_2 x_3$

Con estos tres modelos utilice PRESS y C_p para llegar al mejor entre los tres.

Observación	y	x_1	x_2	x_3
1	0.22200	7.3	0.0	0.0
2	0.39500	8.7	0.0	0.3
3	0.42200	8.8	0.7	1.0
4	0.43700	8.1	4.0	0.2
5	0.42800	9.0	0.5	1.0
6	0.46700	8.7	1.5	2.8
7	0.44400	9.3	2.1	1.0
8	0.37800	7.6	5.1	3.4
9	0.49400	10.0	0.0	0.3
10	0.45600	8.4	3.7	4.1
11	0.45200	9.3	3.6	2.0
12	0.11200	7.7	2.8	7.1
13	0.43200	9.8	4.2	2.0
14	0.10100	7.3	2.5	6.8
15	0.23200	8.5	2.0	6.6
16	0.30600	9.5	2.5	5.0
17	0.09230	7.4	2.8	7.8
18	0.11600	7.8	2.8	7.7
19	0.07640	7.7	3.0	8.0
20	0.43900	10.3	1.7	4.2

8. **Estudio de caso:** Considere el conjunto de datos del ejercicio 12, página 416 (datos del hospital). El conjunto de datos se repite aquí.

Lugar	x_1	x_2	x_3	x_4	x_5	y
1	15.57	2.463	472.92	18.0	4.45	566.52
2	44.02	2.048	1,339.75	9.5	6.92	696.82
3	20.42	3.940	620.25	12.8	4.28	1,033.15
4	18.74	6.505	568.33	36.7	3.90	1,603.62
5	49.20	5.723	1,497.60	35.7	5.50	1,611.37
6	44.92	11.520	1,365.83	24.0	4.60	1,613.27
7	55.48	5.779	1,687.00	43.3	5.62	1,854.17
8	59.28	5.969	1,639.92	46.7	5.15	2,160.55
9	94.38	8.461	2,872.33	78.7	6.18	2,305.58
10	128.02	20.106	3,655.08	180.5	6.15	3,503.93
11	96.00	13.313	2,912.00	60.9	5.88	3,571.89
12	131.42	10.771	3,921.00	103.7	4.88	3,741.40
13	127.21	15.543	3,865.67	126.8	5.50	4,026.52
14	252.90	36.194	7,684.10	157.7	7.00	10,343.81
15	409.20	34.703	12,446.33	169.4	10.78	11,732.17
16	463.70	39.204	14,098.40	331.4	7.05	15,414.94
17	510.22	86.533	15,524.00	371.6	6.35	18,854.45

(a) A continuación se proporciona la salida del SAS PROC REG que suministra una cantidad considerable de información.

Model: MODEL1
 Dependent Variable: Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	490177488.12	98035497.624	237.790	0.0001
Error	11	4535052.3673	412277.46794		
C. Total	16	494712640.49			
Root MSE	642.08938	R-square	0.9908		
Dep Mean	4978.48000	Adj R-sq	0.9867		
C.V.	12.89728				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1962.948156	1071.3616951	1.832	0.0941
X1	1	-15.851675	97.65299019	-0.162	0.8740
X2	1	0.055930	0.02125828	2.631	0.0234
X3	1	1.589624	3.09208349	0.514	0.6174
X4	1	-4.218688	7.17655737	-0.588	0.5685
X5	1	-394.314117	209.63954982	-1.881	0.0867

Variable

Variable	DF	Label
INTERCEP	1	Intercept
X1	1	Average Daily Patient Load
X2	1	Monthly X-Ray Exposures
X3	1	Monthly Occupied Bed Days
X4	1	Eligible Population in the Area/1000
X5	1	Average Length of Patients Stay in Days

Obs.	Dep Var Y	Predict Value	Std Err Predict	Lower95% Mean	Upper95% Mean	Lower95% Predict	Upper95% Predict
1	566.5	775.0	241.232	244.1	1306.0	-734.8	2284.7
2	896.8	740.7	321.140	11.8355	1489.5	-849.4	2330.8
3	1033.2	1103.9	278.512	490.9	1716.9	-436.3	2844.4
4	1603.6	1240.5	298.130	650.3	1830.6	-291.0	2772.0
5	1611.4	1564.4	211.237	1099.5	2029.4	76.6818	3052.12
6	1613.3	2151.3	379.929	1535.2	2767.4	609.6	3693.0
7	1854.2	1689.7	218.998	1207.7	2171.7	196.6	3182.9
8	2160.6	1736.2	468.990	704.0	2788.5	-13.9306	3446.3
9	2305.6	2737.0	290.475	2087.7	3376.3	1185.9	4286.1
10	3503.9	3891.9	585.262	2393.7	4970.0	1769.7	5594.0
11	3671.9	3239.3	199.099	2823.1	3655.5	1766.0	4712.5
12	3741.4	4353.3	329.851	3629.6	5077.1	2785.8	5941.1
13	4026.8	4257.1	314.048	3585.9	4946.3	2683.9	5830.3
14	10343.8	8767.7	252.262	8212.5	9329.0	7249.4	10286.1
15	11702.2	12237.0	873.917	10973.8	13500.2	10341.5	14132.5
16	15414.9	15098.4	585.705	13749.3	16327.5	13125.5	16951.3
17	16654.5	19320.7	698.978	18000.2	20841.2	17386.6	21254.8

[continuación]

Obs.	Residual	Std Err Residual	Student Residual	-2	-1	0	1	2
1	-208.5	595.050	-0.350					
2	-43.8502	550.112	-0.080					
3	-70.7734	578.540	-0.122					
4	363.1	583.424	0.622				*	
5	46.9483	606.347	0.077					
6	-538.0	577.856	-0.931				*	
7	164.5	603.587	0.272					
8	424.3	438.549	0.968				*	
9	-431.4	572.627	-0.753				*	
10	-177.9	264.117	-0.674				*	
11	332.6	613.612	0.542				*	
12	-611.9	551.484	-1.110			**		
13	-230.6	560.046	-0.412					
14	1576.1	590.459	2.669				*****	
15	-504.9	287.918	-1.753			***		
16	376.5	263.111	1.431				**	
17	-466.2	228.700	-2.039			****		
Sum of Residuals			0					
Sum of Squared Residuals			4535052.3673					
Predicted Resid SS (Press)			32195221.776					

Los objetivos son realizar una detección externa y al final determinar cuáles términos del modelo se deben utilizar en el modelo final.

- (a) Comente cuáles otros análisis se deben llevar a cabo. ¿El análisis anterior sugiere otras corridas de cálculo?
 (b) Realice análisis apropiados y escriba sus conclusiones con respecto al modelo final.

9. Una técnica de inundación con bióxido de carbono (CO_2) se utiliza para extraer petróleo crudo. El CO_2 inunda las cavidades de petróleo y lo desplaza. En el experimento, se introducen tubos de flujo en muestras de cavidades de petróleo que contienen una cantidad conocida de petróleo. Con el uso de tres presiones de fluido diferentes y tres diferentes ángulos de inmersión, las cavidades se inundan con CO_2 y se registra el porcentaje de petróleo desplazado.

Considere el modelo

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + \varepsilon_i$$

Ajuste el modelo anterior a los datos y sugiera cualquier cambio al modelo que sea necesario.

Presión lb/pulgada ²	Ángulo de inmersión	Recuperación de petróleo, %
x_1	x_2	y_i
1000	0	60.58
1000	15	72.72
1000	30	79.99
1500	0	66.83
1500	15	80.78
1500	30	89.78
2000	0	69.18
2000	15	80.31
2000	30	91.99

Fuente: G. C. Wang, "Microscopic Investigations of CO_2 Flooding Process", *Journal of Petroleum Technology*, Vol. 34, Núm. 8, agosto de 1982.

Experimentos de un factor: general

13.1 Técnica del análisis de varianza

En el material de estimación y prueba de hipótesis que se cubre en los capítulos 9 y 10 nos restringimos en cada caso a considerar no más de dos parámetros poblacionales. Tal es el caso, por ejemplo, en las pruebas de la igualdad de dos medias poblacionales con el uso de muestras independientes de poblaciones normales con varianza común pero desconocida, donde es necesario obtener una estimación combinada de σ^2 .

Este material que se refiere a la inferencia de dos muestras representa un caso especial de lo que llamamos *problema de un factor*. Por ejemplo, en el ejercicio 17, sección 10.10, se mide el tiempo de sobrevivencia para dos muestras de ratones donde una muestra recibe tratamiento con un nuevo suero contra la leucemia y la otra muestra no recibe tratamiento alguno. En este caso decimos que hay *un factor*, que se denomina *tratamiento*, y el factor es en *dos niveles*. Si se utilizan varios tratamientos en competencia en el proceso de muestreo, se necesitarían más muestras de ratones. En este caso el problema involucraría un factor con más de dos niveles y por ello más de dos muestras.

En el problema de $k > 2$ muestras, se supondrá que hay k muestras de k poblaciones. Un procedimiento muy común que se usa para tratar con la prueba de medias poblacionales se llama **análisis de varianza**.

El análisis de varianza en realidad no es una técnica nueva si el lector siguió el material sobre teoría de la regresión. Utilizamos el enfoque del análisis de varianza para dividir la suma total de cuadrados en una parte que se debe a la regresión y otra debida al error. Además, fuimos capaces, en algunos casos, de dividir de mane-

ra conveniente a *SSR* en componentes significativas para el propósito de probar hipótesis relevantes sobre los parámetros en el modelo. El término *análisis de varianza* describe una técnica por medio de la cual se analiza la variación total o se divide en componentes significativos.

Los problemas de regresión en los que el modelo contiene *variables cuantitativas* (como las hasta aquí discutidas) no son el único tipo en el que el análisis de varianza juega un papel importante. En este capítulo presentamos y estudiamos otros tipos de modelos en los que se utiliza esta técnica. El grado de dificultad del análisis depende de la complejidad del problema.

Suponga que en un experimento industrial un ingeniero está interesado en cómo la absorción media de humedad en concreto varía entre cinco mezclas diferentes de concreto. Las muestras se exponen a la humedad por 48 horas. Se decide que se prueben seis muestras para cada mezcla, por lo que se requiere probar un total de 30 muestras. Los datos se registran en la tabla 13.1.

Tabla 13.1 Absorción de humedad en mezclas de concreto

Mezcla:	1	2	3	4	5	
	551	595	639	417	563	
	457	580	615	449	631	
	450	508	511	517	522	
	731	583	573	438	613	
	499	633	648	415	656	
	632	517	677	555	679	
Total	3320	3416	3663	2791	3664	16,854
Media	553.33	569.33	610.50	465.17	610.67	561.80

El modelo para esta situación se puede considerar como sigue. Hay seis observaciones que se toman de cada una de las cinco poblaciones con medias $\mu_1, \mu_2, \dots, \mu_5$, respectivamente. Deseamos probar

$$H_0: \mu_1 = \mu_2 = \dots = \mu_5,$$

$$H_1: \text{Al menos dos de las medias no son iguales.}$$

Además, nos podemos interesar en realizar comparaciones individuales entre estas cinco medias poblacionales.

En el procedimiento del análisis de varianza, se supone que cualquier variación que exista entre los promedios de las mezclas se atribuye a (1) la variación en la absorción entre las observaciones dentro de los tipos de mezclas y (2) la variación que se debe a los tipos de mezclas; es decir, las que se deben a diferencias en la composición química de las mezclas. Las **variaciones dentro de la mezcla**, por supuesto, son ocasionadas por diversas causas. Quizá las condiciones de humedad y temperatura no se conservaron completamente constantes a lo largo del experimento. Es posible que haya cierta cantidad de heterogeneidad en los lotes de materia prima que se utilizan. De todos modos, consideraremos que la variación dentro de la muestra es una **variación aleatoria** o **al azar**, y parte del objetivo del análisis de varianza es determinar si las diferencias entre las cinco medias muestrales son las que se esperarían debido sólo a la variación aleatoria o si en realidad también hay una contribución de la variación sistemática que se atribuye a los tipos de mezcla. El procedimiento en esencia, entonces, separa la variabilidad total en los siguientes dos componentes:

1. Variabilidad entre mezclas, mediante la medición de la variación sistemática y la aleatoria.
2. Variabilidad dentro de las mezclas, mediante la medición sólo de la variación aleatoria.

Queda entonces la tarea de determinar si el componente 1 es significativamente mayor que el componente 2.

Surgen muchas preguntas críticas en esta etapa con respecto al problema anterior. Por ejemplo, ¿cuántas muestras se deben probar para cada mezcla? Ésta es una pregunta que continuamente obsesiona al practicante. Además, ¿qué pasa si la variación dentro de la mezcla es tan grande que es difícil para un procedimiento estadístico detectar las diferencias sistemáticas? ¿Podemos controlar de manera sistemática las fuentes extrañas de variación y de esta manera eliminarlas de la parte que llamamos variación aleatoria? Intentaremos contestar éstas y otras preguntas en las siguientes secciones.

13.2 La estrategia del diseño experimental

En los capítulos 9 y 10 se cubre la noción de estimación y prueba para el caso de dos muestras bajo la consideración importante de la forma en que se realiza el experimento. Esto cae en la amplia categoría del diseño de experimentos. Por ejemplo, para la prueba que se denomina prueba t combinada que se presentó en el capítulo 10, se supone que los niveles del factor (tratamientos en el ejercicio de los ratones) se asignan de forma aleatoria a las unidades experimentales (ratones). La noción de unidades experimentales se presenta en los capítulos 9 y 10 y se ilustra a través de ejemplos. Puesto de manera simple, las unidades experimentales son las unidades (ratones, pacientes, muestras de concreto, tiempo) que proporcionan la heterogeneidad que produce el error experimental en una investigación científica. La asignación aleatoria elimina el sesgo que podría resultar por la asignación sistemática. El objetivo es distribuir de manera uniforme entre los niveles del factor los riesgos causados por la heterogeneidad de las unidades experimentales. Una asignación aleatoria simula mejor las condiciones que supone el modelo. En la sección 13.8 presentamos el *bloqueo* en los experimentos. La noción de bloqueo se presenta en los capítulos 9 y 10, cuando las comparaciones entre las medias se llevan a cabo con *pareamientos*; es decir, la división de las unidades experimentales en pares homogéneos que se denominan **bloques**. Los niveles del factor o tratamientos se asignan entonces de forma aleatoria dentro de los bloques. El propósito del bloqueo es reducir el error experimental efectivo. En este capítulo naturalmente extendemos el pareamiento a *tamaños de bloques más grandes*, con el análisis de varianza como la herramienta analítica principal.

13.3 Análisis de varianza unilateral: diseño completamente aleatorizado

Se seleccionan muestras aleatorias de tamaño n de cada una de las k poblaciones. Las k poblaciones diferentes se clasifican sobre la base de un solo criterio, como tratamientos o grupos diferentes. Hoy en día, el término **tratamiento** se usa por lo

neral para referirnos a las diversas clasificaciones, ya sea de mezclas diferentes, análisis diferentes, fertilizadores diferentes o regiones del país diferentes. Se supone que las k poblaciones son independientes y normalmente distribuidas con medias $\mu_1, \mu_2, \dots, \mu_k$ y varianza común σ^2 .

Como se indica en la sección 13.2, estas suposiciones son más aceptables mediante la aleatorización. Deseamos derivar métodos apropiados para probar la hipótesis

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k,$$

$$H_1: \text{Al menos dos de las medias no son iguales.}$$

Denotemos con y_{ij} la j -ésima observación del i -ésimo tratamiento y acomodemos los datos como en la tabla 13.2. Aquí, T_i es el total de todas las observaciones en la

Tabla 13.2 k muestras aleatorias

Tratamiento:	1	2	...	i	...	k	
	y_{11}	y_{12}	...	y_{1i}	...	y_{1k}	
	y_{21}	y_{22}	...	y_{2i}	...	y_{2k}	
	\vdots	\vdots		\vdots		\vdots	
	y_{i1}	y_{i2}	...	y_{ii}	...	y_{ik}	
Total	Y_1	Y_2	...	Y_i	...	Y_k	Y
Media	\bar{y}_1	\bar{y}_2	...	\bar{y}_i	...	\bar{y}_k	\bar{y}

muestra del i -ésimo tratamiento, \bar{y}_i es la media de todas las observaciones en la muestra del i -ésimo tratamiento, T_i es el total de las nk observaciones y \bar{y} es la media de todas las nk observaciones. Cada observación se puede escribir en la forma

$$y_{ij} = \mu_i + \varepsilon_{ij}$$

donde ε_{ij} mide la desviación de la j -ésima observación de la i -ésima muestra de la correspondiente media del tratamiento. El término ε_{ij} representa el error aleatorio y juega el mismo papel que los términos de error en los modelos de regresión. Una forma alternativa de esta ecuación y que se prefiere se obtiene al sustituir $\mu_i = \mu + \alpha_i$, sujeta a la restricción $\sum_{i=1}^k \alpha_i = 0$. De aquí podemos escribir

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

donde μ es justo la **media general** de todas las μ_i ; es decir

$$\mu = \frac{\sum_{i=1}^k \mu_i}{k},$$

y α_i se denomina el **efecto** del i -ésimo tratamiento.

La hipótesis nula de que las k medias poblacionales son iguales contra la alternativa de que al menos dos de las medias son diferentes se puede reemplazar ahora por la hipótesis equivalente,

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

$$H_1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$$

Nuestra prueba se basará en una comparación de dos estimaciones independientes de la varianza poblacional común σ^2 . Estas estimaciones se obtendrán al dividir la variabilidad total de nuestros datos, que se representa con la doble sumatoria

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2,$$

en dos componentes.

Teorema 13.1

Identidad de la suma de cuadrados

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2.$$

PRUEBA

$$\begin{aligned} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 &= \sum_{i=1}^k \sum_{j=1}^n [(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2 \\ &= \sum_{i=1}^k \sum_{j=1}^n [(\bar{y}_{i.} - \bar{y}_{..})^2 + 2(\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.}) + (y_{ij} - \bar{y}_{i.})^2] \\ &= \sum_{i=1}^k \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 + 2 \sum_{i=1}^k \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.}) + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2. \end{aligned}$$

El término intermedio es cero, pues

$$\sum_{j=1}^n (y_{ij} - \bar{y}_{i.}) = \sum_{j=1}^n y_{ij} - n\bar{y}_{i.} = \sum_{j=1}^n y_{ij} - n \frac{\sum_{j=1}^n y_{ij}}{n} = 0.$$

La primera suma no tiene a j como subíndice y por tanto se puede escribir como

$$\sum_{i=1}^k \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2.$$

De aquí

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2.$$

Será conveniente en lo que sigue identificar los términos de la identidad de la suma de cuadrados mediante la siguiente notación:

$$SST = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \text{suma total de cuadrados.}$$

$$SSA = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 = \text{suma de cuadrados de tratamientos.}$$

$$SSE = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 = \text{suma de cuadrados del error.}$$

La identidad de la suma de cuadrados se puede representar entonces de manera simbólica con la ecuación

$$SST = SSA + SSE.$$

La identidad anterior expresa como se suma la variación entre tratamientos y dentro del tratamiento para dar la suma total de cuadrados. Sin embargo, se puede obtener mucho conocimiento al investigar el valor esperado de SSA y SSE . Finalmente, desarrollaremos estimaciones de la varianza que formulen la razón a utilizar para probar la igualdad de medias poblacionales.

Como se indicó antes, necesitamos comparar la medición apropiada de la variación entre tratamientos con la variación dentro del tratamiento a fin de detectar diferencias significativas en las observaciones debidas a los efectos del tratamiento. Suponga que consideramos el valor esperado de la suma de cuadrados de tratamientos.

Fórmulas de cálculo para las sumas de cuadrados

Las fórmulas alternativas para las sumas de cuadrados que se comentaron antes a menudo se presentan como métodos para calcular cantidades. El uso de software hace a esta fórmulas considerablemente menos importantes de lo que eran antes de la llegada de la computadora. Por otro lado, la expresión que se muestra en el teorema 13.1 permite al estudiante comprender de manera conceptual la variación expresada por cada suma de cuadrados individual. No obstante, a continuación proporcionamos las fórmulas de cálculo equivalentes.

$$SST = \sum_{i=1}^k \sum_{j=1}^n y_{ij}^2 - \frac{(\sum_i \sum_j y_{ij})^2}{nk}$$

$$SSA = \sum_{i=1}^k \frac{Y_i^2}{k} - \frac{(\sum_i \sum_j y_{ij})^2}{nk}$$

$$SSE = SS_T - SSA$$

Aquí

$$Y_i = \sum_{j=1}^n \frac{y_{ij}}{n},$$

es el total de los valores de respuesta en el i -ésimo tratamiento. El ejercicio 10 de la página 475 pide la demostración de la equivalencia de las expresiones en el teorema 13.1 con las anteriores.

Teorema 13.2

$$E(SSA) = (k - 1)\sigma^2 + n \sum_{i=1}^k \alpha_i^2$$

La demostración del teorema se deja como ejercicio (véase el ejercicio 2 de la página 473).

Si H_0 es verdadera, una estimación de σ^2 , que se basa en $k - 1$ grados de libertad, la proporciona la expresión

Cuadrado medio del tratamiento

$$s_1^2 = \frac{SSA}{k-1}$$

Si H_0 es verdadera y por ello cada α_i en el teorema 13.2 es igual a cero, vemos que

$$E\left(\frac{SSA}{k-1}\right) = \sigma^2,$$

y s_1^2 es un estimador insesgado de σ^2 . Sin embargo, si H_1 es verdadera, tenemos

$$E\left(\frac{SSA}{k-1}\right) = \sigma^2 + \frac{n \sum_{i=1}^k \alpha_i^2}{k-1}$$

y s_1^2 estima a σ^2 más un término adicional, que mide la variación debida a los efectos sistemáticos.

Un segundo e independiente estimador de σ^2 , que se basa en $k(n-1)$ grados de libertad, es la fórmula ya familiar

Cuadrado medio del error

$$s^2 = \frac{SSE}{k(n-1)}$$

La estimación s^2 es insesgada a pesar de que la hipótesis nula sea verdadera o falsa (véase el ejercicio 1 de la página 473). Es importante notar que la identidad de la suma de cuadrados no sólo divide la variabilidad total de los datos, sino también el número total de grados de libertad. Es decir,

$$nk - 1 = k - 1 + k(n - 1).$$

Cuando H_0 es verdadera, la razón $f = \frac{s_1^2}{s^2}$ es un valor de la variable aleatoria F que tiene la distribución F con $k - 1$ y $k(n - 1)$ grados de libertad. Como s_1^2 sobrestima σ^2 cuando H_0 es falsa, tenemos una prueba de una cola con la región crítica completamente en la cola derecha de la distribución.

La hipótesis nula H_0 se rechaza en el nivel de significancia α cuando

$$f > f_{\alpha}[k-1, k(n-1)].$$

Otro enfoque, la aproximación del valor P , sugiere que la evidencia a favor o en contra de H_0 es

$$p = P\{F[k-1, k(n-1)] > f\}.$$

Tabla 13.3 Análisis de varianza para la clasificación unilateral

Fuente de la variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Tratamientos	SSA	$k - 1$	$s_1^2 = \frac{SSA}{k - 1}$	$\frac{s_1^2}{s^2}$
Error	SSE	$k(n - 1)$	$s^2 = \frac{SSE}{k(n - 1)}$	
Total	SST	$nk - 1$		

Los cálculos para un problema de análisis de varianza por lo general se resumen en forma tabular como se muestra en la tabla 13.3.

Ejemplo 13.1 Pruebe la hipótesis $\mu_1 = \mu_2 = \dots = \mu_5$ en el nivel de significancia de 0.05 para los datos de la tabla 13.1 sobre la absorción de la humedad por varios tipos de mezclas de cemento.

SOLUCIÓN

$$H_0: \mu_1 = \mu_2 = \dots = \mu_5.$$

$$H_1: \text{Al menos dos de las medias no son iguales.}$$

$$\alpha = 0.05.$$

Región crítica: $f > 2.76$ con $v_1 = 4$ y $v_2 = 25$ grados de libertad.

Cálculos:

$$SST = (557 - 561.8)^2 + (457 - 561.8)^2 + \dots + (679 - 561.8)^2 = 209,377.$$

$$SSA = 6[(553.33 - 561.8)^2 + (569.33 - 561.8)^2 + \dots + (610.67 - 561.8)^2] \\ = 85,356.$$

$$SSE = 209,377 - 85,356 = 124,021.$$

Estos resultados y los cálculos restantes se exhiben en la tabla 13.4 en el procedimiento SAS ANOVA.

Decisión: rechazar H_0 y concluir que las mezclas no tienen la misma absorción media. El valor P para $f = 4.30$ es menor que 0.01.

Durante el trabajo experimental con frecuencia se pierden algunas de las observaciones que se desean. Los animales de experimentación se mueren, el material experimental se puede dañar y los sujetos humanos pueden abandonar un estudio. El análisis anterior para tamaños iguales de muestras aún será válido al modificar ligeramente las fórmulas de las sumas de cuadrados. Supondremos ahora que las k muestras aleatorias son de tamaño n_1, n_2, \dots, n_k , respectivamente. ■

Suma de cuadrados; tamaños diferentes de muestras

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2, \quad SSA = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2, \quad SSE = SST - SSA.$$

Tabla 13.4 Procedimiento del análisis de varianza

Analysis of Variance Procedure					
Dependent Variable: Y					
Absorption of Moisture					
Source	DF	Sun of Squares	Mean Square	F Value	Pr > F
Model	4	85356.466667	21339.116667	4.30	0.0088
Error	25	124020.333333	4960.813333		
Corrected Total	29	209376.800000			
	R-Square	C.V.	Root MSE		Y Mean
	0.407669	12.53703	70.433041		561.80000
Source	DF	Anova SS	Mean Square	F Value	Pr > F
AGG	4	85356.466667	21339.116667	4.30	0.0088

Los grados de libertad se dividen entonces como antes: $N - 1$ para SST , $k - 1$ para SSA , y $N - 1 - (k - 1) = N - k$ para SSE .

Ejemplo 13.2 Parte de un estudio que se llevó a cabo en el Instituto Politécnico y Universidad Estatal de Virginia se diseñó para medir los niveles de actividad de fosfatasa alcalina en suero (unidades de Bessey-Lowry) de niños con crisis convulsivas que reciben terapia contra convulsiones bajo el cuidado de un médico particular. Se encontraron 45 sujetos para el estudio y se clasificaron en cuatro grupos según el medicamento administrado:

- G-1: Control (no reciben anticonvulsivos y no tienen historial de crisis convulsivas)
- G-2: Fenobarbital
- G-3: Carbamacepina
- G-4: Otros anticonvulsivos

Se determinó el nivel de actividad de la fosfatasa alcalina en suero a partir de muestras sanguíneas obtenidas de cada sujeto y se registran en la tabla 13.5. Pruebe la hipótesis al nivel de significancia de 0.05 de que el nivel promedio de actividad de fosfatasa alcalina en suero es el mismo para los cuatro grupos.

SOLUCIÓN

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1: \text{Al menos dos de las medias no son iguales.}$$

$$\alpha = 0.05$$

Región crítica: $f > 2.836$, por interpolación en la tabla A.6.

Cálculos: $Y_1 = 1460.25$, $Y_2 = 440.36$, $Y_3 = 842.45$, $Y_4 = 707.41$, y $Y_5 = 3450.47$. El análisis de varianza se muestra en la salida del programa MINITAB en la tabla 13.6.

Tabla 13.5 Nivel de actividad de fosfatasa alcalina en suero

	G-1	G-2	G-3	G-4
	49.20	97.50	97.07	62.10
	44.54	105.00	73.40	94.95
	45.80	58.05	68.50	142.50
	95.84	86.60	91.85	53.00
	30.10	58.35	106.60	175.00
	36.50	72.80	0.57	79.50
	82.30	116.70	0.79	29.50
	87.85	45.15	0.77	78.40
	105.00	70.35	0.81	127.50
	95.22	77.40		

Tabla 13.6 Análisis con MINITAB de la tabla 13.5

```

MTB > Oneway C1 C2
One-Way Analysis of Variance

Analysis of Variance for C1
Source      DF      SS      MS      F      P
C2          3      13939   4646   3.57   0.022
Error      41      53376   1302
Total      44      67315

Level      N      Mean     StDev
1          20     73.01    25.75
2          9      48.93    47.11
3          9      93.61    46.57
4          7     101.06    30.76

Pooled StDev = 36.08
MTB >

```

Decisión: rechazar H_0 y concluir que los niveles promedio de actividad de fosfatasa alcalina en suero para los cuatro grupos no son los mismos. El valor P es 0.02. ■

Para concluir nuestra presentación del análisis de varianza para la clasificación unilateral, establecemos las ventajas de elegir tamaños iguales sobre la elección de tamaños diferentes de las muestras. La primera ventaja es que la razón f es insensible a ligeras variaciones de la suposición de varianzas iguales para las k poblaciones cuando las muestras son de tamaño igual. Segunda, la elección de tamaños de muestras iguales minimiza la probabilidad de cometer un error tipo II.

13.4 Pruebas de la igualdad de varias varianzas

Aunque la razón f que se obtiene del procedimiento de análisis de varianza es insensible a desviaciones de la suposición de varianzas iguales para las k poblaciones normales cuando las muestras son de tamaño igual, aún podemos preferir tener pre-

caución y realizar una prueba preliminar para la homogeneidad de las varianzas. Tal prueba en realidad sería aconsejable en el caso de tamaños de muestras diferentes si hay una duda razonable con respecto a la homogeneidad de las varianzas poblacionales. Suponga, por tanto, que deseamos probar la hipótesis nula

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

contra la alternativa

H_1 : No todas las varianzas son iguales.

La prueba que utilizaremos, llamada **prueba de Bartlett**, se basa en una estadística cuya distribución muestral proporciona valores críticos exactos cuando los tamaños muestrales son iguales. Estos valores críticos para tamaños iguales de muestras también se pueden utilizar para obtener aproximaciones altamente precisas a los valores críticos para tamaños diferentes de las muestras.

Primero, calculamos las k varianzas muestrales $s_1^2, s_2^2, \dots, s_k^2$ a partir de muestras de tamaños n_1, n_2, \dots, n_k , con $\sum_{i=1}^k n_i = N$. Segundo, combinamos las varianzas muestrales para obtener la estimación combinada

$$s_p^2 = \frac{\sum_{i=1}^k (n_i - 1)s_i^2}{N - k}$$

Ahora bien,

$$b = \frac{[(s_1^2)^{n_1-1} (s_2^2)^{n_2-1} \dots (s_k^2)^{n_k-1}]^{1/(N-k)}}{s_p^2}$$

es un valor de una variable aleatoria B que tiene la **distribución de Bartlett**. Para el caso especial donde $n_1 = n_2 = \dots = n_k = n$, rechazamos H_0 en el nivel de significancia α si

$$b < b_k(\alpha; n),$$

donde $b_k(\alpha; n)$ es el valor crítico que deja un área de tamaño α en la cola izquierda de la distribución de Bartlett. La tabla A.10 da los valores críticos, $b_k(\alpha; n)$ para $\alpha = 0.01$ y 0.05 ; $k = 2, 3, \dots, 10$; y valores de n seleccionados de 3 a 100.

Cuando los tamaños muestrales son diferentes, la hipótesis nula se rechaza en el nivel de significancia α si

$$b < b_k(\alpha; n_1, n_2, \dots, n_k),$$

donde

$$b_k(\alpha; n_1, n_2, \dots, n_k) = \frac{n_1 b_k(\alpha; n_1) + n_2 b_k(\alpha; n_2) + \dots + n_k b_k(\alpha; n_k)}{N}$$

Como antes, todas las $b(\alpha; n_i)$ para los tamaños muestrales n_1, n_2, \dots, n_k se obtienen de la tabla A.10.

Ejemplo 13.3 Utilice la prueba de Bartlett para probar la hipótesis, en el nivel de significancia 0.01, que las varianzas poblacionales de los cuatro grupos de medicamentos del ejemplo 13.2 son iguales.

SOLUCIÓN

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2.$$

$$H_1: \text{No todas las varianzas son iguales.}$$

$$\alpha = 0.01.$$

Región crítica: con referencia al ejemplo 13.2, tenemos $n_1 = 20$, $n_2 = 9$, $n_3 = 9$, $n_4 = 7$, $N = 45$, y $k = 4$. Por tanto, rechazamos cuando

$$\begin{aligned} b &< b_4(0.01, 20, 9, 9, 7) \\ &= \frac{(20)(0.8586) + (9)(0.6892) + (9)(0.6892) + (7)(0.6045)}{45} \\ &= 0.7513 \end{aligned}$$

Cálculos: calculamos primero

$$s_1^2 = 662.862, \quad s_2^2 = 2219.781, \quad s_3^2 = 2168.434, \quad s_4^2 = 946.032,$$

y después

$$\begin{aligned} s_p^2 &= \frac{(19)(662.862) + (8)(2219.781) + (8)(2168.434) + (6)(946.032)}{41} \\ &= 1301.861. \end{aligned}$$

Ahora

$$\begin{aligned} b &= \frac{[(662.862)^{19}(2219.781)^8(2168.434)^8(946.032)^6]^{1/4}}{1301.861} \\ &= 0.8557 \end{aligned}$$

Decisión: no rechazar la hipótesis y concluir que las varianzas poblacionales de los cuatro grupos de medicamentos no son significativamente diferentes. ■

Aunque la prueba de Bartlett se utiliza con más frecuencia para probar la homogeneidad de las varianzas, se dispone de otros métodos. Un método que se debe a Cochran proporciona un procedimiento simple de cálculo, pero se restringe a situaciones en las que los tamaños muestrales son iguales. La **prueba de Cochran** es útil en particular para detectar si una varianza es mucho más grande que las otras. La estadística que se utiliza es

$$G = \frac{\text{la mayor } S_i^2}{\sum_{i=1}^k S_i^2},$$

y la hipótesis de igualdad de varianzas se rechaza si $g > g_\alpha$, donde el valor de g_α se obtiene de la tabla A.11.

Para ilustrar la prueba de Cochran nos referimos de nuevo a los datos de la tabla 13.1 sobre la absorción de humedad en mezclas de concreto. ¿Se justifica la so-

posición de varianzas iguales cuando realizamos el análisis de varianza del ejemplo 13.1? Encontramos que

$$s_1^2 = 12,134, \quad s_2^2 = 2303, \quad s_3^2 = 3594, \quad s_4^2 = 3319, \quad s_5^2 = 3455.$$

Por tanto,

$$B = \frac{12,134}{24,805} = 0.4892,$$

que no excede del valor de la tabla $g_{0.05} = 0.5065$. De aquí concluimos que la suposición de varianzas iguales es razonable. ■

Ejercicios

1. Muestre que el cuadrado medio del error

$$s^2 = \frac{SSE}{k(n-1)}$$

para el análisis de varianza en una clasificación unilateral es una estimación insesgada de σ^2 .

2. Demuestre el teorema 13.2.

3. Se consideran seis diferentes máquinas para su uso en la fabricación de sellos de caucho. Las máquinas se deben comparar con respecto a la resistencia a la tracción del producto. Se utiliza una muestra aleatoria de cuatro sellos de cada máquina para determinar si la resistencia media a la tracción varía de una máquina a otra. Las siguientes son las mediciones de la resistencia a la tracción en kilogramos por centímetro cuadrado $\times 10^{-2}$:

Máquina

1	2	3	4	5	6
17.5	16.4	20.3	14.6	17.5	18.3
16.9	19.2	15.7	16.7	19.2	16.2
15.8	17.7	17.8	20.8	16.5	17.5
18.6	15.4	18.9	18.9	20.5	20.1

Realice el análisis de varianza al nivel de significancia de 0.05 e indique si las resistencias medias a la tracción difieren de manera significativa para las seis máquinas.

4. Los datos de la siguiente tabla representan el número de horas de alivio que proporcionan cinco marcas diferentes de tabletas contra el dolor de cabeza que se administran a 25 sujetos que sufren fiebres de 38°C o más. Realice el análisis de varianza y pruebe la hipótesis, en el nivel de significancia de 0.05, que el número medio de horas de alivio que proporcionan las tabletas es el mismo para las cinco marcas.

Tableta

A	B	C	D	E
5.2	9.1	3.2	2.4	7.1
4.7	7.1	5.8	3.4	6.6
8.1	8.2	2.2	4.1	9.3
6.2	6.0	3.1	1.0	4.2
3.0	9.1	7.2	4.0	7.6

5. En el artículo *Shelf-Space Strategy in Retailing*, publicado en *Proceedings: Southern Marketing Association*, se investiga el efecto de la altura del anaquel sobre las ventas de comida enlatada para perros en los supermercados. Se realizó un experimento sobre las ventas de una sola marca de comida para perros, a la que se denominará comida para perros Arf, en un pequeño supermercado durante un periodo de ocho días, con tres niveles de altura del anaquel: a las rodillas, a la cintura y a los ojos. Durante cada día se cambió de forma aleatoria en tres ocasiones diferentes el nivel del anaquel. Las secciones restantes de la góndola que contenían la marca dada se llenaron con una mezcla de marcas de comida para perros ya conocidas y desconocidas para los clientes de esta área geográfica particular. Las ventas, en cientos de dólares, de comida para perros Arf por día para las tres alturas del anaquel son las siguientes:

Altura del anaquel

Nivel de la rodilla	Nivel de la cintura	Nivel de los ojos
77	88	85
82	94	85
86	93	87
78	90	81
81	91	80
86	94	79
77	90	87
81	87	93

¿Existe una diferencia significativa en las ventas promedio diarias de esta comida para perros que se basa en la altura del anaquel? Utilice un nivel de significancia de 0.01.

6. La inmovilización de venados cola blanca salvajes mediante fármacos permite a los investigadores la oportunidad de examinar de cerca a los venados y reunir información fisiológica valiosa. En el estudio *Influence of Physical Restraint and Restraint-Facilitating Drugs on Blood Measurements of White-Tailed Deer and Other Selected Mammals* que se realizó en el Instituto Politécnico y Universidad Estatal de Virginia, biólogos de la fauna probaron el tiempo de "caída" (tiempo entre la inyección y la inmovilización) de tres diferentes fármacos inmovilizadores. La inmovilización, en este caso, se define como el punto en el que el animal ya no tiene suficiente control muscular para permanecer parado. Se asignaron treinta venados cola blanca machos a cada uno de tres tratamientos. El grupo A recibió 5 miligramos de cloruro de succinilcolina líquido (SCC); el grupo B recibió 8 miligramos de SCC en polvo; y el grupo C recibió 200 miligramos de hidrocloreuro de fenciclidina. Los tiempos de caída, en minutos, se registraron como sigue:

Grupo:	A	B	C
	11	10	4
	5	7	4
	14	16	6
	7	7	3
	10	7	5
	7	5	6
	23	10	8
	4	10	3
	11	6	7
	11	12	3

Realice un análisis de varianza en el nivel de significancia de 0.01 y determine si el tiempo de caída promedio para los tres fármacos es el mismo o no.

7. Se ha mostrado que el fertilizante fosfato amoniacal de magnesio, $MgNH_4PO_4$, es un proveedor efectivo de los nutrientes necesarios para el crecimiento de las plantas. Los compuestos que suministra este fertilizante son altamente solubles en agua, lo que permite que el fertilizante se aplique de forma directa sobre la superficie del suelo o mezclado con el sustrato de crecimiento durante el proceso de sembrado. En la Universidad George Mason se llevó a cabo un estudio sobre el efecto del fosfato amoniacal de magnesio sobre la altura de los crisantemos (*Effect of Magnesium*

Ammonium Phosphate on Height of Chrysanthemums) para determinar un nivel óptimo posible de fertilización, con base en la respuesta mejorada de crecimiento vertical de los crisantemos. Se dividieron cuarenta plantas de crisantemos en cuatro grupos cada uno de 10 plantas. Cada uno se plantó en una parcela similar que contenía un medio de crecimiento uniforme. Para cada grupo de plantas se agregó una concentración creciente de $MgNH_4PO_4$, medido en gramos por bushel. Los cuatro grupos de plantas se cultivaron bajo condiciones uniformes en un invernadero durante un periodo de cuatro semanas. Los tratamientos y los cambios respectivos en las alturas, medidas en centímetros, se muestran en la siguiente tabla:

Tratamiento			
50 g/bu	100 g/bu	200 g/bu	400 g/bu
13.2	16.0	7.8	21.0
12.4	12.6	14.4	14.8
12.8	14.8	20.0	19.1
17.2	13.0	15.8	15.8
13.0	14.0	17.0	18.0
14.0	23.6	27.0	26.0
14.2	14.0	19.6	21.1
21.6	17.0	18.0	22.0
15.0	22.2	20.2	25.0
20.0	24.4	23.2	18.2

¿Podemos concluir en el nivel de significancia de 0.05 que las diferentes concentraciones de $MgNH_4PO_4$ afectan la altura promedio alcanzada por los crisantemos?

8. Un estudio mide la tasa de absorción de tres tipos diferentes de solventes químicos orgánicos. Estos solventes se utilizan para limpiar partes metálicas industriales labradas y son desechos peligrosos potenciales. Se prueban muestras independientes de solventes de cada tipo y se registran sus tasas de absorción como porcentaje molar. [Véase McClellan, Dietrich y Sincich (1997).]

Aromáticos		Cloroalcanos		Ésteres	
1.06	0.95	1.58	1.12	0.29	0.43
0.79	0.65	1.45	0.91	0.06	0.51
0.82	1.15	0.57	0.83	0.44	0.10
0.89	1.12	1.16	0.43	0.55	0.53
1.05				0.61	0.34

¿Existe una diferencia significativa en la tasa de absorción media para los tres solventes? Utilice un valor P para sus conclusiones.

9. La enzima mitocondrial transhidrogenasa NADPH:NAD de la tenia de rata común (*Hymenolepsis diminuta*) cataliza hidrógeno en la transferencia de NADPH a NAD, lo que produce NADH. Se sabe que esta enzima juega un papel vital en el metabolismo anaeróbico de la tenia, y recientemente se planteó la hipótesis de que puede servir como bomba de intercambio de protones, mediante la transferencia de protones a través de la membrana mitocondrial. En la Universidad Estatal Bowling Green se llevó a cabo el estudio titulado *Effect of Various Substrate Concentrations on the Conformational Variation of the NADPH:NAD Transhydrogenase of Hymenolepsis Diminuta* ("Efecto de varias concentraciones de sustrato sobre la variación de la conformación de la transhidrogenasa NADPH:NAD de la rata común"), el cual se diseñó para evaluar la capacidad de esta enzima para experimentar cambios de conformación o de forma. Los cambios en la actividad específica de la enzima ocasionados por variaciones en la concentración de NADP se podrían interpretar como el apoyo de la teoría del cambio de conformación. La enzima en cuestión se localiza en la membrana interior de la mitocondria de la tenia. Estas tenias se homogeneizaron,

y a través de una serie de centrifugados se aisló la enzima. Después se agregaron varias concentraciones de NADP a la solución de enzima aislada, y la mezcla se incubó en un baño de agua a 56°C durante tres minutos. La enzima se analizó entonces con un espectrómetro de doble haz y se calcularon los siguientes resultados en términos de la actividad específica de la enzima en nanomoles por minuto por miligramo de proteína:

Concentración de NADP (nm)				
0	80	160	360	
11.01	11.38	11.02	6.04	10.31
12.09	10.67	10.67	8.65	8.30
10.55	12.33	11.50	7.76	9.48
11.26	10.08	10.31	10.13	8.89
			9.36	

Pruebe la hipótesis en el nivel de significancia de 0.01 de que la actividad específica promedio es la misma para las cuatro concentraciones.

10. Muestre la equivalencia de las sumas de cuadrados del teorema 13.1 con las fórmulas de cálculo anteriores al teorema 13.2.

13.5 Comparaciones de un solo grado de libertad

El análisis de varianza en una clasificación unilateral, o experimento de un factor, como a menudo se le llama, solamente indica si la hipótesis de medias iguales de tratamientos se puede rechazar o no. Por lo general, un experimentador preferiría que su análisis realizara pruebas más profundas. Como ilustración, en el ejemplo 13.1, al rechazar la hipótesis nula concluimos que las medias no son todas iguales, pero aún no sabemos dónde existen las diferencias entre las mezclas. El ingeniero podría tener la sensación *a priori* de que las mezclas 1 y 2 deben tener propiedades similares de absorción y que lo mismo es cierto para las mezclas 3 y 5. Sin embargo, es de interés estudiar la diferencia entre los dos grupos. Parecería, entonces, apropiado probar la hipótesis

$$H_0: \mu_1 + \mu_2 - \mu_3 - \mu_5 = 0,$$

$$H_1: \mu_1 + \mu_2 - \mu_3 - \mu_5 \neq 0.$$

Notamos que la hipótesis es una función lineal de las medias poblacionales donde los coeficientes suman cero.

Definición 13.1

Cualquier función de la forma

$$w = \sum_{i=1}^k c_i u_i, \quad \text{donde } \sum_{i=1}^k c_i = 0$$

se llama **comparación** o **contraste** de las medias de los tratamientos.

El experimentador a menudo puede realizar comparaciones múltiples mediante la prueba de la significancia de contrastes en las medias de los tratamientos; es decir, por medio de la prueba de una hipótesis de tipo

$$H_0: \sum_{i=1}^k c_i u_i = 0,$$

$$H_1: \sum_{i=1}^k c_i u_i \neq 0,$$

donde $\sum_{i=1}^k c_i = 0$. La prueba se lleva a cabo al calcular primero un contraste similar en las medias muestrales,

$$w = \sum_{i=1}^k c_i \bar{Y}_i.$$

Como $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_k$ son variables aleatorias independientes que tienen distribuciones normales con medias $\mu_1, \mu_2, \dots, \mu_k$ y varianzas $\sigma^2/n_1, \sigma^2/n_2, \dots, \sigma^2/n_k$, respectivamente, el teorema 7.11 nos asegura que w es un valor de la variable aleatoria normal W con media

$$\mu_w = \sum_{i=1}^k c_i \mu_i$$

y varianza

$$\sigma_w^2 = \sigma^2 \sum_{i=1}^k \frac{c_i^2}{n_i}.$$

Por tanto, cuando H_0 es verdadera, $\mu_w = 0$ y, por el ejemplo 7.5, la estadística

$$\frac{W^2}{\sigma_w^2} = \frac{\left(\sum_{i=1}^k c_i \bar{Y}_i \right)^2}{\sigma^2 \sum_{i=1}^k (c_i^2/n_i)}$$

se distribuye como una variable aleatoria ji cuadrada con 1 grado de libertad. Nuestra hipótesis se prueba al nivel de significancia α al calcular

$$f = \frac{\left(\sum_{i=1}^k c_i \bar{Y}_i \right)^2}{s^2 \sum_{i=1}^k (c_i^2/n_i)} = \frac{\left[\sum_{i=1}^k (c_i Y_i/n_i) \right]^2}{s^2 \sum_{i=1}^k (c_i^2/n_i)} = \frac{SSw}{s^2},$$

donde f es un valor de la variable aleatoria F que tiene la distribución F con 1 y $N - k$ grados de libertad y

$$SSw = \frac{\left[\sum_{i=1}^k (c_i Y_i / n_i) \right]^2}{\sum_{i=1}^k (c_i^2 / n_i)}$$

Cuando los tamaños de las muestras son iguales a n ,

$$SSw = \frac{\left(\sum_{i=1}^k c_i Y_i \right)^2}{n \sum_{i=1}^k c_i^2}$$

La cantidad SSw , que se llama **suma de cuadrados de contraste**, indica la parte de SSA que se explica mediante el contraste en cuestión.

Esta suma de cuadrados se utilizará para probar la hipótesis que el contraste

$$\sum_{i=1}^k c_i \mu_i = 0.$$

A menudo es de interés probar contrastes múltiples, en particular contrastes que son linealmente independientes u ortogonales. Como resultado necesitamos la siguiente definición:

Definición 13.2

Se dice que los dos contrastes

$$\omega_1 = \sum_{i=1}^k b_i \mu_i \quad \text{y} \quad \omega_2 = \sum_{i=1}^k c_i \mu_i$$

son **ortogonales** si $\sum_{i=1}^k b_i c_i / n_i = 0$ o, cuando las n_i son todas iguales a n , si

$$\sum_{i=1}^k b_i c_i = 0.$$

Si ω_1 y ω_2 son ortogonales, entonces las cantidades SSw_1 y SSw_2 son componentes de SSA , cada una con un solo grado de libertad. La suma de cuadrados de los tratamientos con $k - 1$ grados de libertad se puede dividir en a lo más $k - 1$ sumas de cuadrados de contraste independientes con un solo grado de libertad que satisfacen la identidad

$$SSA = SSw_1 + SSw_2 + \cdots + SSw_{k-1}$$

si los contrastes son ortogonales entre ellos.

Ejemplo 13.4 Con referencia al ejemplo 13.1, encuentre la suma de cuadrados de contrastes que corresponden a los contrastes ortogonales

$$\omega_1 = \mu_1 + \mu_2 - \mu_3 - \mu_5, \quad \omega_2 = \mu_1 + \mu_2 + \mu_3 + \mu_5 - 4\mu_4,$$

y realice pruebas de significancia apropiadas. En este caso es de interés *a priori* comparar los dos grupos (1, 2) y (3, 5). Un contraste importante e independiente es la comparación entre el conjunto de mezclas (1, 2, 3, 5) y la mezcla 4.

SOLUCIÓN

Es obvio que los dos contrastes son ortogonales, pues

$$(1)(1) + (1)(1) + (-1)(1) + (0)(-4) + (-1)(1) = 0.$$

El segundo contraste indica una comparación entre las mezclas (1, 2, 3 y 5) y la mezcla 4. Se pueden escribir dos contrastes ortogonales adicionales a los primeros dos, a saber:

$$w_3 = \mu_1 - \mu_2 \quad (\text{mezcla 1 contra mezcla 2}),$$

$$w_4 = \mu_3 - \mu_5 \quad (\text{mezcla 3 contra mezcla 5}).$$

De los datos de la tabla 13.1, tenemos

$$SSw_1 = \frac{(3320 + 3416 - 3663 - 3664)^2}{6[(1)^2 + (1)^2 + (-1)^2 + (-1)^2]} = 14,553,$$

$$SSw_2 = \frac{[3320 + 3416 + 3663 + 3664 - 4(2791)]^2}{6[(1)^2 + (1)^2 + (1)^2 + (1)^2 + (-4)^2]} = 70,035.$$

Una tabla de análisis de varianza más extensa se muestra en la tabla 13.7. Notamos que las dos sumas de cuadrados de contraste explican casi todas las sumas de cuadrados de las mezclas. Aunque hay una diferencia significativa entre las mezclas en sus propiedades de absorción, el contraste w_1 es marginalmente significativo. El valor P es 0,009. Sin embargo, el valor f de 14,12 para w_2 es bastante significativo y se rechaza la hipótesis

$$H_0: \mu_1 + \mu_2 + \mu_3 + \mu_5 = 4\mu_4$$

es rechazada.

Los contrastes ortogonales permiten al practicante dividir la variación del tratamiento en componentes independientes. Se dispone de varias elecciones al seleccionar los contrastes ortogonales excepto para el último. Normalmente, el experimentador tendría ciertos contrastes de su interés. Tal es el caso en nuestro ejemplo, donde consideraciones *a priori* sugieren que las mezclas (1, 2) y (3, 5) constituyen grupos distintos con propiedades de absorción diferentes, postura que no es apoyada suficientemente por la prueba de significancia. Sin embargo, la segunda comparación apoya la conclusión de que la mezcla 4 parece "resaltar" entre las demás. En este caso la división completa de SSA no es necesaria, pues dos de las cua-

Tabla 13.7 Análisis de varianza con el uso de contrastes ortogonales

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Mezclas	85,356	4	21,339	4,30
(1, 2) vs. (3, 5)	14,553	1	14,533	2,93
(1, 2, 3, 5) vs. 4	70,035	1	70,035	14,12
Error	124,021	25	4,961	
Total	209,377	29		

tro posibles comparaciones independientes explican la mayor parte de la variación en los tratamientos.

La tabla 13.8 muestra un procedimiento SAS GLM que exhibe un conjunto completo de contrastes ortogonales. Nótese que la suma de cuadrados para los cuatro contrastes se agrega a la suma de cuadrados de la mezcla. Además, que los últimos dos contrastes (1 contra 2 y 3 contra 5) revelan comparaciones insignificantes.

Tabla 13.8 Conjunto de procedimientos ortogonales

General Linear Models Procedure					
Dependent Variable: Y Absorption of Moisture					
Source	DF	Sun of Squares	Mean Square	F Value	Pr > F
Model	4	85356.466667	21339.116667	4.30	0.0088
Error	25	124020.333333	4960.813333		
Corrected Total	29	209376.800000			
		R-Square	C.V.	Root MSE	Y Mean
		0.407669	12.53703	70.433041	561.80000
Source	DF	Type I SS	Mean Square	F Value	Pr > F
AGG	4	85356.466667	21339.116667	4.30	0.0088
Source	DF	Type III SS	Mean Square	F Value	Pr > F
AGG	4	85356.466667	21339.116667	4.30	0.0088
Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
(1,2,3,5) vs 4	1	70035.008333	70035.008333	14.12	0.0009
(1,2) vs (3,5)	1	14553.375000	14553.375000	2.93	0.0991
1 vs 2	1	768.000000	768.000000	0.15	0.6973
3 vs 5	1	0.083333	0.083333	0.00	0.9968

13.6 Comparaciones múltiples

El análisis de varianza es un procedimiento poderoso para probar la homogeneidad de un conjunto de medias. Sin embargo, si rechazamos la hipótesis nula y aceptamos la alternativa establecida —que no todas las medias son iguales— aún no sabemos cuáles de las medias poblacionales son iguales y cuáles son diferentes.

En la sección 13.5 describimos el uso de contrastes ortogonales para hacer comparaciones entre conjuntos de niveles de factor o tratamientos. La noción de ortogonalidad permite al analista realizar pruebas que incluyen contrastes *independientes*. De esta manera la variación entre los tratamientos, SSA, se puede dividir

en componentes de un solo grado de libertad y entonces las proporciones de esta variación se pueden atribuir a contrastes específicos. Sin embargo, hay situaciones en las que el uso de contrastes no es un enfoque apropiado. A menudo es de interés realizar varias comparaciones pareadas (quizá todas las posibles) entre los tratamientos. En realidad, una comparación pareada se puede ver como un contraste simple, a saber, una prueba de

$$H_0: \mu_i - \mu_j = 0,$$

$$H_1: \mu_i - \mu_j \neq 0$$

para toda $i \neq j$. Todas las posibles comparaciones pareadas entre las medias pueden ser muy benéficas cuando no se conocen a priori contrastes complejos particulares. Por ejemplo, en los datos de mezclas de la tabla 13.1, suponga que deseamos probar

$$H_0: \mu_1 - \mu_3 = 0,$$

$$H_1: \mu_1 - \mu_3 \neq 0.$$

La prueba se lleva a cabo a través del uso de una aproximación F , t o de un intervalo de confianza. Con el uso de t , tenemos

$$t = \frac{\bar{y}_1 - \bar{y}_3}{s \sqrt{2/n}}$$

donde s es la raíz cuadrada del cuadrado medio del error y $n = 6$ es el tamaño de la muestra por tratamiento. En este caso

$$t = \frac{553.33 - 610.67}{\sqrt{4961} \sqrt{1/3}} = -1.41.$$

El valor P para la prueba t con 25 grados de libertad es 0.17. Por ello no hay suficiente evidencia para rechazar H_0 .

Relación entre t y F

En lo anterior se mostró el uso de una prueba t combinada de acuerdo con lo que se discutió en el capítulo 10. La estimación combinada proviene del cuadrado medio del error a fin de aprovechar los grados de libertad que se combinan a través de las cinco muestras. Además, probamos un contraste. El lector debe notar que si el valor t se eleva al cuadrado, el resultado es exactamente de la forma del valor de f para una prueba sobre un contraste que se discute en la sección anterior. De hecho,

$$f = \frac{(\bar{y}_1 - \bar{y}_3)^2}{s^2 \left(\frac{1}{6} + \frac{1}{6}\right)} = \frac{(553.33 - 610.67)^2}{4961 \left(\frac{1}{3}\right)} = 1.988,$$

que, por supuesto, es t^2 .

Enfoque del intervalo de confianza para una comparación pareada

Es bastante directo resolver el mismo problema de una comparación pareada (o un contraste) con el uso del enfoque del intervalo de confianza. Claramente, si calculamos un intervalo de confianza de $100(1 - \alpha)\%$ sobre $\mu_1 - \mu_3$, tenemos

$$\bar{y}_1 - \bar{y}_5 \pm t_{\alpha/2, 8} \sqrt{\frac{s^2}{6}}$$

donde $t_{\alpha/2}$ es el punto $100(1 - \alpha/2)\%$ superior de una distribución t con 25 grados de libertad (grados de libertad que provienen de s^2). Esta conexión directa entre pruebas e intervalos de confianza debe ser obvia a partir de las discusiones en los capítulos 9 y 10. La prueba del contraste simple $\mu_1 - \mu_5$ no implica más que observar si el intervalo de confianza anterior abarca o no al cero. Al sustituir los números, tenemos como intervalo de confianza de 95%

$$(553.33 - 610.67) \pm 2.060 \sqrt{4961} \sqrt{\frac{1}{3}} = -57.34 \pm 83.77.$$

De esta manera, como el intervalo abarca al cero, el *contraste no es significativo*. En otras palabras, no encontramos una diferencia significativa entre las medias de las mezclas 1 y 5.

Tasa de error experimento-juicio

Demostramos que se puede realizar un contraste simple (es decir, una comparación de dos medias) a través de una prueba F como se muestra en la sección 13.5, con una prueba t , o mediante el cálculo de un intervalo de confianza sobre la diferencia entre las dos medias. Sin embargo, ocurren serias dificultades cuando el analista intenta hacer muchas o todas las posibles comparaciones pareadas. Para el caso de k medias habrá, por supuesto, $r = k(k - 1)/2$ posibles comparaciones pareadas. Al suponer comparaciones independientes, la *tasa de error experimento-juicio* (es decir, la probabilidad de falso rechazo de al menos una de las hipótesis) está dada por $1 - (1 - \alpha)^r$, donde α es la probabilidad seleccionada de error tipo I para una comparación específica. Claramente, esta medida del error tipo I del experimento-juicio puede ser bastante grande. Por ejemplo, incluso si sólo hay seis comparaciones, digamos, en el caso de cuatro medias, y $\alpha = 0.05$, la tasa de error del experimento-juicio es

$$1 - (0.95)^6 = 0.26.$$

Con la tarea de probar muchas comparaciones pareadas por lo general existe la necesidad de realizar el contraste efectivo en una sola comparación más conservadora. Es decir, con el uso del enfoque del intervalo de confianza, los intervalos de confianza serían mucho más amplios que $\pm t_{\alpha/2} S \sqrt{2/n}$, que se utiliza en el caso donde sólo se hace una sola comparación.

Prueba de Tukey

Hay varios métodos estándar para realizar comparaciones pareadas que apoyen la credibilidad de la tasa de error tipo I. Presentaremos e ilustraremos aquí dos de ellas. La primera, que se denomina **procedimiento de Tukey**, permite la formación de intervalos de confianza de $100(1 - \alpha)\%$ simultáneos para todas las comparaciones pareadas. El método se basa en la distribución del rango *studentizado*. El punto percentil apropiado es una función de α , k , y $v =$ grados de libertad para s^2 . Una lista de puntos porcentuales superiores para $\alpha = 0.05$ se muestra en la tabla A.22. El

método de comparaciones pareadas de Tukey implica encontrar una diferencia significativa entre las medias i y j ($i \neq j$) si $(\bar{y}_i - \bar{y}_j)$ excede $q[\alpha, k, v] s\sqrt{1/n}$.

El procedimiento de Tukey se ilustra con facilidad. Considere un ejemplo hipotético en el que tenemos seis tratamientos en un diseño de un factor completamente aleatorizado con cinco observaciones por tratamiento. Suponga que el cuadrado medio del error que se toma de la tabla de análisis de varianza es $s^2 = 2.45$ (24 grados de libertad). Las medias muestrales son, en orden ascendente,

\bar{y}_2	\bar{y}_3	\bar{y}_1	\bar{y}_3	\bar{y}_6	\bar{y}_4
14.50	16.75	19.84	21.12	22.90	23.20

Con $\alpha = 0.05$, el valor de $q(0.05, 6, 24) = 4.37$. Por ello todas las diferencias absolutas se comparan con

$$4.37 \sqrt{\frac{2.45}{5}} = 3.059.$$

Como resultado, las siguientes representan las medias que se encuentran que son significativamente diferentes con el procedimiento de Tukey:

4 y 1,	4 y 5,	4 y 2,	6 y 1,	6 y 5,
6 y 2,	3 y 5,	3 y 2,	1 y 5,	1 y 2.

Prueba de Duncan

El segundo procedimiento que presentaremos se llama **procedimiento de Duncan o prueba de rango múltiple de Duncan**. Este procedimiento también se basa en la noción general del rango studentizado. El rango de cualquier subconjunto de p medias muestrales debe exceder cierto valor antes de que se encuentre que cualquiera de las p medias es diferente. Este valor se llama **rango de menor significancia** para las p medias y se denota como R_p , donde

$$R_p = r_p \sqrt{\frac{s^2}{n}}.$$

Los valores de la cantidad r_p , que se denomina **rango studentizado de menor significancia**, depende del nivel de significancia que se desea y del número de grados de libertad del cuadrado medio del error. Estos valores se pueden obtener de la tabla A.12 para $p = 2, 3, \dots, 10$ medias.

Para ilustrar el procedimiento de prueba de rango múltiple, consideremos el ejemplo hipotético donde se comparan seis tratamientos con cinco observaciones por tratamiento. Es el mismo ejemplo que se utilizó para ilustrar la prueba de Tukey. Obtenemos R_p al multiplicar cada r_p por 0.70. Los resultados de estos cálculos se resumen como sigue:

p	2	3	4	5	6
r_p	2.919	3.066	3.160	3.226	3.276
R_p	2.043	2.146	2.212	2.258	2.293

Al comparar estos rangos de menor significancia con las diferencias en las medias ordenadas, llegamos a las siguientes conclusiones:

1. Como $\bar{y}_4 - \bar{y}_2 = 8.70 > R_6 = 2.293$, concluimos que \bar{y}_4 y \bar{y}_2 son significativamente diferentes.
2. Al comparar $\bar{y}_4 - \bar{y}_5$ y $\bar{y}_6 - \bar{y}_2$ con R_5 , concluimos que \bar{y}_4 es significativamente mayor que \bar{y}_5 y \bar{y}_6 es significativamente mayor que \bar{y}_2 .
3. Al comparar $\bar{y}_4 - \bar{y}_1$, $\bar{y}_6 - \bar{y}_5$, y $\bar{y}_3 - \bar{y}_2$ con R_4 , concluimos que cada diferencia es significativa.
4. Al comparar $\bar{y}_4 - \bar{y}_3$, $\bar{y}_6 - \bar{y}_1$, $\bar{y}_3 - \bar{y}_5$, y $\bar{y}_1 - \bar{y}_2$ con R_3 , encontramos que todas las diferencias son significativas excepto $\bar{y}_4 - \bar{y}_3$. Por tanto, \bar{y}_4 , \bar{y}_3 , y \bar{y}_6 constituyen un subconjunto de medias homogéneas.
5. Al comparar $\bar{y}_3 - \bar{y}_1$, $\bar{y}_1 - \bar{y}_5$, y $\bar{y}_5 - \bar{y}_2$ con R_2 , concluimos que sólo \bar{y}_3 con \bar{y}_1 no son significativamente diferentes.

Se acostumbra resumir las conclusiones anteriores mediante el dibujo de una línea bajo cualquier subconjunto de medias contiguas que no son significativamente diferentes. De esta forma, tenemos

\bar{y}_2	\bar{y}_5	\bar{y}_1	\bar{y}_3	\bar{y}_6	\bar{y}_4
14.50	16.75	19.84	21.12	22.90	23.20

Es claro que en este caso los resultados de los procedimientos de Tukey y de Duncan son muy similares. El procedimiento de Tukey no detecta una diferencia entre 2 y 5, mientras que el de Duncan sí lo hace.

13.7 Comparación de tratamientos con un control

En muchos problemas científicos y de ingeniería no interesa extraer inferencias con respecto a todas las posibles comparaciones entre las medias de los tratamientos del tipo $\mu_1 - \mu_j$. En su lugar, el experimento a menudo dicta la necesidad de comparar de manera simultánea cada *tratamiento* con un *control*. Un procedimiento de prueba desarrollado por C. W. Dunnett determina diferencias significativas entre cada media del tratamiento y el control, en un solo nivel de significancia α conjunto. Para ilustrar el procedimiento de Dunnett, consideremos los datos experimentales de la tabla 13.9 para la clasificación unilateral donde se estudia el efecto de tres catalizadores sobre el rendimiento de una reacción. Un cuarto tratamiento, sin ningún catalizador, se utiliza como control.

Tabla 13.9 Rendimiento de la reacción

Control	Catalizador 1	Catalizador 2	Catalizador 3
50.7	54.1	52.7	51.2
51.5	53.8	53.9	50.8
49.2	53.1	57.0	49.7
53.1	52.5	54.1	48.0
52.7	54.0	52.5	47.2
$\bar{y}_0 = 51.44$	$\bar{y}_1 = 53.50$	$\bar{y}_2 = 54.04$	$\bar{y}_3 = 49.38$

En general deseamos probar las k hipótesis

$$\left. \begin{array}{l} H_0: \mu_0 = \mu_i \\ H_1: \mu_0 \neq \mu_i \end{array} \right\} i = 1, 2, \dots, k,$$

donde μ_0 representa el rendimiento medio para la población de mediciones en la que se utiliza el control. Se espera que sigan siendo válidas las suposiciones del análisis de varianza acostumbradas, señaladas en la sección 13.2. Para probar las hipótesis nulas especificadas por H_0 contra las alternativas bilaterales para una situación experimental en la que hay k tratamientos, excluido el de control, y n observaciones por tratamiento, primero calculamos los valores

$$d_i = \frac{\bar{y}_i - \bar{y}_0}{\sqrt{2s^2/n}}, \quad i = 1, 2, \dots, k.$$

La varianza muestral s^2 se obtiene, como antes, del cuadrado medio del error en el análisis de varianza. Ahora bien, el régimen crítico para rechazar H_0 , en el nivel de significancia α , se establece mediante la desigualdad

$$|d_i| > d_{\alpha/2}(k, v),$$

donde v es el número de grados de libertad para el cuadrado medio del error. Los valores de la cantidad $d_{\alpha/2}(k, v)$ para una prueba de dos colas se dan en la tabla A.13 (véanse las tablas estadísticas) para $\alpha = 0.05$ y $\alpha = 0.01$ para varios valores de k y v .

Ejemplo 13.5 Para los datos de la tabla 13.9, pruebe las hipótesis mediante la comparación de cada catalizador con el control, con el uso de alternativas bilaterales. Elija $\alpha = 0.05$ como el nivel de significancia conjunto.

SOLUCIÓN

La suma de cuadrados del error con 16 grados de libertad se obtiene de la tabla de análisis de varianza, al usar todos los $k + 1$ tratamientos o mediante el cálculo directo a partir de la fórmula

$$SSE = \sum_{i=0}^k \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=0}^k Y_i^2}{n} = 54,371.960 - 54,335.148 = 36.812.$$

Entonces el cuadrado medio del error es

$$s^2 = \frac{36.812}{16} = 2.30075 \quad \text{y} \quad \sqrt{\frac{2s^2}{n}} = \sqrt{\frac{(2)(2.30075)}{5}} = 0.9593.$$

De aquí

$$d_1 = \frac{53.50 - 51.44}{0.9593} = 2.147,$$

$$d_2 = \frac{54.04 - 51.44}{0.9593} = 2.710,$$

$$d_3 = \frac{49.38 - 51.44}{0.9593} = -2.147.$$

De la tabla A.13 se encuentra que el valor crítico para $\alpha = 0.05$ es

$$d_{\text{crítico}}(3, 16) = 2.59$$

Como $|d_1| < 2.59$, y $|d_2| < 2.59$, concluimos que sólo el rendimiento medio para el catalizador 2 es significativamente diferente del rendimiento medio de la reacción con el uso del control. ■

Muchas aplicaciones prácticas dictan la necesidad de una prueba de una sola cola para comparar tratamientos con un control. En realidad, cuando un farmacólogo se preocupa por la comparación de varias dosis de un medicamento sobre el efecto de reducir el nivel de colesterol, y su control es la dosis cero, es de interés determinar si cada dosis produce una reducción significativamente mayor que la del control. La tabla A.14 muestra los valores críticos de $d_{\alpha}(k, v)$ para alternativas unilaterales.

Ejercicios

1. Considere los datos del ejercicio 5 de la página 524 para realizar pruebas de significancia en los siguientes contrastes:

- (a) B contra A, C, y D;
 (b) C contra A y D;
 (c) A contra D.

2. El Departamento de Nutrición Humana y Alimentos del Instituto Politécnico y Universidad Estatal de Virginia llevó a cabo el estudio *Loss of Nitrogen Through Sweat by Preadolescent Boys Consuming Three Levels of Dietary Protein* ("Pérdida de nitrógeno a través del sudor en preadolescentes que consumen tres niveles de proteína alimenticia") para determinar la pérdida de nitrógeno por transpiración en varios niveles de proteína alimenticia. En el experimento se utilizaron 12 preadolescentes cuyas edades iban de 7 años, 8 meses a 9 años, 8 meses, y considerados clínicamente saludables. Cada niño se sujetó a una de tres dietas controladas en las que se consumieron 29, 54 u 84 gramos de proteína por día. Los siguientes datos que representan la pérdida corporal de nitrógeno por transpiración, en miligramos, se reunieron durante los dos últimos días del período experimental:

Nivel de proteína		
29 gramos	54 gramos	84 gramos
190	318	390
266	295	321
270	271	396
	438	399
	402	

(a) Realice un análisis de varianza en el nivel de significancia de 0.05 para mostrar que las pérdidas medias de nitrógeno por transpiración en los tres niveles de proteína son diferentes.

(b) Utilice un contraste con un solo grado de libertad con $\alpha = 0.05$ para comparar la pérdida media de nitrógeno por transpiración para niños que consumen 29 gramos de proteína por día contra niños que consumen 54 y 84 gramos de proteína.

3. El propósito del estudio *The Incorporation of a Chelating Agent into a Flame Retardant Finish of a Cotton Flannelette and the Evaluation of Selected Fabric Properties* que se realizó en el Instituto Politécnico y Universidad Estatal de Virginia fue la evaluación del uso de un agente gelatinoso como parte del terminado para retardar la flama en una franela de algodón mediante la determinación de sus efectos en la inflamabilidad después de que se lava la tela bajo condiciones específicas. Se prepararon dos baños, uno con celulosa carboximetil y otro sin ésta. Se lavaron 12 piezas de tela cinco veces en el baño I, y otras 12 piezas de tela se lavaron 10 veces en el baño I. Esto se repitió con el uso de 24 piezas adicionales de tela en el baño II. Después de las lavadas, se midieron las longitudes de tela que se quemó y los tiempos de quemado. Por conveniencia, definamos los siguientes tratamientos:

- Tratamiento 1: 5 lavadas en el baño I;
 Tratamiento 2: 5 lavadas en el baño II;
 Tratamiento 3: 10 lavadas en el baño I;
 Tratamiento 4: 10 lavadas en el baño II.

Los tiempos de quemado, en segundos, se registraron como sigue:

Tratamiento			
1	2	3	4
13.7	6.2	27.2	18.2
23.0	5.4	16.8	8.8
15.7	5.0	12.9	14.5
25.5	4.4	14.9	14.7
15.8	5.0	17.1	17.1
14.8	3.3	13.0	13.9
14.0	16.0	10.8	10.6
29.4	2.5	13.5	5.8
9.7	1.6	25.5	7.3
14.0	3.9	14.2	17.7
12.3	2.5	27.4	18.3
12.3	7.1	11.5	9.9

(a) Realice un análisis de varianza con un nivel de significancia de 0.01 y determine si hay algunas diferencias significativas entre las medias de los tratamientos.

(b) Utilice contrastes de un solo grado de libertad con $\alpha = 0.01$ para comparar el tiempo medio de quemado del tratamiento 1 contra el tratamiento 2 y también el tratamiento 3 contra el 4.

4. Utilice la prueba de rango múltiple de Duncan, con un nivel de significancia de 0.05, para analizar las medias de las cinco marcas diferentes de tabletas contra el dolor de cabeza del ejercicio 4 en la página 473.

5. Para los datos que se utilizan en el ejercicio 1 de la página 485, realice una prueba de rango múltiple de Duncan con un nivel de significancia de 0.01, para determinar cuáles laboratorios difieren, en promedio, en sus análisis.

6. Se realiza una investigación para determinar la fuente de reducción en el rendimiento de cierto producto químico. Se sabe que la pérdida en el rendimiento ocurre en el licor madre, es decir, el material que se elimina en la etapa de filtración. Se considera que diferentes mezclas del material original tendrán como resultado diferentes reducciones de la producción en la etapa del licor madre. Los siguientes son resultados del porcentaje de reducción para tres lotes en cada una de las cuatro mezclas preseleccionadas:

Mezcla			
1	2	3	4
25.6	25.2	20.8	31.6
24.3	28.6	26.7	29.8
27.9	24.7	22.2	34.3

(a) Realice el análisis de varianza en el nivel de significancia $\alpha = 0.05$.

(b) Use la prueba de rango múltiple de Duncan para determinar cuáles mezclas difieren.

7. En el estudio *An Evaluation of the Removal Method for Estimating Benthic Populations and Diversity* realizado por el Instituto Politécnico y Universidad Estatal de Virginia en el río Jackson, se utilizan cinco procedimientos de muestreo diferentes para determinar el conteo de especies. Se seleccionan 20 muestras al azar y cada uno de los cinco procedimientos de muestreo se repite cuatro veces. Se registran los conteos de especies a continuación:

Procedimiento de muestreo				
Agotamiento	Hess-modificado	Surber	Eliminación de sustrato de Kicknet	Kicknet
85	75	31	43	17
55	45	20	21	10
40	35	9	15	8
77	67	37	27	15

(a) ¿Hay una diferencia significativa en el conteo promedio de especies para los diferentes procedimientos de muestreo? Utilice un nivel de significancia de 0.01.

(b) Use la prueba de rango múltiple de Duncan con $\alpha = 0.05$ para encontrar cuáles procedimientos de muestreo difieren.

8. La estructura financiera de una empresa se refiere a la forma en que se dividen los activos de la empresa en capital y deuda, y la ventaja financiera se refiere al porcentaje de activos financiados por deuda. En el artículo *The Effect of Financial Leverage on Return*, Tai Ma del Instituto Politécnico y Universidad Estatal de Virginia afirma que la ventaja financiera se puede utilizar para aumentar la tasa de rendimiento sobre el capital. Dicho de otra manera, los accionistas reciben rendimientos más altos sobre el capital con el mismo monto de inversión mediante el uso de la ventaja financiera. Los siguientes datos muestran las tasas de rendimiento sobre el capital con el uso de tres diferentes niveles de ventaja financiera y un nivel de control cero (deuda cero) para 24 empresas seleccionadas al azar:

Ventaja financiera			
Control	Baja	Media	Alta
2.1	6.2	9.6	10.3
5.6	4.0	8.0	6.9

Ventaja financiera (continuación)

Control	Baja	Media	Alta
3.0	8.4	5.5	7.8
7.8	2.8	12.6	5.8
5.2	4.2	7.0	7.2
2.6	5.0	7.8	12.0

Fuente: Machinery Industry Survey de S & P, 1973.

- (a) Realice el análisis de varianza en el nivel de significancia de 0.05.
 (b) Use la prueba de Duncan en el nivel de significancia de 0.01 para determinar si las tasas medias de rendimiento sobre el capital en los niveles bajo, medio y alto de ventaja financiera son más altos que en el nivel de control.

9. En el siguiente experimento biológico se usan cuatro concentraciones de cierto químico para reforzar el crecimiento en centímetros de cierto tipo de planta con el tiempo. Se utilizan cinco plantas en cada concentración y se mide el crecimiento de cada planta. Se toman los siguientes datos de crecimiento. También se aplica un control (ningún químico).

Control	Concentración			
	1	2	3	4
6.8	8.2	7.7	6.9	5.9
7.3	8.7	8.4	5.8	6.1
6.3	9.4	8.6	7.2	6.9
6.9	9.2	8.1	6.8	5.7
7.1	8.6	8.0	7.4	6.1

Utilice la prueba bilateral de Duncan en el nivel de significancia de 0.05 para comparar de manera simultánea las concentraciones con el control.

10. La siguiente tabla (A. Hald, *Statistical Theory with Engineering Applications*, John Wiley & Sons, Nueva York, 1952) da las resistencias de ruptura de muestras de 340 de alambres tomados de nueve cables que serán utilizados para una red de alto voltaje. Cada cable se forma con 12 alambres. Queremos conocer si las resistencias medias de los alambres en los

Cable	1	2	3	4	5	6	7	8	9			
1	5	13	-5	-2	-10	-6	-5	0	-3	-2	-7	-5
2	11	13	-8	8	-3	-12	-12	10	5	-6	-12	-10
3	0	-10	-15	-12	-2	-8	-5	0	-4	-1	-5	-11
4	-2	4	2	10	-5	-8	-12	0	-5	-3	-3	0
5	7	1	5	0	10	6	5	2	0	-1	-10	-2
6	1	0	-5	-4	-1	0	2	5	1	-2	6	7
7	-1	0	2	1	-4	-2	7	5	1	0	-4	2
8	-1	0	7	5	10	-8	1	2	-3	6	0	5
9	-2	6	7	8	15	11	-7	7	10	7	8	1

nueve cables son las mismas. Si los cables son diferentes, ¿cuáles difieren?

11. Se sospecha que la temperatura del medio en la que se activan las baterías afecta su vida activa. Se prueban treinta baterías homogéneas, seis a cada una de cinco temperaturas, y se obtienen los datos que se muestran abajo (vida activada en segundos). Analice e interprete los datos. (C. R. Hicks, *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, Nueva York, 1973.)

	Temperatura (°C)				
	0	25	50	75	100
55	60	70	72	65	
55	61	72	72	66	
57	60	73	72	60	
54	60	68	70	64	
54	60	77	68	65	
56	60	77	69	65	

12. Los siguientes datos son valores de presión (psi) en un resorte de torsión para varios ajustes del ángulo entre las vueltas del resorte en una posición libre:

	Ángulo de las vueltas del resorte (°)				
	67	71	75	79	83
83	84	86	89	90	
85	85	87	90	92	
	85	87	90		
	86	87	91		
	86	88			
	87	88			
		88			
		88			
		88			
		89			
		90			

Calcule un análisis de varianza unilateral para este experimento y establezca su conclusión con respecto al efecto del ángulo sobre la presión en el resorte. (C. R. Hicks, *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart and Winston, Nueva York, 1973).

13. La información del impreso en la siguiente página da la prueba de Duncan con el uso del PROC ANOVA en el SAS para los datos de la mezcla en el ejemplo E3.1. Proporcione conclusiones con respecto a las comparaciones pareadas.

14. Realice una prueba de Duncan para comparaciones pareadas para los datos del ejercicio 8 de la página 474. Discuta los resultados.

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: Y

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate

Alpha= 0.05 df= 25 MSE= 4960.815

Number of Means	2	3	4	5
Critical Range	89.79	87.97	90.69	92.61

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	AGG
A	610.67	6	5
A			
A	610.50	6	3
A			
A	569.33	6	2
A			
A	553.33	6	1
B	485.17	6	4

13.8 Comparación de un conjunto de tratamientos en bloques

En la sección 13.2 presentamos la idea de bloqueo, es decir, aislar conjuntos de unidades experimentales que son razonablemente homogéneos y asignar de forma aleatoria tratamientos a estas unidades. Ésta es una extensión del concepto de "preamiento" que se presenta en los capítulos 9 y 10 y se hace para reducir el error experimental, pues las unidades en bloques tienen características que son más comunes que las unidades de bloques diferentes.

El lector no debe ver a los bloques como un segundo factor, aunque ésta es una forma atractiva de visualizar el diseño. De hecho, el factor principal (tratamientos) aún lleva la carga principal del experimento. Las unidades experimentales son todavía la fuente de error, justo como en el diseño completamente aleatorizado. Simplemente tratamos los conjuntos de estas unidades de manera más sistemática cuando se lleva a cabo el bloqueo. De esta forma, decimos que hay restricciones en la aleatorización. Por ejemplo, para un experimento químico que se diseña para determinar si hay una diferencia en el rendimiento medio de la reacción entre cuatro catalizadores, las muestras de materiales a ser probados se extraen de los mismos lotes de materia prima, mientras que otras condiciones, como la temperatura y concentración de reactantes, se mantienen constantes. En este caso la hora del día para las corridas experimentales podría representar las unidades experimentales, y si el experimentador considera que posiblemente podría haber un ligero efecto de la hora, *aleatorizaría* la asignación de los catalizadores a las corridas para contrarrestar la posible tendencia. Este tipo de estrategia experimental es el **diseño completamente aleatorizado**. Como segundo ejemplo de tal diseño, considere un experimento para comparar cuatro métodos de medición de una propiedad física particular de una sustancia fluida. Suponga que el proceso de muestreo es destructivo; es decir, una vez que se mide una muestra

de la sustancia mediante un método, no se puede medir otra vez con cualquiera de los otros métodos. Si se decide que se realizarán cinco mediciones para cada método, entonces se seleccionan 20 muestras del material de un lote grande al azar y se utilizan en el experimento para comparar los cuatro dispositivos de medición. Las unidades experimentales son las muestras seleccionadas al azar. Cualquier variación de una muestra a otra aparecerá en la variación del error, como se mide con s^2 en el análisis.

Si la variación debida a la heterogeneidad en las unidades experimentales es tan grande que la sensibilidad para detectar diferencias de tratamiento se reduce debido a un valor aumentado de s^2 , un plan mejor sería "bloquear" la variación debida a estas unidades y reducir de esta forma la variación ajena a la que presentan bloques más pequeños o más homogéneos. Por ejemplo, suponga que en la ilustración anterior del catalizador se sabe *a priori* que definitivamente hay un efecto significativo día a día sobre el rendimiento y que podemos medir el rendimiento para cuatro catalizadores en un día dado. En lugar de asignar los cuatro catalizadores a las 20 corridas de prueba completamente al azar, elegimos, digamos, cinco días y corremos cada uno de los cuatro catalizadores cada día, asignando de forma aleatoria el catalizador a las corridas dentro de los días. De esta manera se elimina del análisis la variación diaria y en consecuencia el error experimental, que aún incluye alguna tendencia de la hora *dentro de los días*, representa de forma más precisa la variación debida al azar. Cada día se denomina **bloque**.

El ejemplo clásico, que utiliza un diseño de bloques aleatorizado, es un experimento agrícola en el que se comparan diferentes fertilizantes según su capacidad de aumentar el rendimiento de una cosecha particular. En lugar de asignar los fertilizantes al azar a muchas parcelas en un área grande de composición de suelo variable, se deben asignar los fertilizantes a bloques más pequeños compuestos de parcelas homogéneas. La variación entre estos bloques, que es probablemente más significativa comparada con la uniformidad de las parcelas dentro de un bloque, se elimina entonces del error experimental en el análisis de varianza.

El diseño más directo de bloques aleatorizados es uno en el que asignamos al azar cada tratamiento una vez a cada bloque. Tal disposición experimental se llama **diseño de bloques completos aleatorizados**, cada bloque constituye una sola réplica de los tratamientos.

13.9 Diseño de bloques completos aleatorizados

Una disposición típica para el diseño de bloques completos aleatorizados con el uso de tres mediciones en cuatro bloques es el siguiente:

Bloque 1	Bloque 2	Bloque 3	Bloque 4
t_2	t_1	t_3	t_2
t_1	t_2	t_2	t_1
t_3	t_3	t_1	t_3

Las t denotan la asignación a los bloques de cada uno de los tres tratamientos. Por supuesto, la asignación real de los tratamientos a unidades dentro de los bloques se realiza al azar. Una vez que el experimento termina, los datos se pueden registrar como en el siguiente arreglo de 3×4 :

Tratamiento	Bloque: 1	2	3	4
1	y_{11}	y_{12}	y_{13}	y_{14}
2	y_{21}	y_{22}	y_{23}	y_{24}
3	y_{31}	y_{32}	y_{33}	y_{34}

donde y_{11} representa la respuesta que se obtiene al utilizar el tratamiento 1 en el bloque 1, y_{12} representa la respuesta que se obtiene al usar el tratamiento 1 en el bloque 2, ..., y y_{34} representa la respuesta que se obtiene al utilizar el tratamiento 3 en el bloque 4.

Generalicemos y consideremos el caso de k tratamientos que se asignan a b bloques. Los datos se pueden resumir como se muestra en el arreglo rectangular de $k \times b$ de la tabla 13.10. Se supondrá que las y_{ij} , $i = 1, 2, \dots, k$ y $j = 1, 2, \dots, b$, son va-

Tabla 13.10 Arreglo de $k \times b$ para un diseño de bloques completos aleatorizados

Tratamiento	Bloque: 1	2	...	j	...	b	Total	Media
1	y_{11}	y_{12}	...	y_{1j}	...	y_{1b}	$T_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2b}	$T_{2.}$	$\bar{y}_{2.}$
...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ib}	$T_{i.}$	$\bar{y}_{i.}$
...
k	y_{k1}	y_{k2}	...	y_{kj}	...	y_{kb}	$T_{k.}$	$\bar{y}_{k.}$
Total	$T_{.1}$	$T_{.2}$...	$T_{.j}$...	$T_{.b}$	T	\bar{y}
Media	$\bar{y}_{.1}$	$\bar{y}_{.2}$...	$\bar{y}_{.j}$...	$\bar{y}_{.b}$		\bar{y}

lores de variables aleatorias independientes que tienen distribuciones normales con medias μ_{ij} y varianza común σ^2 . En la tabla 13.10 definimos

$T_{i.}$ = suma de las observaciones para el i -ésimo tratamiento,

$T_{.j}$ = suma de las observaciones para el j -ésimo bloque,

T = suma de todas las bk observaciones,

$\bar{y}_{i.}$ = media de las observaciones para el i -ésimo tratamiento,

$\bar{y}_{.j}$ = media de las observaciones para el j -ésimo bloque,

\bar{y} = media de todas las bk observaciones.

Representemos con $\mu_{i.}$ el promedio (en lugar del total) de las b medias poblacionales para el i -ésimo tratamiento. Es decir,

$$\mu_{i.} = \frac{\sum_{j=1}^b \mu_{ij}}{b}$$

De manera similar, el promedio de las medias poblacionales para el j -ésimo bloque, $\mu_{.j}$, se define como

$$\mu_{.j} = \frac{\sum_{i=1}^k \mu_{ij}}{k}$$

y el promedio de las bk medias poblacionales, μ , se define como

$$\mu = \frac{\sum_{i=1}^k \sum_{j=1}^b \mu_{ij}}{bk}$$

Para determinar si parte de la variación en nuestras observaciones se debe a diferencias entre los tratamientos, consideramos la prueba

$$H_0^1: \mu_1 = \mu_2 = \dots = \mu_k = \mu$$

H_1^1 : No todas las μ_i son iguales.

Cada observación se puede escribir en la forma

$$y_{ij} = \mu_{ij} + e_{ij}$$

donde e_{ij} mide la desviación del valor observado y_{ij} de la media poblacional μ_{ij} . La forma de esta ecuación que se prefiere se obtiene al sustituir

$$\mu_{ij} = \mu + \alpha_i + \beta_j$$

donde α_i es, como antes, el efecto del i -ésimo tratamiento y β_j es el efecto del j -ésimo bloque. Se supone que los efectos del tratamiento y del bloque se pueden sumar. De aquí, podemos escribir

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

Nótese que el modelo parece el de la clasificación unilateral, la diferencia esencial es la introducción del efecto del bloque β_j . El concepto básico es muy semejante al de la clasificación unilateral excepto que debemos tomar en cuenta en el análisis el efecto adicional debido a los bloques, pues ahora controlamos de manera sistemática la variación en *dos direcciones*. Si imponemos ahora las restricciones

$$\sum_{i=1}^k \alpha_i = 0 \quad \text{y} \quad \sum_{j=1}^b \beta_j = 0,$$

entonces

$$\mu_i = \frac{\sum_{j=1}^b (\mu + \alpha_i + \beta_j)}{b} = \mu + \alpha_i \quad \text{y} \quad \mu_j = \frac{\sum_{i=1}^k (\mu + \alpha_i + \beta_j)}{k} = \mu + \beta_j$$

La hipótesis nula de que las k medias de tratamiento μ_i son iguales, y por tanto iguales a μ , es equivalente ahora a probar la hipótesis

$$H_0^2: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

H_1^2 : Al menos una de las α_i no es igual a cero.

Cada una de las pruebas de tratamientos se basará en una comparación de estimaciones independientes de la varianza poblacional común σ^2 . Estas estimaciones se obtendrán al dividir la suma total de cuadrados de nuestros datos en tres componentes por medio de la siguiente identidad.

Teorema 13.3

Identidad de la suma de cuadrados

$$\sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y})^2 = b \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 + k \sum_{j=1}^b (\bar{y}_j - \bar{y})^2 + \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2.$$

Se deja la demostración al lector.

La identidad de la suma de cuadrados se puede representar de manera simbólica con la ecuación

$$SST = SSA + SSB + SSE,$$

donde

$$SST = \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y})^2 = \text{suma total de los cuadrados,}$$

$$SSA = b \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 = \text{suma de cuadrados de tratamientos,}$$

$$SSB = k \sum_{j=1}^b (\bar{y}_j - \bar{y})^2 = \text{suma de cuadrados de bloques,}$$

$$SSE = \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 = \text{suma de cuadrados del error.}$$

Al seguir el procedimiento que se señala en el teorema 13.2, donde interpretamos la suma de cuadrados como funciones de las variables aleatorias independientes, $Y_{11}, Y_{12}, \dots, Y_{kb}$, podemos mostrar que los valores esperados de las sumas de cuadrados de tratamiento, bloque y error están dadas por

$$E(SSA) = (k-1)\sigma^2 + b \sum_{i=1}^k \alpha_i^2,$$

$$E(SSB) = (b-1)\sigma^2 + k \sum_{j=1}^b \beta_j^2,$$

$$E(SSE) = (b-1)(k-1)\sigma^2.$$

Como en el caso del problema de un factor, tenemos el cuadrado medio del tratamiento

$$s_1^2 = \frac{SSA}{k-1}.$$

Si los efectos de tratamiento $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$, s_1^2 es un estimador insesgado de σ^2 . Sin embargo, si los efectos del tratamiento no son todos cero, tenemos

$$E\left(\frac{SSA}{k-1}\right) = \sigma^2 + \frac{b \sum_{i=1}^k \alpha_i^2}{k-1}.$$

y s_1^2 sobrestima σ^2 . Una segunda estimación de σ^2 , basada en $b - 1$ grados de libertad, es

$$s_2^2 = \frac{SSB}{b-1}$$

La estimación s_2^2 es una estimación insesgada de σ^2 si los efectos de bloque $\beta_1 = \beta_2 = \dots = \beta_b = 0$. Si los efectos de bloque no son todos cero, entonces

$$E\left(\frac{SSB}{b-1}\right) = \sigma^2 + \frac{k \sum \beta_i^2}{b-1}$$

y s_2^2 sobrestimarán σ^2 . Una tercera estimación de σ^2 , basada en $(k-1)(b-1)$ grados de libertad e independiente de s_1^2 y s_2^2 , es

$$s^2 = \frac{SSE}{(k-1)(b-1)}$$

que es insesgada sin importar la verdad o falsedad de cualquiera de las hipótesis nulas.

Para probar la hipótesis nula de que los efectos de tratamiento son todos iguales a cero, calculamos la razón

$$f_1 = \frac{s_1^2}{s^2}$$

que es un valor de la variable aleatoria F_1 que tiene una distribución F con $k-1$ y $(k-1)(b-1)$ grados de libertad cuando la hipótesis nula es verdadera. La hipótesis nula se rechaza en el nivel de significancia α cuando

$$f_1 > f_{\alpha}[k-1, (k-1)(b-1)]$$

En la práctica, primero calculamos SST , SSA y SSB , y después, con el uso de la identidad de suma de cuadrados, obtenemos SSE por sustracción. Los grados de libertad asociados con SSE por lo general también se obtienen por sustracción; es decir,

$$(k-1)(b-1) = bk - 1 - (k-1) - (b-1)$$

Los cálculos en un problema de análisis de varianza para un diseño de bloques completos aleatorizados se puede resumir como se muestra en la tabla 13.11.

Tabla 13.11 Análisis de varianza para el diseño de bloques completos aleatorizados

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Tratamientos	SSA	$k-1$	$s_1^2 = \frac{SSA}{k-1}$	$f_1 = \frac{s_1^2}{s^2}$
Bloques	SSB	$b-1$	$s_2^2 = \frac{SSB}{b-1}$	
Error	SSE	$(k-1)(b-1)$	$s^2 = \frac{SSE}{(b-1)(k-1)}$	
Total	SST	$bk-1$		

Ejemplo 13.6 Se consideran cuatro diferentes máquinas, M_1 , M_2 , M_3 , y M_4 , para el ensamblaje de un producto particular. Se decide que se utilizarán seis operadores diferentes en un experimento de bloques aleatorizado para comparar las máquinas. Las máquinas se asignan en orden aleatorio a cada operador. La operación de las máquinas requiere destreza física y se anticipa que habrá una diferencia entre los operadores en la rapidez con la que operan las máquinas (tabla 13.12). Se registra la cantidad de tiempo (en segundos) para ensamblar el producto:

Tabla 13.12 Tiempo, en segundos, para ensamblar el producto

Máquina	Operador:	1	2	3	4	5	6	Total
1		42.5	39.3	39.6	39.9	42.9	43.6	247.8
2		39.8	40.1	40.5	42.3	42.5	43.1	248.3
3		40.2	40.5	41.3	43.4	44.9	45.1	255.4
4		41.3	42.2	43.5	44.2	45.9	42.3	259.4
Total		163.8	162.1	164.9	169.8	176.2	174.1	1010.9

Pruebe la hipótesis H_0 , en el nivel de significancia de 0.05, de que las máquinas funcionan a la misma tasa media de velocidad.

SOLUCIÓN

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0 \text{ (los efectos de máquina son cero),}$$

$$H_1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$$

Las fórmulas de sumas de cuadrados que se muestran en la página 492 y los grados de libertad se utilizan para producir el análisis de varianza de la tabla 13.13. El valor $f = 3.34$ es significativo en $P = 0.048$. Si se usa $\alpha = 0.05$ al menos como un criterio aproximado, se concluye que las máquinas no funcionan a la misma tasa media de velocidad. ■

Tabla 13.13 Análisis de varianza para los datos de la tabla 13.12

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Máquinas	15.93	3	5.31	3.34
Operadores	42.09	5	8.42	
Error	23.84	15	1.59	
Total	81.86	23		

Comentarios adicionales con respecto al bloqueo

En el capítulo 10 presentamos un procedimiento para comparar medias cuando las observaciones son *pareadas*. El procedimiento implica "sustraer" el efecto debido al par homogéneo y trabajar así con diferencias. Éste es un caso especial de un diseño de bloques completos aleatorizados con $k = 2$ tratamientos. Las n unidades homogéneas a las que se asignan los tratamientos toman el papel de bloques.

Si existe heterogeneidad en las unidades experimentales, el experimentador no se debe equivocar al creer que siempre es ventajoso reducir el error experimental a través del uso de pequeños bloques homogéneos. En realidad, puede haber casos donde no sería deseable realizar el bloqueo. El propósito al reducir la varianza del error es aumentar la *sensibilidad* de la prueba para detectar diferencias en las medias de los tratamientos. Esto se refleja en la potencia del procedimiento de prueba. (La potencia del procedimiento de prueba del análisis de varianza se presenta con mayor extensión en la sección 13.12). La potencia para detectar ciertas diferencias entre las medias de los tratamientos aumenta con un descenso en la varianza del error. Sin embargo, la potencia también resulta afectada por los grados de libertad con los que se estima esta varianza, y el bloqueo reduce los grados de libertad que están disponibles de $k(b-1)$ para la clasificación unilateral a $(k-1)(b-1)$. Por lo que se puede perder potencia por el bloqueo si no hay una reducción significativa en la varianza del error.

Interacción entre bloques y tratamientos

Otra suposición importante implícita al escribir el modelo para un diseño de bloques completos aleatorizados es que los efectos de tratamiento y de bloque se suponen aditivos. Esto es equivalente a establecer que

$$\mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'} \quad \text{o} \quad \mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'}$$

para todo valor de i, i', j, j' . Es decir, la diferencia entre las medias poblacionales para los bloques j y j' es la misma para cada tratamiento y la diferencia entre las medias poblacionales para los tratamientos i e i' es la misma para cada bloque. Las líneas paralelas de la figura 13.1(a) ilustran un conjunto de respuestas medias para las que los efectos de tratamiento y de bloque son aditivos, mientras que las líneas que se cruzan en la figura 13.1(b) muestran una situación en la que se dice que los efectos de tratamiento y de bloque **interactúan**. Con referencia al ejemplo 13.6, si el operador 3 es 0.5 segundos más rápido en promedio que el operador 2 cuando se utiliza la máquina 1, entonces el operador 3 aún será 0.5 segundos más rápido en promedio que el operador 2 cuando se utiliza la máquina 2, 3 o 4. En muchos experimentos la suposición de aditividad no se conserva y el análisis de la sección 13.8 conduce a conclusiones erróneas. Suponga, por ejemplo, que el operador

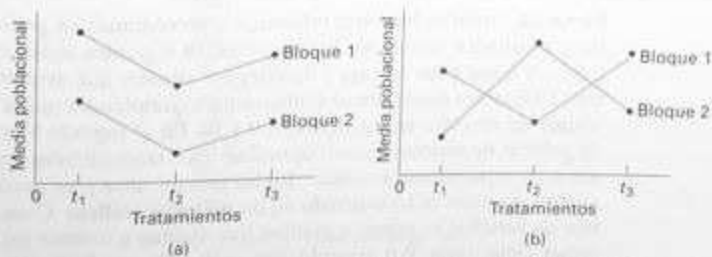


Figura 13.1 Medias poblacionales para (a) efectos aditivos y (b) efectos de interacción.

3 es 0.5 segundos más rápido en promedio que el operador 2 cuando se usa la máquina 1 pero es 0.2 segundos más lento en promedio que el operador 2 cuando se usa la máquina 2. Ahora los operadores y las máquinas interactúan.

Una inspección de la tabla 13.12 sugiere la presencia de una posible interacción. Esta interacción aparente puede ser real o se puede deber al error experimental. El análisis del ejemplo 13.6 se basa en la suposición de que la interacción aparente se debe por completo al error experimental. Si la variabilidad total de nuestros datos se debe en parte a un efecto de interacción, esta fuente de variación permanece como parte de la suma de cuadrados de los errores, lo que ocasiona que el cuadrado medio del error sobrestime σ^2 , y con ello aumente la probabilidad de cometer un error tipo II. Suponemos, en efecto, un modelo incorrecto. Si denotamos con $(\alpha\beta)_{ij}$ el efecto de interacción del i -ésimo tratamiento y el j -ésimo bloque, podemos escribir un modelo más apropiado en la forma

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij},$$

en el que imponemos las restricciones adicionales

$$\sum_{i=1}^k (\alpha\beta)_{ij} = \sum_{j=1}^b (\alpha\beta)_{ij} = 0.$$

Se puede verificar ahora con facilidad que

$$E \left[\frac{SSE}{(b-1)(k-1)} \right] = \sigma^2 + \frac{\sum_{i=1}^k \sum_{j=1}^b (\alpha\beta)_{ij}^2}{(b-1)(k-1)}.$$

De esta forma se observa que el cuadrado medio del error es una estimación sesgada de σ^2 cuando se ignora la interacción existente. Parecería necesario en este momento llegar a un procedimiento para la detección de interacción para casos donde se sospecha que ésta existe. Tal procedimiento requiere de disponibilidad de una estimación insesgada e independiente de σ^2 . Desafortunadamente, el diseño de bloques aleatorizados no se presta a tal prueba a menos que se altere el plan experimental. Este tema se discute en forma más extensa en el capítulo 14.

13.10 Métodos gráficos y diagnósticos posteriores

En varios capítulos hacemos referencia a procedimientos gráficos que muestran datos y resultados analíticos. En los primeros capítulos utilizamos los diagramas de tronco y hojas y los de caja y despliegues visuales que ayudan a resumir las muestras. Utilizamos diagnósticos similares para comprender mejor los datos en dos problemas de muestra en los capítulos 9 y 10. En el capítulo 9 introducimos la noción de gráficas de residuos (residuos ordinarios o studentizados) para detectar violaciones a las suposiciones estándar. En los últimos años gran parte de la atención en el análisis de datos se ha centrado en los **métodos gráficos**. Como la regresión, el análisis de varianza se presta a gráficas que ayudan a resumir los datos así como a detectar violaciones. Por ejemplo, una gráfica simple de las observaciones absolutas alrededor de cada media de tratamiento puede dar al analista la percepción de va-

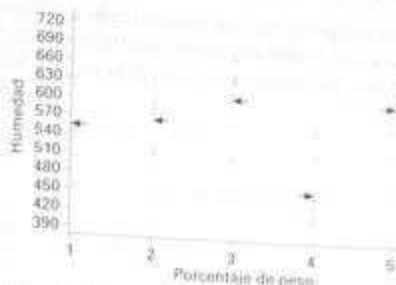


Figura 13.2 Gráfica de los datos alrededor de la media para los datos de las mezclas de la tabla 13.1.

riabilidad entre las medias muestrales y dentro de las muestras. La figura 13.2 representa tal gráfica para los datos de las mezclas de la tabla 13.1. Por la apariencia de la gráfica se puede obtener incluso una idea gráfica acerca de cuáles mezclas (si hay alguna) destacan de las otras. Es claro que la mezcla 4 resalta de las demás. Las mezclas 3 y 5 en realidad forman un grupo homogéneo, como lo hacen las mezclas 1 y 2.

Como en el caso de la regresión, los residuos pueden ser útiles en el análisis de varianza para proporcionar un diagnóstico que pueda detectar violaciones de las suposiciones. Para formar los residuos, sólo necesitamos considerar el modelo del problema de un factor, a saber

$$y_{ij} = \mu_j + \epsilon_{ij}$$

Es sencillo determinar que la estimación de μ_j es \bar{y}_j . De aquí $\bar{y}_j = \bar{y}_j$ y por ello el j -ésimo residuo es $y_{ij} - \bar{y}_j$. Esto se extiende con facilidad al modelo de bloques completos aleatorizados. Puede ser instructivo tener graficados los residuos para cada mezcla a fin de obtener alguna idea con respecto a la suposición de varianza homogénea. Esta gráfica se muestra en la figura 13.3.

Las tendencias en gráficas como éstas pueden revelar dificultades en algunas situaciones, en particular cuando la violación de una suposición particular es gráfica. En el caso de la figura 13.3 los residuos parecen indicar que las varianzas *dentro del tratamiento* son razonablemente homogéneas excepto la mezcla 1. Existe alguna evidencia gráfica de que la varianza para la mezcla 1 es mayor que el resto.

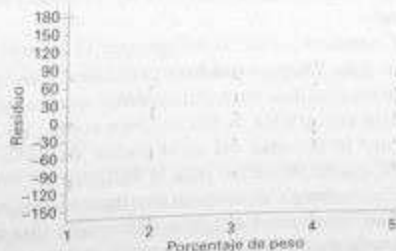


Figura 13.3 Gráficas de residuos para cinco mezclas, con el uso de los datos de la tabla 13.1.

El bloque completo aleatorizado es otro diseño en el que las gráficas pueden hacer que el analista se sienta confiado con un "cuadro ideal", o quizá resalte las dificultades. Recuerde que el modelo para el bloque completo aleatorizado es

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad \begin{array}{l} i = 1, 2, \dots, k, \\ j = 1, 2, \dots, b, \end{array}$$

con las restricciones impuestas

$$\sum_{i=1}^k \alpha_i = 0, \quad \sum_{j=1}^b \beta_j = 0.$$

Para determinar lo que en realidad constituye un residuo, considere que

$$\alpha_i = \mu_i - \mu, \quad \beta_j = \mu_j - \mu$$

y que μ se estima con \bar{y} , μ_i se estima con \bar{y}_i , y μ_j se estima mediante \bar{y}_j . Como resultado, el valor ajustado o pronosticado \hat{y}_{ij} está dado por

$$\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j = \bar{y}_i + \bar{y}_j - \bar{y}$$

y de esta manera el residuo en la observación (i, j) está dado por

$$y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}$$

Nótese que \hat{y}_{ij} , el valor ajustado, es una estimación de la media μ_{ij} . Esto es consistente con la división de la variabilidad que se da en el teorema 13.3, donde la suma de cuadrados del error es

$$SSE = \sum_i \sum_j (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2.$$

La presentación visual en el bloque completo aleatorizado implica graficar los residuos por separado para cada tratamiento y para cada bloque. El analista debe esperar variabilidad aproximadamente igual si es válida la suposición de varianzas homogéneas. El lector debe recordar que en el capítulo 12 presentamos gráficas donde los residuos se grafican con el propósito de detectar las especificaciones erróneas en el modelo. En el caso del bloque completo aleatorizado, las especificaciones erróneas serias del modelo se pueden relacionar con nuestra suposición de aditividad (es decir, ninguna interacción). Si no hay interacción, debe aparecer un patrón aleatorio.

Considere los datos del ejemplo 13.6, en el que los tratamientos son cuatro máquinas y los bloques son los operadores. Las figuras 13.4 y 13.5 muestran las gráficas de los residuos para tratamientos separados y bloques separados. La figura 13.6 muestra una gráfica de los residuos contra los valores ajustados. La figura 13.4 revela que la varianzas del error puede no ser la misma para todas las máquinas. Lo mismo puede ser cierto para la varianzas del error en cada uno de los seis operadores. Sin embargo, dos residuos extraordinariamente grandes parecen producir la dificultad aparente. La figura 13.6 revela una gráfica de los residuos que muestra evidencia razonable de un comportamiento aleatorio. Sin embargo, sobresalen los residuos grandes antes señalados.



Figura 13.4 Gráfica de los residuos de las cuatro máquinas para los datos del ejemplo 13.8. Los asteriscos indican residuos múltiples.

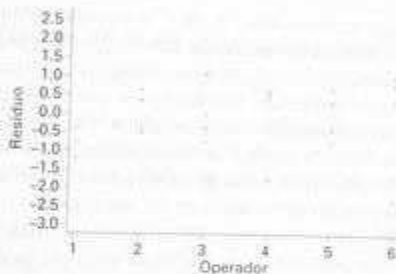


Figura 13.5 Gráficas de los residuos de los seis operadores para los datos del ejemplo 13.6. El asterisco indica residuos múltiples.

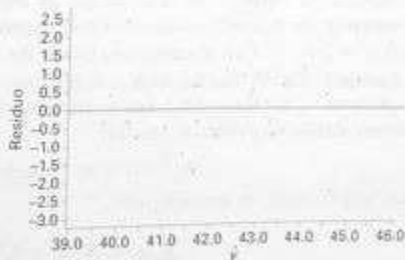


Figura 13.6 Residuos graficados contra los valores ajustados para los datos del ejemplo 13.6.

13.11 Cuadrados latinos

El diseño en bloques aleatorizados es muy efectivo para reducir el error experimental al eliminar una fuente de variación. Otro diseño que es útil en particular para controlar dos fuentes de variación, mientras reduce el número de combinaciones de tratamientos que se requieren, se llama **cuadrados latinos**. Suponga que nos interesan

los rendimientos de cuatro variedades de trigo con el uso de cuatro diferentes fertilizantes por un periodo de cuatro años. El número total de combinaciones de tratamientos para un diseño completamente aleatorizado sería 64. Al seleccionar el mismo número de categorías para los tres criterios de clasificación, podemos seleccionar un diseño de cuadrados latinos y realizar el análisis de varianza con el uso de los resultados de sólo 16 combinaciones de tratamientos. Un cuadrado latino típico, que se selecciona al azar de todos los posibles cuadrados 4×4 , es el siguiente:

Fila	Columna:	1	2	3	4
1		A	B	C	D
2		D	A	B	C
3		C	D	A	B
4		B	C	D	A

Las cuatro letras, A, B, C y D, representan las cuatro variedades de trigo que se denominan **tratamientos**. Los renglones y columnas, representadas por los cuatro fertilizantes y los cuatro años, respectivamente, son las dos fuentes de variación que deseamos controlar. Vemos ahora que cada tratamiento ocurre con exactitud una vez en cada renglón y cada columna. Con dicho arreglo balanceado el análisis de varianza permite separar la variación debida a los diferentes fertilizantes y a los diferentes años de la suma de cuadrados del error y con ello obtener una prueba más precisa de las diferencias en las capacidades de rendimiento de las cuatro variedades de trigo. Cuando existe interacción entre cualesquiera de las fuentes de variación, los valores f en el análisis de varianza ya no son válidos. En este caso, el diseño de cuadrados latinos no sería apropiado.

Ahora generalizamos y consideramos un cuadrado latino $r \times r$ donde y_{ijk} denota una observación en el i -ésimo renglón y la j -ésima columna que corresponden a la k -ésima letra. Nótese que una vez que se especifican i y j para un cuadrado latino particular, de manera automática conocemos la letra dada para k . Por ejemplo, cuando $i = 2$ y $j = 3$ en el cuadrado latino de 4×4 anterior, tenemos $k = B$. Por ello k es una función de i y j . Si α_i y β_j son los efectos del i -ésimo renglón y la j -ésima columna, τ_k el efecto del k -ésimo tratamiento, μ la media general y ϵ_{ijk} el error aleatorio, entonces podemos escribir

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \epsilon_{ijk},$$

donde imponemos las restricciones

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \tau_k = 0.$$

Como antes, se supone que las y_{ijk} son valores de variables aleatorias independientes que tienen distribuciones normales con medias

$$\mu_{ijk} = \mu + \alpha_i + \beta_j + \tau_k$$

y varianza común σ^2 .

La hipótesis a probar es la siguiente:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_r = 0,$$

$$H_1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$$

Esta prueba se basará en una comparación de estimaciones independientes de σ^2 proporcionada al dividir la suma total de cuadrados de nuestros datos en cuatro componentes por medio de la siguiente identidad. Se pide al lector proporcionar la demostración en el ejercicio 13 de la página 505.

Teorema 13.4

Identidad de la suma de cuadrados

$$\begin{aligned} \sum_i \sum_j \sum_k (y_{ijk} - \bar{y})^2 &= r \sum_i (\bar{y}_i - \bar{y})^2 + r \sum_j (\bar{y}_j - \bar{y})^2 \\ &+ r \sum_k (\bar{y}_k - \bar{y})^2 + \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_i - \bar{y}_j - \bar{y}_k + 2\bar{y})^2. \end{aligned}$$

De forma simbólica, escribimos la identidad de la suma de cuadrados como

$$SST = SSR + SSC + SSTr + SSE,$$

donde *SSR* y *SSC* se llaman suma de cuadrados del **renglón** y suma de cuadrados de la **columna**, respectivamente; *SSTr* se llama la suma de cuadrados del **tratamiento** y *SSE* es la suma de cuadrados del **error**. Los grados de libertad se dividen de acuerdo con la identidad

$$r^2 - 1 = (r - 1) + (r - 1) + (r - 1) + (r - 1)(r - 2).$$

Al dividir cada una de las sumas de cuadrados del lado derecho de la identidad de la suma de cuadrados entre su número de grados de libertad correspondiente, obtenemos las cuatro estimaciones independientes

$$s_1^2 = \frac{SSR}{r-1}, \quad s_2^2 = \frac{SSC}{r-1}, \quad s_3^2 = \frac{SSTr}{r-1}, \quad s^2 = \frac{SSE}{(r-1)(r-2)}$$

de σ^2 . Si se interpretan las sumas de cuadrados como funciones de variables aleatorias independientes, no es difícil verificar que

$$E(S_1^2) = E\left[\frac{SSR}{r-1}\right] = \sigma^2 + \frac{r \sum \alpha_i^2}{r-1}, \quad E(S_2^2) = E\left[\frac{SSC}{r-1}\right] = \sigma^2 + \frac{r \sum \beta_j^2}{r-1},$$

$$E(S_3^2) = E\left[\frac{SSTr}{r-1}\right] = \sigma^2 + \frac{r \sum \tau_i^2}{r-1}, \quad E(S^2) = E\left[\frac{SSE}{(r-1)(r-2)}\right] = \sigma^2.$$

El análisis de varianza (tabla 13.14, página 502) indica la prueba *F* apropiada para los tratamientos.

Ejemplo 13.7 Para ilustrar el análisis de un diseño de cuadrados latinos, regresemos al experimento donde las letras A, B, C y D representan cuatro variedades de trigo; los renglones representan cuatro diferentes fertilizantes; y las columnas 4 años diferentes. Los datos de la tabla 13.15 son los rendimientos para las cuatro variedades.

Tabla 13.14 Análisis de varianza para un cuadrado latino $r \times r$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Reglón	SSR	$r - 1$	$\sigma_1^2 = \frac{SSR}{r - 1}$	$f = \frac{\sigma_1^2}{\sigma^2}$
Columnas	SSC	$r - 1$	$\sigma_2^2 = \frac{SSC}{r - 1}$	
Tratamientos	$SSTr$	$r - 1$	$\sigma_3^2 = \frac{SSTr}{r - 1}$	
Error	SSE	$(r - 1)(r - 2)$	$\sigma^2 = \frac{SSE}{(r - 1)(r - 2)}$	
Total	SST	$r^2 - 1$		

Tabla 13.15 Rendimientos del trigo (kilogramos por parcela)

Tratamiento de fertilizante	1981	1982	1983	1984
t_1	A 70	B 75	C 68	D 81
t_2	D 66	A 59	B 55	C 63
t_3	C 59	D 66	A 79	B 42
t_4	B 41	C 57	D 39	A 55

des de trigo, medidas en kilogramos por parcela. Se supone que las diversas fuentes de variación no interactúan. Utilice un nivel de significancia de 0.05 para probar la hipótesis H_0 : no hay diferencia en los rendimientos promedio de las cuatro variedades de trigo.

SOLUCIÓN

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0,$$

H_1 : Al menos una de las α_i no es igual a cero.

Se utiliza la presentación de sumas de cuadrados y grados de libertad de la tabla 13.14. Las fórmulas de las sumas de cuadrados aparecen en la página 501. Aquí, por supuesto, la tabla del análisis de varianza (tabla 13.16) debe reflejar la variabilidad debida al fertilizante, años y tipos de tratamiento. La $f = 2.02$ es sobre 3 y 6 grados de libertad. El valor P de aproximadamente 0.2 es en realidad demasiado grande para concluir que las variedades de trigo afectan de manera significativa el rendimiento. ■

Tabla 13.16 Análisis de varianza para los datos de la tabla 13.15

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Fertilizante	1557			
Año	418	3	519,000	
Tratamientos	264	3	139,333	
Error	261	3	88,000	2,02
Total	2500	6	43,500	
		15		

Ejercicios

1. Muestre que la fórmula de cálculo para SSB , en el análisis de varianza del diseño de bloques completos aleatorizados, es equivalente al término correspondiente en la identidad del teorema 13.3.

2. Para el diseño de bloques aleatorizados con k tratamientos y b bloques, muestre que

$$E(SSB) = (b-1)\sigma^2 + k \sum_{j=1}^k \beta_j^2$$

3. Se utilizan cuatro clases de fertilizante f_1, f_2, f_3 , y f_4 para estudiar el rendimiento de frijoles. El suelo se divide en tres bloques, cada uno de los cuales contiene cuatro parcelas homogéneas. Los rendimientos en kilogramos por parcela y los tratamientos correspondientes son los siguientes:

Bloque 1

$$\begin{matrix} f_1 = 42.7 \\ f_2 = 48.5 \\ f_3 = 32.8 \\ f_4 = 39.3 \end{matrix}$$

Bloque 2

$$\begin{matrix} f_1 = 50.9 \\ f_2 = 50.0 \\ f_3 = 38.0 \\ f_4 = 40.2 \end{matrix}$$

Bloque 3

$$\begin{matrix} f_1 = 51.1 \\ f_2 = 46.3 \\ f_3 = 51.9 \\ f_4 = 53.5 \end{matrix}$$

(a) Realice un análisis de varianza en el nivel de significancia de 0.05 con el uso del modelo de bloques completos aleatorizados.

(b) Utilice contrastes con un solo grado de libertad y un nivel de significancia de 0.01 para comparar los fertilizantes (f_1, f_3) contra (f_2, f_4) y f_3 contra f_1 .

4. Se compara el rendimiento de tres variedades de papas. El experimento se lleva a cabo mediante la asignación de cada variedad de forma aleatoria a tres parcelas de igual tamaño en cada uno de cuatro lugares distintos. Se registran los siguientes rendimientos para las variedades A, B y C en 100 kilogramos por parcela:

Lugar 1	Lugar 2	Lugar 3	Lugar 4
B 13	C 21	C 9	A 11
A 18	A 20	B 12	C 10
C 12	B 23	A 14	B 17

Realice un análisis de varianza de bloques aleatorizados para probar la hipótesis de que no hay diferencia en las capacidades de rendimiento de las tres variedades de papas. Utilice un nivel de significancia de 0.05.

5. Los siguientes datos son los porcentajes de aditivos externos medidos por cinco analistas para tres marcas similares de jalea de fresa, A, B, y C.

Analista 1	Analista 2	Analista 3	Analista 4	Analista 5
B 2.7	C 7.5	B 2.8	A 1.7	C 8.1
C 3.6	A 1.6	A 2.7	B 1.9	A 2.0
A 3.8	B 5.2	C 6.4	C 2.6	B 4.8

Realice un análisis de varianza y pruebe la hipótesis, en el nivel de significancia de 0.05, de que el porcentaje de aditivos externos es el mismo para las tres marcas de jalea.

6. Los siguientes datos representan las calificaciones finales de matemáticas, inglés, francés y biología, que obtienen cinco estudiantes:

Estudiante	Materia			
	Matemáticas	Inglés	Francés	Biología
1	68	57	73	61
2	83	94	91	86
3	72	81	63	59
4	55	73	77	66
5	92	68	75	87

Utilice un nivel de significancia de 0.05 para probar la hipótesis de que los cursos son de igual dificultad.

7. En el estudio *The Periphyton of the South River, Virginia: Mercury Concentration, Productivity, and Autotrophic Index Studies*, realizado por el Departamento de Ciencias Ambientales del Instituto Politécnico y Universidad Estatal de Virginia, se midió la concentración total de mercurio en sólidos totales de periphyton en seis estaciones diferentes en seis días diferentes. Se registraron los siguientes datos:

Fecha	Estación					
	CA	CB	E1	E2	E3	E4
8 de abril	0.45	3.24	1.33	2.04	3.93	5.93
23 de junio	0.10	0.10	0.99	4.31	9.92	6.49
1 ^o de julio	0.25	0.25	1.65	3.13	7.39	4.43
8 de julio	0.09	0.06	0.92	3.66	7.88	6.24
15 de julio	0.15	0.16	2.17	3.50	8.82	5.39
23 de julio	0.17	0.39	4.30	2.91	5.50	4.29

Utilice un nivel de significancia de 0.01 para probar la hipótesis de que la concentración media de mercurio es la misma para las seis estaciones.

8. Una planta de energía nuclear produce una gran cantidad de calor que por lo general se descarga en sistemas acuáticos. Este calor eleva la temperatura del sistema acuático, lo que tiene como resultado una mayor concentración de clorofila *a*, que a su vez amplía la estación de crecimiento. Para estudiar este efecto, cada mes se recogen muestras de agua en tres estaciones durante un periodo de 12 meses. La estación A se encuentra cerca de una descarga potencial de agua caliente, la estación C se ubica lo más lejos posible de la descarga y la estación B se halla a la mitad del camino entre las estaciones A y C. Se registran las siguientes concentraciones de clorofila *a*. Realice un análisis de varianza y pruebe la hipótesis, en el nivel de significancia de 0.05, de que no hay diferencia en las concentraciones medias de clorofila *a* en las tres estaciones.

Mes	Estación		
	A	B	C
Enero	9.867	3.723	4.470
Febrero	14.035	8.416	11.100
Marzo	10.700	20.723	4.470
Abril	13.853	9.168	8.010
Mayo	7.067	4.778	34.080
Junio	11.670	9.145	8.990
Julio	7.357	8.463	3.350
Agosto	3.558	4.086	4.500
Septiembre	4.210	4.233	6.830
Octubre	3.630	2.320	5.800
Noviembre	2.953	3.843	3.490
Diciembre	2.640	3.610	3.020

9. En un estudio que se lleva a cabo en el Departamento de Salud y Educación Física del Instituto Politécnico y Universidad Estatal de Virginia se asignan tres dietas por un periodo de tres días a cada uno de seis sujetos en un diseño de bloques aleatorizado. A los sujetos, que juegan el papel de bloques, se les asignan las siguientes tres dietas en orden aleatorio:

Dieta 1: mezcla de grasa y carbohidratos.

Dieta 2: alta en grasa.

Dieta 3: alta en carbohidratos.

Al final del periodo de tres días cada sujeto se coloca en un aparato para caminata y se mide el tiempo de agotamiento, en segundos. Se registran los siguientes datos:

Dieta	Sujeto					
	1	2	3	4	5	6
1	84	35	91	57	56	45
2	91	48	71	45	61	61
3	122	53	110	71	91	122

Realice el análisis de varianza, mediante la separación de la suma de cuadrados de dietas, sujetos y errores. Utilice un nivel de significancia de 0.01 para determinar si hay diferencias significativas entre las dietas.

10. El personal de Silvicultura utiliza arsénicos orgánicos como pesticidas. La cantidad de arsénico que se acumula en el cuerpo cuando se expone a estos pesticidas es un grave problema de salud. Es importante que la cantidad de exposición se determine con rapidez de modo que un trabajador de campo con un alto nivel de arsénico se pueda separar del trabajo. En un experimento reportado en el artículo, *A Rapid Method for the Determination of Arsenic Concentrations in Urine at Field Locations*, publicado en *Si-*

Amer. Ind. Hyg. Assoc. J. (vol. 37, 1976); muestras de arena de cuatro miembros del Servicio Forestal se dividieron en tres muestras iguales de modo que cada individuo se pudiera analizar en busca de arsénico por un laboratorio universitario, un químico con un equipo portátil y por un empleado forestal después de una breve orientación. Se registraron los siguientes niveles de arsénico, en partes por millón:

Individuo	Analista		
	Empleado	Químico	Laboratorio
1	0.05	0.05	0.04
2	0.05	0.05	0.04
3	0.04	0.04	0.03
4	0.15	0.17	0.10

Realice un análisis de varianza y pruebe la hipótesis, en el nivel de significancia de 0.05, de que no hay diferencia en los niveles de arsénico de los tres métodos de análisis.

11. Científicos del Departamento de Patología de Plantas del Tecnológico de Virginia diseñaron un experimento en el que se aplicaron cinco tratamientos diferentes a seis lugares diferentes de un huerto de manzanas para determinar si había diferencias significativas en el crecimiento entre los tratamientos. Los tratamientos 1 a 4 representan diferentes herbicidas y el tratamiento 5 representa un control. El periodo de crecimiento fue de mayo a noviembre de 1982, y se registró el nuevo tamaño, medido en centímetros, para las muestras seleccionadas de los seis lugares en el huerto:

Tratamiento	Lugares					
	1	2	3	4	5	6
1	455	72	61	215	695	501
2	622	82	444	170	437	134
3	695	56	50	443	701	373
4	607	650	493	257	490	262
5	388	263	185	103	518	622

Realice un análisis de varianza, mediante la separación de las sumas de cuadrados de tratamientos, lugares y errores. Utilice un nivel de significancia de 0.01 para probar la hipótesis de que no hay diferencias entre las medias de los tratamientos.

12. En el artículo *Self-Control and Therapist Control in the Behavioral Treatment of Overweight Women*, publicado en *Behavioral Research and Therapy* (vol. 10, 1972), se estudian los efectos sobre el cambio de

peso de mujeres obesas de dos tratamientos de reducción y un tratamiento de control. Los dos tratamientos de reducción incluidos fueron, respectivamente, un programa de reducción de peso autoinducido y un programa de reducción controlado por el terapeuta. Cada uno de los 10 sujetos se asignó a los tres programas de tratamiento en orden aleatorio y se midió su pérdida de peso. Se registraron los siguientes cambios de peso:

Sujeto	Tratamiento		
	Control	Autoinducido	Terapeuta
1	1.00	-2.25	-10.50
2	3.75	-6.00	-13.50
3	0.00	-2.00	0.75
4	-0.25	-1.50	-4.50
5	-2.25	-3.25	-6.00
6	-1.00	-1.80	4.00
7	-1.00	-10.75	-12.25
8	3.75	-0.75	-2.75
9	1.50	0.00	-6.75
10	-0.50	-3.75	-7.00

Realice un análisis de varianza y pruebe la hipótesis, en el nivel de significancia de 0.01, de que no hay diferencia en las pérdidas medias de peso para los tres tratamientos.

13. Verifique la identidad de la suma de cuadrados del teorema 13.4 en la página 501.

14. Para el diseño de cuadrados latinos de $r \times r$ muestre que:

$$E(SSTr) = (r-1)\sigma^2 + r \sum_{j=1}^r \tau_j^2$$

15. El departamento de matemáticas de una universidad desea evaluar las capacidades de enseñanza de cuatro profesores. A fin de eliminar cualquier efecto debido a los diferentes cursos de matemáticas y los diferentes horarios, se decide realizar un experimento con el uso de un diseño de cuadrados latinos en que las letras A, B, C y D representan a los cuatro diferentes profesores. Cada profesor enseña una sección de cada uno de cuatro diferentes cursos programados en cada uno de los cuatro diferentes horarios durante el día. Los datos en la siguiente página muestran las calificaciones asignadas por estos profesores a 16 estudiantes de aproximadamente igual capacidad. Utilice un nivel de significancia de 0.05 para probar la hipótesis de que los diferentes profesores no tienen efecto en las calificaciones.

Horario	Curso				
	Algebra	Geometría	Estadística	Cálculo	
1	A 84	B 79	C 63	D 97	
2	B 91	C 82	D 80	A 93	
3	C 59	D 70	A 77	B 80	
4	D 75	A 91	B 75	C 68	

16. Una empresa fabricante quiere investigar los efectos de cinco aditivos de color en el tiempo de fraguado de una mezcla de concreto nueva. Las variaciones en el tiempo de fraguado se pueden esperar de los cambios diarios en la temperatura y humedad y también de los diferentes trabajadores que preparan los

Trabajador	Día				
	1	2	3	4	5
1	D 10.7	E 10.3	B 11.2	A 10.9	C 10.5
2	E 11.3	C 10.5	D 12.0	B 11.5	A 10.3
3	A 11.8	B 10.9	C 10.5	D 11.3	E 7.5
4	B 14.1	A 11.6	E 11.0	C 11.7	D 11.5
5	C 14.5	D 11.5	A 11.5	E 12.7	B 10.9

moldes de prueba. Para eliminar estas fuentes externas de variación se utiliza un diseño de cuadrado latino de 5×5 en el que las letras A, B, C, D y E representan los cinco aditivos. Los tiempos de fraguado, en horas, para los 25 moldes se muestran en la anterior tabla.

En el nivel de significancia de 0.05, ¿podemos decir que los aditivos de color tienen algún efecto en el tiempo de fraguado de la mezcla de concreto?

17. En el libro *Design of Experiments for the Quality Improvement*, publicado por la Asociación de Estándares Japonesa (1989), se realiza un estudio sobre la cantidad de tinte necesaria para obtener el mejor color para cierto tipo de tela. Las tres cantidades de tinte: $\frac{1}{2}\%$ owl* ($\frac{1}{2}\%$ del peso de la tela), 1% owl y 3% owl se administran cada uno a dos diferentes plantas. La densidad de color de la tela se observa entonces cuatro veces para cada nivel de tinte en cada planta.

	Cantidad de tinte					
	1		2		3	
Planta 1	5.2	6.0	12.3	10.5	22.4	17.8
	5.9	5.9	12.4	10.9	22.5	18.4
Planta 2	6.5	5.5	14.5	11.8	29.0	23.2
	6.4	5.9	16.0	13.6	28.7	24.0

Realice un análisis de varianza para probar la hipótesis en el nivel de significancia de 0.05, de que no hay diferencia en la densidad del color de una tela para los tres niveles de tinte. Considere las plantas como bloques.

13.12 Modelos de efectos aleatorios

A lo largo de este capítulo tratamos procedimientos del análisis de varianza en los que el objetivo principal es estudiar el efecto sobre alguna respuesta de ciertos tratamientos fijos o predeterminados. Los experimentos en los que los tratamientos o niveles de tratamiento son preseleccionados por el experimentador en oposición a ser elegidos de manera aleatoria se llaman **experimentos de efectos fijos** o **experimentos modelo I**. Para el modelo de efectos fijos, las inferencias se hacen sólo en los tratamientos particulares que se usan en el experimento.

Con frecuencia es importante que el experimentador sea capaz de extraer inferencias acerca de una población de tratamientos por medio de un experimento en el que los tratamientos que se utilizan se elijan de forma aleatoria de la población.

* owl, of the weight of a fabric (del peso de la tela)

Por ejemplo, un biólogo se puede interesar en si hay o no una varianza significativa en alguna característica fisiológica debida al tipo de animal. Los tipos de animales que se utilizan en realidad en el experimento se eligen entonces aleatoriamente y representan los efectos de tratamiento. Un químico se puede interesar en estudiar el efecto de laboratorios analíticos en el análisis químico de una sustancia. No se preocupa de laboratorios particulares sino más bien de una población grande de laboratorios. Puede entonces seleccionar un grupo de laboratorios al azar y asignar muestras a cada uno para análisis. La inferencia estadística entonces implicaría 1) probar si los laboratorios contribuyen o no a una varianza diferente de cero para los resultados analíticos y 2) la estimación de la varianza debida a los laboratorios y la varianza dentro de los laboratorios.

El modelo de efectos aleatorios unilateral, con frecuencia denominado **modelo II**, se escribe como el modelo de efectos fijos pero los términos tienen diferentes significados. La respuesta

$$y_{ij} = \mu + \alpha_i + e_{ij}$$

es ahora un valor de la variable aleatoria

$$Y_{ij} = \mu + A_i + E_{ij}$$

con $i = 1, 2, \dots, k$ y $j = 1, 2, \dots, n$, donde las A_i se distribuyen de forma normal e independiente con media cero y varianza σ_a^2 y son independientes de las E_{ij} . Como para el modelo de efectos fijos, las E_{ij} también se distribuyen de forma normal e independiente con media cero y varianza σ^2 . Nótese que para un experimento modelo II, la variable aleatoria $\sum_{i=1}^k A_i$ toma el valor $\sum_{i=1}^k \alpha_i$ y la restricción de que estas α_i sumen cero ya no se aplica.

Teorema 13.5

Para el modelo de análisis de varianza unilateral de efectos aleatorios

$$E(SSA) = (k-1)\sigma^2 + n(k-1)\sigma_a^2 \quad \text{y} \quad E(SSE) = k(n-1)\sigma^2.$$

PRUEBA Del modelo

$$Y_{ij} = \mu + A_i + E_{ij}$$

obtenemos

$$\bar{Y}_{i.} = \mu + A_i + \bar{E}_{i.}, \quad \bar{Y}_{.j} = \mu + \bar{A} + \bar{E}_{.j}.$$

De aquí

$$SSA = n \sum_{i=1}^k (\bar{Y}_{i.} - \bar{Y}_{.j})^2 = n \sum_{i=1}^k [(A_i - \bar{A}) + (\bar{E}_{i.} - \bar{E}_{.j})]^2$$

y

$$\begin{aligned} E(SSA) &= n \sum_{i=1}^k E(A_i^2) - nkE(\bar{A}^2) + n \sum_{i=1}^k E(\bar{E}_{i.}^2) - nkE(\bar{E}_{.j}^2) \\ &= nk\sigma_a^2 - n\sigma_a^2 + k\sigma^2 - \sigma^2 = (k-1)\sigma^2 + n(k-1)\sigma_a^2. \end{aligned}$$

Al seguir los mismos pasos que arriba, también encontramos que

$$SSE = \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2 = \sum_{i=1}^k \sum_{j=1}^n (E_{ij} - \bar{E}_i)^2$$

y por tanto

$$E(SSE) = \sum_{i=1}^k \sum_{j=1}^n E(E_{ij}^2) - n \sum_{i=1}^k E(\bar{E}_i^2) = nk\sigma^2 - k\sigma^2 = k(n-1)\sigma^2.$$

La tabla 13.17 muestra los cuadrados medios esperados para un experimento modelo I y uno modelo II. Los cálculos para un experimento modelo II se llevan a cabo exactamente de la misma manera que para el experimento modelo I. Es decir, las columnas de sumas de cuadrados, grados de libertad y cuadrados medios en una tabla de análisis de varianza son las mismas para ambos modelos.

Tabla 13.17 Cuadrados medios esperados para el experimento de un factor

Fuente de variación	Grados de libertad	Cuadrados medios	Cuadrados medios esperados	
			Modelo I	Modelo II
Tratamientos	$k-1$	s_1^2	$\sigma^2 + \frac{n \sum_{i=1}^k \alpha_i^2}{k-1}$	$\sigma^2 + n\sigma_a^2$
Error	$k(n-1)$	s^2	σ^2	σ^2
Total	$nk-1$			

Para el modelo de efectos aleatorios, la hipótesis de que los efectos de los tratamientos son todos cero se escribe

$$H_0: \sigma_a^2 = 0,$$

$$H_1: \sigma_a^2 \neq 0,$$

la cual dice que los diferentes tratamientos no contribuyen en nada a la variabilidad de la respuesta. Es obvio de la tabla 13.17 que s_1^2 y s^2 son estimaciones de σ^2 cuando H_0 es verdadera y que la razón

$$f = \frac{s_1^2}{s^2}$$

es un valor de la variable aleatoria F que tiene la distribución F con $k-1$ y $k(n-1)$ grados de libertad. La hipótesis nula se rechaza en el nivel de significancia α cuando

$$f > f_\alpha [k-1, k(n-1)].$$

En muchos estudios científicos y de ingeniería el interés no se centra en la prueba F . El científico sabe que el efecto aleatorio, en realidad, tiene un efecto significativo. Lo que es más importante es la estimación de los diversos componentes de la varianza. Esto produce un sentido de *clasificación* en términos de cuáles factores producen la mayor variabilidad y en qué cantidad. En el contexto presente puede

ser de interés cuantificar cuánto más grande es el *componente de varianza de un solo factor* que el que se produce al azar (variación aleatoria).

La tabla 13.17 también se puede utilizar para estimar los *componentes de la varianza* σ^2 y σ_a^2 . Como y_1^2 estima $\sigma^2 + n\sigma_a^2$ y s^2 estima a σ^2 ,

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_a^2 = \frac{y_1^2 - s^2}{n}$$

Ejemplo 13.8 Los siguientes datos son observaciones codificadas del rendimiento de un proceso químico, en el que se utilizan cinco lotes de materia prima seleccionados al azar:

Lote:	1	2	3	4	5	
	9.7	10.4	15.9	8.6	9.7	
	5.6	9.6	14.4	11.1	12.8	
	8.4	7.3	8.3	10.7	8.7	
	7.9	6.8	12.8	7.6	13.4	
	8.2	8.8	7.9	6.4	8.3	
	7.7	9.2	11.6	5.9	11.7	
	8.1	7.6	9.8	8.1	10.7	
Total	55.6	59.7	80.7	58.4	75.3	329.7

Muestre que el componente de varianza del lote es significativamente mayor que cero y obtenga su estimación.

SOLUCIÓN

Las sumas de cuadrados total, del lote y de los errores son

$$SST = 194.64, \quad SSA = 72.60, \quad SSE = 194.64 - 72.60 = 122.04,$$

Éstos resultados, con los cálculos restantes, se muestran en la tabla 13.18. La razón f es significativa en el nivel $\alpha = 0.05$, lo que indica que la hipótesis de un componente cero del lote se rechaza. Una estimación del componente de la varianza del lote es

$$\hat{\sigma}_a^2 = \frac{18.15 - 4.07}{7} = 2.01.$$

Tabla 13.18 Análisis de varianza para el ejemplo 13.8

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Lotes	72.60	4	18.15	4.46
Error	122.04	30	4.07	
Total	194.64	34		

Diseño de bloques aleatorizado con bloques aleatorios

En un experimento de bloques completo aleatorizado donde los bloques representan días, es imaginable que el experimentador quiera que los resultados se apliquen no sólo a los días reales que se utilizan en el análisis sino a cada día del año. Enton-

ces seleccionaría al azar los días en los cuales realizar el experimento así como los tratamientos y utilizaría el modelo de efectos aleatorios

$$Y_{ij} = \mu + A_i + B_j + E_{ij}, \quad i = 1, 2, \dots, k \quad \text{y} \quad j = 1, 2, \dots, b,$$

siendo A_i , B_j , y E_{ij} variables aleatorias independientes con medias cero y varianzas σ_a^2 , σ_b^2 , y σ^2 , respectivamente. Los cuadrados medios esperados para un diseño de bloques completos aleatorizados modelo II se obtienen mediante el uso del mismo procedimiento que el del problema de un solo factor, y se presentan junto con los del experimento modelo I en la tabla 13.19.

Tabla 13.19 Cuadrados medios esperados para el diseño de bloques completos aleatorizados

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrados medios esperados	
			Modelo I	Modelo II
Tratamientos	$k - 1$	s_1^2	$\sigma^2 + \frac{b \sum a_i^2}{k-1}$	$\sigma^2 + b\sigma_a^2$
Bloques	$b - 1$	s_2^2	$\sigma^2 + \frac{k \sum \beta_j^2}{b-1}$	$\sigma^2 + k\sigma_b^2$
Error	$(k-1)(b-1)$	s^2	σ^2	σ^2
Total	$bk - 1$			

De nuevo los cálculos para la suma de cuadrados individual y grados de libertad son idénticos a los del modelo de efectos fijos. La hipótesis

$$H_0: \sigma_a^2 = 0,$$

$$H_1: \sigma_a^2 \neq 0$$

se verifica al calcular

$$f = \frac{s_1^2}{s^2}$$

y se rechaza H_0 cuando $f > f_{\alpha}[k-1, (b-1)(k-1)]$.

Las estimaciones de los componentes de la varianza son

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_a^2 = \frac{s_1^2 - s^2}{b}, \quad \hat{\sigma}_b^2 = \frac{s_2^2 - s^2}{k}$$

Para el diseño de cuadrados latinos, el modelo de efectos aleatorios se escribe

$$Y_{ijk} = \mu + A_i + B_j + T_k + E_{ijk}$$

$i = 1, 2, \dots, r$; $j = 1, 2, \dots, r$; y $k = A, B, C, \dots$, siendo A_i , B_j , T_k , y E_{ijk} variables aleatorias independientes con medias cero y varianzas σ_a^2 , σ_b^2 , σ_t^2 , y σ^2 , respectivamente. La derivación de los cuadrados medios esperados para un diseño de cuadrados latinos modelo II es directa y, por comparación, los presentamos junto con los del experimento modelo I en la tabla 13.20.

Tabla 13.20 Cuadrados medios esperados para un diseño de cuadrados latinos

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrados medios esperados	
			Modelo I	Modelo II
Renglones	$r - 1$	$\hat{\sigma}_1^2$	$\sigma^2 + \frac{r \sum \alpha_i^2}{r - 1}$	$\sigma^2 + r\sigma_\alpha^2$
Columnas	$r - 1$	$\hat{\sigma}_2^2$	$\sigma^2 + \frac{r \sum \beta_j^2}{r - 1}$	$\sigma^2 + r\sigma_\beta^2$
Tratamientos	$r - 1$	$\hat{\sigma}_3^2$	$\sigma^2 + \frac{r \sum \tau_k^2}{r - 1}$	$\sigma^2 + r\sigma_\tau^2$
Error	$\frac{(r - 1)(r - 2)}{r^2 - 1}$	$\hat{\sigma}^2$	σ^2	σ^2
Total	$r^2 - 1$			

Las pruebas de hipótesis relativas a los diversos componentes de la varianza se realizan mediante el cálculo de las relaciones de los cuadrados medios apropiados como se indica en la tabla 13.20, y al comparar con los valores f correspondientes de la tabla A.6.

13.13 Enfoque de regresión para el análisis de varianza

Hasta aquí, tratamos los modelos de regresión y los modelos del análisis de varianza como dos temas separados y sin relación. Aunque éste es el enfoque aceptado para tratar con estos procedimientos en un nivel elemental, se puede tratar un modelo de análisis de varianza como un caso especial de un modelo de regresión lineal múltiple. En esta sección mostramos la relación entre los dos modelos e indicamos cómo las técnicas del análisis de varianza se pueden desarrollar a través de un enfoque de regresión.

Suponga que consideramos dos modelos, el modelo de regresión lineal múltiple,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i,$$

y el modelo de análisis de varianza de clasificación unilateral,

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} = \mu_j + \varepsilon_{ij}.$$

Tradicionalmente, los dos se presentan como métodos para manejar diferentes problemas prácticos, siendo el modelo de regresión un medio de llegar a un procedimiento para predecir alguna respuesta como función de una o más variables cuantitativas independientes, y el modelo del análisis de varianza para llegar a pruebas de significancia sobre medias poblacionales múltiples. Sin embargo, cualquier modelo matemático que es lineal en los parámetros, como el modelo del análisis de varianza, se puede considerar un caso especial del modelo de regresión lineal múltiple. Podemos utilizar una notación matricial convencional para describir cómo se expresa cada observación como función de los parámetros para los dos modelos. Para el modelo de regresión,

$$y = X\beta + \varepsilon,$$

o de forma más explícita,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

donde el vector y a la izquierda del signo de igualdad es el arreglo de respuestas en el experimento. La matriz X ya se describió y utilizó en la sección 12.3. El vector β es el vector de los parámetros que aparecen en el modelo, y el vector e completa el modelo por la adición del error aleatorio. El lector recordará que las estimaciones de mínimos cuadrados b_0, b_1, \dots, b_k de los parámetros $\beta_0, \beta_1, \dots, \beta_k$ se obtienen al resolver la ecuación

$$Ab = g,$$

donde $A = X'X$ es una matriz no singular y $g = X'y$ es un vector cuyos elementos son sumas de productos de los elementos en las columnas de X y los elementos en el vector y .

De esta manera las estimaciones están dadas por

$$b = (X'X)^{-1}X'y.$$

Considere ahora el modelo de análisis de varianza en forma matricial:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n} \\ \vdots \\ y_{k1} \\ y_{k2} \\ \vdots \\ y_{kn} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ \vdots \\ e_{1n} \\ e_{21} \\ e_{22} \\ \vdots \\ e_{2n} \\ \vdots \\ e_{k1} \\ e_{k2} \\ \vdots \\ e_{kn} \end{bmatrix}$$

De nuevo, cada observación se expresa como función de los parámetros. Aquí la muy importante matriz X , la matriz de las condiciones experimentales, consiste en unos y ceros. Formulaciones similares se pueden escribir para el modelo de bloques completos aleatorizados.

Aplicamos la aproximación de mínimos cuadrados al modelo de análisis de varianza unilateral. Las ecuaciones normales son

$$\begin{bmatrix} nk & n & n & \cdots & n \\ n & n & 0 & \cdots & 0 \\ n & 0 & n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & 0 & 0 & \cdots & n \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \vdots \\ \hat{\alpha}_k \end{bmatrix} = \begin{bmatrix} Y \\ Y_1 \\ Y_2 \\ \vdots \\ Y_k \end{bmatrix}$$

En esta etapa es simple ilustrar por qué se hace una distinción en la presentación de los dos modelos. Las últimas k columnas de la matriz A para el modelo de análisis de varianza se suman a la primera columna, y así la matriz es **singular**, lo que implica que no hay una solución única para las ecuaciones de estimación. Esto inicialmente parece un serio inconveniente en lo que al modelo concierne. De hecho, decimos que los parámetros en el modelo no son **estimables**. El lector recordará que se realizaron pruebas de significancia de las medias poblacionales,

$$\mu_1 = \mu + \alpha_1, \quad \mu_2 = \mu + \alpha_2, \quad \dots, \quad \mu_k = \mu + \alpha_k$$

y, al formular el procedimiento de prueba, se aplicó la restricción lineal

$$\sum_{i=1}^k \alpha_i = 0$$

De esta forma las α_i toman el papel de desviaciones (más o menos) de las medias de tratamiento o poblacionales a partir de la media general μ . Probar la igualdad de las medias poblacionales se vuelve equivalente a probar que las α_i ($i = 1, 2, \dots, k$) son todas cero.

Con la restricción de que las α_i sumen cero, las ecuaciones de estimación se pueden resolver y dan

$$\hat{\mu} = \frac{Y}{nk} = \bar{y}, \quad \hat{\alpha}_i = \frac{Y_i}{n} - \frac{Y}{nk} = \bar{y}_i - \bar{y}, \quad i = 1, 2, \dots, k.$$

Aunque estas estimaciones no son únicas, pues dependen de la restricción que se aplica a las α_i , nos dan una base para utilizar el procedimiento de regresión general que se señala en la sección 12.6 para determinar si la eliminación de las α_i del modelo aumenta de manera significativa la suma de cuadrados del error, y con ello nos proporciona, en el contexto de la regresión, una prueba de hipótesis de efectos de tratamiento no significativos.

Si aproximamos el problema de la prueba de hipótesis para el modelo de análisis de varianza unilateral con los procedimientos de regresión múltiple del capítulo 12, podemos comenzar por calcular la suma de cuadrados de la regresión para los parámetros $\alpha_1, \alpha_2, \dots, \alpha_k$. Estos parámetros toman el mismo papel que los coeficientes $\beta_1, \beta_2, \dots, \beta_k$ en el modelo de regresión lineal múltiple. Calcularíamos entonces la suma de cuadrados de la regresión

$$R(\alpha_1, \alpha_2, \dots, \alpha_k) = SSR$$

$$= b_0 g_0 + b_1 g_1 + \dots + b_k g_k - \frac{\left(\sum_{i=1}^k \sum_{j=1}^n Y_{ij}\right)^2}{nk}$$

$$= b_1 g_1 + \dots + b_k g_k = \hat{\alpha}_1 g_1 + \dots + \hat{\alpha}_k g_k$$

El lado derecho de las ecuaciones de estimación da $g_1 = Y_1, g_2 = Y_2, \dots, g_k = Y_k$. De aquí

$$R(\alpha_1, \alpha_2, \dots, \alpha_k) = \sum_{i=1}^k \left(\frac{Y_i}{n} - \frac{Y}{nk}\right) Y_i = \sum_{i=1}^k \frac{Y_i^2}{n} - \frac{Y^2}{nk} = SSA$$

con $k-1$ grados de libertad en lugar de k . Se pierde un grado de libertad debido a la restricción lineal única que se impone a los efectos de tratamiento. La suma de cuadrados de los errores con

$$(nk - 1) - (k - 1) = k(n - 1)$$

grados de libertad está dada por

$$SSE = SST - R(\alpha_1, \alpha_2, \dots, \alpha_k) = SST - SSA$$

que es idéntica a la expresión que se desarrolló antes en este capítulo.

La hipótesis de que la regresión sobre las α_i es insignificante, es decir, $\alpha_i = 0$ para todas las i se prueba al formar la relación

$$f = \frac{R(\alpha_1, \alpha_2, \dots, \alpha_k)/(k-1)}{SSE/k(n-1)} = \frac{SSA/(k-1)}{s^2}$$

Un valor de $f > f_{\alpha}[(k-1), k(n-1)]$ implica que la regresión aumenta de forma significativa, y en consecuencia la suma de cuadrados del error disminuye significativamente al incluir en el modelo los efectos de tratamiento.

El enfoque de regresión para los modelos tipo análisis de varianza se puede extender al diseño de bloques completos aleatorizados y al de cuadrados latinos que se presentan en las secciones 13.8 y 13.9 y también a los diseños factoriales del capítulo 14.

13.14 Potencia de las pruebas de análisis de varianza

Como se indicó antes, el investigador a menudo se preocupa por el problema de no saber qué tan grande elegir la muestra. Al planear un diseño completamente aleatorizado de un factor con n observaciones por tratamiento, el objetivo principal es probar la hipótesis de igualdad de las medias de tratamientos.

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_k = 0,$$

H_1 : Al menos una de las α_i no es igual a cero.

Muy a menudo, sin embargo, la varianza del error experimental σ^2 es tan grande que el procedimiento de prueba será insensible a las diferencias reales entre las k medias de tratamientos. En la sección 13.3 los valores esperados de los cuadrados medios para el modelo unilateral están dados por

$$E(S_1^2) = E\left(\frac{SSA}{k-1}\right) = \sigma^2 + \frac{n \sum_{i=1}^k \alpha_i^2}{k-1} \quad E(S^2) = E\left(\frac{SSE}{k(n-1)}\right) = \sigma^2.$$

Así, para una desviación dada de la hipótesis nula H_0 , medida por

$$n \sum_{i=1}^k \alpha_i^2 / (k-1),$$

valores grandes de σ^2 disminuyen la posibilidad de obtener un valor $f = s_1^2/s^2$ que esté en la región crítica para la prueba. La sensibilidad de la prueba describe la capacidad del procedimiento para detectar diferencias en las medias poblacionales y se mide por la potencia de la prueba (véase la sección 10.2), la cual es simplemente $1 - \beta$, donde β es la probabilidad de aceptar una hipótesis falsa. Podemos interpretar la potencia de nuestras pruebas de análisis de varianza, entonces, como la probabilidad de que la estadística F esté en la región crítica cuando, de hecho, la hipótesis nula es falsa y las medias de tratamientos difieren. Para la prueba del análisis de varianza unilateral, la potencia, $1 - \beta$, es

$$\begin{aligned} 1 - \beta &= P\left[\frac{S_1^2}{S^2} > f_\alpha(v_1, v_2) \text{ cuando } H_1 \text{ es verdadera}\right] \\ &= P\left[\frac{S_1^2}{S^2} > f_\alpha(v_1, v_2) \text{ cuando } \sum_{i=1}^k \alpha_i^2 \neq 0\right]. \end{aligned}$$

El término $f_\alpha(v_1, v_2)$ es, es, por supuesto, el punto crítico de la cola superior de la distribución F con v_1 y v_2 grados de libertad. Para valores dados de $\sum_{i=1}^k \alpha_i^2 / (k-1)$ y σ^2 , la potencia se puede aumentar mediante el uso de un tamaño muestral más grande n . El problema se convierte en el de diseñar el experimento con un valor de n tal que los requerimientos de potencia se cumplan. Por ejemplo, podemos requerir que para valores específicos de $\sum_{i=1}^k \alpha_i^2 \neq 0$ y σ^2 , la hipótesis se rechace con probabilidad 0.9. Cuando la potencia de la prueba es baja, limita de manera severa el alcance de las inferencias que se pueden extraer de los datos experimentales.

Caso de efectos fijos

En el análisis de varianza la potencia depende de la distribución de la razón F bajo la hipótesis alternativa de que las medias de tratamientos difieren. Por tanto, en el caso del modelo de efectos fijos unilateral, requerimos la distribución de S_1^2/S^2 cuando, de hecho,

$$\sum_{i=1}^k \alpha_i^2 \neq 0.$$

Por supuesto, cuando la hipótesis es verdadera, $\alpha_i = 0$ para $i = 1, 2, \dots, k$, y la estadística sigue la distribución F con $k - 1$ y $N - k$ grados de libertad. Si $\sum_{i=1}^k \alpha_i^2 \neq 0$, la razón sigue una **distribución F no central**.

La variable aleatoria básica de la F no central se denota con F . Sea $f_{\alpha}^*(v_1, v_2, \lambda)$ un valor de F con parámetros v_1 , v_2 , y λ . Los parámetros v_1 y v_2 de la distribución son los grados de libertad asociados con S_1^2 y S^2 , respectivamente, y λ se llama **parámetro de no centralidad**. Cuando $\lambda = 0$, la F no central simplemente se reduce a la distribución F ordinaria con v_1 y v_2 grados de libertad.

Para el análisis de varianza unifactorial de efectos fijos con tamaños muestrales n_1, n_2, \dots, n_k definimos

$$\lambda = \frac{\sum_{i=1}^k n_i \alpha_i^2}{2\sigma^2}$$

Si tenemos tablas de la F no central a nuestra disposición, la potencia para detectar una alternativa particular se obtiene al evaluar la siguiente probabilidad:

$$1 - \beta = P \left[\frac{S_1^2}{S^2} > f_{\alpha}(k-1, N-k) \text{ cuando } \lambda = \frac{\sum_{i=1}^k n_i \alpha_i^2}{2\sigma^2} \right]$$

$$= P[F^* > f_{\alpha}(k-1, N-k)]$$

Aunque la F no central se define normalmente en términos de λ , es más conveniente, con propósitos de tabulación, trabajar con

$$\phi^2 = \frac{2\lambda}{v_1 + 1}$$

La tabla A.15 muestra gráficas de la potencia del análisis de varianza como función de ϕ para varios valores de v_1 , v_2 , y el nivel de significancia α . Estas **gráficas de potencia** se pueden usar no sólo para los modelos de efectos fijos que se presentan en este capítulo, sino también para los modelos multifactor del capítulo 14. Resta ahora dar un procedimiento por el que el parámetro de no centralidad λ , y por ello ϕ , se pueda encontrar para estos casos de efectos fijos.

El parámetro de no centralidad λ se puede escribir en términos de los **valores esperados del cuadrado medio del numerador** de la razón F en el análisis de varianza. Tenemos

$$\lambda = \frac{v_1 [E(S_1^2)]}{2\sigma^2} - \frac{v_1}{2}$$

y así

$$\phi^2 = \frac{[E(S_1^2) - \sigma^2]}{\sigma^2} \frac{v_1}{v_1 + 1}$$

En la tabla 13.21 se muestran expresiones para λ y ϕ^2 para el modelo unilateral, el diseño de bloques completos aleatorizados y el diseño de cuadrados latinos.

Tabla 13.21 Parámetro de no centralidad λ y ϕ^2 para el modelo de efectos fijos

	Clasificación unilateral	Bloque completo aleatorizado	Cuadrado latino
λ :	$\frac{\sum_i n_i \alpha_i^2}{2\sigma^2}$	$\frac{b \sum_i \alpha_i^2}{2\sigma^2}$	$\frac{r \sum_i \tau_i^2}{2\sigma^2}$
ϕ^2 :	$\frac{\sum_i \mu_i \alpha_i^2}{k\sigma^2}$	$\frac{b \sum_i \alpha_i^2}{k\sigma^2}$	$\frac{\sum_i \tau_i^2}{r\sigma^2}$

Nótese de la tabla A.15 que, para valores dados de v_1 y v_2 , la potencia de la prueba aumenta con el aumento de los valores de ϕ . El valor de λ depende, por supuesto, de σ^2 , y en un problema práctico a menudo se puede necesitar sustituir el cuadrado medio del error como una estimación al determinar ϕ^2 .

Ejemplo 13.9 En un experimento de bloques aleatorizados se comparan cuatro tratamientos en seis bloques, lo que tiene como resultado 15 grados de libertad para el error. ¿Son suficientes seis bloques si la potencia de nuestra prueba para detectar diferencias entre las medias de tratamiento, en el nivel de significancia 0.05, debe ser al menos 0.8 cuando las medias reales son $\mu_1 = 5.0$, $\mu_2 = 7.0$, $\mu_3 = 4.0$, y $\mu_4 = 4.0$? Una estimación de σ^2 para utilizarse en el cálculo de la potencia está dada por $\sigma^2 = 2.0$.

SOLUCIÓN

Recuerde que las medias de tratamiento están dadas por $\mu_i = \mu + \alpha_i$. Si recurrimos a la restricción $\sum_{i=1}^4 \alpha_i = 0$, tenemos

$$\mu = \frac{\sum_{i=1}^4 \mu_i}{4} = 5.0,$$

y entonces $\alpha_1 = 0$, $\alpha_2 = 2.0$, $\alpha_3 = -1.0$, y $\alpha_4 = -1.0$. Por tanto,

$$\phi^2 = \frac{b \sum_{i=1}^k \alpha_i^2}{k\sigma^2} = \frac{(6)(6)}{(4)(2)} = 4.5,$$

de la que obtenemos $\phi = 2.121$. Con el uso de la tabla A.15, se encuentra que la potencia es aproximadamente 0.89 y por ello se cumplen los requerimientos de potencia. Esto significa que si el valor de $\sum_{i=1}^4 \alpha_i^2 = 6$ y $\sigma^2 = 2.0$, el uso de seis bloques tendrá como resultado el rechazo de la hipótesis de medias de tratamiento iguales con probabilidad 0.89. ■

Caso de efectos aleatorios

En el caso de efectos fijos, el cálculo de la potencia requiere el uso de la distribución F no central. Este no es el caso en el modelo de efectos aleatorios. De hecho, la potencia se calcula de manera muy simple mediante el uso de las tablas F estándar. Considere, por ejemplo, el modelo unilateral de efectos aleatorios, con n observaciones por tratamiento, con la hipótesis

$$H_0: \sigma_a^2 = 0,$$

$$H_1: \sigma_a^2 \neq 0.$$

Cuando H_1 es verdadera, la razón

$$f = \frac{SSA/[k(k-1)(\sigma^2 + n\sigma_a^2)]}{SSE/[k(n-1)\sigma^2]} = \frac{S_1^2}{S^2(1 + n\sigma_a^2/\sigma^2)}$$

es un valor de la variable aleatoria F que tiene la distribución F con $k-1$ y $k(n-1)$ grados de libertad. El problema se convierte, entonces, en el de determinar la probabilidad de rechazar H_0 bajo la condición de que el componente real de la varianza de tratamiento $\sigma_a^2 \neq 0$. Tenemos entonces

$$\begin{aligned} 1 - \beta &= P\left\{ \frac{S_1^2}{S^2} > f_{\alpha}[(k-1), k(n-1)] \text{ cuando } H_1 \text{ es verdadera} \right\} \\ &= P\left\{ \frac{S_1^2}{S^2(1 + n\sigma_a^2/\sigma^2)} > \frac{f_{\alpha}[(k-1), k(n-1)]}{1 + n\sigma_a^2/\sigma^2} \right\} \\ &= P\left\{ F > \frac{f_{\alpha}[(k-1), k(n-1)]}{1 + n\sigma_a^2/\sigma^2} \right\}. \end{aligned}$$

Nótese que conforme n aumenta, el valor $f_{\alpha}[(k-1), k(n-1)]/(1 + n\sigma_a^2/\sigma^2)$ se aproxima a cero, lo que tiene como resultado un aumento en la potencia de la prueba. Una ilustración de la potencia para esta clase de situación se muestra en

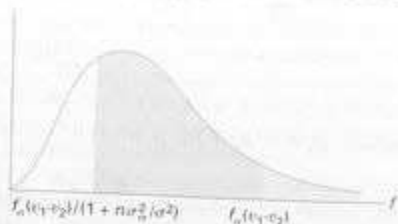


Figura 13.7 Potencia para el análisis de varianza unilateral de efectos aleatorios.

la figura 13.7. El área de sombreado más claro es el nivel de significancia α , mientras que toda el área sombreada es la potencia de la prueba.

Ejemplo 13.10 Suponga que en un problema de un factor es de interés probar la significancia del componente de varianza σ_a^2 . Se utilizarán cuatro tratamientos en el

experimento, con cinco observaciones por tratamiento, ¿Cuál será la probabilidad de rechazar la hipótesis $\sigma_a^2 = 0$, cuando de hecho el componente de la varianza de tratamiento es $(3/4)\sigma^2$?

SOLUCIÓN

Al usar un nivel de significancia $\alpha = 0.05$, tenemos

$$1 - \beta = P\left\{F > \frac{f_{0.05}(3, 16)}{1 + (5)(3)/4}\right\} = P\left\{F > \frac{f_{0.05}(3, 16)}{4.75}\right\} = P\left(F > \frac{3.24}{4.75}\right) \\ = P(F > 0.682).$$

Con el teorema 8.7 y después con la tabla A-7c de *Introduction to Statistics*, 3a. ed., de Dixon y Massey (véase la bibliografía), vemos que

$$1 - \beta = 0.58.$$

Por tanto, sólo cerca de 58% de las veces el procedimiento detectará un componente de varianza de $(3/4)\sigma^2$.

13.15 Estudio de caso

Al personal del Departamento de Química del Tecnológico de Virginia se le pidió analizar un conjunto de datos que se produjo al comparar cuatro métodos diferentes de análisis de aluminio en cierta mezcla sólida encendedor. Para obtener un amplio rango de laboratorios de análisis involucrados, se utilizaron cinco laboratorios en el experimento. Estos laboratorios se seleccionaron debido a que por lo general realizan estos tipos de análisis. Veinte muestras de material encendedor que contienen 2.70% de aluminio se asignaron al azar, cuatro a cada laboratorio, y se dieron instrucciones sobre cómo realizar el análisis químico con el uso de los cuatro métodos. Los datos recogidos son los siguientes:

Método	Laboratorio					Media
	1	2	3	4	5	
A	2.67	2.69	2.62	2.66	2.70	2.668
B	2.71	2.74	2.69	2.70	2.77	2.722
C	2.76	2.76	2.70	2.76	2.81	2.758
D	2.65	2.69	2.60	2.64	2.73	2.662

Los laboratorios no se consideran como efectos aleatorios, pues no se seleccionaron al azar a partir de una población grande de laboratorios. Los datos se analizaron como un diseño de bloques completos aleatorizados. Se buscan gráficas de estos datos para determinar si es apropiado un modelo aditivo tipo:

$$y_{ij} = \mu + m_i + l_j + e_{ij}$$

en otras palabras, un modelo con efectos aditivos. El bloque aleatorizado no es apropiado cuando existe interacción entre laboratorios y métodos. Considere la gráfica que se muestra en la figura 13.8. Aunque esta gráfica es un poco difícil de interpretar debido a que cada punto es una sola observación, no parece haber una interacción apreciable entre métodos y laboratorios.

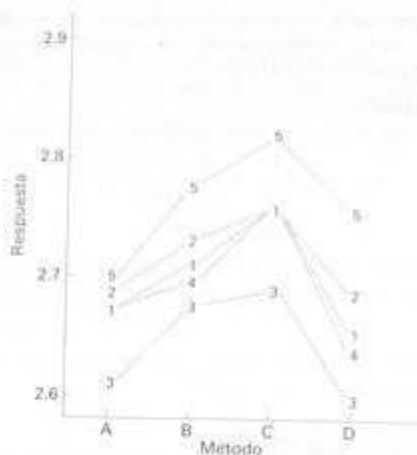


Figura 13.8 Gráfica de interacción para los datos del estudio 13.15.

Gráficas de residuos

Las gráficas de residuos se usan como indicaciones de diagnóstico con respecto a la suposición de varianza homogénea. La figura 13.9 muestra una gráfica de residuos contra métodos analíticos. La variabilidad que se describe en los residuos parece ser extraordinariamente homogénea. Para completar, en la figura 13.10 se muestra una gráfica de probabilidad normal de los residuos.

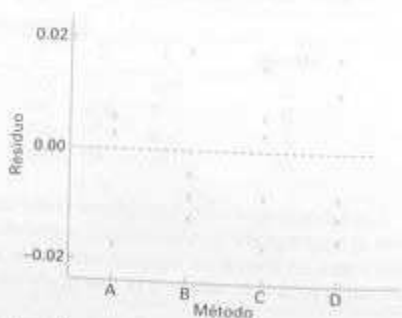


Figura 13.9 Gráfica de residuos contra método para los datos del estudio 13.15. (Hay dos observaciones ocultas.)

Las gráficas de residuos no muestran dificultad con la suposición de errores normales ni con la varianza homogénea. El SAS PROC GLM se utilizó para realizar el análisis de varianza. La figura 13.11 muestra el impreso comentado de computador.

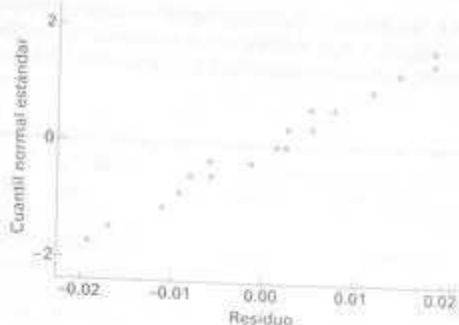


Figura 13.10 Gráfica de probabilidad normal de los residuos para los datos del estudio 13.15. (Hay dos observaciones ocultas.)

```

SAS
General Linear Models Procedure
Class Level Information
Class      Levels      Values
TREATMENT  4              A B C D
BLOCK      5              1 2 3 4 5
Number of observations in data set = 20

Dependent Variable: Y

Source      DF      Sum of Squares      Mean Square      F Value      Pr > F
Model       7      0.0534050           0.0076293        42.38        0.0001
Error       12     0.0021700           0.0001808
Corrected   19     0.0555750
Total

Source      DF      Mean Square      F Value      Pr > F
TREATMENT  3      0.0314550        0.0104850        57.78        0.0001
BLOCK      4      0.0219500        0.0054875
Root MSE   0.0134
F Mean     0.702500

General Linear Models Procedure
Observation  Observed Value  Predicted Value  Residual
1            2.6700000      2.6630000       0.0070000
2            2.7300000      2.7170000       0.0130000
3            2.7400000      2.7520000      -0.0120000
4            2.6500000      2.6570000      -0.0070000
5            2.6900000      2.6455000       0.0445000
6            2.7400000      2.7395000       0.0005000
7            2.7600000      2.7730000      -0.0130000
8            2.6900000      2.6795000       0.0105000
9            2.6700000      2.6800000      -0.0100000
10           2.6900000      2.6720000       0.0180000
11           2.7000000      2.7000000       0.0000000
12           2.6000000      2.6120000      -0.0120000
13           2.6400000      2.6550000      -0.0150000
14           2.7000000      2.7095000      -0.0095000
15           2.7400000      2.7445000      -0.0045000
16           2.6400000      2.6495000      -0.0095000
17           2.7000000      2.7100000      -0.0100000
18           2.7700000      2.7780000      -0.0080000
19           2.8100000      2.8000000       0.0100000
20           2.7300000      2.7120000       0.0180000

```

Figura 13.11 Impreso del SAS para los datos del estudio 13.15.

ra. Los valores f y F calculados indican una diferencia significativa entre los métodos analíticos. A este análisis puede continuar un análisis de comparación múltiple para determinar donde están las diferencias entre los métodos.

Ejercicios

1. Los siguientes datos muestran el efecto de cuatro operadores, elegidos al azar, sobre la producción de una máquina particular:

Operador		1	2	3	4
1	2	175.4	168.5	170.1	175.2
1	2	171.7	162.7	173.4	175.7
1	2	173.0	165.0	175.7	180.1
1	2	170.5	164.1	170.7	183.7

(a) Realice un análisis de varianza modelo II en el nivel de significancia 0.05.

(b) Calcule una estimación del componente de la varianza del operador y el componente de la varianza del error experimental.

2. Suponga un modelo de efectos aleatorios, muestre que

$$E(SSB) = (b-1)\sigma^2 + k(b-1)\sigma_p^2$$

para el diseño de bloques completos aleatorizados.

3. Se lleva a cabo un experimento en el que se comparan cuatro tratamientos en cinco bloques. Se generan los siguientes datos:

Tratamiento	Bloque				
	1	2	3	4	5
1	12.8	10.6	11.7	10.7	11.0
2	11.7	14.2	11.8	9.9	13.8
3	11.5	14.7	13.6	10.7	15.9
4	12.6	16.5	15.4	9.6	17.1

(a) Con la suposición de un modelo de efectos aleatorios, pruebe la hipótesis en el nivel de significancia 0.05 de que no hay diferencia entre las medias de tratamientos.

(b) Calcule estimaciones para los componentes de la varianza de tratamiento y de bloque.

4. Suponga un modelo de efectos aleatorios, muestre que

$$E(SST) = (r-1)(\sigma^2 + r\sigma_a^2)$$

para el diseño de cuadrados latinos.

5. (a) Con el uso de un enfoque de regresión para el diseño de bloques completos aleatorizados, obtenga las ecuaciones normales $\mathbf{Ab} = \mathbf{g}$ en forma matricial.

(b) Muestre que

$$R(\beta_1, \beta_2, \dots, \beta_p | \alpha_1, \alpha_2, \dots, \alpha_k) = SSB$$

6. En el ejercicio 1, si nos interesa probar la significancia del componente de la varianza del operador, ¿tenemos muestras suficientemente grandes para asegurar con una probabilidad tan grande como 0.95 un componente de varianza significativa si la σ_p^2 real es $1.5\sigma^2$? Si no, ¿cuántas corridas se necesitan por operador? Utilice un nivel de significancia de 0.05.

7. Si se supone un modelo de efectos fijos en el ejercicio 3 y se utiliza una prueba de nivel $\alpha = 0.05$, ¿cuántos bloques se necesitan a fin de que aceptemos la hipótesis de igualdad de medias de tratamientos con probabilidad 0.1, cuando de hecho,

$$\frac{\sum_{i=1}^4 \sigma_i^2}{\sigma^2} = 2.0?$$

8. Verifique los valores dados para A y σ^2 en la tabla 13.20 para el diseño de bloques completos aleatorizados.

9. En la prueba de anticuerpos de VIH en muestras sanguíneas de pacientes, un espectrómetro determina la densidad óptica de cada muestra. La densidad típica se mide como la absorbencia de luz en una longitud de onda particular. La muestra sanguínea es positiva si excede cierto valor tope que se determina mediante las muestras de control para esa prueba. Los investigadores se interesan en la comparación de la variabilidad del laboratorio para los valores de control positivos. Los datos representan valores de control positivos para 10 diferentes pruebas en cuatro laboratorios seleccionados al azar.

Corrida	Laboratorio			
	1	2	3	4
1	0.888	1.065	1.325	1.232
2	0.983	1.226	1.069	1.127
3	1.047	1.332	1.219	1.051
4	1.087	0.958	0.958	0.897
5	1.125	0.816	0.819	1.222
6	0.997	1.015	1.140	1.125
7	1.025	1.071	1.222	0.990
8	0.969	0.905	0.995	0.875
9	0.898	1.140	0.928	0.930
10	1.018	1.051	1.322	0.775

- (a) Escriba un modelo apropiado para este experimento.
 (b) Estime el componente de varianza del laboratorio y la varianza dentro de los laboratorios.
10. Cinco "vaciados" de metales tienen cinco muestras del núcleo cada una analizada para la cantidad de un enfriamiento. Los datos para los cinco vaciados seleccionados al azar son los siguientes:

Núcleo	1	2	3	4	5
1	0.98	0.85	1.12	1.21	1.00
2	1.02	0.92	1.68	1.19	1.21
3	1.57	1.16	0.99	1.32	0.93
4	1.25	1.43	1.26	1.08	0.86
5	1.16	0.99	1.05	0.94	1.41

- (a) La intención es que los vaciados sean idénticos. Pruebe el componente de la varianza del "vaciado" es cero. Extraiga conclusiones.
 (b) Muestre un ANOVA completo junto con una estimación de la varianza dentro del vaciado.

11. Una compañía textil teje cierta tela en un número grande de telares. Los administradores desean que los telares fuesen homogéneos de modo que su tela sea de resistencia uniforme. Se sospecha que puede haber una variación significativa en la resistencia entre los telares. Considere los siguientes datos para los cuatro telares seleccionados al azar. Cada observación es una determinación de la resistencia de la tela en libras por pulgada cuadrada.

Telar	1	2	3	4
	99	97	94	93
	97	96	95	94
	97	92	90	90
	96	98	92	92

- (a) Escriba un modelo para el experimento.
 (b) ¿El componente de la varianza del telar difiere de forma significativa de cero?
 (c) Comente la sospecha.

Ejercicios de repaso

1. Se lleva a cabo un análisis en el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia junto con el Departamento de Silvicultura. Se aplica cierto tratamiento a un conjunto de cepas de árbol. Se utiliza el químico Garlon con el propósito de regenerar las raíces de las cepas. Se utiliza un pulverizador con cuatro niveles de concentración de Garlon. Después de un periodo se observa la altura de los retoños. Trate los siguientes datos como un análisis de varianza de un factor. Pruebe si la concentración de Garlon tiene un impacto significativo en la altura de los retoños. Utilice $\alpha = 0.05$.

Nivel de Garlon	1	2	3	4
1	2.87	2.31	3.91	2.04
2	3.27	2.66	3.15	2.00
3	2.39	1.91	2.89	1.89
4	3.05	0.91	2.43	0.01

2. Considere los datos de las mezclas del ejemplo 13.1. Realice una prueba de Bartlett para determinar si hay heterogeneidad de la varianza entre las mezclas.
3. En 1983 el Departamento de Ciencias Lácteas en el Instituto Politécnico y Universidad Estatal de Virginia realizó un experimento para estudiar el efecto de las raciones de alimento, diferenciadas por fuente

- de proteína, sobre la producción promedio diaria de la vacas. Se utilizaron cinco raciones en el experimento. Se usó un cuadrado latino de 5×5 en el que los renglones representan diferentes vacas y las columnas diferentes periodos de lactancia. Los siguientes datos, registrados en kilogramos, se analizaron en el Centro de Consulta Estadística del Tecnológico de Virginia. ¿Podemos concluir en el nivel de significancia de 0.01 que las raciones con diferentes fuentes de proteína tienen un efecto sobre la producción promedio de leche diaria de las vacas?

Periodos de lactancia	Vacas				
	1	2	3	4	5
1	A 33.1	B 34.4	C 26.4	D 34.6	E 33.9
2	C 30.7	D 28.7	E 24.9	A 28.8	B 28.0
3	D 28.7	E 28.8	A 20.0	B 31.9	C 22.7
4	E 31.4	A 22.3	B 18.7	C 31.0	D 21.3
5	B 28.9	C 22.3	D 15.8	E 30.9	A 19.0

4. Se utilizan tres catalizadores en un proceso químico con un control (no catalizador) incluido. Los siguientes son datos de rendimiento del proceso:

Control	Catalizador:	1	2	3
74.5		77.5	81.5	78.1
76.1		82.0	82.3	80.2
75.9		80.6	81.4	81.5
78.1		84.9	79.5	83.0
76.2		81.0	83.0	82.1

Utilice la prueba de Dunnett en el nivel de significancia $\alpha = 0.01$ para determinar si se obtiene un rendimiento significativamente alto con el catalizador que sin este.

5. Se utilizan cuatro laboratorios para realizar análisis químicos. Muestras del mismo material se mandan a los laboratorios para su análisis como parte del estudio para determinar si, en promedio, dan los mismos resultados. Los resultados analíticos para los cuatro laboratorios son los siguientes:

Laboratorio:	A	B	C	D
	58.7	62.7	55.9	60.7
	61.4	64.5	56.1	60.3
	60.9	63.1	57.3	60.9
	59.1	59.2	55.2	61.4
	58.2	60.3	58.1	62.3

(a) Use la prueba de Bartlett para mostrar que las varianzas dentro del laboratorio no son significativamente diferentes en el nivel de significancia de 0.05.

(b) Realice el análisis de varianzas y proporcione conclusiones con respecto a los laboratorios.

6. Utilice la prueba de Bartlett en el nivel de significancia de 0.01 para probar la homogeneidad de las varianzas en el ejercicio 9 de la página 475.

7. Use la prueba de Cochran en el nivel de significancia de 0.01 para probar la homogeneidad de las varianzas en el ejercicio 6 de la página 474.

8. Use la prueba de Bartlett en el nivel de significancia de 0.05 para probar la homogeneidad de las varianzas en el ejercicio 8 de la página 474.

9. Personal del Departamento de Ciencias Animales del Instituto Politécnico y Universidad Estatal de Virginia diseña un experimento con el propósito de estudiar un tratamiento de urea y amoníaco acuoso de la paja de trigo. El propósito es mejorar el valor

nutricional de las ovejas macho. Los tratamientos de dieta son: control; urea en el alimento; paja tratada con amoníaco; paja tratada con urea. Se usaron 24 ovejas en el experimento y se separaron de acuerdo con su peso relativo. Había seis ovejas en cada grupo homogéneo. A cada una de las seis en cada grupo cada una de las cuatro dietas en orden aleatorio. Para cada una de las 24 ovejas se mide el porcentaje de materia seca digerida. Los datos son los siguientes:

Dieta	Grupo por peso (bloques)					
	1	2	3	4	5	6
Control	32.68	36.22	36.36	40.95	34.99	33.80
Urea en el alimento	35.90	38.73	37.55	34.64	37.36	34.35
Tratado con amoníaco	49.43	53.50	52.86	49.00	47.20	49.76
Tratado con urea	46.58	42.82	45.41	43.08	43.81	47.40

(a) Utilice un análisis tipo bloque aleatorizado para probar las diferencias entre las dietas. Use $\alpha = 0.05$.

(b) Utilice la prueba de Dunnett para comparar las tres dietas con el control. Use $\alpha = 0.05$.

10. En un conjunto de datos que analiza el personal del Departamento de Bioquímica del Instituto Politécnico y Universidad Estatal de Virginia se dieron tres dietas a un grupo de ratas a fin de estudiar el efecto de cada una en el zinc dietético residual en el flujo sanguíneo. Se asignan al azar cinco ratas preñadas a cada grupo de dieta y a cada una se le da la dieta en el día 22 de embarazo. Se mide la cantidad de zinc en partes por millón. Los datos son los siguientes:

1	0.50	0.42	0.65	0.47	0.44
Dieta 2	0.42	0.40	0.73	0.47	0.69
3	1.06	0.82	0.72	0.72	0.82

Determine si hay una diferencia significativa en el zinc dietético residual entre las tres dietas. Use $\alpha = 0.05$. Realice un ANOVA unilateral.

11. Se realiza un estudio para comparar el rendimiento para tres marcas de gasolina en competencia. Se seleccionan cuatro modelos diferentes de automóvil de tamaño variable. Los datos, en millas por galón, son los siguientes. El orden de prueba es al azar para cada modelo.

Modelo	Marca de gasolina		
	A	B	C
A	32.4	35.6	38.7
B	28.8	28.6	29.9
C	36.5	37.6	39.1
D	34.4	36.2	37.9

- (a) Analice la necesidad del uso de más de un modelo.
 (b) Considere el ANOVA del siguiente impreso del SAS. ¿Importa la marca de la gasolina?
 (c) ¿Cuál marca de gasolina seleccionaría? Consulte el resultado de la prueba de Duncan.

Analysis of Variance Procedure

Dependent Variable: MILEAGE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	153.2508333	30.6516667	24.66	0.0006
Error	6	7.4583333	1.2430556		
Corrected Total	11	160.7091667			

	R-Square	C.V.	Root MSE	MILEAGE Mean
	0.953591	3.218448	1.1149240	34.641867

Source	DF	Anova SS	Mean Square	F Value	Pr > F
MODEL	3	130.3481867	43.4497222	34.96	0.0003
BRAND	2	22.9016667	11.4508333	9.21	0.0148

Duncan's Multiple Range Test for variable: MILEAGE

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha= 0.05 df= 6 MSE= 1.243056

Number of Means = 2 3
 Critical Range 1.829 1.999

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	BRAND
A	36.4000	4	C
A	34.5000	4	B
B	33.0250	4	A

12. Una compañía que troquea empaques de hojas de caucho, plástico y cocho quiere comparar el número medio de empaques producidos por hora para tres tipos de material. Se eligen al azar dos máquinas troqueladoras como bloques. Los datos representan el número de empaques (en miles) producidos por hora. El impreso del análisis es el siguiente.

- (a) ¿Por qué se eligen como bloques las máquinas troqueladoras?
 (b) Grafique las seis medias para las combinaciones de máquina y material.

- (c) ¿Hay un solo material que sea el mejor?
 (d) ¿Existe una interacción entre tratamientos y bloques? Si es así, ¿la interacción ocasiona alguna dificultad seria para llegar a una conclusión apropiada? Explique.

Material del empaque: 1 = cocho 2 = caucho 3 = plástico

Máquina troqueladora	Material del empaque					
	1	2	3	1	2	3
A	4.31	4.27	4.40	3.36	3.42	3.49
B	3.93	3.81	3.99	3.91	3.86	3.87

Analysis of Variance Procedure

Dependent Variable: Y						
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F	
Model	5	1.68120778	0.33624156	76.52	<0.0001	
Error	12	0.05273333	0.00439444			
Corrected Total	17	1.73394111				
R-Square		C.V.	Root MSE			Y Mean
	0.969588	1.734095	0.0662906			3.8227778

Source	DF	Anova SS	Mean Square	F Value	Pr > F
MACHINE	1	0.10125000	0.10125000	23.04	0.0004
MATERIAL	2	0.81194444	0.40597222	92.38	<0.0001
MACHINE*MATERIAL	2	0.76803333	0.38401667	87.39	<0.0001

Level of MACHINE	Level of MATERIAL	N	Y	
			Mean	SD
1	1	3	4.3266667	0.06658328
1	2	3	3.4200000	0.06000000
1	3	3	3.9466667	0.06027714
2	1	3	3.9133333	0.08291573
2	2	3	3.8533333	0.05507571
2	3	3	3.4766667	0.05507571

Level of MACHINE	N	Y	
		Mean	SD
1	9	3.8977778	0.39796800
2	9	3.7477778	0.21376259

Level of MATERIAL	N	Y	
		Mean	SD
1	6	4.1200000	0.23765521
2	6	3.6366667	0.24287171
3	6	3.7186667	0.26255793

Experimentos factoriales

14.1 Introducción

Considere una situación donde es de interés estudiar el efecto de **dos factores**, *A* y *B*, sobre alguna respuesta. Por ejemplo, en un experimento químico descartaríamos variar de manera simultánea la presión de reacción y el tiempo de reacción y estudiar el efecto de cada uno en el rendimiento. En un experimento biológico es de interés estudiar el efecto del tiempo de secado y la temperatura sobre la cantidad de sólidos (porcentaje por peso) que quedan en muestras de levadura. Como en el capítulo 13, el término **factor** se utiliza en un sentido general para denotar cualquier característica del experimento que se puede variar como temperatura, tiempo o presión de una prueba a otra. Definimos los **niveles** de un factor como los valores reales que se utilizan en el experimento.

Para cada uno de estos casos es importante no sólo determinar si los dos factores tienen una influencia en la respuesta, sino también si hay una interacción significativa entre los dos factores. En lo que respecta a la terminología, el experimento que aquí se describe es una clasificación en dos direcciones o un experimento de dos factores y el diseño experimental puede ser un diseño completamente aleatorizado, en el que las diversas combinaciones de tratamiento se asignan de forma aleatoria a todas las unidades experimentales, o un diseño en bloques completos aleatorizados, en el que las combinaciones de factores se asignan aleatoriamente a bloques. En el caso del ejemplo de la levadura, las diversas combinaciones de tratamiento de temperatura y tiempo de secado se asignarían aleatoriamente a las muestras de levadura si utilizamos un diseño completamente aleatorizado.

Muchos de los conceptos que se estudian en el capítulo 13 se extienden en este capítulo a dos y tres factores. El principal avance en este material es el uso del diseño completamente aleatorizado con un *experimento factorial*. Un experimento factorial en dos factores implica pruebas experimentales (o una sola prueba) en todas las combinaciones de los factores. Como ilustración, en el ejemplo del tiempo de secado-temperatura con, digamos, tres niveles de cada uno y $n = 2$ corridas en cada una de las nueve combinaciones, tenemos un *factorial de dos factores en un diseño completamente aleatorizado*. Ningún factor es un factor de bloque; nos interesa cómo ambas influyen en el porcentaje de sólidos en las muestras y si interactúan o no. El biólogo tendría disponibles entonces 18 muestras físicas de material, las cuales son unidades experimentales. Estas se asignarían entonces de forma aleatoria a las 18 combinaciones (nueve combinaciones de tratamiento, cada una duplicada).

Antes de profundizar en los detalles analíticos, sumas de cuadrados, etcétera, puede ser de interés para el lector observar la conexión obvia entre lo que describimos y la situación con el problema de un factor. Considere el experimento de la levadura. La explicación de los grados de libertad ayuda al lector o al analista a visualizar el alcance. Se deben ver inicialmente las nueve combinaciones de tratamiento como si representaran un factor con nueve niveles (ocho grados de libertad). De esta manera, una observación inicial a los grados de libertad da

Combinaciones de tratamiento	8
Error	9
Total	17

Efectos principales e interacción

En realidad, el experimento se podría analizar como se describe en esta tabla. Sin embargo, la prueba F para combinaciones probablemente no dé al analista la información que desea, a saber, la que considera el papel de la temperatura y el tiempo de secado. Tres tiempos de secado tienen 2 grados de libertad asociados, tres temperaturas tienen dos grados de libertad. Los factores principales, temperatura y tiempo de secado, se denominan **efectos principales**. Los efectos principales representan cuatro de los ocho grados de libertad para *combinaciones de factores*. Los cuatro grados de libertad adicionales se asocian con *interacciones* entre los dos factores. Como resultado, el análisis incluye:

Combinaciones	8
Temperatura	2
Tiempo de secado	2
Interacción	4
Error	9
Total	17

Recuerde del capítulo 13 que los factores en un análisis de varianza se pueden ver como fijos o aleatorios lo que depende del tipo de inferencia que se desea y de cómo se eligen los niveles. Aquí debemos considerar efectos fijos, efectos aleatorios, e incluso casos donde los efectos están mezclados. Se debe poner mucha atención en los cuadrados medios esperados cuando avancemos a estos temas. En la siguiente sección ponemos más atención al concepto de interacción.

14.2 Interacción y el experimento de dos factores

En el modelo de bloques aleatorizados que se presentó anteriormente se supuso que en cada bloque se toma una observación de cada tratamiento. Si la suposición del modelo es correcta, es decir, si los bloques y los tratamientos son los únicos efectos reales y no existe la interacción, el valor esperado del cuadrado medio del error es la varianza del error experimental σ^2 . Suponga, sin embargo, que ocurre interacción entre los tratamientos y bloques como indica el modelo

$$y_{ij} = \mu + \alpha_j + \beta_i + (\alpha\beta)_{ij} + x_{ij}$$

de la sección 13.9. El valor esperado del cuadrado medio del error está dado entonces por

$$E\left[\frac{SSE}{(b-1)(k-1)}\right] = \sigma^2 + \frac{\sum_{i=1}^b \sum_{j=1}^k (\alpha\beta)_{ij}^2}{(b-1)(k-1)}$$

Los efectos de tratamiento y bloques no aparecen en el cuadrado medio del error esperado, pero los efectos de la interacción sí. De esta forma, si hay interacción en el modelo, el cuadrado medio del error refleja la variación debida al error experimental más una contribución de interacción y, para este plan experimental, no hay forma de separarlos.

Desde el punto de vista del experimentador parecería necesario llegar a una prueba de significancia sobre la existencia de interacción mediante la separación de la variación del error real de los debidos a la interacción. Los efectos principales, A y B , toman un significado diferente en presencia de interacción. En el ejemplo biológico anterior el efecto que tiene el tiempo de secado sobre la cantidad de sólidos que quedan en la levadura podría muy bien depender de la temperatura a la que se exponen las muestras. En general, podría haber situaciones experimentales en las que el factor A tiene un efecto positivo sobre la respuesta en un nivel del factor B , mientras en un nivel diferente del factor B el efecto de A es negativo. Utilizamos aquí el término **efecto positivo** para indicar que el rendimiento o respuesta aumenta conforme los niveles de un factor dado aumentan de acuerdo con algún orden definido. En el mismo sentido un **efecto negativo** corresponde a una disminución en el rendimiento para niveles crecientes del factor. Considere, por ejemplo, los siguientes datos hipotéticos tomados para dos factores en tres niveles:

A	B	b_1	b_2	b_3	Total
a_1		4.4	8.8	8.2	18.4
a_2		7.5	8.5	2.4	18.4
a_3		9.7	7.9	0.8	18.4
Total		21.6	25.2	8.4	55.2

Claramente, el efecto de A es positivo en b_1 y negativo en b_3 . Estas diferencias en los niveles de A en diferentes niveles de B son de interés para el experimentador, pero una prueba de significancia común sobre el factor A produciría un valor de cero para SSA , pues los totales para cada nivel de A son todos de la misma magnitud. Decimos, entonces, que la presencia de interacción **encubre** el efecto del factor A .

Por tanto, si consideramos la influencia promedio de A , sobre los tres niveles de B , no hay efecto. Sin embargo, lo más probable es que esto no sea lo pertinente para el experimentador.

Antes de sacar algunas conclusiones finales que resulten de las pruebas de significancia sobre los efectos principales y los efectos de la interacción, el experimentador primero debe observar si la prueba de interacción es significativa o no. Si la interacción no es significativa, entonces los resultados de las pruebas sobre los efectos principales sí lo son. Sin embargo, si la interacción debe ser significativa, entonces sólo las pruebas sobre los efectos principales que resultan ser significativa, tienen sentido. Efectos principales no significativos en presencia de interacción podrían ser el resultado del encubrimiento y dictar la necesidad de observar la influencia de cada factor en niveles fijos del otro.

El error experimental y de interacción están separados en el experimento de dos factores sólo si se toman observaciones múltiples en las diversas combinaciones de tratamiento. Para eficiencia máxima, debe haber el mismo número n de observaciones en cada combinación. Éstas deben ser réplicas reales, no sólo mediciones repetidas. Por ejemplo, en el caso de la ilustración de la levadura, si tomamos $n = 2$ observaciones en cada combinación de temperatura y tiempo de secado, debe haber dos muestras separadas y no solamente mediciones repetidas de la misma muestra. Esto permite que la variabilidad debida a las unidades experimentales aparezca en el "error", por lo que la variación no es sólo error de medición.

14.3 Análisis de varianza de dos factores

Para presentar las fórmulas generales del análisis de varianza de un experimento de dos factores con el uso de observaciones repetidas en un diseño totalmente aleatorizado, consideraremos el caso de n réplicas de las combinaciones del tratamiento que se determinan por a niveles del factor A y b niveles del factor B . Las observaciones se pueden clasificar por medio de un arreglo rectangular donde los renglones representan los niveles del factor A y las columnas los niveles del factor B . Cada combinación de tratamiento define una celda en nuestro arreglo. De esta manera tenemos ab celdas, cada una de las cuales contiene n observaciones. Denotemos la k -ésima observación tomada en el i -ésimo nivel del factor A y el j -ésimo nivel del factor B con y_{ijk} . Las abn observaciones se muestran en la tabla 14.1.

Las observaciones en la celda (ij) -ésima constituyen una muestra aleatoria de tamaño n de una población que se supone distribuida normalmente con media μ_{ij} y varianza σ^2 . Todas las ab poblaciones se suponen con la misma varianza σ^2 . Definamos los siguientes símbolos útiles, algunos de los cuales se utilizan en la tabla 14.1:

Y_{ij} = suma de observaciones en la (ij) -ésima celda.

$Y_{i.}$ = suma de observaciones para el i -ésimo nivel del factor A .

$Y_{.j}$ = suma de observaciones para el j -ésimo nivel del factor B .

$Y_{..}$ = suma de todas las abn observaciones.

y_{ij} = media de las observaciones en la (ij) -ésima celda.

$\bar{y}_{i.}$ = media de las observaciones para el i -ésimo nivel del factor A .

$\bar{y}_{.j}$ = media de las observaciones para el j -ésimo nivel del factor B .

$\bar{y}_{..}$ = media de todas las abn observaciones.

Tabla 14.1 Experimento de dos factores con n réplicas

A	B:	1	2	...	b	Total	Media			
1		y_{111}	y_{112}	...	y_{11b}	$Y_{1.}$	$\bar{y}_{1.}$			
		y_{121}	y_{122}	...	y_{12b}					
		\vdots	\vdots	...	\vdots					
2		y_{211}	y_{212}	...	y_{21b}	$Y_{2.}$	$\bar{y}_{2.}$			
		y_{221}	y_{222}	...	y_{22b}					
		\vdots	\vdots	...	\vdots					
\vdots		\vdots	\vdots	...	\vdots	\vdots	\vdots			
	a		y_{a11}	y_{a12}	...			y_{a1b}	$Y_{a.}$	$\bar{y}_{a.}$
			y_{a21}	y_{a22}	...			y_{a2b}		
		\vdots	\vdots	...	\vdots					
Total		$Y_{.1}$	$Y_{.2}$...	$Y_{.b}$	$Y_{..}$	$\bar{y}_{..}$			
Media		$\bar{y}_{.1}$	$\bar{y}_{.2}$...	$\bar{y}_{.b}$		$\bar{y}_{..}$			

Cada observación en la tabla 14.1 se puede escribir en la forma

$$y_{ijk} = \mu_{ij} + e_{ijk}$$

donde e_{ijk} mide las desviaciones de los valores y_{ijk} observados en la (ij) -ésima celda de la media poblacional μ_{ij} . Si denotamos con $(\alpha\beta)_{ij}$ el efecto de la interacción del i -ésimo nivel del factor A y el j -ésimo nivel del factor B , con α_i el efecto del i -ésimo nivel del factor A , con β_j el efecto del j -ésimo nivel del factor B y con μ la media general, podemos escribir

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

y entonces:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

sobre la que imponemos las restricciones:

$$\sum_{i=1}^a \alpha_i = 0, \quad \sum_{j=1}^b \beta_j = 0, \quad \sum_{i=1}^a (\alpha\beta)_{ij} = 0, \quad \sum_{j=1}^b (\alpha\beta)_{ij} = 0.$$

Las tres hipótesis a probar son las siguientes:

- $H_0^1: \alpha_1 = \alpha_2 = \dots = \alpha_a = 0$
 $H_1^1: \text{Al menos una de las } \alpha_i \text{ no es igual a cero.}$
- $H_0^2: \beta_1 = \beta_2 = \dots = \beta_b = 0$
 $H_1^2: \text{Al menos una de las } \beta_j \text{ no es igual a cero.}$
- $H_0^3: (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{ab} = 0$
 $H_1^3: \text{Al menos una de las } (\alpha\beta)_{ij} \text{ no es igual a cero.}$

Cada una de estas pruebas se basará en una comparación de estimaciones independientes de σ^2 que proporciona la división de la suma total de cuadrados de nuestros datos en cuatro componentes por medio de la siguiente identidad.

Teorema 14.1 *Identidad de suma de cuadrados*

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y})^2 &= bn \sum_{i=1}^a (\bar{y}_{i.} - \bar{y})^2 + an \sum_{j=1}^b (\bar{y}_{.j} - \bar{y})^2 \\ &\quad + n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y})^2 \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ijk})^2 \end{aligned}$$

De manera simbólica, escribimos la identidad de suma de cuadrados como

$$SST = SSA + SSB + SS(AB) + SSE,$$

donde SSA y SSB se llaman suma de cuadrados para los efectos principales de A y B , respectivamente, $SS(AB)$ se llama suma de cuadrados de la interacción para A y B , y SSE es la suma de cuadrados del error. Los grados de libertad se dividen de acuerdo con la identidad

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1).$$

Formación de cuadrados medios

Si dividimos cada una de las sumas de cuadrados del lado derecho de la identidad de suma de cuadrados entre su correspondiente número de grados de libertad, obtenemos las cuatro estadísticas

$$s_1^2 = \frac{SSA}{a - 1}, \quad s_2^2 = \frac{SSB}{b - 1}, \quad s_3^2 = \frac{SS(AB)}{(a - 1)(b - 1)}, \quad s_4^2 = \frac{SSE}{ab(n - 1)}$$

Todas estas varianzas estimadas son estimaciones independientes de σ^2 bajo la condición de que no hay efectos α_i , β_j , y, por supuesto, $(\alpha\beta)_{ij}$. Si interpretamos la suma de cuadrados como funciones de las variables aleatorias independientes $y_{(1)}, y_{(2)}, \dots, y_{(abn)}$, no es difícil verificar que

$$E(S_1^2) = E\left[\frac{SSA}{a - 1}\right] = \sigma^2 + \frac{nb \sum_{i=1}^a \alpha_i^2}{a - 1},$$

$$E(S_2^2) = E\left[\frac{SSB}{b - 1}\right] = \sigma^2 + \frac{na \sum_{j=1}^b \beta_j^2}{b - 1}.$$

$$E(S_1^2) = E\left[\frac{SS(AB)}{(a-1)(b-1)}\right] = \sigma^2 + \frac{n \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2}{(a-1)(b-1)}$$

$$E(S^2) = E\left[\frac{SSE}{ab(n-1)}\right] = \sigma^2$$

de donde de inmediato observamos que las cuatro estimaciones de σ^2 son insesgadas cuando H_0 , H_0' , y H_0'' son verdaderas.

Para probar la hipótesis H_0' , de que los efectos de los factores A son todos iguales a cero, calculamos la razón

$$f_1 = \frac{s_1^2}{s^2}$$

que es un valor de la variable aleatoria F_1 que tiene la distribución F con $a-1$ y $ab(n-1)$ grados de libertad cuando H_0' es verdadera. La hipótesis nula se rechaza en el nivel de significancia α cuando $f_1 > f_{\alpha}[a-1, ab(n-1)]$. De manera similar, para probar la hipótesis H_0'' , de que todos los efectos del factor B son iguales a cero, calculamos la razón

$$f_2 = \frac{s_2^2}{s^2}$$

que es un valor de la variable aleatoria F_2 que tiene la distribución F con $b-1$ y $ab(n-1)$ grados de libertad cuando H_0'' es verdadera. La hipótesis se rechaza en el nivel de significancia α cuando $f_2 > f_{\alpha}[b-1, ab(n-1)]$. Finalmente, para probar la hipótesis H_0 , que todos los efectos de interacción son iguales a cero, calculamos la razón

$$f_3 = \frac{s_3^2}{s^2}$$

que es un valor de la variable aleatoria F_3 que tiene la distribución F con $(a-1)(b-1)$ y $ab(n-1)$ grados de libertad cuando H_0 es verdadera. Concluimos que la interacción está presente cuando

$$f_3 > f_{\alpha}[(a-1)(b-1), ab(n-1)]$$

Como se indica en la sección 14.2, se aconseja realizar la prueba de la interacción antes de intentar hacer inferencias sobre los efectos principales. Si la interacción no es significativa, hay en realidad evidencia de que las pruebas sobre los efectos principales son interpretables. Sin embargo, una interacción significativa podría implicar muy bien que los datos se deben analizar de una manera algo diferente —quizá mediante la observación de los efectos del factor A a niveles fijos del factor B , etcétera.

Los cálculos en un problema de análisis de varianza, para un experimento de dos factores con n réplicas, por lo general se resumen como en la tabla 14.2.

Ejemplo 14.1 En un experimento que se lleva a cabo para determinar cuál de tres diferentes sistemas de misiles es preferible, se mide la tasa de consumo del propul-

Tabla 14.2 Análisis de varianza para el experimento de dos factores con n replicas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	f calculada
Efecto principal				
A	SSA	$a - 1$	$s_1^2 = \frac{SSA}{a - 1}$	$f_1 = \frac{s_1^2}{s^2}$
B	SSB	$b - 1$	$s_2^2 = \frac{SSB}{b - 1}$	$f_2 = \frac{s_2^2}{s^2}$
Interacciones de dos factores				
AB	$SS(AB)$	$(a - 1)(b - 1)$	$s_3^2 = \frac{SS(AB)}{(a - 1)(b - 1)}$	$f_3 = \frac{s_3^2}{s^2}$
Error	SSE	$ab(n - 1)$	$s^2 = \frac{SSE}{ab(n - 1)}$	
Total	SST	$abn - 1$		

por para 24 encendidos estáticos. Se utilizan cuatro tipos diferentes de propulsores. El experimento da observaciones duplicadas de las tasas de consumo en cada combinación de los tratamientos. Los datos, después de codificarlos, se dan en la tabla 14.3. Pruebe las siguientes hipótesis: (a) H_0^a : no hay diferencia en las tasas medias de consumo del propulsor cuando se utilizan los diferentes sistemas de misiles; (b) H_0^b : no hay diferencia en las tasas de consumo medio del propulsor de los cuatro tipos de propulsores; (c) H_0^c : no hay interacción entre los diferentes sistemas de misiles y los diferentes tipos de propulsores.

Tabla 14.3 Tasas de consumo de los propulsores

Sistema de misil	Tipo de propulsor	b_1	b_2	b_{12}	b_4
a_1		34.0	30.1	29.8	29.0
		32.7	32.8	26.7	28.9
a_2		32.0	30.2	28.7	27.6
		33.2	29.8	28.1	27.8
a_3		28.4	27.3	29.7	28.8
		29.3	28.9	27.3	29.1

SOLUCIÓN

- (a) H_0^a : $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

(b) H_0^b : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$.

(c) H_0^c : $(\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{34} = 0$.
- (a) H_0^a : Al menos una de las α_i es diferente de cero.

(b) H_0^b : Al menos una de las β_j es diferente de cero.

(c) H_0^c : Al menos una de las $(\alpha\beta)_{ij}$ es diferente de cero.

Las fórmulas de sumas de cuadrados se utilizan como se describe en la página 532. El análisis de varianza se muestra en la tabla 14.4.

Se remite al lector a un procedimiento de análisis de varianza del SAS para el análisis de los datos de la tasa de consumo. Nótese cómo el "modelo" (11 grados de libertad) se prueba inicialmente y el sistema, el tipo y la interacción del sistema por tipo se prueban por separado. La prueba f del modelo ($P = 0.0030$) prueba la acumulación de los dos efectos principales y la interacción.

Analysis of Variance Procedure					
Dependent Variable: RATE					
Propellant Burning Rate					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	76.76833333	6.97893939	5.82	0.0030
Error	12	14.91000000	1.24250000		
Corrected Total	23	91.67833333			
	R-Square	C.V.	Root MSE		RATE Mean
	0.837366	3.766854	1.146748		29.591667
Source	DF	Anova SS	Mean Square	F Value	Pr > F
SYSTEM	2	14.52333333	7.26166667	5.84	0.0169
TYPE	3	40.08166667	13.36055556	10.75	0.0010
SYSTEM*TYPE	6	22.16333333	3.69388889	2.97	0.0512

Figura 14.1 Impreso del SAS del análisis de los datos de tasa de consumo de propulsor de la tabla 14.3.

Tabla 14.4 Análisis de varianza para los datos de la tabla 14.3

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Sistema de misiles	14.52	2	7.26	5.85
Tipo de propulsor	40.08	3	13.36	10.77
Interacción	22.17	6	3.70	2.98
Error	14.91	12	1.24	
Total	91.68	23		

- Rechazar H_0 y concluir que los diversos sistemas de misiles tienen como resultado diferentes tasas medias de consumo del propulsor. El valor P es aproximadamente 0.017.
- Rechazar H_0 y concluir que las tasas medias de consumo del propulsor no son las mismas para los cuatro tipos de propulsor. El valor P es menor que 0.0010.
- La interacción es apenas insignificante en el nivel 0.05, pero el valor P de aproximadamente 0.0512 indicaría que la interacción se debe tomar en serio.

En este punto se debe hacer algún tipo de interpretación de la interacción. Se debe enfatizar que la significancia estadística de un efecto principal simplemente indica que estas *medias marginales son significativamente diferentes*. Sin embargo, considere la siguiente tabla de promedios en dos direcciones:

	b_1	b_2	b_3	b_4	Promedio
a_1	66.7	62.9	56.5	57.9	30.5
a_2	65.2	60.0	56.8	55.4	29.7
a_3	57.7	56.2	57.0	57.9	28.6
Promedio	31.6	29.9	28.4	28.5	

Es evidente que existe más información importante en el cuerpo de la tabla —tendencias inconsistentes con la tendencia que describen los promedios marginales—. Esta tabla en realidad sugiere que el efecto del tipo de propulsor depende del sistema que se utilice. Por ejemplo, para el sistema 3 el efecto del tipo de propulsor no parece ser importante, aunque tiene un efecto grande si se utiliza el sistema 1 o el sistema 2. Esto explica la interacción "significativa". Se revelará más con respecto a esta interacción más adelante. ■

Ejemplo 14.2 Con referencia al ejemplo 14.1, elija dos contrastes ortogonales para dividir la suma de cuadrados para los sistemas de misiles en componentes de un solo grado de libertad para utilizarlos en la comparación de los sistemas 1 y 2 con el 3 y el sistema 1 contra el 2.

SOLUCIÓN

El contraste para comparar los sistemas 1 y 2 con el 3 es

$$w_1 = \mu_1 + \mu_2 - 2\mu_3$$

Un segundo contraste, ortogonal a w_1 , para comparar el sistema 1 con el 2, está dado por $w_2 = \mu_1 - \mu_2$. Las sumas de cuadrados de un solo grado de libertad son

$$SSw_1 = \frac{[244.0 + 237.4 - (2)(228.8)]^2}{(8)[(1)^2 + (1)^2 + (-2)^2]} = 11.80$$

y

$$SSw_2 = \frac{(244.0 - 237.4)^2}{(8)[(1)^2 + (-1)^2]} = 2.72$$

Nótese que $SSw_1 + SSw_2 = SSA$, como se esperaba. Los valores f calculados que corresponden a w_1 y w_2 son, respectivamente,

$$f_1 = \frac{11.80}{1.24} = 9.5 \quad \text{y} \quad f_2 = \frac{2.72}{1.24} = 2.2$$

Comparado con el valor crítico $f_{0.05}(1, 12) = 4.75$, encontramos que f_1 es significativo. De hecho, el valor P es menor que 0.01. De esta forma el primer contraste indica que la hipótesis

$$H_0: \frac{\mu_1 + \mu_2}{2} = \mu_3$$

se rechaza. Como $f_2 < 4.75$, las tasas medias de consumo del primer y segundo sistema no son significativamente diferentes. ■

Impacto de la interacción significativa

Si la hipótesis de no interacción en el ejemplo 14.1 es verdadera, podemos hacer las comparaciones *generales* del ejemplo 14.2 con respecto a nuestros sistemas de misiles en lugar de comparaciones separadas para cada propulsor. De manera similar, podemos hacer comparaciones generales entre los propulsores en lugar de comparaciones separadas para cada sistema de misiles. Por ejemplo, podemos comparar los propulsores 1 y 2 con 3 y 4 y también el propulsor 1 contra el propulsor 2. Las razones f que resultan, cada una con 1 y 12 grados de libertad, resultan ser 24.86 y 7.41, respectivamente, y ambas son bastante significativas en el nivel 0.05.

De los promedios de los propulsores parece haber evidencia de que el propulsor 1 da la tasa media de consumo más alta. Un experimentador prudente podría ser algo cauto al hacer conclusiones globales en un problema como éste, donde la razón f para la interacción está apenas por debajo del valor crítico 0.05. Por ejemplo, la evidencia total, 31.6 contra 29.9 sobre el promedio para los dos propulsores, en realidad indica que el propulsor 1 es superior, en términos de una tasa de consumo más alta, que el propulsor 2. Sin embargo, si nos restringimos al sistema 3, donde tenemos un promedio de 28.85 para el propulsor 1 en oposición a 28.10 para el propulsor 2, parece haber poca o ninguna diferencia entre estos dos propulsores. De hecho, parece haber una estabilización de las tasas de consumo para los diferentes propulsores si operamos con el sistema 3. Hay en realidad evidencia global que indica que el sistema 1 da una tasa de consumo más alta que el sistema 3, pero si nos restringimos al propulsor 4, esta conclusión no parece ser válida.

El analista puede realizar una prueba t simple con el uso de las tasas de consumo promedio en el sistema 3 a fin de mostrar evidencia concluyente de que la interacción produce grandes dificultades al permitir amplias conclusiones sobre los efectos principales. Considere una comparación del propulsor 1 contra el propulsor 2 sólo con el uso del sistema 3. Si nos apropiamos de una estimación de σ^2 del análisis global, es decir, utilizamos $s^2 = 1.24$ con 12 grados de libertad, podemos usar

$$|t| = \frac{0.75}{\sqrt{2s^2/n}} = \frac{0.75}{\sqrt{1.24}} = 0.68,$$

que ni siquiera está cerca de ser significativa. Esta ilustración sugiere que se debe tener cautela acerca de la interpretación estricta de los efectos principales en presencia de interacción.

14.4 Análisis gráfico en el problema de dos factores

Muchos de los tipos de representación gráfica que se sugieren en los problemas de un factor en realidad se aplican al caso de dos factores. Las gráficas de dos dimensiones de medias de celdas o medias de combinaciones de tratamientos pueden proporcionar una idea de la presencia de interacciones entre los dos factores. Además, una gráfica de residuos contra valores ajustados puede proporcionar una indicación de si la suposición de varianzas homogéneas es válida o no. A menudo, por supuesto, una violación de la suposición de varianzas homogéneas implica un aumento en la va-

rianza del error conforme los números de respuesta se hacen más grandes. Como resultado, esta gráfica puede señalar la violación.



Figura 14.2 Gráfica de las medias de celdas para los datos del ejemplo 14.1. Los números representan sistemas de misiles.

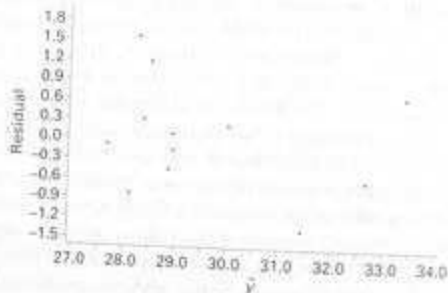


Figura 14.3 Gráfica de los residuos de los datos del ejemplo 14.1.

La figura 14.2 muestra la gráfica de medias de celda en el caso de la ilustración del propulsor del sistema de misiles del ejemplo 14.1. Nótese como se muestra gráficamente (en este caso) la falta de paralelismo. Nótese lo plano de la parte de la figura que muestra el efecto del propulsor en el sistema 3. Esto sugiere una interacción entre los factores. La figura 14.3 muestra la gráfica de los residuos contra los valores ajustados para los mismos datos. No hay un signo aparente de dificultad con la suposición de varianzas homogéneas.

Ejercicios

1. Se lleva a cabo un experimento para estudiar el efecto de la temperatura y tipo de horno sobre la vida de un componente particular que se prueba. En el experimento se utilizan cuatro tipos de horno y tres niveles de temperatura. Se asignan de forma aleatoria 24 piezas, dos a cada combinación de tratamientos, y

se registran los siguientes resultados de la página 538. Con el uso de un nivel de significancia de 0.05, pruebe la hipótesis de que

(a) diferentes temperaturas no tienen efecto sobre la vida del componente;

Temperatura (grados)	Horno			
	O ₁	O ₂	O ₃	O ₄
500	227	214	225	260
	221	259	236	229
550	187	181	232	246
	208	179	198	273
600	174	198	178	206
	202	194	213	219

(b) diferentes hornos no tienen efecto sobre la vida del componente;

(c) el tipo de horno y la temperatura no interactúan.

2. El Departamento de Nutrición Humana y Alimentos del Instituto Politécnico y Universidad Estatal de Virginia lleva a cabo el estudio *Vitamin C Retention in Reconstituted Frozen Orange Juice* para determinar la estabilidad de la vitamina C en el concentrado de jugo de naranja reconstituido congelado que se almacena en un refrigerador por un periodo superior a una semana. Se prueban tres tipos de concentrado de jugo de naranja congelado en tres periodos diferentes. Los periodos se refieren al número de días desde la mezcla del jugo de naranja hasta que se prueba. Los resultados, en miligramos de ácido ascórbico por litro, se registran a continuación:

Marca	Tiempo (días)					
	0	3	7			
Richfood	52.6	54.2	49.4	49.2	42.7	48.8
	49.8	46.5	42.8	53.2	40.4	47.6
Scaled-Sweet	56.0	48.0	48.8	44.0	49.2	44.0
	49.6	48.4	44.0	42.4	42.0	43.2
Minute Maid	52.5	52.0	48.0	47.0	48.5	43.3
	51.8	53.6	48.2	49.6	45.2	47.6

Utilice un nivel de significancia de 0.05 para probar la hipótesis de que

(a) no hay diferencia en el contenido de ácido ascórbico entre las diferentes marcas de concentrado de jugo de naranja;

(b) no hay diferencia en el contenido de ácido ascórbico para los diferentes tiempos;

(c) las marcas de concentrado de jugo de naranja y el número de días del momento en que se mezcla el jugo hasta su prueba no interactúan.

3. Se estudia el comportamiento de tres camadas de titanio bajo dos condiciones ambientales en una prueba de fatiga. Las calificaciones de error para las 48 ratas se registran a continuación:

Ambiente	Camada					
	Brillante	Mezclada	Lenta			
Libre	28	12	33	63	101	94
	22	23	36	14	33	56
	25	10	41	76	122	83
Restringido	36	86	22	58	35	23
	72	32	60	89	136	120
	48	93	35	126	38	153
	25	31	83	110	64	128
	91	19	99	118	87	140

Utilice un nivel de significancia de 0.01 para probar la hipótesis de que:

(a) no hay diferencia en las calificaciones de error para los diferentes ambientes;

(b) no hay diferencia en las calificaciones de error para las diferentes camadas;

(c) los ambientes y las camadas de ratas no interactúan.

4. La fatiga por corrosión en metales se define como la acción simultánea de la tensión cíclica y el ataque químico sobre una estructura metálica. Una técnica que se utiliza con amplitud para minimizar el daño por fatiga por corrosión en aluminio implica la aplicación de un recubrimiento protector. En un estudio que lleva a cabo el Departamento de Ingeniería Mecánica del Instituto Politécnico y Universidad Estatal de Virginia, se utilizan diferentes niveles de humedad

Bajo: 20–25% de humedad relativa

Medio: 55–60% de humedad relativa

Alto: 86–91% de humedad relativa

y tres tipos de recubrimientos superficiales

Sin recubrimiento: ningún recubrimiento

Anodizado: recubrimiento de óxido

ánodico a base de

ácido sulfúrico

Conversión: recubrimiento de conversión química de cromo.

Los datos de fatiga por corrosión, expresados en miles de ciclos antes de la falla, se registran a continuación:

Recubrimiento	Humedad relativa					
	Bajo	Medio	Alto			
Sin recubrimiento	361	469	314	522	1344	1216
	466	937	244	739	1027	1097
	1089	1357	261	134	663	1011
Anodizado	114	1032	322	471	78	466
	1236	92	306	130	387	107
	533	211	68	398	130	327
Conversión	130	1482	252	874	586	524
	841	529	105	785	402	751
	1595	754	847	573	846	529

(a) Realice un análisis de varianza con $\alpha = 0.05$ para probar la significancia de los efectos principales y de interacción.

(b) Utilice la prueba de rango múltiple de Duncan en el nivel de significancia 0.05 para determinar qué niveles de humedad tienen como resultado diferente daño debido a la fatiga por corrosión.

5. Se lleva a cabo un estudio en el Departamento de Salud Física, Educación y Recreación del Instituto Politécnico y Universidad Estatal de Virginia para determinar cuáles músculos se necesitan sujetar a un programa de acondicionamiento a fin de mejorar el desempeño en el servicio tendido en el juego de tenis. Se prueban cinco diferentes músculos:

- 1: deltoides anterior
- 2: pectoral mayor
- 3: deltoides posterior
- 4: deltoides medio
- 5: tríceps

en cada uno de tres sujetos, y el experimento se realiza tres veces para cada combinación de tratamiento. Los datos del electromiograma, que se registran durante el servicio, son los siguientes:

Sujeto	Músculo				
	1	2	3	4	5
1	32	5	58	10	19
	59	1.5	61	10	20
	38	2	66	14	23
2	63	10	64	45	43
	60	9	78	61	61
	50	7	78	71	42
3	43	41	26	63	61
	54	43	29	46	85
	47	42	23	55	95

Utilice un nivel de significancia de 0.01 para probar la hipótesis de que

- (a) diferentes sujetos tienen mediciones iguales del electromiograma;
- (b) los diferentes músculos no tienen efecto en las mediciones del electromiograma;

(c) los sujetos y los tipos de músculo no interactúan.

6. Se lleva a cabo un experimento para aumentar la capacidad de adhesión de productos de caucho. Se fabrican 16 productos con el nuevo aditivo y otros 16 sin éste. Las capacidades de adhesión se registran a continuación.

	Temperatura (°C)			
	50	60	70	80
Sin aditivos	2.3	3.4	3.8	3.9
Con aditivos	2.9	3.7	3.9	3.2
	3.1	3.6	4.1	3.0
	3.2	3.2	3.8	2.7
	4.3	3.8	3.9	3.5
Con aditivos	3.9	3.8	4.0	3.6
	3.9	3.9	3.7	3.8
	4.2	3.5	3.6	3.9

Realice un análisis de varianza para probar la significancia de los efectos principales y de interacción.

7. Se sabe que la tasa de extracción de cierto polímero depende de la temperatura de reacción y de la cantidad de catalizador que se utiliza. Se lleva a cabo un experimento en cuatro niveles de temperatura y cinco niveles del catalizador, y se registra la tasa de extracción en la siguiente tabla.

	Cantidad de catalizador				
	0.5%	0.6%	0.7%	0.8%	0.9%
50°C	38	45	57	50	37
	41	47	59	61	58
60°C	44	56	70	73	61
	43	57	69	72	58
70°C	44	56	70	73	61
	47	60	67	61	59
80°C	49	62	70	62	53
	47	65	55	69	58

Desarrolle un análisis de varianza. Pruebe la significancia de los efectos principales y de interacción.

14.5 Experimentos de tres factores

En esta sección consideramos un experimento con tres factores A , B y C , en los niveles a , b y c , respectivamente, en un diseño experimental completamente aleatorio. Suponga nuevamente que tenemos n observaciones para cada una de las abc combinaciones de tratamientos. Procederemos a señalar las pruebas de significancia para los tres efectos principales e interacciones que se involucran. Se espera que el

lector pueda utilizar después la descripción que se da aquí para generalizar el análisis a $k > 3$ factores.

El modelo para el experimento de tres factores es

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + e_{ijk}$$

$$i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, c; \quad \text{y } l = 1, 2, \dots, n$$

donde α_i , β_j , y γ_k son los efectos principales; $(\alpha\beta)_{ij}$, $(\alpha\gamma)_{ik}$, y $(\beta\gamma)_{jk}$ son los efectos de interacción de dos factores que tienen la misma interpretación que en el experimento de dos factores. El término $(\alpha\beta\gamma)_{ijk}$ se llama **efecto de interacción de tres factores**, término que representa una no aditividad de las $(\alpha\beta)_{ij}$ sobre los diferentes niveles del factor C . Como antes, la suma de todos los efectos principales es cero y la suma sobre cualquier subíndice de los efectos de interacción de dos y tres factores es cero. En muchas situaciones experimentales estas interacciones de orden superior son insignificantes y sus cuadrados medios reflejan sólo la variación aleatoria, pero señalaremos el análisis en su detalle más general.

De nuevo, a fin de que se puedan realizar pruebas de significancia válidas, debemos suponer que los errores y valores de variables aleatorias independientes y normalmente distribuidas, cada una con media cero y varianza común σ^2 .

La filosofía general con respecto al análisis es la misma que la que se presentó para los experimentos de uno y dos factores. La suma de cuadrados se divide en ocho términos, cada uno de los cuales representa una fuente de variación de la que obtenemos estimaciones independientes de σ^2 cuando todos los efectos principales y efectos de interacción son cero. Si los efectos de cualquier factor dados o interacción no son todos cero, entonces el cuadrado medio estimará la varianza del error más un componente debido al efecto sistemático en cuestión.

Aunque enfatizamos la interpretación de los impresos comentados de computadora en esta sección en lugar de preocuparnos del cálculo laborioso de la suma de cuadrados, ofrecemos a continuación las sumas de cuadrados para los tres efectos principales e interacciones. Nótese la expansión obvia del problema de dos a tres factores.

$$SSA = bcn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 \quad SS(AB) = cn \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$SSB = acn \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2 \quad SS(AC) = bn \sum_{i=1}^a \sum_{k=1}^c (\bar{y}_{ik.} - \bar{y}_{i..} - \bar{y}_{.k.} + \bar{y}_{...})^2$$

$$SSC = abn \sum_{k=1}^c (\bar{y}_{.k.} - \bar{y}_{...})^2 \quad SS(BC) = an \sum_{j=1}^b \sum_{k=1}^c (\bar{y}_{.jk.} - \bar{y}_{.j.} - \bar{y}_{.k.} + \bar{y}_{...})^2$$

$$SS(ABC) = n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c (\bar{y}_{ijk.} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{.k.} + \bar{y}_{i.} + \bar{y}_{.j.} + \bar{y}_{.k.} - \bar{y}_{...})^2$$

$$SST = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{l=1}^n (y_{ijkl} - \bar{y}_{...})^2 \quad SSE = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{l=1}^n (y_{ijkl} - \bar{y}_{ijk.})^2$$

A continuación se definen los promedios en las fórmulas anteriores:

$\bar{y}_{...}$ = promedio de todas las $abcn$ observaciones

$\bar{y}_{i..}$ = promedio de las observaciones para el i -ésimo nivel del factor A.

- $\bar{y}_{.j}$ = promedio de las observaciones para el j -ésimo nivel del factor B
 $\bar{y}_{.k}$ = promedio de las observaciones para el k -ésimo nivel del factor C
 $\bar{y}_{i.}$ = promedio de las observaciones para el i -ésimo nivel de A y el j -ésimo nivel de B
 $\bar{y}_{.jk}$ = promedio de las observaciones para el i -ésimo nivel de A y el k -ésimo nivel de C
 $\bar{y}_{.jk}$ = promedio de las observaciones para el j -ésimo nivel de B y el k -ésimo nivel de C
 \bar{y}_{ijk} = promedio de las observaciones para la ijk -ésima combinación de tratamientos.

En la tabla 14.5 se resumen los cálculos en una tabla de análisis de varianza para un problema de tres factores con n corridas replicadas en cada combinación de factor.

Tabla 14.5 Análisis de varianza para un experimento de tres factores con n réplicas

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Efecto principal:				
A	SSA	$a - 1$	s_a^2	$f_a = \frac{s_a^2}{s^2}$
B	SSB	$b - 1$	s_b^2	$f_b = \frac{s_b^2}{s^2}$
C	SSC	$c - 1$	s_c^2	$f_c = \frac{s_c^2}{s^2}$
Interacción de dos factores:				
AB	$SS(AB)$	$(a - 1)(b - 1)$	s_{ab}^2	$f_{ab} = \frac{s_{ab}^2}{s^2}$
AC	$SS(AC)$	$(a - 1)(c - 1)$	s_{ac}^2	$f_{ac} = \frac{s_{ac}^2}{s^2}$
BC	$SS(BC)$	$(b - 1)(c - 1)$	s_{bc}^2	$f_{bc} = \frac{s_{bc}^2}{s^2}$
Interacción de tres factores:				
ABC	$SS(ABC)$	$(a - 1)(b - 1)(c - 1)$	s_{abc}^2	$f_{abc} = \frac{s_{abc}^2}{s^2}$
Error	SSE	$abc(n - 1)$	s^2	
Total	SST	$abcn - 1$		

Para el experimento de tres factores con una sola corrida experimental por combinación podemos utilizar el análisis de la tabla 14.5 al hacer $n = 1$ y usar la suma de cuadrados de la interacción ABC para SSE . En este caso suponemos que todos los efectos de interacción $(\alpha\beta\gamma)_{ijk}$ son iguales a cero por lo que

$$E\left[\frac{SS(ABC)}{(a-1)(b-1)(c-1)}\right] = \sigma^2 + \frac{n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c (\alpha\beta\gamma)_{ijk}^2}{(a-1)(b-1)(c-1)} = \sigma^2.$$

Es decir, $SS(ABC)$ representa la variación debida sólo al error experimental. Por eso su cuadrado medio proporciona una estimación insesgada de la varianza del error. Con $n = 1$ y $SSE = SS(ABC)$, la suma de cuadrados del error se encuentra al res-

tar las sumas de los cuadrados de los efectos principales y las interacciones de dos factores de la suma total de cuadrados.

Ejemplo 14.3 En la producción de un material particular son de interés tres variables: *A*, el efecto del operador (tres operadores); *B*, el catalizador que se utiliza en el experimento (tres catalizadores); y *C*, el tiempo de lavado del producto que sigue al proceso de enfriado (15 minutos y 20 minutos). Se realizan tres corridas en cada combinación de factores. Se considera que se deben estudiar todas las interacciones entre los factores. Los resultados codificados son los siguientes:

A (operador)	Tiempo de lavado, C					
	15 minutos B (catalizador)			20 minutos B (catalizador)		
	1	2	3	1	2	3
1	10.7	10.3	11.2	10.9	10.5	12.2
	10.8	10.2	11.6	12.1	11.1	11.7
	11.3	10.5	12.0	11.5	10.3	11.0
2	11.4	10.2	10.7	9.8	12.6	10.8
	11.8	10.9	10.5	11.3	7.5	10.2
	11.5	10.5	10.2	10.9	9.9	11.5
3	13.6	12.0	11.1	10.7	10.2	11.9
	14.1	11.6	13.0	11.7	11.5	11.6
	14.5	11.5	11.5	12.7	10.9	12.2

Haga un análisis de varianza para probar los efectos significativos.

SOLUCIÓN

La figura 14.4 de la página 544 muestra un impreso de computadora para el análisis de varianza de los datos del ejemplo 14.3. Se utiliza el paquete SAS PROC GLM. Ninguna de las interacciones muestra un efecto significativo en el nivel $\alpha = 0.05$. Sin embargo, el valor *P* para *BC* es 0.0610. Por ello no se debe ignorar. Los efectos del operador y del catalizador son significativos, mientras que el efecto del tiempo de lavado no es significativo. ■

Impacto de la interacción

Se debe discutir más el ejemplo 14.3, en particular al tratar con el efecto que tiene la interacción entre el catalizador y el tiempo de lavado en la prueba sobre el efecto principal del tiempo de lavado (factor *C*). Recuerde nuestra discusión en la sección 14.2. Se dieron ejemplos de cómo la presencia de la interacción puede cambiar la interpretación que hacemos con respecto a los efectos principales. En el ejemplo 14.3 la interacción *BC* es significativa en el nivel 0.06. Suponga, sin embargo, que observamos una tabla de medias en dos direcciones.

	Tiempo de lavado C	
	15 min.	20 min.
1	12.19	11.29
Catalizador, B 2	10.86	10.50
3	11.09	11.46
Medias	11.38	11.08

General Linear Models Procedure					
Class Level Information					
Class	Levels	Values			
A	3	1 2 3			
B	3	1 2 3			
C	2	1 2			
Number of observations in data set = 54					
General Linear Models Procedure					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	17	41.579259	2.445839	4.07	0.0002
Error	36	21.613333	0.600370		
Corrected Total	53	63.192593			
Root MSE					
0.7746					
V. Mean					
11.27613					
Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	13.982593	6.991296	11.64	0.0001
B	2	10.182593	5.091296	8.68	0.0010
C	1	1.185185	1.185185	1.97	0.1666
A*B	4	4.774074	1.193519	1.99	0.1172
A*C	2	2.933704	1.466852	2.43	0.1027
B*C	2	3.633704	1.816852	3.03	0.0410
A*B*C	4	4.907407	1.226852	2.04	0.1049
Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	2	13.982593	6.991296	11.64	0.0001
B	2	10.182593	5.091296	8.68	0.0010
C	1	1.185185	1.185185	1.97	0.1666
A*B	4	4.774074	1.193519	1.99	0.1172
A*C	2	2.933704	1.466852	2.43	0.1027
B*C	2	3.633704	1.816852	3.03	0.0410
A*B*C	4	4.907407	1.226852	2.04	0.1049

Figura 14.4 Impreso de computadora para el ejemplo 14.3.

Es claro por qué se encuentra que el tiempo de lavado no es significativo. Un analista poco cuidadoso puede obtener la impresión que el tiempo de lavado se puede eliminar de cualquier estudio futuro en que se mida la producción. Sin embargo, es obvio cómo cambia el efecto del tiempo de lavado de un efecto negativo para el primer catalizador al que parece ser un efecto positivo para el tercer catalizador. Si solamente nos fijamos en los datos para el catalizador 1, una simple comparación entre las medias en los dos tiempos de lavado producirá una estadística t simple:

$$t = \frac{12.19 - 11.29}{\sqrt{0.6(2/9)}} = 2.5,$$

que es significativa en niveles menores que 0.02. Por ello se puede ignorar por completo un efecto negativo importante del tiempo de lavado para el catalizador 1 si el analista hace la interpretación amplia incorrecta de la razón F insignificante sobre el tiempo de lavado.

14.6 Modelos multifactoriales específicos

Describimos el modelo de tres factores y su análisis en la forma más general mediante la inclusión de todas las interacciones posibles en el modelo. Por supuesto, hay

muchas situaciones donde se sabe *a priori* que el modelo no debe contener ciertas interacciones. Podemos entonces tomar ventaja de este conocimiento al combinar o reunir las sumas de cuadrados que corresponden a interacciones insignificantes con las sumas de cuadrados del error para formar un nuevo estimador para σ^2 con un mayor número de grados de libertad. Por ejemplo, en un experimento metalúrgico diseñado para estudiar el efecto sobre el espesor de la película de tres variables importantes del proceso, suponga que se sabe que el factor *A*, concentración ácida, no interactúa con los factores *B* y *C*. Las sumas de cuadrados *SSA*, *SSB*, *SSC* y *SS(BC)* se calculan con el uso de los métodos que se describen en la sección 14.5. Todos los cuadrados medios para los efectos restantes estimarán ahora de manera independiente la varianza del error σ^2 . Por tanto, formamos nuestro nuevo cuadrado medio del error al combinar *SS(AB)*, *SS(AC)*, *SS(ABC)* y *SSE*, junto con los correspondientes grados de libertad. El denominador que resulta para las pruebas de significancia es entonces el cuadrado medio del error dado por

$$s^2 = \frac{SS(AB) + SS(AC) + SS(ABC) + SSE}{(a-1)(b-1) + (a-1)(c-1) + (a-1)(b-1)(c-1) + abc(n-1)}$$

Del cálculo, por supuesto, se obtiene la suma combinada de cuadrados y los grados de libertad combinados por sustracción una vez que se calculan *SST* y las sumas de cuadrados para los efectos existentes. La tabla de análisis de varianza tomaría entonces la forma de la tabla 14.6.

Tabla 14.6 Análisis de varianza con el factor *A* sin interacción

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	<i>f</i> calculada
Efecto principal:				
<i>A</i>	<i>SSA</i>	$a - 1$	s_A^2	$f_1 = \frac{s_A^2}{s^2}$
<i>B</i>	<i>SSB</i>	$b - 1$	s_B^2	$f_2 = \frac{s_B^2}{s^2}$
<i>C</i>	<i>SSC</i>	$c - 1$	s_C^2	$f_3 = \frac{s_C^2}{s^2}$
Interacción de dos factores:				
<i>BC</i>	<i>SS(BC)</i>	$(b-1)(c-1)$	s_{BC}^2	$f_4 = \frac{s_{BC}^2}{s^2}$
Error	<i>SSE</i>	<u>Sustracción</u>	s^2	
Total	<i>SST</i>	$abcn - 1$		

En nuestro análisis del experimento de dos factores en la sección 14.3 se utilizó un diseño completamente aleatorizado. Al interpretar los niveles del factor *A* en la tabla 14.6 como bloques diferentes, tenemos entonces el procedimiento de análisis de varianza para un experimento de dos factores en un diseño por bloques aleatorizados. Por ejemplo, si interpretamos los operadores en el ejemplo 14.3 como bloques y no suponemos interacción alguna entre los bloques y los otros dos factores, el análisis de varianza toma la forma de la tabla 14.7 en lugar de la de la tabla 14.6. El lector puede verificar que el cuadrado medio del error también es

$$s^2 = \frac{4.78 + 2.92 + 4.89 + 21.62}{4 + 2 + 4 + 36} = 0.74,$$

lo que demuestra la combinación de las sumas de cuadrados para los efectos de interacción inexistentes.

Tabla 14.7. Análisis de varianza para un experimento de dos factores en un diseño de bloques aleatorizado

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Bloques	13.98	2	6.99	
Efecto principal:				
<i>B</i>	10.18	2	5.09	6.88
<i>C</i>	1.18	1	1.18	1.59
Interacción de dos factores:				
<i>BC</i>	3.64	2	1.82	2.46
Error	34.21	46	0.74	
Total	63.19	53		

Ejercicios

1. Los siguientes datos se toman en un estudio que incluye tres factores *A*, *B* y *C*, todos efectos fijos.

	<i>C</i> ₁			<i>C</i> ₂			<i>C</i> ₃		
	<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₃	<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₃	<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₃
<i>A</i> ₁	15.0	14.8	15.9	16.8	14.2	13.2	15.8	15.5	19.2
	18.5	13.6	14.8	15.4	12.9	11.6	14.3	13.7	13.5
	22.1	12.2	13.6	14.3	13.0	10.1	13.0	12.6	11.1
<i>A</i> ₂	11.3	17.2	16.1	18.9	15.4	12.4	12.7	17.3	7.8
	14.6	15.5	14.7	17.3	17.0	13.6	14.2	15.8	11.5
	18.2	14.2	13.4	16.1	18.6	15.2	15.9	14.6	12.2

- (a) Realice pruebas de significancia sobre todas las interacciones al nivel $\alpha = 0.05$.
 (b) Realice pruebas de significancia sobre los efectos principales en el nivel $\alpha = 0.05$.
 (c) Proporcione una explicación de cómo una interacción significativa encubre el efecto del factor *C*.

2. El método de fluorescencia de rayos X es una herramienta analítica importante para determinar la concentración de material en los propulsores sólidos de misiles. En el artículo *An X-ray Fluorescence Method for Analyzing Polybutadiene Acrylic Acid (PBAA) Propellants*, Quarterly Report, RK-TR-62-1, Army Ordnance Missile Command (1962), se postula que el proceso de mezcla del propulsor y el tiempo del análisis tienen una influencia en la homogeneidad del material y por ello sobre la precisión de las mediciones de la intensidad de rayos X. Se llevó a cabo un

experimento con el uso de tres factores: *A*, condiciones de mezclado (4 niveles); *B*, el tiempo del análisis (2 niveles); y *C*, el método de carga del propulsor en los soportes de la muestra (caliente y temperatura ambiente). Se registraron los siguientes datos, que representan el análisis en porcentaje de peso de peróxido de amonio en un propulsor particular.

<i>A</i>	Método de carga, <i>C</i>					
	Caliente		Temperatura ambiente			
	<i>B</i> ₁	<i>B</i> ₂	<i>B</i> ₁		<i>B</i> ₂	
1	38.62	38.45	39.82	39.82	37.20	38.64
	38.02	38.75	39.78	39.72	37.57	37.75
2	37.67	37.81	39.53	39.56	37.85	37.91
	37.57	37.75	39.76	39.25	37.51	37.21
3	37.74	37.42	39.60	39.49	37.58	37.79
	37.58	37.79	39.62	39.45	37.52	37.60
4	37.15	37.55	39.63	39.38	37.51	37.91
	37.51	37.91	39.67	39.00		

Realice un análisis de varianza con $\alpha = 0.01$ para probar la significancia de los efectos principal y de interacción.

3. La fatiga por corrosión en metales se define como la acción simultánea de tensión cíclica y ataque químico sobre una estructura metálica. En el estudio *Effect of Humidity and Several Surface Coatings on the Fatigue Life of 2024-T351 Aluminum Alloy* realizado por el Departamento de Ingeniería Mecánica del Instituto Politécnico y Universidad Estatal de Virginia, se utilizó una técnica que implica la aplicación de un recubrimiento protector de cromo para minimizar el daño de fatiga por corrosión en aluminio. Se utilizaron tres factores en la investigación con cinco réplicas para cada combinación de tratamientos: recubrimiento, en dos niveles, y humedad y esfuerzo cortante, ambos con tres niveles. Los datos de fatiga, que se registran en miles de ciclos antes de la falla, son los siguientes:

		Esfuerzo cortante (psi)			
Recubrimiento	Humedad	13,000	17,000	20,000	
Aluminio sin recubrimiento:	Baja: (20-25% RH)	4,580	5,252	361	
		10,126	897	466	
		1,341	1,465	1,069	
	Media: (50-60% RH)	6,414	2,694	469	
		3,549	1,017	937	
		2,858	799	314	
	Alta: (86-91% RH)	8,829	3,471	244	
		10,914	685	261	
		4,067	810	522	
			2,595	3,409	739
	Aluminio cromado:	Baja: (20-25% RH)	6,489	1,862	1,344
			5,248	2,710	1,027
6,816			2,632	663	
Media: (50-60% RH)		5,860	2,131	1,216	
		5,901	2,470	1,097	
		5,395	4,035	130	
Alta: (86-91% RH)		2,768	2,022	841	
		1,821	914	1,595	
		3,604	2,036	1,482	
		4,106	3,524	529	
Media: (50-60% RH)		4,833	1,847	252	
		7,414	1,684	105	
	10,022	3,042	847		
Alta: (86-91% RH)	7,463	4,482	874		
	21,906	996	755		
	3,287	1,319	586		
		5,200	929	402	
		5,493	1,263	846	
		4,145	2,236	524	
		3,336	1,392	751	

(a) Realice un análisis de varianza con $\alpha = 0.01$ para probar la significancia de los efectos principales y de interacción.

(b) Haga una recomendación para combinaciones de los tres factores que resulten en bajo daño por fatiga.

4. Considere una situación experimental que incluya los factores A , B y C , donde suponemos un modelo de efectos fijos en tres direcciones de la forma

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijk}$$

Todas las demás interacciones se consideran inexistentes o insignificantes. Los datos se registran a continuación:

	B_1			B_2		
	C_1	C_2	C_3	C_1	C_2	C_3
A_1 :	4.0	3.4	3.9	4.4	3.1	3.1
	4.9	4.1	4.3	3.4	3.5	3.7
A_2 :	3.6	2.8	3.1	2.7	2.9	3.7
	3.9	3.2	3.5	3.0	3.2	4.2
A_3 :	4.8	3.3	3.6	3.6	2.9	2.9
	3.7	3.8	4.2	3.8	3.3	3.5
A_4 :	3.6	3.2	3.2	2.2	2.9	3.6
	3.9	2.8	3.4	3.5	3.2	4.3

(a) Realice una prueba de significancia de la interacción BC en el nivel $\alpha = 0.05$.

(b) Realice pruebas de significancia de los efectos principales A , B y C con el uso de un cuadrado medio del error combinado en el nivel $\alpha = 0.05$.

5. Las copiatoras electrónicas hacen copias al fijar tinta negra en papel, con el uso de electricidad estática. Calentar y fijar la tinta en el papel constituye la etapa final del proceso de copiado. La potencia del fijado durante este proceso final determina la calidad de la copia. Se establece que la temperatura, el estado de la superficie del rodillo de fijación y la dureza del rodillo de presión influyen en la potencia de fijación de la copiadora. Se realiza un experimento con tratamientos que consisten en una combinación de estos tres factores en cada uno de los tres niveles. Los datos en la siguiente página muestran la potencia de fijación para cada combinación de tratamientos. Realice un análisis de varianza con $\alpha = 0.05$ para probar la significancia de los efectos principales y de interacción.

6. Para un estudio de la dureza de rellenos dentales de oro, se asignan combinaciones de tres métodos de condensación y dos tipos de oro a cinco dentistas elegidos al azar. Se mide la dureza. [Véase Haaglin, Masteller y Tukey (1991).]

Temperatura	Estado de la superficie del rodillo de fijación	Dureza del rodillo de presión (hrc)		
		20	40	60
Temperatura baja:	Suave:	0.52	0.54	0.60
		0.44	0.52	0.55
		0.57	0.65	0.78
		0.53	0.56	0.68
		0.64	0.79	0.49
	Media:	0.59	0.73	0.48
		0.58	0.79	0.74
		0.64	0.78	0.50
		0.67	0.58	0.55
		0.77	0.68	0.65
	Dura:	0.74	0.57	0.57
		0.65	0.59	0.58
		0.46	0.31	0.56
		0.40	0.49	0.42
		0.58	0.48	0.49
Temperatura media:	Suave:	0.37	0.66	0.49
		0.60	0.66	0.64
		0.43	0.57	0.54
		0.62	0.72	0.74
		0.61	0.56	0.56
	Media:	0.53	0.53	0.56
		0.65	0.45	0.66
		0.66	0.59	0.71
		0.56	0.47	0.67
		0.52	0.54	0.65
	Dura:	0.44	0.52	0.49
		0.57	0.65	0.65
		0.53	0.56	0.52
		0.53	0.53	0.49
		0.65	0.45	0.48
Temperatura alta:	Suave:	0.66	0.59	0.74
		0.56	0.47	0.50
		0.43	0.48	0.55
		0.43	0.31	0.65
		0.47	0.43	0.57
	Media:	0.44	0.27	0.58
		0.43	0.48	0.55
		0.43	0.31	0.65
		0.47	0.43	0.57
		0.44	0.27	0.58

Haga que el dentista juegue el papel de bloque.

(a) Establezca el modelo apropiado con las suposiciones.

(b) ¿Existe una interacción significativa entre el método de condensación y el tipo de material de relleno de oro?

(c) ¿Existe un método de condensación que parezca ser el mejor? Explique. Los datos son los siguientes:

Bloque de dentista	Método	Tipo	
		Hoja de oro	Goldent
1	1	792	824
	2	772	772
	3	782	803
2	1	803	803
	2	752	772
	3	715	707
3	1	715	724
	2	792	715
	3	762	606
4	1	673	946
	2	657	743
	3	690	245
5	1	634	715
	2	649	724
	3	724	627

7. Considere las combinaciones de tres factores en la eliminación de mugre de cargas estándar de lavandería. El primer factor es la marca del detergente, *X, Y* o *Z*. El segundo factor es el tipo de detergente, líquido o en polvo. El tercer factor es la temperatura del agua, caliente o tibia. El experimento se replica tres veces. La respuesta es el porcentaje de eliminación de mugre. Los datos son los siguientes:

Marca	Tipo	Temperatura		
X	Polvo	Caliente	85	88
		Tibia	82	83
	Líquido	Caliente	78	75
		Tibia	75	73
Y	Polvo	Caliente	90	92
		Tibia	88	89
	Líquido	Caliente	78	76
		Tibia	76	77
Z	Polvo	Caliente	85	87
		Tibia	76	74
	Líquido	Caliente	60	70
		Tibia	55	57

(a) ¿Existen efectos de interacción significativos en el nivel $\alpha = 0.05$?

(b) ¿Hay diferencias significativas entre las marcas de detergente?

(c) ¿Cuál combinación de factores preferiría utilizar?

14.7 Experimentos factoriales modelos II y III

En un experimento de dos factores con efectos aleatorios tenemos el modelo II:

$$Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + E_{ijk},$$

$$i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n,$$

donde A_i , B_j , $(AB)_{ij}$, y E_{ijk} son variables aleatorias independientes con medias cero y varianzas σ_a^2 , σ_b^2 , σ_{ab}^2 , y σ^2 , respectivamente. La suma de cuadrados para experimentos del modelo II se calculan exactamente de la misma forma que para experimentos del modelo I. Nos interesa ahora probar hipótesis de la forma

$$H_0: \sigma_a^2 = 0, \quad H_0': \sigma_b^2 = 0, \quad H_0'': \sigma_{ab}^2 = 0,$$

$$H_1: \sigma_a^2 \neq 0, \quad H_1': \sigma_b^2 \neq 0, \quad H_1'': \sigma_{ab}^2 \neq 0,$$

donde el denominador en la razón f no necesariamente es el cuadrado medio del error. El denominador apropiado se puede determinar al examinar los valores esperados de los diversos cuadrados medios. Éstos se muestran en la tabla 14.8.

Tabla 14.8 Cuadrados medios esperados para un experimento de dos factores del modelo II

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
A	$a - 1$	s_a^2	$\sigma^2 + n\sigma_a^2 + bn\sigma_b^2$
B	$b - 1$	s_b^2	$\sigma^2 + n\sigma_a^2 + an\sigma_b^2$
AB	$(a - 1)(b - 1)$	s_{ab}^2	$\sigma^2 + n\sigma_{ab}^2$
Error	$ab(n - 1)$	s^2	σ^2
Total	$abn - 1$		

De la tabla 14.8 vemos que H_0 y H_0' se prueban con el uso de s_a^2 en el denominador de la razón f , mientras que H_0'' se prueba con el uso de s^2 en el denominador. Las estimaciones insesgadas de los componentes de la varianza son

$$\hat{\sigma}^2 = s^2, \quad \hat{\sigma}_{ab}^2 = \frac{s_{ab}^2 - s^2}{n}, \quad \hat{\sigma}_a^2 = \frac{s_a^2 - s^2}{bn}, \quad \hat{\sigma}_b^2 = \frac{s_b^2 - s^2}{an}.$$

Los cuadrados medios esperados para el experimento de tres factores con efectos aleatorios en un diseño completamente aleatorizado se muestran en la tabla 14.9. Es evidente de los cuadrados medios esperados de la tabla 14.9 que se pueden formar razones f apropiadas para probar todos los componentes de la varianza de la interacción de dos y tres factores. Sin embargo, para probar una hipótesis de la forma

$$H_0: \sigma_a^2 = 0$$

$$H_1: \sigma_a^2 \neq 0,$$

parece no haber una razón f apropiada a menos que encontremos uno o más de los componentes de la varianza de interacción de dos factores no significativos. Suponga, por ejemplo, que comparamos s_{ab}^2 con s^2 y encontramos que σ_{ab}^2 es insignificante. Podemos argumentar entonces que el término σ_a^2 se debe eliminar de todos los

Tabla 14.9 Cuadrados medios esperados para un experimento de tres factores modelo II

Fuente de variación	Grados de libertad	Cuadrado medio	Cuadrado medio esperado
A	$a - 1$	s_1^2	$\sigma^2 + n\sigma_{ab}^2 + cn\sigma_{ac}^2 + bn\sigma_{bc}^2 + bcn\sigma_{abc}^2$
B	$b - 1$	s_2^2	$\sigma^2 + n\sigma_{ab}^2 + cn\sigma_{ac}^2 + an\sigma_{bc}^2 + acn\sigma_{abc}^2$
C	$c - 1$	s_3^2	$\sigma^2 + n\sigma_{ab}^2 + bn\sigma_{bc}^2 + an\sigma_{ac}^2 + abn\sigma_{abc}^2$
AB	$(a - 1)(b - 1)$	s_4^2	$\sigma^2 + n\sigma_{ab}^2 + cn\sigma_{ac}^2$
AC	$(a - 1)(c - 1)$	s_5^2	$\sigma^2 + n\sigma_{ac}^2 + bn\sigma_{bc}^2$
BC	$(b - 1)(c - 1)$	s_6^2	$\sigma^2 + n\sigma_{bc}^2 + an\sigma_{ac}^2$
ABC	$(a - 1)(b - 1)(c - 1)$	s_7^2	$\sigma^2 + n\sigma_{abc}^2$
Error	$abc(n - 1)$	s^2	σ^2
Total	$abcn - 1$		

cuadrados medios esperados de la tabla 14.9; entonces la razón s_1^2/s_4^2 proporciona una prueba de la significancia del componente σ_{ab}^2 de la varianza. Por tanto, si probamos las hipótesis con respecto a los componentes de la varianza de los efectos principales, es necesario primero investigar la significancia de los componentes de interacción de dos factores. Una prueba aproximada derivada por Satterthwaite se puede utilizar cuando se encuentra que ciertos componentes de la varianza de la interacción de dos factores son significativos y por ello deben permanecer como parte del cuadrado medio esperado.

Ejemplo 14.4 En un estudio para determinar cuáles son las fuentes importantes de variación en un proceso industrial, se toman tres medidas de la producción de tres operadores que se eligen al azar y cuatro lotes de materia prima elegidos aleatoriamente. Se decide que se debe realizar una prueba de significancia en el nivel 0.05 para determinar si los componentes de la varianza debidos a los lotes, operadores e interacciones son significativos. Además, se deben calcular estimaciones de los componentes de la varianza. Los datos son los siguientes; la respuesta está en porcentaje del peso:

Operador	Lote:	1	2	3	4
1		66.9	68.3	69.0	69.3
		68.1	67.4	69.8	70.9
		67.2	67.7	67.5	71.4
2		66.3	68.1	69.7	69.4
		65.4	66.9	68.8	69.6
		65.8	67.6	69.2	70.0
3		65.6	66.0	67.1	67.9
		66.3	66.9	66.2	68.4
		65.2	67.3	67.4	68.7

SOLUCIÓN

Las sumas de cuadrados se encuentran de la manera acostumbrada, con los siguientes resultados:

$$SST(\text{total}) = 84.5564, \quad SSA(\text{operators}) = 18.2106$$

$$SSB (\text{lotes}) = 50.1564.$$

$$SS(AB) (\text{interacción}) = 5.5161.$$

$$SSE (\text{error}) = 10.6733.$$

Todos los demás cálculos se realizan y muestran en la tabla 14.10. Como

$$f_{0.05}(2, 6) = 5.14, \quad f_{0.05}(3, 6) = 4.76, \quad \text{y} \quad f_{0.05}(6, 24) = 2.51.$$

encontramos que los componentes de varianza de operador y lote son significativos. Aunque la varianza de interacción no es significativa en el nivel $\alpha = 0.05$, el valor F es 0.095. Las estimaciones de los componentes de la varianza del efecto principal son

$$\hat{\sigma}_a^2 = \frac{9.1053 - 0.9194}{12} = 0.68, \quad \hat{\sigma}_b^2 = \frac{16.7188 - 0.9144}{9} = 1.76.$$

Tabla 14.10 Análisis de varianza para el ejemplo 14.4

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Operadores	18.2106	2	9.1053	9.90
Lotes	50.1564	3	16.7188	18.18
Interacción	5.5161	6	0.9194	2.07
Error	10.6733	24	0.4447	
Total	84.5564	35		

Experimento modelo III (modelo mezclado)

Existen situaciones donde el experimento dicta la suposición de un **modelo mezclado**, es decir, una mezcla de efectos aleatorios y fijos. Por ejemplo, para el caso de dos factores podemos tener

$$Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + E_{ijk}$$

$$i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, n.$$

Las A_i pueden ser variables aleatorias independientes, independientes de E_{ijk} , y las B_j pueden ser efectos fijos. La naturaleza mezclada del modelo requiere que los términos de interacción sean variables aleatorias. Como resultado, las hipótesis relevantes son de la forma

$$H_0^a: \sigma_a^2 = 0, \quad H_0^b: B_j = 0, \quad i = 1, 2, \dots, a, \quad H_0^{ab}: \sigma_{ab}^2 = 0,$$

$$H_1^a: \sigma_a^2 \neq 0, \quad H_1^b: B_j \neq 0, \quad H_1^{ab}: \sigma_{ab}^2 \neq 0.$$

De nuevo, los cálculos de la suma de cuadrados son idénticos a los de las situaciones fijas y del modelo II, y la prueba f está dictada por los cuadrados medios esperados. La tabla 14.11 de la página 552 proporciona los cuadrados medios esperados para el problema de dos factores modelo II.

De la naturaleza de los cuadrados medios esperados es claro que la **prueba sobre el efecto aleatorio emplea el cuadrado medio del error s^2** como denominador, mientras que la **prueba sobre el efecto fijo** utiliza el cuadrado medio de la interacción. Suponga que consideramos ahora tres factores. Aquí, por supuesto debemos

Tabla 14.11 Cuadrados medios esperados para un experimento de dos factores modelo III

Factor	Cuadrado medio esperado
A (aleatorio)	$\sigma^2 + bn\sigma_{\alpha}^2$
B (fijo)	$\sigma^2 + n\sigma_{\beta}^2 + \frac{n\sum B_i^2}{b-1}$
AB	$\sigma^2 + n\sigma_{\alpha\beta}^2$
Error	σ^2

Tabla 14.12 Cuadrados medios esperados para experimentos factoriales de tres factores modelo III

	A aleatoria	A aleatoria, B aleatoria
A	$\sigma^2 + bn\sigma_{\alpha}^2$	$\sigma^2 + bn\sigma_{\alpha}^2 + bn\sigma_{\alpha}^2$
B	$\sigma^2 + cn\sigma_{\beta}^2 + acn \sum_{j=1}^c \beta_j^2 / (b-1)$	$\sigma^2 + cn\sigma_{\beta}^2 + acn\sigma_{\beta}^2$
C	$\sigma^2 + bn\sigma_{\gamma}^2 + abn \sum_{i=1}^c \frac{C_i^2}{(c-1)}$	$\sigma^2 + n\sigma_{\alpha\gamma}^2 + an\sigma_{\beta\gamma}^2 + bn\sigma_{\gamma}^2 + abn \sum_{i=1}^c \frac{C_i}{(c-1)}$
AB	$\sigma^2 + cn\sigma_{\alpha\beta}^2$	$\sigma^2 + cn\sigma_{\alpha\beta}^2$
AC	$\sigma^2 + bn\sigma_{\alpha\gamma}^2$	$\sigma^2 + n\sigma_{\alpha\gamma}^2 + bn\sigma_{\gamma}^2$
BC	$\sigma^2 + n\sigma_{\beta\gamma}^2 + an \sum_{j=1}^c \sum_{k=1}^c \frac{(BC)_{jk}^2}{(b-1)(c-1)}$	$\sigma^2 + n\sigma_{\beta\gamma}^2 + an\sigma_{\beta\gamma}^2$
ABC	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$
Error	σ^2	σ^2

tomar en cuenta la situación donde un factor es fijo y la situación en la que dos factores son fijos. La tabla 14.12 cubre ambas situaciones.

Nótese que en el caso de A aleatoria, todos los efectos tienen pruebas apropiadas. Pero en el caso de A y B aleatorias, el efecto principal C se debe probar con el uso de un procedimiento tipo Satterthwaite similar al experimento modelo II.

14.8 Elección del tamaño de la muestra

Nuestro estudio de experimentos factoriales a lo largo de este capítulo se restringe al uso de un diseño completamente aleatorizado con la excepción de la sección 14.6 donde demostramos el análisis de un experimento de dos factores en un diseño de bloques aleatorizado. El diseño completamente aleatorizado es fácil de preparar y el análisis es simple de realizar; sin embargo, se debe utilizar sólo cuando el número de combinaciones de tratamiento es pequeña y el material experimental es homogéneo. Aunque el diseño por bloques aleatorizado es ideal para dividir un grupo grande de unidades heterogéneas en subgrupos de unidades homogéneas, por lo general es difícil obtener bloques uniformes con suficientes unidades a las que se pueda asignar un número grande de combinaciones de tratamiento. Esta desventaja se puede evitar al elegir un diseño del catálogo de **diseños incompletos de bloque**. Estos diseños permiten investigar diferencias entre t tratamientos acomodados en b bloques, cada uno con k unidades experimentales, donde $k < t$. El lector puede consultar a Box, Hunter y Hunter para más detalles.

Una vez que se selecciona un diseño completamente aleatorizado se debe decidir si el número de réplicas es suficiente para producir pruebas en el análisis de varianza con potencia alta. Si no, se deben agregar réplicas adicionales, que a su vez pueden necesitar un diseño completo de bloques aleatorizados. Si se comienza con un diseño de bloques aleatorizado, sería necesario aún determinar si el número de bloques es suficiente para producir pruebas de potencia. Básicamente, entonces, regresamos a la cuestión del tamaño de la muestra.

La potencia de una prueba de efectos fijos para un tamaño de muestra dado se encuentra de la tabla A.15 al calcular el parámetro de no centralidad λ y la función ϕ que se presenta en la sección 13.14. En la tabla 14.13 se dan expresiones para λ y ϕ^2 para experimentos de efectos fijos de dos y de tres factores.

Tabla 14.13 Parámetro de no centralidad λ y ϕ^2 para modelos de dos y tres factores

	Experimentos de dos factores		Experimentos de tres factores		
	A	B	A	B	C
λ :	$\frac{bn \sum_{i=1}^a \alpha_i^2}{2\sigma^2}$	$\frac{an \sum_{j=1}^b \beta_j^2}{2\sigma^2}$	$\frac{bcn \sum_{i=1}^a \alpha_i^2}{2\sigma^2}$	$\frac{acn \sum_{j=1}^b \beta_j^2}{2\sigma^2}$	$\frac{abn \sum_{k=1}^c \gamma_k^2}{2\sigma^2}$
ϕ^2 :	$\frac{bn \sum_{i=1}^a \alpha_i^2}{a\sigma^2}$	$\frac{an \sum_{j=1}^b \beta_j^2}{b\sigma^2}$	$\frac{bcn \sum_{i=1}^a \alpha_i^2}{a\sigma^2}$	$\frac{acn \sum_{j=1}^b \beta_j^2}{b\sigma^2}$	$\frac{abn \sum_{k=1}^c \gamma_k^2}{c\sigma^2}$

Los resultados de la sección 13.14 para el modelo de efectos aleatorios se puede extender con facilidad a los modelos de dos y tres factores. Una vez más el procedimiento general se basa en los valores de los cuadrados medios esperados. Por ejemplo, si probamos $\sigma_a^2 = 0$ en un experimento de dos factores mediante el cálculo de la razón s_1^2/s_2^2 (véase la tabla 14.8), entonces

$$f = \frac{s_1^2/(\sigma^2 + n\sigma_{a0}^2 + bn\sigma_a^2)}{s_2^2/(\sigma^2 + n\sigma_{a0}^2)}$$

es un valor de la variable aleatoria F que tiene la distribución F con $a - 1$ y $(a - 1)(b - 1)$ grados de libertad, y la potencia de la prueba es

$$\begin{aligned} 1 - \beta &= P\left\{ \frac{S_1^2}{S_2^2} > f_\alpha[(a - 1), (a - 1)(b - 1)] \text{ cuando } \sigma_a^2 \neq 0 \right\} \\ &= P\left\{ F > \frac{f_\alpha[(a - 1), (a - 1)(b - 1)](\sigma^2 + n\sigma_{a0}^2)}{\sigma^2 + n\sigma_{a0}^2 + bn\sigma_a^2} \right\} \end{aligned}$$

Ejercicios

1. Para estimar los diferentes componentes de variabilidad en un proceso de filtración, se mide el porcentaje de pérdida de material en el licor madre para 12 condiciones experimentales, tres corridas en cada

condición. Se seleccionan al azar tres filtros y cuatro operadores para el experimento, lo que tiene como resultado las siguientes mediciones:

Filtro	Operador			
	1	2	3	4
1	16.2	15.9	15.6	14.9
	16.8	15.1	15.9	15.2
	17.1	14.5	16.1	14.9
2	16.6	16.0	16.1	15.4
	16.9	16.3	16.0	14.6
	16.8	16.5	17.2	15.9
3	16.7	16.5	16.4	16.1
	16.9	16.9	17.4	15.4
	17.1	16.8	16.9	15.6

(a) Pruebe la hipótesis de componente de varianza de no interacción entre filtros y operadores en el nivel de significancia de $\alpha = 0.05$.

(b) Pruebe la hipótesis de que los operadores y los filtros no tienen efecto en la variabilidad del proceso de filtración en el nivel de significancia $\alpha = 0.05$.

(c) Estime los componentes de la varianza debida a filtros, operadores y error experimental.

2. Suponga un experimento modelo II para el ejercicio 2 de la página 539, estime los componentes de la varianza para la marca del concentrado de jugo de naranja, el número de días desde que el jugo se mezcla hasta que se prueba y para el error experimental.

3. Considere el siguiente análisis de varianza para un experimento modelo II:

Fuente de variación	Grados de libertad	Cuadrado medio
A	3	140
B	1	480
C	2	325
AB	3	15
AC	6	24
BC	2	18
ABC	6	2
Error	24	5
Total	47	

Pruebe la significancia de los componentes de la varianza entre todos los efectos principales y los efectos de interacción en el nivel de significancia de 0.01

(a) con el uso de una estimación combinada del error cuando sea apropiado;

(b) sin combinar las sumas de los cuadrados de los efectos insignificantes.

4. ¿Son suficientes dos observaciones para cada combinación de tratamientos en el ejercicio 4 de la página 547 si la potencia de nuestra prueba para detec-

tar diferencias entre los niveles del factor C en el nivel de significancia de 0.05 debe ser al menos 0.8, cuando $\gamma_1 = -0.2$, $\gamma_2 = 0.4$, y $\gamma_3 = -0.2$? Utilice la misma estimación combinada de σ^2 que se utiliza en el análisis de varianza.

5. Con el uso de las estimaciones de los componentes de la varianza en el ejercicio 1, evalúe la potencia cuando probamos el componente de la varianza debido a que los filtros son cero.

6. Un contratista de la defensa se interesa en estudiar un proceso de inspección para detectar la falla o *fatiga de partes de transformadores*. Se utilizan tres niveles de inspección por tres inspectores elegidos al azar. Se utilizan cinco lotes para cada combinación en el estudio. En los datos se dan los niveles del factor. La respuesta es en fallas por cada 1000 piezas.

(a) Escriba un modelo apropiado, con suposiciones.

(b) Utilice el análisis de varianza para probar la hipótesis apropiada para inspector, nivel de inspección e interacción.

Inspector	Nivel de inspección		
	Inspección militar completa	Especificación militar reducida	Comercial
A	7.50	7.06	6.15
	7.42	6.17	5.52
	5.85	5.65	5.44
	5.89	5.30	5.48
	5.35	5.02	5.98
B	7.58	7.68	6.17
	6.52	5.86	6.20
	6.54	5.28	5.44
	5.64	5.38	5.75
	5.12	4.87	5.68
C	7.70	7.19	6.21
	6.82	6.19	5.66
	6.42	5.85	5.36
	5.39	5.35	5.90
	5.35	5.01	6.12

7. Un fabricante de pintura de látex para casas (marca A) desea mostrar que su pintura es más robusta para el material que se pinta que sus dos competidores más cercanos. La respuesta es el tiempo, en años, hasta que se pica. El estudio incluye tres marcas de pintura y tres materiales que se eligen al azar. Se utilizan dos piezas de material para cada combinación. Los datos están en la página siguiente.

(a) ¿Cómo se llama este tipo de modelo?

(b) Analice los datos con el uso del modelo apropiado.

	Marca de pintura					
	A		B		C	
Material A	5.50	5.15	4.75	4.60	5.10	5.20
B	5.60	5.55	5.50	5.60	5.40	5.50
C	5.40	5.48	5.05	4.95	4.50	4.55

(c) ¿El fabricante de la marca A apoyó su afirmación con los datos?

8. El administrador de una planta desea mostrar que la producción de una tela tejida en su planta no depende del operador de la máquina o de la hora del día, y es consistentemente alta. Se eligen para el estudio cuatro operadores al azar y tres horas del día también al azar. La producción se mide en yardas producidas por minuto. Las muestras se toman en tres días seleccionados al azar. Los datos son los siguientes:

		Operador			
		1	2	3	4
-1	9.5	9.8	9.8	10.0	
	9.8	10.1	10.3	9.7	
	10.0	9.6	9.7	10.2	
Hora 2	10.2	10.1	10.2	10.3	
	9.9	9.8	9.8	10.1	
	9.5	9.7	9.7	9.9	
3	10.5	10.4	9.9	10.0	
	10.2	10.2	10.3	10.1	
	9.3	9.8	10.2	9.7	

Ejercicios de repaso

1. El Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia se encargó de analizar un conjunto de datos que tomó el personal en el Departamento de Nutrición Humana y Alimentos en el que fue de interés estudiar los efectos del tipo de harina y el porcentaje de endulzante sobre ciertos atributos físicos de un tipo de pastel. Se utilizaron harina común y harina para pasteles y el porcentaje de endulzante se varió en cuatro niveles. Los siguientes datos muestran información sobre la densidad específica de muestras de pastel. Se prepararon tres pasteles en cada una de las combinaciones de ocho factores.

(a) Trate el análisis como un análisis de varianza de dos factores. Pruebe para las diferencias entre el tipo de harina. Pruebe las diferencias entre la concentración de endulzante.

- (a) Escriba el modelo apropiado.
 (b) Estime los componentes de la varianza para operador y hora.
 (c) Saque conclusiones.

9. Un ingeniero de proceso quiere determinar si el ajuste de potencia de las máquinas que se utilizan para llenar cierto tipo de cajas de cereal tiene como resultado un efecto significativo sobre el peso real del producto. El estudio consiste en tres tipos de cereal elegidos de forma aleatoria fabricados por la compañía y tres ajustes de potencia fijos. El peso se mide en cuatro cajas de cereal seleccionadas al azar en cada combinación. El peso que se desea es 400 gramos. Los datos son los siguientes:

		Tipo de cereal					
		1		2		3	
Ajuste de potencia	Baja	395	390	392	392	402	405
		401	400	394	401	399	399
	Actual	396	399	390	392	404	403
		400	402	395	502	400	399
Alta		410	408	404	406	415	412
		408	407	401	400	413	415

- (a) Proporcione el modelo apropiado y liste las suposiciones que se formulan.
 (b) ¿Hay un efecto significativo debido al ajuste de la potencia?
 (c) ¿Hay un componente de la varianza significativo debido al tipo de cereal?

		Harina	
		Común	Para pastel
Concentración de endulzante	0	0.90	0.91
		0.87	0.90
		0.90	0.80
50		0.86	0.88
		0.89	0.82
		0.91	0.83
75		0.93	0.86
		0.88	0.85
		0.87	0.80
100		0.79	0.86
		0.82	0.85
		0.80	0.85

(b) Discuta el efecto de la interacción, si hay alguna. Proporcione valores F para todas las pruebas.

2. Se realizó un experimento en el Departamento de Ciencias de los Alimentos del Instituto Politécnico y Universidad Estatal de Virginia. El objetivo fue caracterizar la textura de ciertos tipo de pescado de la familia de los arenques. También se estudió el efecto de los tipos de salsa que se utilizan en la preparación del pescado. La respuesta en el experimento fue el "valor de la textura" medido con una máquina que rebanaba el pescado. Los siguientes son datos de los valores de la textura:

	Sábalo sin curar	Sábalo curado	Arenque
Crema ácida	27.6	64.0	107.0
	57.4	66.9	83.9
	47.8	66.5	110.4
	71.1	66.8	93.4
	53.8	53.8	83.1
Salsa envinada	49.8	48.3	88.0
	31.0	62.2	95.2
	11.8	54.6	108.2
	35.1	43.6	86.7
	16.1	41.8	105.2

(a) Realice un análisis de varianza. Determine si hay una interacción entre el tipo de salsa y el tipo de pescado o no.

(b) Con base en sus resultados de la parte (a) y sobre pruebas F de los efectos principales, determine si hay una diferencia en la textura debido al tipo de salsa, y determine si hay una diferencia significativa en los tipos de pescado.

3. Se realiza un estudio para determinar si las condiciones de humedad tienen un efecto en la fuerza que se requiere para romper piezas de plástico engomado. Se prueban tres tipos de plástico con el uso de cuatro diferentes niveles de humedad. Los resultados en kilogramos, son los siguientes:

Tipo de plástico	Humedad			
	30%	50%	70%	90%
A	39.0	33.1	33.8	33.0
	42.8	37.8	30.7	32.9
B	36.9	27.2	29.7	28.5
	41.0	26.8	29.1	27.9
C	27.4	29.2	26.7	30.9
	30.3	29.9	32.0	31.5

(a) Suponga un experimento modelo 1, realice un análisis de varianza y pruebe la hipótesis de no interacción entre humedad y tipo de plástico en el nivel de significancia de 0.05.

(b) Con el uso sólo de los plásticos A y B y el valor de s^2 de la parte (a), una vez más pruebe la presencia de interacción en el nivel de significancia de 0.05.

(c) Utilice una comparación de un solo grado de libertad y el valor de s^2 de la parte (a) para comparar, en el nivel de significancia de 0.05, la fuerza que se requiere con 30% de humedad contra 50%, 70% y 90% de humedad.

(d) Con el uso sólo del plástico C y el valor de s^2 de la parte (a), repita la parte (c).

4. Personal del Departamento de Ingeniería de Materiales del Instituto Politécnico y Universidad Estatal de Virginia lleva a cabo un experimento para estudiar los efectos de los factores ambientales sobre la estabilidad de cierto tipo de aleación cobre-níquel. La respuesta básica es la vida de fatiga del material. Los factores son nivel de esfuerzo y ambiente. Los datos son los siguientes:

	Nivel de esfuerzo		
	Bajo	Medio	Alto
Hidrógeno seco	11.08	13.12	14.18
	10.98	13.04	14.90
	11.24	13.37	15.10
Ambiente (95%)	10.75	12.73	14.15
	10.52	12.87	14.42
	10.45	12.95	14.25

(a) Realice un análisis de varianza para probar la interacción entre los factores. Utilice $\alpha = 0.05$.

(b) Con base en la parte (a), realice un análisis sobre los dos efectos principales y extraiga conclusiones. Utilice una aproximación de valor P para sacar conclusiones.

5. En el experimento del ejercicio 1, también el volumen del pastel se utilizó como respuesta. Las unidades son pulgadas cúbicas. Pruebe la interacción entre los factores y discuta los efectos principales. Suponga que ambos factores son efectos fijos. Los datos están en la página siguiente.

6. Una válvula de control necesita ser muy sensible al voltaje de entrada, con lo que genera un buen voltaje de salida. Un ingeniero gira las perillas de control para cambiar el voltaje de entrada. En el libro *SN:K_e*

		Harina	
		Común	Para pastel
0		4.48	4.12
		3.98	4.92
		4.42	5.10
50		3.68	5.00
		5.04	4.26
		3.72	4.34
75		3.92	4.82
		3.82	4.34
		4.06	4.40
100		3.26	4.32
		3.80	4.18
		3.40	4.30

En el libro *SN-Ratio for the Quality Evaluation* publicado por la Asociación de Estándares Japonesa (1988), se realiza un estudio sobre cómo estos tres factores (posición relativa de las perillas de control, rango de control de las perillas y voltaje de entrada) afectan la sensibilidad de una válvula de control. Los factores y sus niveles se muestran a continuación. Los datos muestran la sensibilidad de una válvula de control.

Factor A: posición relativa de las perillas de control: centro -0.5, centro y centro +0.5.

Factor A	Factor B	Factor C		
		C ₁	C ₂	C ₃
A ₁	B ₁	151	151	151
		135	135	138
	B ₂	178	180	181
		171	173	174
	B ₃	204	205	206
		190	190	192
A ₂	B ₁	156	158	158
		148	149	150
	B ₂	183	183	185
		168	170	172
	B ₃	210	211	213
		204	203	204
A ₃	B ₁	161	162	163
		145	148	148
	B ₂	189	191	192
		182	184	183
	B ₃	215	216	217
		202	203	205

Factor B: rango de control de las perillas: 2, 4.5, y 7 (mm).

Factor C: voltaje de entrada:

100, 120, y 150 (V).

Realice un análisis de varianza con $\alpha = 0.05$ para probar la significancia de los efectos principal y de interacción.

7. Para estimar el grado de suspensión de un polietileno de suspensión, extraemos polietileno mediante el uso de un solvente y comparamos el contenido de gel (proporción de gel). Este método se llama método de estimación de la proporción de gel. En el libro *Design of Experiments for the Quality Improvement* publicado por la Asociación de Estándares Japonesa (1989), se realiza un estudio sobre la relación entre la cantidad de gel generado y tres factores (solvente, temperatura de extracción y tiempo de extracción) que influyen en la cantidad de gel. Los datos son los siguientes:

Solvente	Temperatura	Tiempo		
		4	8	24
Etanol	120	94.0	93.8	91.1
		94.0	94.2	90.5
	80	95.3	94.9	92.5
		95.1	95.3	92.4
Tolueno	120	94.6	93.6	91.1
		94.5	94.1	91.0
	80	95.4	95.6	92.1
		95.4	96.0	92.1

Realice un análisis de varianza con $\alpha = 0.05$ para probar la significancia de los efectos principal y de interacción.

8. En el libro *SN-Ratio for the Quality Evaluation* publicado por la Asociación de Estándares Japonesa (1988), se realiza un estudio sobre cómo afecta la presión de aire en las llantas la capacidad de maniobra de un automóvil. Se comparan tres presiones de aire diferentes sobre tres diferentes superficies de manejo. Las tres presiones de aire son llantas izquierdas y derechas infladas a 6 kgf/cm², llantas izquierdas infladas a 6 kgf/cm², y llantas derechas infladas a 3 kgf/cm². Las tres superficies de manejo son: asfalto, asfalto seco y cemento seco. Se observa el radio de giro de un vehículo de prueba dos veces en cada nivel de presión de las llantas para cada una de las tres diferentes superficies de manejo.

	Presión de aire en las llantas					
	1	2	3	4	5	6
Asfalto	44.0	25.5	34.2	37.2	27.4	42.8
Asfalto seco	31.9	33.7	31.8	27.6	43.7	38.2
Cemento seco	27.3	39.5	46.6	28.1	35.5	34.6

Realice un análisis de varianza de los datos anteriores. Comente la interpretación de los efectos principal y de interacción.

9. El fabricante de cierta marca de café secado por congelación desea acortar el tiempo de proceso sin poner en peligro la integridad del producto. Quiere utilizar tres temperaturas para la cámara de secado y cuatro tiempos de secado. El tiempo actual de secado es tres horas a una temperatura de -15°C . La respuesta de sabor es un promedio de las calificaciones de cuatro jueces profesionales. La calificación está en

una escala de 1 a 10 donde 10 es el mejor. Los datos son los que se muestran en la tabla contigua.

(a) ¿Qué tipo de modelo se debe utilizar? Establezca las suposiciones.

(b) Analice los datos de forma apropiada.

(c) Escriba un breve reporte al vicepresidente a cargo y haga una recomendación para la fabricación futura de este producto.

	Temperatura					
	-20°C	-15°C	-10°C	-5°C	0°C	5°C
1 h	9.60	9.63	9.55	9.50	9.40	9.43
1½ h	9.75	9.73	9.60	9.61	9.55	9.48
2 h	9.82	9.93	9.81	9.78	9.50	9.52
3 h	9.78	9.81	9.80	9.75	9.55	9.58

Experimentos 2^k factoriales y fracciones

15.1 Introducción

En casi cualquier estudio experimental donde se aplican procedimientos estadísticos a un conjunto de datos científicos, los métodos implican realizar ciertas operaciones o cálculos sobre la información muestral, seguidos de la extracción de inferencias acerca de la población o poblaciones que se estudian. A menudo existen características del experimento que se sujetan al control del experimentador, cantidades como el tamaño de la muestra, el número de niveles de los factores, combinación de tratamientos a utilizar, etcétera. Estos *parámetros experimentales* con frecuencia pueden tener un gran efecto sobre la precisión con la que se prueban las hipótesis o se realiza la estimación.

Ya hemos expuesto ciertos conceptos de diseño experimental. El plan de muestreo para la prueba *t* simple de la media de una población normal y también el análisis de varianza implican la asignación aleatoria de tratamientos preseleccionados a unidades experimentales. El diseño de bloques aleatorizados, donde los tratamientos se asignan a unidades dentro de bloques relativamente homogéneos, implica una aleatorización restringida.

En este capítulo ponemos especial atención a diseños experimentales en los que el plan experimental requiere el estudio del efecto sobre una respuesta de *k* factores, cada uno en dos niveles. Estos por lo general se conocen como **experimentos 2^k**. A menudo denotamos los niveles como "alto" y "bajo", aunque esta notación puede ser arbitraria en el caso de variables cualitativas. El diseño factorial completo requiere que cada nivel de todos los factores ocurra con cada nivel de todos los demás factores, lo que da un total de 2^k combinaciones de tratamientos.

Simbolismo para las combinaciones de los factores

Denotamos los niveles más altos de los factores A, B, C, \dots con las letras a, b, c, \dots y los niveles más bajos de cada factor con la notación (1). En presencia de otras letras omitimos el símbolo (1). Por ejemplo, la combinación de tratamientos en un experimento 2² que contiene los niveles altos de los factores B y C y los niveles bajos de los factores A y D se escribe simplemente como bc . La combinación de tratamientos que consiste en el nivel bajo de todos los factores en el experimento se denota con el símbolo (1). En el caso de un experimento 2³ las ocho posibles combinaciones de tratamientos son (1), a, b, c, ab, ac, bc y abc .

El experimento factorial permite que el efecto de todos y cada uno de los factores se estime y pruebe de manera independiente por medio del análisis de varianza acostumbrado. Además, los efectos de interacción se evalúan con facilidad.

A menudo el uso de estos diseños de dos niveles es natural cuando los datos del análisis que resultan se prestan a una *aproximación por regresión*. Aquí, por supuesto, los niveles de diseño son realmente continuos y el científico se interesa en la construcción de un modelo así como en la realización de una prueba de hipótesis.

La desventaja con el experimento factorial, por supuesto, es la excesiva cantidad de experimentación que se puede requerir. Por ejemplo, si se desea estudiar el efecto de ocho variables, se requieren, $2^8 = 256$ combinaciones de los tratamientos. En muchos casos podemos obtener considerable información mediante el uso de sólo una fracción de las corridas experimentales. Este tipo de diseño se llama **diseño factorial fraccionado**.

La intención de estos diseños fraccionados es reducir el tamaño del experimento total mientras se retiene información importante relacionada con los factores.

15.2 Análisis de varianza

Considere inicialmente un plan factorial 2² en el que hay n observaciones experimentales por combinación de tratamientos. Al extender nuestra notación anterior, interpretamos ahora los símbolos (1), a, b y ab como los *resultados totales* para cada una de las cuatro combinaciones de tratamientos. La tabla 15.1 presenta una tabla de dos direcciones de estos resultados totales.

Tabla 15.1 Experimento factorial 2²

		B	Media
A	(1)	b	$\frac{b + (1)}{2n}$
	a	ab	$\frac{ab + a}{n}$
Media		$\frac{a + (1)}{2n}$	$\frac{ab + b}{2n}$

Definamos los siguientes contrastes entre los totales de los tratamientos:

$$\begin{aligned}\text{Contraste } A &= ab + a - b - (1), \\ \text{Contraste } B &= ab - a + b - (1), \\ \text{Contraste } AB &= ab - a - b + (1).\end{aligned}$$

Es evidente que habrá exactamente un contraste de un solo grado de libertad para las medias de cada factor A y B , el cual escribiremos como

Cálculos de efectos

$$w_A = \frac{ab + a - b - (1)}{2n} = \frac{\text{Contraste } A}{2n}$$

y

$$w_B = \frac{ab + b - a - (1)}{2n} = \frac{\text{Contraste } B}{2n}.$$

Se ve que la cantidad w_A es la diferencia entre la respuesta media en los niveles bajo y alto del factor A . De hecho, llamamos a w_A el **efecto principal** de A . De manera similar, w_B es el efecto principal del factor B . Se observa la interacción aparente en los datos al inspeccionar la diferencia entre $ab - b$ y $a - (1)$ o entre $ab - a$ y $b - (1)$ en la tabla 15.1. Si, por ejemplo,

$$ab - a = b - (1) \quad \text{o} \quad ab - a - b + (1) = 0,$$

una línea que conecta las respuestas para cada nivel del factor A en el nivel alto del factor B será aproximadamente paralela a una línea que conecta la respuesta para cada nivel del factor A en el nivel bajo del factor B . Las líneas no paralelas de la figura 15.1



Figura 15.1 Respuesta que sugiere una interacción aparente.

sugieren la presencia de interacción. Para probar si esta interacción aparente es significativa se construye un tercer contraste en los totales de tratamiento ortogonales a los contrastes del efecto principal, que se llama **efecto de interacción**, mediante la evaluación de

$$w_{AB} = \frac{ab - a - b + (1)}{2n} = \frac{\text{Contraste } AB}{2n}.$$

Cálculo de la suma de cuadrados

Nos aprovechamos del hecho que en el factorial 2² o en este caso en el experimento factorial 2^k general, cada efecto principal y efecto de interacción tiene asociado un solo grado de libertad. Por tanto, podemos escribir 2^k - 1 contrastes ortogonales de un solo grado de libertad en las combinaciones de tratamientos, donde cada uno representa la variación debida a algún efecto principal o de interacción. De esta manera, bajo las suposiciones acostumbradas de independencia y normalidad en el modelo experimental, podemos realizar pruebas para determinar si el contraste refleja la variación sistemática o sólo la variación casual o aleatoria. Las sumas de cuadrados para cada contraste se encuentran al seguir los procedimientos que se dan en la sección 13.5. Al escribir

$$T_1 = b + (1), \quad T_2 = ab + a, \quad c_1 = -1, \quad \text{y} \quad c_2 = 1,$$

donde T_1 y T_2 son los totales de $2n$ observaciones, tenemos

$$SSA = SSw_A = \frac{\left(\sum_{i=1}^2 c_i T_i\right)^2}{2n \sum_{i=1}^2 c_i^2} = \frac{[ab + a - b - (1)]^2}{2^2 n} = \frac{(\text{contraste } A)^2}{2^2 n}$$

con 1 grado de libertad. De manera similar, encontramos que

$$SSB = \frac{[ab + b - (1) - a]^2}{2^2 n} = \frac{(\text{contraste } B)^2}{2^2 n}$$

y

$$SS(AB) = \frac{[ab + (1) - a - b]^2}{2^2 n} = \frac{(\text{contraste } AB)^2}{2^2 n}$$

cada una con 1 grado de libertad, mientras que la suma de cuadrados de los errores, con $2^2(n - 1)$ grados de libertad, se obtiene por sustracción a partir de la fórmula

$$SSE = SST - SSA - SSB - SS(AB).$$

Al calcular las sumas de cuadrados para los efectos principales A y B y el efecto de interacción AB , es conveniente presentar los resultados totales de las combinaciones de tratamientos junto con los signos algebraicos apropiados para cada contraste, como en la tabla 15.2. Los efectos principales se obtienen como comparaciones simples entre los niveles bajo y alto. Por tanto, asignamos un signo positivo a la com-

Tabla 15.2 Signos para los contrastes en un experimento 2² factorial

Combinación de tratamientos	Efecto factorial		
	A	B	AB
(1)	-	-	+
a	+	-	-
b	-	+	-
ab	+	+	+

binación de tratamientos que está en el nivel alto de un factor dado y un signo negativo a la combinación de tratamientos en el nivel bajo. Los signos negativo y positivo para el efecto de interacción se obtienen al multiplicar los signos correspondientes de los contrastes de los factores en interacción.

Factorial 2^3

Consideremos ahora un experimento con el uso de los tres factores, A , B y C , con niveles (1), a ; (1), b ; y (1), c , respectivamente. Éste es un experimento 2^3 factorial que da las ocho combinaciones de tratamientos (1), a , b , c , ab , ac , bc , y abc . Las combinaciones de tratamientos y los signos algebraicos apropiados para cada contraste que se utilizan al calcular las sumas de cuadrados para los efectos principales y para los efectos de interacción se presentan en la tabla 15.3.

Tabla 15.3 Signos para los contrastes en un experimento factorial 2^3

Combinación de tratamientos	Efecto factorial (simbólico)						
	A	B	C	AB	AC	BC	ABC
(1)	-	-	-	+	+	+	-
a	+	-	-	-	-	+	+
b	-	+	-	+	+	-	+
c	-	-	+	+	-	-	+
ab	+	+	-	+	-	-	-
ac	+	-	+	-	+	-	-
bc	-	+	+	-	-	+	-
abc	+	+	+	+	+	+	+

Una inspección de la tabla 15.3 revela que para el experimento 2^3 cualesquiera dos contrastes entre los siete son ortogonales y por tanto los siete efectos se evalúan en forma independiente. La suma de cuadrados para, digamos, la interacción ABC con 1 grado de libertad, está dada por

$$SS(ABC) = \frac{[abc + a + b + c - (1) - ab - ac - bc]^2}{2^3 n}$$

y el efecto de interacción ABC está dado por

$$w(ABC) = \frac{abc + a + b + c - (1) - ab - ac - bc}{4n}$$

Para un experimento 2^4 factorial las sumas de cuadrados de un solo grado de libertad para los efectos principales y los efectos de interacción se obtienen al elevar al cuadrado los contrastes apropiados en los totales de los tratamientos y dividir entre $2^4 n$, donde n es el número de réplicas de las combinaciones de tratamientos.

Como antes, un efecto siempre se calcula al restar la respuesta promedio en el nivel "bajo" de la respuesta promedio en el nivel "alto". El nivel alto y bajo para los efectos principales es bastante claro. El nivel alto y bajo simbólico para las interacciones es evidente a partir de información como la de la tabla 15.3.

La propiedad de ortogonalidad tiene la misma importancia aquí como la que tiene para el material sobre comparaciones que se presenta en el capítulo 13. La ortogonalidad de los contrastes implica que los efectos estimados y por ello las sumas de cuadrados son independientes. Esta independencia se ilustra con facilidad en un experimento 2³ factorial si los resultados, con el factor A en su nivel alto, se incrementan en una cantidad x en la tabla 15.3. Sólo el contraste A conduce a una suma de cuadrados mayor, pues el efecto x anula la formación de los seis contrastes restantes como resultado de los dos signos positivos y los dos negativos asociados con las combinaciones de tratamientos en las que A está en el nivel alto.

Existen ventajas adicionales que produce la ortogonalidad. Éstas se señalarán más adelante cuando presentemos el experimento 2^k factorial en situaciones de regresión.

15.3 Experimento 2^k factorial sin réplicas

En 2^k factorial completo a menudo puede implicar una considerable experimentación, en particular cuando k es grande. Como resultado, a menudo no se permite la repetición de cada combinación de factores. Si todos los efectos, y todas las interacciones, se incluyen en el modelo del experimento, no se permiten grados de libertad para el error. A menudo, cuando k es grande, el analista de datos combinará las sumas de cuadrados y los grados de libertad correspondientes para las interacciones de orden alto que se saben o se suponen insignificantes. Esto producirá pruebas F para los efectos principales y para las interacciones de orden más bajo.

Graficación de diagnóstico con experimentos 2^k factoriales sin réplicas

La graficación de la probabilidad normal puede ser una metodología muy útil para determinar la importancia relativa de los efectos en un experimento con factores a dos niveles razonablemente grande cuando no hay réplica. Este tipo de gráfica de diagnóstico puede ser útil en particular cuando el analista de datos está indeciso en combinar interacciones de orden alto por temor de que algunos de los efectos combinados en el "error" puedan en verdad ser efectos reales y no sólo aleatorios. El lector debe tener en cuenta que todos los efectos que no son reales (es decir, que son estimaciones de cero independientes) siguen una distribución normal con media cercana a cero y varianzas constante. Por ejemplo, en un experimento 2³ factorial, todos los efectos (tenga en mente $n = 1$) son de la forma

$$w(AB) = \frac{\text{contraste}}{8} = \bar{y}_H - \bar{y}_L,$$

donde \bar{y}_H es el promedio de ocho corridas experimentales independientes en el nivel alto o "+" y \bar{y}_L es el promedio de ocho corridas independientes en el nivel bajo o "-". De esta manera la varianzas de cada contraste es $\text{Var}(\bar{y}_H - \bar{y}_L) = \sigma^2/4$. Para cualesquier efectos reales, $E(\bar{y}_H - \bar{y}_L) \neq 0$. Por ello la graficación de probabilidad normal debe revelar efectos "significativos" como los que caen fuera de la línea recta que describe las realizaciones de variables aleatorias normales independientes distribuidas de forma idéntica.

La gráfica de probabilidad puede tomar una de muchas formas. Se remite al lector al capítulo 8, donde se presentan por primera vez estas gráficas. Se puede uti-

lizar la gráfica cuantil-cuantil normal empírica. También se puede emplear el procedimiento de graficación que hace uso de la prueba de probabilidad normal. Además, hay varios tipos diferentes de gráficas de probabilidad normal de diagnóstico. En resumen, las gráficas de efecto de diagnóstico son como sigue.

Gráficas de efecto de probabilidad para experimentos 2^k factorial sin réplicas

1. Calcule los efectos como

$$\text{efecto} = \frac{\text{contraste}}{2^{k-1}}$$

2. Construya una gráfica de probabilidad normal de todos los efectos.
3. Los efectos que caen fuera de la línea recta se deben considerar como efectos reales.

Los siguientes son comentarios adicionales respecto a la graficación de probabilidad de los efectos. Primero, el analista de datos se puede sentir bastante frustrado si los utiliza en un pequeño experimento. Es probable que la graficación proporcione resultados satisfactorios cuando existe *dispersión de efectos*—muchos efectos que ciertamente no son reales. Esta dispersión será evidente en experimentos grandes donde las interacciones de orden alto probablemente no son reales.

15.4 Estudio de caso

Muchas compañías fabricantes en Estados Unidos y en el extranjero utilizan partes moldeadas como componentes de un proceso. Con frecuencia la contracción es un problema principal. A menudo, una matriz moldeada para una pieza se construye más grande que lo nominal para permitir que la pieza se contraiga. En la siguiente situación experimental se produce una nueva matriz, y finalmente es importante encontrar los ajustes apropiados del proceso para minimizar la contracción. En el siguiente experimento, los valores de respuesta son desviaciones de los nominales (es decir, contracciones). Los factores y niveles son los siguientes:

	Niveles codificados:	
	-1	+1
A. Velocidad de inyección (pies/s)	1.0	2.0
B. Temperatura de moldeado (°C)	100	150
C. Presión de moldeado (psi)	500	1000
D. Contrapresión (psi)	75	120

El propósito del experimento es determinar cuáles efectos (efectos principales y de interacción) influyen en la contracción. El experimento se considera un experimento preliminar de investigación del que se pueden determinar los factores para un análisis más completo. Además, se desea que se pueda determinar una idea sobre cómo los factores importantes tienen impacto en la contracción. Los siguientes son los datos de un experimento 2^4 factorial sin réplicas:

Combinación de factores	Respuesta (cm × 10 ³)	Combinación de factores	Respuesta (cm × 10 ³)
(1)	72.68	<i>d</i>	73.52
<i>a</i>	71.74	<i>ad</i>	75.97
<i>b</i>	76.09	<i>bd</i>	74.28
<i>ab</i>	93.19	<i>abd</i>	92.87
<i>c</i>	71.25	<i>cd</i>	79.34
<i>ac</i>	70.59	<i>acd</i>	75.12
<i>bc</i>	70.92	<i>bcd</i>	79.67
<i>abc</i>	104.96	<i>abcd</i>	97.80

Inicialmente, se calcularon los efectos y se colocaron en una gráfica de probabilidad normal. Los efectos calculados son los siguientes:

$$\begin{array}{lll}
 A = 10.5613, & BD = -2.2787, & B = 72.4463, \\
 C = 2.4138, & D = 2.1438, & AB = 11.4038, \\
 AC = 1.2613, & AD = -1.8238, & BC = 3.8163, \\
 CD = 1.4088, & ABC = 2.8588, & ABD = -1.7813, \\
 ACD = -3.0438, & BCD = -0.4788, & ABCD = -1.3063
 \end{array}$$

La gráfica de probabilidad normal se muestra en la figura 15.2. La gráfica parece implicar que los efectos *A*, *B* y *AB* destacan como importantes. Las conclusiones preliminares son las siguientes:

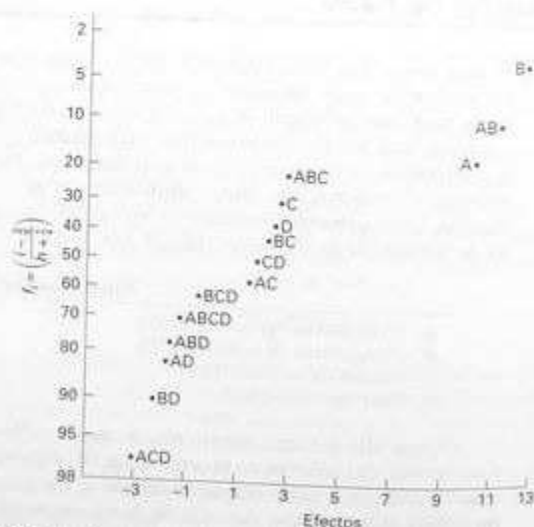


Figura 15.2 Gráfica de probabilidad normal de los efectos para el estudio de caso 15.5

1. Un aumento en la velocidad de inyección de 1.0 a 2.0 aumenta la contracción.
2. Un aumento en la temperatura de moldeado de 100°C a 200°C aumenta la contracción.
3. Hay una interacción entre la velocidad de inyección y la temperatura de moldeado; aunque ambos efectos principales son importantes, es crucial que comprendamos el impacto de la interacción de los dos factores.

Análisis con el cuadrado medio del error combinado: impreso de computadora comentado

Puede ser de interés observar un análisis de varianza de los datos de moldeado por inyección con interacciones de orden alto combinadas para formar un cuadrado medio del error. Las interacciones de orden tres y cuatro se combinan. La figura 15.3 muestra la salida del SAS PROC GLM. El análisis de varianza revela en esencia la misma conclusión que la de la gráfica de probabilidad normal.

```

General Linear Model Procedure
Dependent Variable: Y

```

Source	DF	Sum of Squares	Mean Square	F Value	Pr>F
Model	10	1684.2375	168.42375	9.37	0.0017
Error	5	90.1806	18.03612		
Corrected Total	15	1774.4181			

Source	DF	Type III SS	Mean Square	F Value	Pr>F
A	1	446.16001	446.16001	24.74	0.0002
B	1	119.63656	119.63656	6.76	0.0020
C	1	23.30474	23.30474	1.29	0.2672
D	1	18.38866	18.38866	1.02	0.3590
A*B	1	520.18206	520.18206	28.44	0.0010
A*C	1	6.36301	6.36301	0.35	0.5764
A*D	1	13.30426	13.30426	0.74	0.4297
B*C	1	13.19504	13.19504	0.73	0.4314
B*D	1	20.77061	20.77061	1.15	0.3302
C*D	1	7.93831	7.93831	0.44	0.5364

Source	DF	Type III SS	Mean Square	F Value	Pr>F
A	1	446.16001	446.16001	24.74	0.0002
B	1	119.63656	119.63656	6.76	0.0020
C	1	23.30474	23.30474	1.29	0.2672
D	1	18.38866	18.38866	1.02	0.3590
A*B	1	520.18206	520.18206	28.44	0.0010
A*C	1	6.36301	6.36301	0.35	0.5764
A*D	1	13.30426	13.30426	0.74	0.4297
B*C	1	13.19504	13.19504	0.73	0.4314
B*D	1	20.77061	20.77061	1.15	0.3302
C*D	1	7.93831	7.93831	0.44	0.5364

Parameter	Estimate	t for H0: Parameter = 0	Pr>T	Std Error of Estimate
INTERCEPT	79.49937500	75.35	0.0001	1.06172520
A	5.28062500	4.97	0.0042	1.06172520
B	5.02312500	5.86	0.0002	1.06172520
C	5.20687500	1.14	0.3052	1.06172520
D	1.07187500	1.03	0.3390	1.06172520
A*B	5.70187500	5.37	0.0030	1.06172520
A*C	0.43062500	0.39	0.5764	1.06172520
A*D	0.91187500	-0.86	0.4297	1.06172520
B*C	0.40682500	0.84	0.4314	1.06172520
B*D	-1.13927500	-1.07	0.3302	1.06172520
C*D	0.70637500	0.64	0.5364	1.06172520

Figura 15.3 Impreso del SAS para los datos del estudio de caso 15.4.

Las pruebas y valores P que se muestran en la figura 15.3 requieren una interpretación. Un valor P significativo sugiere que el efecto difiere significativamente de cero. Las pruebas sobre los efectos principales (que en presencia de interacciones se pueden considerar como los efectos promediados sobre los niveles de los demás factores) indican significancia para los efectos A y B . Los signos de los efectos también son importantes. Un aumento en los niveles de bajo a alto de A , la velocidad de inyección, tiene como resultado una contracción aumentada. Lo mismo es cierto para B . Sin embargo, debido a la interacción significativa AB , las interpretaciones de los efectos principales se pueden ver como tendencias de los niveles de los otros factores. El impacto de la interacción significativa AB se comprende mejor al utilizar una tabla de dos vías de las medias.

Interpretación de la interacción de dos factores

Como se podría esperar, una tabla de dos vías de las medias debe facilitar la interpretación de la interacción AB . Considere la siguiente:

A (velocidad)	2	73.355	97.205
	1	74.1975	75.240
		100	150
		B (temperatura)	

Nótese que la media muestral grande a alta velocidad y alta temperatura crea la interacción significativa. La contracción aumenta en forma no aditiva. La temperatura de moldeado parece tener un efecto positivo a pesar del nivel de la velocidad. Pero el efecto es mayor a alta velocidad. El efecto de la velocidad es muy ligero a baja temperatura pero es claramente positivo a alta temperatura de moldeado. Para controlar la contracción a un nivel bajo se debe evitar el uso de alta velocidad de inyección y alta temperatura de moldeado de manera simultánea.

Ejercicios

1. Los siguientes datos se obtienen de un experimento 2³ factorial replicado tres veces.

Combinación de tratamientos	Réplica 1	Réplica 2	Réplica 3
(1)	12	19	10
a	15	20	16
b	24	16	17
ab	23	17	27
c	17	25	21
ac	16	19	19
bc	24	23	29
abc	28	25	20

Evalúe las sumas de cuadrados para todos los efectos factoriales mediante el método de contraste.

2. En un experimento que lleva a cabo el Departamento de Ingeniería Minera del Instituto Politécnico y Universidad Estatal de Virginia para estudiar un sistema particular de filtración para carbón, se agrega un coagulante a una solución en un tanque que contiene carbón y lodo, el cual se coloca entonces en un sistema de recirculación a fin de que se pueda lavar el carbón. Se hacen variar tres factores en el proceso experimental:

Factor A: porcentaje de sólidos que circulan inicialmente en el sobreflujo.

Factor *B*: tasa de flujo del polímero
 Factor *C*: pH del tanque

La cantidad de sólidos en el flujo inferior del sistema de purificación determina qué tan limpio queda el carbón. Se usan dos niveles de cada factor y se realizan dos corridas experimentales para cada una de las $2^3 = 8$ combinaciones. Las respuestas, porcentaje de sólidos por peso, en el flujo inferior del sistema de circulación se especifican en la siguiente tabla:

Combinación de tratamientos	Respuesta	
	Réplica 1	Réplica 2
(1)	4.65	5.81
<i>a</i>	21.42	21.35
<i>b</i>	12.66	12.56
<i>ab</i>	18.27	16.62
<i>c</i>	7.93	7.88
<i>ac</i>	13.18	12.87
<i>bc</i>	6.51	6.26
<i>abc</i>	18.23	17.83

Suponga que todas las interacciones son potencialmente importantes; realice un análisis completo de los datos. Utilice valores *P* en su conclusión.

3. En un experimento metalúrgico se desea probar el efecto de cuatro factores y sus interacciones sobre la concentración (porcentaje del peso) de un compuesto fosforoso particular en material fundido. Las variables son *A*, porcentaje fosforoso en la refinación; *B*, porcentaje de material refundido; *C*, tiempo de fundición; y *D*, tiempo de retención. Los cuatro factores se hacen variar en un experimento 2^4 factorial con dos piezas fundidas que se toman en cada combinación de factores. Las 32 piezas fundidas se hacen en orden aleatorio. Las siguientes tablas muestran los datos y una tabla de ANOVA.

Combinación de tratamientos	% de peso del compuesto fosforoso		
	Réplica 1	Réplica 2	Total
(1)	30.3	28.6	58.9
<i>a</i>	28.5	31.4	59.9
<i>b</i>	24.5	25.6	50.1
<i>ab</i>	25.9	27.2	53.1
<i>c</i>	24.8	23.4	48.2
<i>ac</i>	26.9	23.8	50.7
<i>bc</i>	24.8	27.8	52.6
<i>abc</i>	22.2	24.9	47.1
<i>d</i>	31.7	33.5	65.2
<i>ad</i>	24.6	26.2	50.8
<i>bd</i>	27.6	30.6	58.2
<i>abd</i>	26.3	27.8	54.1
<i>cd</i>	29.9	27.7	57.6
<i>acd</i>	26.8	24.2	51.0
<i>bcd</i>	26.4	24.9	51.3
<i>abcd</i>	26.9	29.1	56.0
Total	428.1	436.9	865.0

Fuente de variación	Efectos	Suma de cuadrados	Grados de libertad	Cuadrado medio	<i>f</i> calculada	Valor <i>P</i>
Efecto principal:						
<i>A</i>	-1.2600	11.52	1	11.52	4.68	
<i>B</i>	-1.2250	12.01	1	12.01	4.88	0.0459
<i>C</i>	-2.2250	39.61	1	39.61	16.10	0.0010
<i>D</i>	1.4875	17.70	1	17.70	7.20	0.0163
Interacción de dos factores:						
<i>AB</i>	0.9875	7.80	1	7.80	3.17	0.0939
<i>AC</i>	-0.6125	3.00	1	3.00	1.22	0.2857
<i>AD</i>	-1.3250	14.05	1	14.05	5.71	0.0295
<i>BC</i>	1.1875	11.28	1	11.28	4.59	0.0480
<i>BD</i>	0.6250	3.13	1	3.13	1.27	0.2760
<i>CD</i>	0.7000	3.92	1	3.92	1.59	0.2249
Interacción de tres factores:						
<i>ABC</i>	-0.5500	2.42	1	2.42	0.98	0.3560
<i>ABD</i>	1.7375	24.15	1	24.15	9.82	0.0064
<i>ACD</i>	1.4875	17.70	1	17.70	7.20	0.0163
<i>BCD</i>	-0.8625	5.95	1	5.95	2.42	0.1394
Interacción de cuatro factores:						
<i>ABCD</i>	0.7000	3.92	1	3.92	1.59	0.2249
Error		<u>39.36</u>	<u>16</u>	2.46		
Total		217.51	31			

Discuta los efectos de los factores y sus interacciones sobre la concentración del compuesto fosforoso.

4. Se lleva a cabo un experimento preliminar para estudiar los efectos de cuatro factores y sus interacciones sobre la producción de la operación de cierta máquina. Se realizan dos pruebas en cada una de las combinaciones de tratamientos a fin de tener una medida del error experimental puro. Se utilizan dos niveles de cada factor, con lo que se obtienen los datos que se muestran en la tabla adjunta. Realice pruebas sobre todos los efectos principales e interacciones en el nivel de significancia de 0.05.

5. En el estudio *An X-Ray Fluorescence Method for Analyzing Polybutadiene-Acrylic Acid (PBAA) Propellants*, Quarterly Reports RK-TR-62-1, Army Ordnance Missile Command se realiza un experimento para determinar si existe una diferencia significativa en la cantidad de aluminio que se alcanza en el análisis entre ciertos niveles de algunas variables de proceso o no. Las variables y los datos se dan en la siguiente página.

Combinación de tratamientos	Réplica 1	Réplica 2
(1)	7.9	9.6
<i>a</i>	9.1	10.2
<i>b</i>	8.6	5.8
<i>c</i>	10.4	12.0
<i>d</i>	7.1	8.3
<i>ab</i>	11.1	12.3
<i>ac</i>	16.4	15.5
<i>ad</i>	7.1	8.7
<i>bc</i>	12.6	15.2
<i>bd</i>	4.7	5.8
<i>cd</i>	7.4	10.9
<i>abc</i>	21.9	21.9
<i>abd</i>	9.8	7.8
<i>acd</i>	13.8	11.2
<i>bcd</i>	10.2	11.1
<i>abcd</i>	12.8	14.3

- A: tiempo de mezcla
 nivel 1—2 horas
 nivel 2—4 horas
- B: velocidad de aspa
 nivel 1—36 rpm
 nivel 2—78 rpm
- C: condición de nitrógeno que pasa por el propulsor
 nivel 1—seco
 nivel 2—72% humedad relativa
- D: estado físico del propulsor
 nivel 1—no curado
 nivel 2—curado

Observación	Estado físico	Tiempo de mezcla	Velocidad de aspa	Condición del nitrógeno	Aluminio
1	1	1	2	2	16.3
2	1	2	2	2	16.0
3	1	1	1	1	16.2
4	1	2	1	2	16.1
5	1	1	1	2	16.0
6	1	2	1	1	16.0
7	1	2	2	1	15.5
8	1	1	2	1	15.9
9	2	1	2	2	16.7
10	2	2	2	2	16.1
11	2	1	1	1	16.3
12	2	2	1	2	15.8
13	2	1	1	2	15.9
14	2	2	1	1	15.9
15	2	2	2	1	15.6
16	2	1	2	1	15.8

Se registran los datos en la siguiente tabla. Suponga que todas las interacciones de tres y cuatro factores son insignificantes; analice los datos. Utilice un nivel de significancia de 0.05.

15.5 Experimentos factoriales en bloques incompletos

El experimento 2^k factorial se presta a la división en *bloques incompletos*. Para un experimento de k factores, a menudo es útil usar un diseño en 2^b bloques ($p < k$) cuando no se pueden aplicar todas las 2^k combinaciones de tratamientos bajo condiciones homogéneas. La desventaja con este arreglo experimental es que ciertos efectos se sacrifican por completo como resultado del bloqueo; la cantidad de eliminados depende del número de bloques que se requieran. Por ejemplo, suponga que las ocho combinaciones de tratamientos en un experimento 2^3 factorial se deben realizar en dos bloques de tamaño 4. Suponga, además, que se está dispuesto a sacrificar la interacción ABC . Note los "signos del contraste" en la tabla 15.3. Un arreglo razonable es

Bloque 1	Bloque 2
(1)	a
ab	b
ac	c
bc	abc

Concepto de confusión

Si se supone el modelo acostumbrado con el efecto de bloque aditivo, este efecto se cancela en la formación de los contrastes sobre todos los efectos excepto para el ABC . Para ilustrar, designemos con x la contribución al resultado debida a la diferencia entre bloques. Al escribir los resultados como

Bloque 1	Bloque 2
(1)	$a + x$
ab	$b + x$
ac	$c + x$
bc	$abc + x$

observamos que el contraste ABC y también el contraste que compara los dos bloques están dados por

$$\begin{aligned} \text{Contraste } ABC &= (abc + x) + (c + x) + (b + x) + (a + x) - (1) - ab - ac - bc \\ &= abc + a + b + c - (1) - ab - ac - bc + 4x. \end{aligned}$$

Por tanto, medimos el efecto ABC más el efecto del bloque y no hay manera de evaluar el efecto de interacción ABC independientemente de los bloques. Decimos entonces que la interacción ABC está **confundida por completo con los bloques**. Por necesidad, se sacrifica la información sobre ABC . Por otro lado, el efecto del bloque se cancela en la formación de todos los demás contrastes. Por ejemplo, el contraste A está dado por

$$\begin{aligned} \text{Contraste } A &= (abc + x) + (a + x) + ab + ac - (b + x) - (c + x) - bc - (1) \\ &= abc + a + ab + ac - b - c - bc - (1), \end{aligned}$$

como en el caso del diseño completamente aleatorizado. Decimos que los efectos A , B , C , AB , AC y BC son *ortogonales a los bloques*. Por lo general, para un experimento 2^k factorial en 2^r bloques, el número de efectos confundidos con bloques es 2^{r-1} , que es equivalente a los grados de libertad para los bloques.

2^k factorial en dos bloques

Cuando se deben usar dos bloques con un 2^k factorial, se elige un efecto, por lo general una interacción de orden superior, como el **contraste de definición**. Este efecto se confundirá con bloques. Los $2^k - 2$ efectos adicionales son ortogonales al contraste de definición y por ello con los bloques.

Suponga que representamos el contraste de definición como $A^{\gamma_1} B^{\gamma_2} C^{\gamma_3} \dots$, donde γ_i es 0 o 1. Esto genera la expresión

$$L = \gamma_1 + \gamma_2 + \dots + \gamma_k,$$

que a su vez se evalúa para cada una de las 2^k combinaciones de tratamientos al hacer γ_i igual a 0 o 1 de acuerdo a si la combinación de tratamientos contiene el i -ésimo factor en su nivel alto o bajo. Los valores L se reducen entonces (módulo 2) a 0 o 1 y de aquí se determina a cuál bloque se asignan las combinaciones de tratamientos. En otras palabras, las combinaciones de tratamientos se dividen en dos bloques de acuerdo con si los valores L dejan un residuo de 0 o 1 cuando se dividen entre 2.

Ejemplo 15.2 Determine los valores de L (módulo 2) para un experimento 2^3 factorial cuando el contraste de definición es ABC .

SOLUCIÓN

Con ABC como contraste de definición, tenemos

$$L = \gamma_1 + \gamma_2 + \gamma_3,$$

que se aplica a cada combinación de tratamientos como sigue:

$$(1): L = 0 + 0 + 0 = 0 = 0 \quad (\text{módulo } 2)$$

$$a: L = 1 + 0 + 0 = 1 = 1 \quad (\text{módulo } 2)$$

$$b: L = 0 + 1 + 0 = 1 = 1 \quad (\text{módulo } 2)$$

$$ab: L = 1 + 1 + 0 = 2 = 0 \quad (\text{módulo } 2)$$

$$c: L = 0 + 0 + 1 = 1 = 1 \quad (\text{módulo } 2)$$

$$ac: L = 1 + 0 + 1 = 2 = 0 \quad (\text{módulo } 2)$$

$$bc: L = 0 + 1 + 1 = 2 = 0 \quad (\text{módulo } 2)$$

$$abc: L = 1 + 1 + 1 = 3 = 1 \quad (\text{módulo } 2).$$

El arreglo en bloques, en el que se confunde ABC , es como antes,

Bloque 1	Bloque 2
(1)	a
ab	b
ac	c
bc	abc

Los efectos A , B , C , AB , AC y BC y las sumas de cuadrados se calculan de la manera acostumbrada, ignorando los bloques.

Nótese que este arreglo es el mismo esquema de bloques que resultaría de asignar las combinaciones de factores con signo "+" para el contraste ABC a un bloque y las combinaciones de factores con signo "-" para el contraste ABC al otro bloque.

El bloque que contiene la combinación de tratamientos (1) en el ejemplo 15.2 se llama **bloque principal**. Este bloque forma un grupo algebraico con respecto a la multiplicación cuando los exponentes se reducen a la base módulo 2. Por ejemplo, la propiedad de cerradura es válida, pues

$$(ab)(bc) = ab^2c = ac, \quad (ab)(ab) = a^2b^2 = (1),$$

etcétera: ■

2^3 factorial en cuatro bloques

Si se pide al experimentador asignar las combinaciones de tratamientos a cuatro bloques, éste elige dos contrastes de definición. Un tercer efecto, que se conoce como su **interacción generalizada**, se confunde de manera automática con los bloques, es-

los tres efectos corresponden a los tres grados de libertad para los bloques. El procedimiento para la construcción del diseño se explica mejor por medio de un ejemplo. Suponga que se decide que para un 2⁴ factorial AB y CD son los contrastes de definición. El tercer efecto confundido, su interacción generalizada, se forma al multiplicar los dos módulos 2 iniciales. De esta manera, el efecto

$$(AB)(CD) = ABCD$$

también se confunde con los bloques. Construimos el diseño al calcular las expresiones

$$L_1 = \gamma_1 + \gamma_2 \quad (AB),$$

$$L_2 = \gamma_3 + \gamma_4 \quad (CD)$$

módulo 2 para cada una de las 16 combinaciones de tratamientos para generar el siguiente esquema de bloques:

Bloque 1	Bloque 2	Bloque 3	Bloque 4
(1) <i>ab</i> <i>cd</i> <i>abcd</i>	<i>a</i> <i>b</i> <i>acd</i> <i>bcd</i>	<i>c</i> <i>abc</i> <i>d</i> <i>abd</i>	<i>ac</i> <i>bc</i> <i>ad</i> <i>bd</i>
$L_1 = 0$ $L_2 = 0$	$L_1 = 1$ $L_2 = 0$	$L_1 = 0$ $L_2 = 1$	$L_1 = 1$ $L_2 = 1$

Se puede utilizar un procedimiento de atajo para construir los bloques restantes después de que se genera el bloque principal. Comenzamos al colocar cualquier combinación de tratamientos que no esté en el bloque principal en un segundo bloque y construimos el bloque al multiplicar (módulo 2) por las combinaciones de tratamiento en el bloque principal. En el ejemplo anterior el segundo, tercero y cuarto bloques se generan como sigue:

Bloque 2	Bloque 3	Bloque 4
$a(1) = a$ $a(ab) = b$ $a(cd) = acd$ $a(abcd) = bcd$	$c(1) = c$ $c(ab) = abc$ $c(cd) = d$ $c(abcd) = abd$	$ac(1) = ac$ $ac(ab) = bc$ $ac(cd) = ad$ $ac(abcd) = bd$

El análisis para el caso de cuatro bloques es bastante simple. Todos los efectos que son ortogonales a los bloques (los diferentes a los contrastes de definición) se calculan en la forma acostumbrada. De hecho, se puede utilizar la técnica de Yates en todo el experimento, pero las sumas de cuadrados para los tres efectos confundidos se suman entonces para formar la suma de cuadrados debida a los bloques.

2^k factorial en 2^p bloques

El esquema general para el experimento 2^k factorial en 2^p bloques no es difícil. Seleccionamos p contrastes de definición de modo que ninguno sea la interacción generalizada de cualesquiera dos en el grupo. Como hay $2^p - 1$ grados de libertad para los bloques, tenemos $2^p - 1 - p$ efectos adicionales confundidos con los bloques. Por ejemplo, en un experimento 2^6 factorial en ocho bloques, podemos elegir ACF , $BCDE$ y $ABDF$ como contrastes de definición. Entonces

$$(ACF)(BCDE) = ABDEF,$$

$$(ACF)(ABDF) = \overline{BCD},$$

$$(BCDE)(ABDF) = ACEF,$$

$$(ACF)(BCDE)(ABDF) = E$$

son los cuatro efectos adicionales que se confunden con los bloques. Éste no es un esquema de bloques deseable, pues uno de los efectos confundidos es el efecto principal E . El diseño se construye al evaluar

$$L_1 = \gamma_1 + \gamma_3 + \gamma_6,$$

$$L_2 = \gamma_2 + \gamma_5 + \gamma_4 + \gamma_6,$$

$$L_3 = \gamma_1 + \gamma_2 + \gamma_4 + \gamma_6$$

y asignar tratamientos en combinaciones a los bloques de acuerdo con el siguiente esquema:

$$\text{Bloque 1: } L_1 = 0, L_2 = 0, L_3 = 0$$

$$\text{Bloque 2: } L_1 = 0, L_2 = 0, L_3 = 1$$

$$\text{Bloque 3: } L_1 = 0, L_2 = 1, L_3 = 0$$

$$\text{Bloque 4: } L_1 = 0, L_2 = 1, L_3 = 1$$

$$\text{Bloque 5: } L_1 = 1, L_2 = 0, L_3 = 0$$

$$\text{Bloque 6: } L_1 = 1, L_2 = 0, L_3 = 1$$

$$\text{Bloque 7: } L_1 = 1, L_2 = 1, L_3 = 0$$

$$\text{Bloque 8: } L_1 = 1, L_2 = 1, L_3 = 1.$$

Aquí también se aplica el procedimiento de atajo que se ilustra para el caso de cuatro bloques. Por ello podemos construir los siete bloques restantes a partir del bloque principal.

Ejemplo 15.3 Es de interés estudiar el efecto de cinco factores sobre alguna respuesta con la suposición de que las interacciones que incluyen tres, cuatro y cinco de los factores son insignificantes. Debemos dividir las 32 combinaciones de tratamientos en cuatro bloques con el uso de los contrastes de definición $BCDE$ y $ABCD$. De esta forma

$$(BCDE)(ABCD) = AE$$

también se confunden con los bloques. El diseño experimental y las observaciones se dan en la tabla 15.7.

La asignación de combinaciones de tratamientos a unidades experimentales dentro de los bloques es, por supuesto, aleatoria. Mediante la combinación de las interacciones no confundidas de tres, cuatro y cinco factores para formar el término de error, realice el análisis de varianza para los datos de la tabla 15.7.

Tabla 15.7 Datos para un experimento 2⁵ en cuatro bloques

Bloque 1	Bloque 2	Bloque 3	Bloque 4
(1) = 30.6	a = 32.4	b = 32.6	e = 30.7
bc = 31.5	abc = 32.4	c = 31.9	bce = 31.7
bd = 32.4	abd = 32.1	d = 33.3	bde = 32.2
cd = 31.5	acd = 35.3	bcd = 33.0	cde = 31.8
abe = 32.8	be = 31.5	ac = 32.0	ab = 32.0
ace = 32.1	ce = 32.7	abce = 33.1	ac = 33.1
aide = 32.4	de = 33.4	abde = 32.9	ad = 32.2
abcde = 31.8	bide = 32.9	acde = 35.0	abcd = 32.3

SOLUCIÓN

Se calculan las sumas de cuadrados para cada uno de los 31 contrastes y se encuentra que la suma de cuadrados de los bloques es

$$SS(\text{bloques}) = SS(ABCD) + SS(BCDE) + SS(AE) = 7.538.$$

El análisis de varianza se da en la tabla 15.8. Ninguna de las interacciones de dos factores son significativas en el nivel $\alpha = 0.05$ cuando se comparan con $f_{0.05}(1, 14) = 4.60$. Los efectos principales *A* y *D* son significativos y ambos proporcionan efectos positivos sobre la respuesta conforme vamos del nivel bajo al alto. ■

Tabla 15.8 Análisis de varianza para los datos de la tabla 15.7

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	<i>f</i> calculada
Efecto principal:				
<i>A</i>	3.251	1	3.251	6.32
<i>B</i>	0.320	1	0.320	0.62
<i>C</i>	1.361	1	1.361	2.64
<i>D</i>	4.061	1	4.061	7.89
<i>E</i>	0.005	1	0.005	0.01
Interacción de dos factores:				
<i>AB</i>	1.531	1	1.531	2.97
<i>AC</i>	1.125	1	1.125	2.18
<i>AD</i>	0.320	1	0.320	0.62
<i>BC</i>	1.201	1	1.201	2.35
<i>BD</i>	1.711	1	1.711	3.32
<i>BE</i>	0.020	1	0.020	0.04
<i>CD</i>	0.045	1	0.045	0.09
<i>CE</i>	0.001	1	0.001	0.002
<i>DE</i>	0.001	1	0.001	0.002
Bloques (<i>ABCD, BCDE, AE</i>):	7.538	3	2.513	
Error	7.208	14	0.515	

Es posible confundir cualquier efecto con bloques por medio de los métodos que se describen en la sección 15.6. Suponga que consideramos un experimento 2^3 factorial en dos bloques con tres réplicas completas. Si ABC se confunde con los bloques en las tres réplicas, podemos proceder como antes y determinar sumas de cuadrados de un solo grado de libertad para todos los efectos principales y efectos de interacción de dos factores. La suma de cuadrados para los bloques tiene 5 grados de libertad, lo que deja $23 - 5 - 6 = 12$ grados de libertad para el error.

Confundamos ahora ABC en una réplica, AC en la segunda y BC en la tercera. El plan para este tipo de experimento sería el siguiente:

Bloque		Bloque		Bloque	
1	2	1	2	1	2
abc	ab	abc	ab	abc	ab
a	ac	ac	bc	bc	ac
b	bc	b	a	a	b
c	(1)	(1)	c	(1)	c
Réplica 1 <i>ABC confundido</i>		Réplica 2 <i>AC confundido</i>		Réplica 3 <i>BC confundido</i>	

Se dice que los efectos ABC , AC y BC están **parcialmente confundidos con los bloques**. Estos tres efectos se pueden estimar a partir de dos de las tres réplicas. La proporción $2/3$ sirve como una medición de la magnitud de la confusión. Esta proporción da la cantidad de información disponible sobre el efecto parcialmente confundido en relación con la disponible en un efecto no confundido.

La presentación del análisis de varianza se da en la tabla 15.9. Las sumas de cuadrados para los bloques y para los efectos no confundidos A , B , C y AB se encuentran de la manera acostumbrada. Las sumas de cuadrados para AC , BC y ABC se calculan a partir de las dos réplicas en las que el efecto particular no se confunde. Se debe ser cuidadoso y dividir entre 16 en lugar de 24 cuando se obtienen las sumas de cuadrados para los efectos parcialmente confundidos; pues sólo utilizamos 16 observaciones. En la tabla 15.9 se insertan grados de libertad primados como recordatorio de que estos efectos están parcialmente confundidos y requieren de cálculos especiales.

Tabla 15.9. Análisis de varianza con confusión parcial

Fuente de variación	Grados de libertad
Bloques	5
A	1
B	1
C	1
AB	1
AC	1'
BC	1'
ABC	1'
Error	11
Total	23

Ejercicios

1. En un experimento 2² factorial con tres réplicas, muestre el arreglo en bloques e indique por medio de una tabla de análisis de varianza los efectos a probar y sus grados de libertad, cuando la interacción AB se confunde con los bloques.

2. Los siguientes datos codificados representan la resistencia de cierto tipo de envoltura normal para pan que se produce bajo 16 diferentes condiciones, estas últimas representan dos niveles de cada una de cuatro variables del proceso. Se introduce un efecto del operador en el modelo, pues fue necesario obtener la mitad de las corridas experimentales con el operador 1 y la mitad con el operador 2. Se consideró que los operadores tienen un efecto sobre la calidad del producto.

Operador 1

(1) = 18.8
ab = 16.5
ac = 17.8
bc = 17.3
d = 13.5
abd = 17.6
acd = 18.5
bcd = 17.6

Operador 2

a = 14.7
b = 15.1
c = 14.7
abc = 19.0
ad = 16.9
bd = 17.5
cd = 18.2
abcd = 20.1

(a) Con la suposición de que todas las interacciones son insignificantes, realice pruebas de significancia para los factores A , B , C y D . Utilice un nivel de significancia de 0.05.

(b) ¿Cuál interacción se confunde con los operadores?

3. Divida las combinaciones de tratamientos de un experimento 2⁴ factorial en cuatro bloques mediante la confusión de ABC y ABD . ¿Cuál efecto adicional también se confunde con los bloques?

4. Se lleva a cabo un experimento para determinar la resistencia a la ruptura de cierta aleación que contiene cinco metales, A , B , C , D y E . Se utilizan dos porcentajes diferentes de cada metal para formar las 2⁵ = 32 diferentes aleaciones. Como sólo ocho aleaciones se pueden probar en un día determinado, el experimento se realiza en un periodo de cuatro días durante el que los efectos $ABDE$ y AE se confunden con los días. Los datos experimentales se registran a continuación:

Combinación de tratamientos	Resistencia a la ruptura	Combinación de tratamientos	Resistencia a la ruptura
(1)	21.4	e	29.5
a	32.5	ae	31.3
b	28.1	be	33.0
ab	25.7	abe	23.7
c	34.2	ce	26.1
ac	34.0	ace	25.9
bc	23.5	bce	35.2
abc	24.7	abce	30.4
d	32.6	de	28.5
ad	29.0	ade	36.2
bd	30.1	bde	24.7
abd	27.3	abde	29.0
cd	22.0	cde	31.3
acd	35.8	acde	34.7
bcd	26.8	bcede	26.8
abcd	36.4	abcde	23.7

(a) Establezca el esquema de bloques para los cuatro días.

(b) ¿Cuál efecto adicional se confunde con los días?

(c) Obtenga las sumas de cuadrados para todos los efectos principales.

5. Mediante la confusión de ABC en dos réplicas y AB en la tercera, muestre el arreglo de bloques y la tabla de análisis de varianza para un experimento 2³ factorial con tres réplicas. ¿Cuál es la información relativa a los efectos confundidos?

6. Se realiza el siguiente experimento para estudiar los efectos principales y todas las interacciones. Se utilizan cuatro factores en dos niveles cada uno. El experimento se replica y se necesitan dos bloques en cada réplica. Los datos son los siguientes:

Réplica 1

Bloque 1	Bloque 2
(1) = 17.1	a = 15.5
d = 16.8	b = 14.8
ab = 16.4	c = 16.2
ac = 17.2	ad = 17.2
bc = 16.8	bd = 18.3
abd = 18.1	cd = 17.3
abd = 19.1	abc = 17.7
bcd = 18.4	abcd = 19.2

Replica II

Bloque 3	Bloque 4
(1) = 18,7	a = 17,0
ab = 18,6	b = 17,1
ac = 18,5	c = 17,2
ad = 18,7	d = 17,6
bc = 18,9	abc = 17,5
bd = 17,0	abd = 18,3
cd = 18,7	acd = 18,4
abcd = 19,8	bcd = 18,3

- (4) ¿Cuál efecto se confunde con los bloques en la primera réplica del experimento? ¿En la segunda réplica?
- (5) Realice un análisis de varianza apropiado que muestre las pruebas sobre todos los efectos principa-

les y las interacciones. Utilice un nivel de significancia de 0,05.

7. Considere un experimento 2^3 donde las corridas experimentales se realizan en cuatro máquinas diferentes. Utilice las máquinas como bloques y suponga que todos los efectos principales e interacciones de dos factores pueden ser importantes.

(a) ¿Cuáles corridas se deben realizar en cada una de las cuatro máquinas?

(b) ¿Cuáles efectos se confunden con los bloques?

8. Suponga que se puede realizar una segunda réplica del experimento en el ejercicio 7.

(a) ¿Una segunda réplica del esquema de bloques del ejercicio 7 sería la mejor elección?

(b) Si la respuesta a la parte (a) es no, proporcione el arreglo para una mejor elección para la segunda réplica.

(c) ¿Cuál concepto utilizó en su selección del diseño?

15.7 Experimentos factoriales en un ajuste de regresión

En muchas situaciones prácticas el científico tiene la tarea de construir un modelo de regresión con las variables de regresión controladas en el proceso de observación. En los capítulos 11 y 12 las variables regresoras o independientes en el análisis de regresión no necesariamente se suponen sujetas a un experimento diseñado. Obviamente, en muchos procesos químicos y biológicos y en otros medios científicos y de ingeniería, los niveles de las variables de regresión se pueden y se deben controlar. En esta sección mostramos algunas ventajas y discutimos el papel del experimento 2^3 factorial en la situación de regresión múltiple.

Recuerde que el modelo de regresión que se emplea en la sección 12.3 se puede escribir en notación matricial como

$$y = X\beta + \epsilon$$

La matriz X se denomina **matriz del modelo**. Suponga, por ejemplo, que se emplea un experimento 2^3 factorial con las variables

Temperatura	150°C	200°C
Humedad	15%	20%
Presión (psi)	1000	1500

Los niveles ya familiares +1, -1 se pueden generar a través del siguiente centrado y escalado a unidades de diseño:

$$x_1 = \frac{\text{temperatura} - 175}{25}, \quad x_2 = \frac{\text{humedad} - 17.5}{2.5}, \quad x_3 = \frac{\text{presión} - 1250}{250}$$

Como resultado la matriz \mathbf{X} se vuelve:

$\mathbf{X} =$	x_1	x_2	x_3	Identificación de diseño
$\begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$				(1)
				<i>a</i>
				<i>b</i>
				<i>c</i>
				<i>ab</i>
				<i>ac</i>
				<i>bc</i>
				<i>abc</i>

Se ve ahora que los contrastes que se ilustran y presentan en la sección 15.2 están directamente relacionados con los coeficientes de regresión. Nótese que todas las columnas de la matriz \mathbf{X} en nuestro ejemplo 2³ son ortogonales. Como resultado, el cálculo de los coeficientes de regresión que se describe en la sección 12.3 es

$$b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \left(\frac{1}{8}\mathbf{I}\right)\mathbf{X}'\mathbf{y}$$

$$= \frac{1}{8} \begin{bmatrix} a + ab + ac + abc + (1) + b + c + bc \\ a + ab + ac + abc - (1) - b - c - bc \\ b + ab + bc + abc - (1) - a - c - ac \\ c + ac + bc + abc - (1) - a - b - ab \end{bmatrix}$$

donde *a*, *ab*, etcétera, son mediciones de la respuesta.

Se puede observar que la noción de *efectos principales calculados* que se enfatizan a lo largo de este capítulo con 2^k factoriales se relaciona con los coeficientes en un modelo de regresión ajustado cuando los factores son cuantitativos. De hecho, para un 2^k con, digamos, *n* corridas experimentales por punto del diseño, la relación entre efectos y coeficientes de regresión es la siguiente:

$$\text{Efecto} = \frac{\text{contraste}}{2^{k-1}(n)}$$

$$\text{Coeficiente de regresión} = \frac{\text{contraste}}{2^k(n)} = \frac{\text{efecto}}{2^k}$$

Esta relación debe tener sentido para el lector pues un coeficiente de regresión b_j es una tasa promedio del cambio en la respuesta *por unidad de carga* en x_j . Por supuesto, conforme se va de -1 a $+1$ en x_j (bajo a alto), la variable de diseño cambia en dos unidades.

Ejemplo 15.4 Considere un experimento donde un ingeniero desea ajustar una regresión lineal del rendimiento y contra el tiempo de retención x_1 y el tiempo de relajación x_2 en cierto sistema químico. Todos los demás factores se mantienen fijos. Los datos en las unidades naturales son los siguientes:

Tiempo de retención (h)	Tiempo de relajación(h)	Rendimiento (%)
0.5	0.10	28
0.8	0.10	39
0.5	0.20	32
0.8	0.20	46

Estime el modelo de regresión lineal.

SOLUCIÓN

Como resultado, el modelo de regresión ajustado es

$$\hat{y} = b_0 + b_1x_1 + b_2x_2.$$

Las unidades de diseño son

$$x_1 = \frac{\text{tiempo de retención} - 0.65}{0.15}, \quad x_2 = \frac{\text{tiempo de relajación} - 0.15}{0.05}$$

y la matriz \mathbf{X} es

$$\mathbf{X} = \begin{bmatrix} & x_1 & x_2 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

con los coeficientes de regresión

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} \frac{(1) + a + b + ab}{4} \\ \frac{a + ab - (1) - b}{4} \\ \frac{b + ab - (1) - a}{4} \end{bmatrix} = \begin{bmatrix} 36.25 \\ 6.25 \\ 2.75 \end{bmatrix}$$

De esta manera la ecuación de regresión por mínimos cuadrados es

$$\hat{y} = 36.25 + 6.25x_1 + 2.75x_2.$$

Este ejemplo proporciona una ilustración del uso del experimento factorial de dos niveles en un ajuste de regresión. Las cuatro corridas experimentales en el diseño 2² se utilizan para calcular una ecuación de regresión, con la interpretación obvia de los coeficientes de regresión. El valor $b_1 = 6.25$ representa el aumento estimado en la respuesta (porcentaje de rendimiento) por cambio de la *unidad de diseño* (0.15 horas) en el tiempo de retención. El valor $b_2 = 2.75$ representa una tasa de cambio similar para el tiempo de relajación. ■

Interacción en el modelo de regresión

Los contrastes de interacción que se presentan en la sección 15.3 tienen interpretaciones precisas en el contexto de la regresión. De hecho, las interacciones se explican en los modelos de regresión por los términos de producto. Como ilustración, en el ejemplo 15.4, el modelo con interacción es

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$$

con b_0, b_1, b_2 como antes y

$$b_{12} = \frac{ab + (1) - a - b}{4} = \frac{46 + 28 - 39 - 32}{4} = 0.75.$$

De esta manera la ecuación de regresión que expresa *dos efectos principales lineales e interacción* es

$$\hat{y} = 36.25 + 6.25x_1 + 2.75x_2 + 0.75x_1x_2.$$

El contexto de regresión proporciona un marco en el que el lector debe comprender mejor la ventaja de la ortogonalidad que goza el 2² factorial. En la sección 15.2 se presentan los méritos de la ortogonalidad desde el punto de vista del *análisis de varianzas* de los datos en un experimento 2². Se señaló que la ortogonalidad entre los efectos conduce a la independencia entre las sumas de cuadrados. Por supuesto, la presencia de las variables de regresión en realidad no regula el uso del análisis de varianzas. De hecho, se realizan pruebas *F* justo como se describe en la sección 15.2. Por supuesto se debe hacer una distinción. En el caso del ANOVA, las hipótesis se desarrollan de las medias poblacionales mientras que en el caso de la regresión las hipótesis incluyen coeficientes de regresión.

Por ejemplo, considere el diseño experimental del ejercicio 2 de la página 568. Cada factor es continuo y se supone que los niveles son

$A(x_1)$:	20%	40%
$B(x_2)$:	5 lbs	10 lbs
$C(x_3)$:	5	5.5

tenemos, para los niveles del diseño,

$$x_1 = \frac{\text{sólidos} - 30}{10}, \quad x_2 = \frac{\text{tasa de flujo} - 7.5}{2.5}, \quad x_3 = \frac{\text{pH} - 5.25}{0.25}$$

Suponga que es de interés ajustar un modelo de regresión múltiple en el que se consideran todos los coeficientes lineales e interacciones disponibles. Además, es de interés para el ingeniero proporcionar alguna idea de cuáles niveles del factor *maximizarán* la purificación (es decir, *maximizar* la respuesta). Este problema será el tema de un estudio de caso en la siguiente sección.

15.8 Estudio de caso: experimento de purificación de carbón*

La figura 15.4 representa un impreso de computadora comentado para el análisis de regresión del modelo ajustado

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

SAS

Dependent Variable: Y

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	7	490.23494	70.03357	254.430	0.0001
Error	14	2.20205	0.15729		
C Total	15	492.43704			

Root MSE	0.52665	R-square	0.4955
Dep Mean	12.75166		
C.V.	4.11624		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob> T
INTERCEPT	1	12.75167	0.13116217	97.222	0.0001
A (X ₁)	1	4.719375	0.13116217	35.961	0.0001
B (X ₂)	1	0.865625	0.13116217	6.600	0.0002
C (X ₃)	1	-1.435625	0.13116217	-10.943	0.0001
AB (X ₁ X ₂)	1	-0.599375	0.13116217	-4.570	0.0010
AC (X ₁ X ₃)	1	0.520145	0.13116217	3.967	0.0034
BC (X ₂ X ₃)	1	0.005625	0.13116217	0.043	0.4668
ABC (X ₁ X ₂ X ₃)	1	0.230625	0.13116217	1.7607	0.0901

Obs	Dep Var		Std Err	Lower 95%		Upper 95%	
	Y	Predict		Mean	Predict	Mean	Predict
1	4.6500	5.2300	0.371	4.3745	4.0655	5.7452	6.7150
2	5.8100	5.2300	0.371	4.3745	4.0655	5.7452	6.7150
3	7.9300	7.4050	0.371	7.0495	6.7405	8.4332	9.3060
4	7.8000	7.4050	0.371	7.0495	6.7405	8.4332	9.3060
5	12.4600	12.4100	0.371	11.7545	11.4655	13.1250	14.0910
6	12.5600	12.4100	0.371	11.7545	11.4655	13.1250	14.0910
7	6.5100	6.3650	0.371	5.9295	5.6405	7.4032	7.6660
8	6.2600	6.3650	0.371	5.9295	5.6405	7.4032	7.6660
9	21.4200	21.3650	0.371	20.9895	20.6405	22.4032	22.6660
10	21.3500	21.3650	0.371	20.9895	20.6405	22.4032	22.6660
11	13.1600	13.0750	0.371	12.1695	11.8805	13.5432	14.0060
12	13.1600	13.0750	0.371	12.1695	11.8805	13.5432	14.0060
13	18.2700	17.4450	0.371	16.5395	16.1005	18.9632	19.9660
14	18.4600	17.4450	0.371	16.5395	16.1005	18.9632	19.9660
15	18.2700	18.0300	0.371	17.1745	16.8855	18.5462	19.5110
16	18.2700	18.0300	0.371	17.1745	16.8855	18.5462	19.5110

Figura 15.4 Impreso del SAS para los datos del estudio de caso 15.8.

*Véase el ejercicio 2, página 568.

donde x_1 , x_2 y x_3 son el porcentaje de sólidos, la tasa de flujo y el pH del sistema. El programa de computadora que se utiliza es el SAS PROC REG. Note los elementos de estimaciones de parámetros, error estándar y valores t en el impreso. Las estimaciones de parámetros representan los coeficientes en el modelo. Todos los coeficientes del modelo son significativos excepto el término x_2x_3 (interacción). Note también que los residuos, intervalos de confianza e intervalos de predicción aparecen como se presentan en el material sobre regresión de los capítulos 11 y

Diseño ortogonal

En situaciones experimentales donde resulta apropiado ajustar modelos que son lineales en las variables de diseño y que posiblemente deban incluir interacciones de términos de producto, existen ventajas que se obtienen a partir del *diseño ortogonal* de dos niveles, o arreglo ortogonal. Por diseño ortogonal queremos decir ortogonalidad entre las columnas de la matriz X . Por ejemplo, considere la matriz X para el 2^3 factorial del ejemplo 15.4. Nótese que las tres columnas son mutuamente ortogonales. La matriz X para el 2^3 factorial también contiene columnas ortogonales. El 2^3 factorial con interacciones daría una matriz X del tipo

$$X = \begin{matrix} & x_1 & x_2 & x_3 & x_1x_2 & x_2x_3 & x_1x_3 & x_1x_2x_3 \\ \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

Resulta simple verificar que el 2^3 es ortogonal para el modelo que contiene interacciones. La ortogonalidad tiene en cuenta las sumas de cuadrados independientes y las estimaciones independientes de coeficientes. Sin embargo, ¿cuál es el impacto de la ortogonalidad sobre las varianzas de estos coeficientes? El siguiente teorema subraya la utilidad de un diseño ortogonal a este respecto.

Teorema 15.1

Considere un modelo de regresión que contiene términos de efectos principales lineales y efectos de interacción entre los términos lineales. También, suponga que los rangos de las variables y el número de corridas experimentales son fijos. Las varianzas de estimaciones por mínimos cuadrados de los coeficientes se minimizan para los diseños que permiten dos niveles en cada variable en los extremos (± 1 en unidades de diseño), y columnas ortogonales en la matriz X .

Los *diseños ortogonales* de dos niveles a menudo se llaman diseños de *varianza mínima*. Los ejemplos principales de diseños de varianza mínima provienen de la familia 2^k factorial. Por ahora debe ser claro que el 2^k factorial proporciona estimacio-

nes óptimas no sólo para coeficientes lineales (debido a la ortogonalidad) sino también para coeficientes de interacción. Por ejemplo, un experimento 2^3 tiene en cuenta las estimaciones óptimas (en términos de la varianza) $\beta_0, \beta_1, \beta_2, \beta_3, \beta_{12}, \beta_{13}, \beta_{23}, \beta_{123}$ en un modelo del tipo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3.$$

Nótese que hay ocho coeficientes a estimar en el modelo anterior y el diseño tiene en cuenta ocho *puntos de diseño*. De hecho, cada columna en la tabla 15.3 representa un contraste o *una pieza de información* disponible. De esta manera todos los 8 grados de libertad disponibles a partir de los puntos de diseño se utilizan para estimar los coeficientes. Como resultado, *ninguna información de falta de ajuste está disponible*. La definición 15.1 introduce la terminología que se utiliza para describir tal diseño.

Definición 15.1

Se dice que un diseño de regresión es **saturado** cuando el número de puntos del diseño (combinaciones de factores distintas) es igual al número de términos en el modelo de regresión.

Hay, por supuesto, algún peligro al diseñar el ajuste de un modelo con un plan tal que no permanezcan grados de libertad para la falta de ajuste. La mayor parte de los planificadores prudentes de experimentos tiene en cuenta que algunos grados de libertad estén disponibles para el error de réplica (error puro) y para la falta de ajuste.

Ejemplo 15.5 Considere el experimento en el estudio de caso que se analizó antes en esta sección. Suponga que el ingeniero desea ajustar un modelo de regresión lineal a los datos. Es decir, no se colocan interacciones en el modelo. Se dispone de 16 observaciones en total sobre el porcentaje de sólidos en el flujo inferior. Señale los grados de libertad para el análisis.

SOLUCIÓN

Hay 15 grados de libertad en total (además de la media). El modelo es

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

El esquema de los grados de libertad es

Fuente	d.f.	
Regresión	3	
Falta de ajuste	4	$(x_1, x_2, x_3, x_1 x_2, x_1 x_3, x_2 x_3)$
Error (puro)	8	
Total	15	

Los ocho grados de libertad para el error puro se obtienen a partir de las *corridas duplicadas* en cada punto del diseño. Los grados de libertad de falta de ajuste se pueden ver como la diferencia entre el número de puntos distintos de diseño y el número total de términos del modelo: en este caso hay ocho puntos y cuatro términos del modelo. ■

Error estándar de coeficientes y pruebas *t*

En las secciones anteriores mostramos cómo el diseñador de un experimento puede explotar la noción de ortogonalidad para diseñar un experimento de regresión con coeficientes que alcanzan varianza mínima en una base por costo. Debemos ser capaces de hacer uso de nuestra exposición a la regresión en la sección 12.4 para calcular estimaciones de varianzas de coeficientes y de allí sus errores estándar. También es de interés notar la relación entre la estadística *t* sobre un coeficiente y la estadística *F* que se describe e ilustra en los capítulos anteriores.

Recuerde de la sección 12.4 que las varianzas y covarianzas de los coeficientes aparecen en A^{-1} , o en términos de la notación actual, la *matriz de varianzas-covarianza* de los coeficientes es

$$\sigma^2 A^{-1} = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}.$$

En el caso del experimento 2^k factorial, las columnas de \mathbf{X} son mutuamente ortogonales, lo que impone una estructura muy especial. En general, para el 2^k podemos escribir

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_2 & \cdots & x_k & x_1 x_2 & \cdots \\ \pm 1 & \pm 1 & \pm 1 & \cdots & \pm 1 & \pm 1 & \cdots \end{bmatrix}'$$

donde cada columna contiene 2^k entradas o 2^k*n*, donde *n* es el número de réplicas en cada punto de diseño. Entonces la formación de $\mathbf{X}'\mathbf{X}$ da

$$\mathbf{X}'\mathbf{X} = 2^k n \mathbf{I}_p,$$

donde \mathbf{I} es la matriz identidad de dimensión *p*, el número de parámetros del modelo.

Ejemplo 15.6 Considere un 2³ con corridas duplicadas que se ajustan al modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3,$$

Proporcione expresiones para los errores estándar de las estimaciones por mínimos cuadrados de $b_0, b_1, b_2, b_3, b_{12}, b_{13}$ y b_{23} .

SOLUCIÓN

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

con cada unidad vista como *repetida* (es decir, se duplica cada observación). Como resultado,

$$\mathbf{X}'\mathbf{X} = 16\mathbf{I}_7.$$

De esta forma

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{6}\mathbf{I}_7.$$

De lo anterior debe ser claro que las varianzas de todos los coeficientes para un 2^4 factorial con n corridas en cada punto de diseño son

$$\text{Var}(b_i) = \frac{\sigma^2}{2^4 n},$$

y, por supuesto, todas las covarianzas son cero. Como resultado, los errores estándar de los coeficientes se calculan como

$$s_{b_i} = s \sqrt{\frac{1}{2^4 n}},$$

donde s se encuentra a partir de la raíz cuadrada del cuadrado medio del error (con optimismo, obtenido de la réplica adecuada). Así, en nuestro caso con el 2^4 ,

$$s_{b_i} = s\left(\frac{1}{4}\right). \quad \blacksquare$$

Ejemplo 15.7 Considere el experimento metalúrgico del ejemplo 15.1. Suponga que el modelo ajustado es

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 \\ + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4.$$

¿Cuáles son los errores estándar de los coeficientes de regresión por mínimos cuadrados?

SOLUCIÓN

Los errores estándar de todos los coeficientes para el 2^4 factorial son iguales y están dados como

$$s_{b_i} = s \sqrt{\frac{1}{2^4 n}},$$

que en este ejemplo es

$$s_{b_i} = s \sqrt{\frac{1}{(16)(2)}}.$$

En este caso el cuadrado medio del error puro está dado por $s^2 = 2.46$ (16 grados de libertad). De esta manera

$$s_{b_i} = 0.28.$$

Los errores estándar de los coeficientes se pueden utilizar para construir estadísticas t de todos los coeficientes. Estos valores t se relacionan con las estadísticas F en el análisis de varianza. Demostramos ya que una estadística F de un coeficiente, con el uso del 2^4 factorial, es

$$F = \frac{(\text{contraste})^2}{(2^4)(n)s^2}.$$

Estas son las estadísticas F , por ejemplo, en la tabla 15.8. Es fácil verificar que si escribimos

$$t = \frac{b_i}{s_N}, \quad \text{donde} \quad b_i = \frac{\text{contraste}}{2^k n}$$

donde

$$t^2 = \frac{(\text{contraste})^2}{s^2 2^k n} = F.$$

Como resultado, la relación acostumbrada se mantiene entre las estadísticas t de los coeficientes y los valores F . Como se podría esperar, la única diferencia en el uso de la t o la F al evaluar la significancia yace en el hecho de que la estadística t indica el signo o dirección del efecto del coeficiente.

Parecería que el plan 2^k factorial manejaría muchas situaciones prácticas en las que se ajusta el modelo de regresión. Se pueden adaptar términos lineales y de interacción, que proporcionen estimaciones óptimas de todos los coeficientes (desde un punto de vista de la varianza). Sin embargo, cuando k es grande, el número de puntos de diseño que se requieren es muy grande. A menudo, partes del diseño total se pueden utilizar y aún permitir la ortogonalidad con todas sus ventajas. Estos diseños se presentan en la sección 15.9, a continuación. ■

15.9 Experimentos factoriales fraccionados

El experimento 2^k factorial se puede volver bastante demandante, en términos del número de unidades experimentales que se requieren, cuando k es grande. Una de las ventajas reales con este plan experimental es que permite un grado de libertad para cada interacción. Sin embargo, en muchas situaciones experimentales, se sabe que ciertas interacciones son insignificantes, y de esta manera sería un desperdicio de esfuerzo experimental utilizar el experimento factorial completo. De hecho, el experimentador puede tener una restricción económica que no permita tomar observaciones en todas las 2^k combinaciones de tratamientos. Cuando k es grande, con frecuencia podemos hacer uso de un **experimento factorial fraccionado** donde quizá la mitad, un cuarto o incluso un octavo del plan factorial total en realidad se lleva a cabo.

Construcción de la fracción 1/2

La construcción del diseño de media réplica es idéntico a la asignación del experimento 2^k factorial a dos bloques. Comenzamos con la selección de un contraste de definición que se sacrifica por completo. Construimos después los dos bloques en consecuencia y elegimos cualquiera de ellos como plan experimental.

Considere un experimento 2^k factorial donde deseamos utilizar media réplica. Se elige el contraste de definición $ABCD$ y de esta forma un plan experimental apropiado sería seleccionar el bloque principal que consiste en las siguientes combinaciones de tratamientos:

$$\{(1), ab, ac, ad, bc, bd, cd, abcd\}.$$

Con este plan, tenemos contrastes en todos los efectos excepto $ABCD$. Claramente,

$$\text{Contraste } A = ab + ac + ad + abcd - (1) - bc - bd - cd,$$

$$\text{Contraste } AB = abcd + ab + (1) + cd - ac - ad - bc - bd,$$

con expresiones similares para los contrastes de los efectos principales y efectos de interacción restantes. Sin embargo, con no más de ocho de las 16 observaciones en nuestro diseño fraccionado, sólo siete de los 14 contrastes no confundidos son ortogonales.

Alias en la fracción 1/2

Considere, por ejemplo, el contraste CD dado por

$$\text{Contraste } CD = abcd + cd + (1) + ab - ac - ad - bc - bd.$$

Observe que éste también es el contraste de un solo grado de libertad para AB . La palabra "alias" identifica dos efectos factoriales que tienen el mismo contraste. Por tanto, AB y CD son alias. En los experimentos 2^k factorial, el alias de cualquier efecto factorial es la interacción generalizada con el contraste de definición. Por ejemplo, si $ABCD$ es el contraste de definición, entonces el alias de A es $A(ABCD) = BCD$. Se puede ver entonces que la estructura completa de alias en media réplica de un experimento 2^k factorial, con el uso de $ABCD$ como contraste de definición, es (el símbolo = implica alias de)

$$A = BCD, \quad AB = CD,$$

$$B = ACD, \quad AC = BD,$$

$$C = ABD, \quad AD = BC,$$

$$D = ABC.$$

Sin evidencia complementaria no hay forma de explicar cuál de dos efectos alias proporciona en realidad la influencia en la respuesta. En un sentido *comparan un grado de libertad*. En esto yace la desventaja de los experimentos factoriales fraccionados. Tienen su mayor aplicación cuando k es bastante grande y hay algún conocimiento previo con respecto a las interacciones. En el ejemplo que se presenta, los efectos principales se pueden estimar si se sabe que las interacciones de tres factores son insignificantes. Para propósitos de prueba el único procedimiento, en ausencia de una medición externa del error experimental o de una réplica del experimento, sería combinar las sumas de cuadrados asociadas con las interacciones de dos factores. Esto, por supuesto, es deseable sólo si estas interacciones representan efectos insignificantes.

Construcción de la fracción 1/4

La construcción de la fracción 1/4 o cuarto de réplica es idéntica al procedimiento por el que se asignan 2^o combinaciones de tratamientos a cuatro bloques. Esto implica el sacrificio de dos contrastes de definición junto con su interacción generalizada. Cualquiera de los cuatro bloques resultantes sirve como un conjunto apropiado de corridas experimentales. Cada efecto tiene tres alias, que están dados por la interacción generalizada con los tres contrastes de definición. Suponga que en una fracción 1/4 de un experimento 2^o factorial, utilizamos *ACEF* y *BDEF* como contrastes de definición lo que tiene como resultado

$$(ACEF)(BDEF) = ABCD$$

que también se sacrifica. Al utilizar $L_1 = 0$, $L_2 = 0$ (módulo 2), donde

$$L_1 = \gamma_1 + \gamma_3 + \gamma_5 + \gamma_6, \quad L_2 = \gamma_2 + \gamma_4 + \gamma_5 + \gamma_6,$$

tenemos un conjunto apropiado de corridas experimentales:

$$\{(1), abcd, ef, abcdef, cde, cdf, abe, abf, acef, bdef, ac, bd, adf, ade, bcf, bce\}$$

y la estructura de alias para los efectos principales se escribe como

$$\begin{aligned} A &= CEF = ABDEF = BCD, & D &= ACDEF = BEF = ABC, \\ B &= ABCEF = DEF = ACD, & E &= ACF = BDF = ABCDE, \\ C &= AEF = BCDEF = ABD, & F &= ACE = BDE = ABCDF, \end{aligned}$$

cada uno con un sólo grado de libertad. Para las interacciones de dos factores,

$$\begin{aligned} AB &= BCEF = ADEF = CD, & AF &= CE = ABDE = BCDF, \\ AC &= EF = ABCDEF = BD, & BE &= ABCF = DF = ACDE, \\ AD &= CDEF = ABEF = BC, & BF &= ABCE = DE = ACDF, \\ AE &= CF = ABDF = BCDE, \end{aligned}$$

Aquí, por supuesto, hay algunos alias entre las interacciones de dos factores. Los 2 grados de libertad restantes los tienen en cuenta los siguientes grupos:

$$ADF = CDE = ABE = BCF, \quad ABF = BCE = ADE = CDF.$$

Es evidente que siempre se debe estar consciente de cuál es la estructura de alias para un experimento fraccionado antes de que finalmente se recomiende el plan experimental. La elección apropiada de los contrastes de definición es importante, pues dicta la estructura de alias. Por ejemplo, si se quisiera estudiar los efectos principales y todas las interacciones de dos factores en un experimento que incluye ocho factores y se sabe que las interacciones que incluyen tres o más factores son insignificantes, un diseño muy práctico sería uno en el que los contrastes de definición son *ACEGH* y *BDEFGH*, lo que tiene como resultado un tercero,

$$(ACEGH)(BDEFGH) = ABCDF.$$

Todos los efectos principales y las interacciones de dos factores no son alias entre sí y por tanto son estimables. El análisis de varianza contendría lo siguiente:

Efectos principales:	8	grados de libertad simples
Interacción de dos factores:	28	grados de libertad simples
Error:	27	grados de libertad combinados
Total:	63	grados de libertad

Para factoriales fraccionados de $1/8$ y mayores, se generaliza el método de construcción del diseño. Por supuesto, la relación de alias puede ser bastante extensa. Por ejemplo, con una fracción de $1/8$, cada efecto tiene siete alias. El diseño se construye mediante la selección de tres contrastes de definición como si se construyeran ocho bloques. Se sacrifican cuatro efectos adicionales y cualquiera de los ocho bloques se puede utilizar de forma apropiada como diseño. En la sección 15.10 se proporciona información adicional sobre fracciones mayores.

15.10 Análisis de experimentos factoriales fraccionados

La dificultad de realizar pruebas formales de significancia con el uso de los datos de experimentos factoriales fraccionados radica en la determinación del término de error apropiado. A menos que se disponga de datos de experimentos anteriores, el error debe provenir de una combinación de contrastes que representen efectos que se presumen insignificantes.

Las sumas de cuadrados para efectos individuales se encuentran mediante el uso de esencialmente el mismo procedimiento que se da para el factorial completo. Se puede formar un contraste en las combinaciones de tratamientos por medio de la construcción de la tabla acostumbrada de signos positivos y negativos. Por ejemplo, para media réplica de un experimento 2^3 factorial, con ABC como contraste de definición, se presenta en la tabla 15.10 un posible conjunto de combinaciones de tratamientos y los signos algebraicos apropiados para cada contraste que se utilizan al calcular los efectos y las sumas de cuadrados para los diversos efectos.

Tabla 15.10 Signos para los contrastes en media réplica de un experimento 2^3 factorial

Combinación de tratamientos	Efecto factorial						
	A	B	C	AB	AC	BC	ABC
a	+	-	-	-	-	+	+
b	-	+	-	+	-	-	+
c	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+

Nótese que en la tabla 15.10 los contrastes A y BC son idénticos, lo que ilustra la relación de alias. También, $B = AC$ y $C = AB$. En esta situación tenemos tres contrastes ortogonales que representan los 3 grados de libertad disponibles. Si se obtienen dos observaciones para cada una de las cuatro combinaciones de tratamientos, tendremos entonces una estimación de la varianza del error con 4 grados de libertad. Al suponer que los efectos de interacción son insignificantes, podemos probar la significancia de todos los efectos principales.

Un ejemplo de efecto y suma de cuadrados correspondiente es

$$w_A = \frac{a - b - c + abc}{2n} \quad SSA = \frac{(a - b - c + abc)^2}{2^2 n}$$

En general, la suma de cuadrados de un solo grado de libertad para cualquier efecto en una fracción 2^{-k} de un experimento 2^k factorial ($k > p$) se obtiene al elevar al cuadrado los contrastes en los totales de los tratamientos seleccionados y dividir entre $2^{k-p}n$, donde n es el número de réplicas de estas combinaciones de tratamientos.

Ejemplo 15.8 Suponga que deseamos utilizar media réplica para estudiar los efectos de cinco factores; cada uno a dos niveles, sobre alguna respuesta y se sabe que cualquiera que sea el efecto de cada factor, será constante para cada nivel de los otros factores. En otras palabras, no hay interacciones. Sea el contraste de definición $ABCDE$, lo que ocasiona que los efectos principales tengan una relación de alias con las interacciones de cuatro factores. La combinación de contrastes que incluyen interacciones proporciona $15 - 5 = 10$ grados de libertad para el error. Realice un análisis de varianza de los siguientes datos, pruebe la significancia en el nivel 0.05 de todos los efectos principales.

Tratamiento	Respuesta	Tratamiento	Respuesta
<i>a</i>	11.3	<i>bcid</i>	14.1
<i>b</i>	15.6	<i>abc</i>	14.2
<i>c</i>	12.7	<i>acd</i>	11.7
<i>d</i>	10.4	<i>ade</i>	9.4
<i>e</i>	9.2	<i>bee</i>	16.2
<i>abc</i>	11.0	<i>bde</i>	13.9
<i>abd</i>	8.9	<i>cde</i>	14.7
<i>acd</i>	9.6	<i>abcde</i>	13.2

SOLUCIÓN

Las sumas de cuadrados y efectos para los efectos principales son

$$SSA = \frac{(11.3 - 15.6 - \dots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-17.5)^2}{16} = 19.14,$$

$$\text{efecto de } A = \frac{17.5}{8} = 2.19,$$

$$SSB = \frac{(-11.3 + 15.6 - \dots - 14.7 + 13.2)^2}{2^{5-1}} = \frac{(18.1)^2}{16} = 20.48,$$

$$\text{efecto de } B = \frac{18.1}{8} = 2.26,$$

$$SSC = \frac{(-11.3 - 15.6 + \dots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(10.3)^2}{16} = 6.63,$$

$$\text{efecto de } C = \frac{10.3}{8} = 1.31,$$

$$SSD = \frac{(-11.3 - 15.6 - \dots + 14.7 + 13.2)^2}{2^{5-1}} = \frac{(-7.7)^2}{16} = 3.71,$$

$$\text{efecto de } D = \frac{-7.7}{8} = -0.96,$$

$$SS(E) = \frac{(-11.3 - 15.6 - \dots + 14.7 + 13.2)^2}{2^{3-1}} = \frac{(8.9)^2}{16} = 4.95,$$

$$\text{efecto de } E = \frac{8.9}{8} = 1.11.$$

Aquí el factor E se encierra dentro de paréntesis para evitar confusión con la suma de cuadrados del error. Todos los demás cálculos y pruebas de significancia se resumen en la tabla 15.11. Las pruebas indican que el factor A tiene un efecto significativo negativo sobre la respuesta, mientras que el factor B tiene un efecto positivo. Los factores C , D y E no son significativos en el nivel 0.05. ■

Tabla 15.11 Análisis de varianza para los datos de media réplica de un experimento 2^3 factorial

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	f calculada
Efecto principal				
A	19.14	1	19.14	6.21
B	20.48	1	20.48	6.65
C	6.63	1	6.63	2.15
D	3.71	1	3.71	1.20
E	4.95	1	4.95	1.61
Error	30.83	10	3.08	
Total	85.74	15		

Ejercicios

1. Liste los alias para los diversos efectos en un experimento 2^3 factorial cuando el contraste de definición es $ACDE$.

2. (a) Obtenga una fracción de $1/2$ de un diseño 2^3 factorial con el uso de BCD como el contraste de definición. (b) Divida la fracción $1/2$ en dos bloques de cuatro unidades cada uno mediante la confusión de ABC . (c) Muestre la tabla del análisis de varianza (fuentes de variación y grados de libertad) para probar todos los efectos no confundidos, suponga que todos los efectos de interacción son insignificantes.

3. Construya una fracción de $1/4$ de un diseño 2^3 factorial con el uso de $ABCD$ y $BDEF$ como los contrastes de definición. Muestre cuáles efectos tienen relación de alias con los seis efectos principales.

4. (a) Con el uso de los contrastes de definición $ABCE$ y $ABDF$, obtenga una fracción de $1/4$ de un diseño de 2^3 . (b) Muestre el análisis de varianza (fuentes de variación y grados de libertad) para todas las pruebas apropiadas; suponga que E y F no interactúan y que todas las interacciones de tres factores y superiores son insignificantes.

5. Se hacen variar siete factores en dos niveles en un experimento que incluye sólo 16 pruebas. Se utiliza una fracción de $1/8$ de un experimento 2^7 factorial con los contrastes de definición ACD , BEF y CEG . Los datos son los siguientes:

Combinación de tratamientos	Respuesta	Combinación de tratamientos	Respuesta
(1)	31.6	acg	31.1
ad	28.7	cdg	32.0
$abce$	33.1	beg	32.8
$cdef$	33.6	$adefg$	35.3
$acef$	33.7	efg	32.4
bcd	34.2	$abdeg$	35.3
$abdf$	32.5	$bcdg$	35.6
bf	27.8	$abcf$	35.1

Realice un análisis de varianza para los siete efectos principales, suponga que las interacciones son insignificantes. Utilice un nivel de significancia de 0.05.

6. Se realiza un experimento para que un ingeniero pueda obtener una idea de la influencia de la temperatura de sellado *A*, la temperatura de la bolsa de enfriamiento *B*, el porcentaje de aditivo de polietileno *C* y la presión *D* sobre la resistencia del sello (en gramos por pulgada) de una envoltura normal para pan. Se utiliza una fracción de 1/2 de un experimento 2⁴ factorial con *ABCD* como contraste de definición. Los datos son los siguientes:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	Respuesta
-1	-1	-1	-1	6.6
1	-1	-1	1	6.9
-1	1	-1	1	7.9
1	1	-1	-1	6.1
-1	-1	1	1	9.2
1	-1	1	-1	6.8
-1	1	1	-1	10.4
1	1	1	1	7.3

Realice un análisis de varianza de los efectos principales y de las interacciones de dos factores, suponga que todas las interacciones de tres factores y superiores son insignificantes. Utilice $\alpha = 0.05$.

7. En un experimento que se llevó a cabo en el Departamento de Ingeniería Mecánica y se analizó en el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia, un sensor detecta una carga eléctrica cada vez que las aspas de una turbina realizan una rotación. El sensor mide entonces la amplitud de la corriente eléctrica. Seis factores son: rpm *A*, temperatura *B*, distancia entre aspas *C*, distancia entre aspas y carcasa *D*, posición de la entrada *E* y posición de detección *F*. Se utiliza una fracción de 1/4 de un experimento 2⁶ factorial, con los contrastes de definición *ABCE* y *BCDF*. Los datos son los siguientes:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	Respuesta
-1	-1	-1	-1	-1	-1	3.89
1	-1	-1	-1	1	-1	10.46
-1	1	-1	-1	1	1	25.98
1	1	-1	-1	-1	1	39.88
-1	-1	1	-1	1	1	61.88
1	-1	1	-1	-1	1	3.22
-1	1	1	-1	-1	-1	8.94
-1	1	1	-1	1	-1	20.29
-1	-1	1	1	-1	1	32.07
1	-1	1	1	1	1	50.76
-1	1	-1	1	1	-1	2.80
-1	1	-1	1	-1	-1	8.15

[continuación]

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	Respuesta
-1	-1	1	1	1	-1	16.80
1	-1	1	1	-1	-1	25.47
-1	1	1	1	-1	1	44.44
1	1	1	1	1	1	2.45

Realice un análisis de varianza de los efectos principales y de las interacciones de dos factores, suponga que todas las interacciones de tres factores y superiores son insignificantes. Utilice $\alpha = 0.05$.

8. En el estudio *Durability of Rubber to Steel Adhesively Bonded Joints* realizado en el Departamento de Ciencias Ambientales y Mecánica y analizado en el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia, un experimentador mide el número de fallas en un sello adhesivo. Se postuló que la concentración de agua de mar *A*, temperatura *B*, pH *C*, voltaje *D* y resistencia *E* influyen en las fallas de un sello adhesivo. Se utiliza una fracción 1/2 de un experimento 2⁵ factorial, con *ABCDE* como contraste de definición. Los datos son los siguientes:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	Respuesta
-1	-1	-1	-1	1	462
1	-1	-1	-1	-1	746
-1	1	-1	-1	-1	714
1	1	-1	-1	1	1070
-1	-1	1	-1	-1	474
1	-1	1	-1	1	832
-1	1	1	-1	1	764
1	1	1	-1	-1	1087
-1	-1	-1	1	-1	522
1	-1	-1	1	1	854
-1	1	-1	1	1	773
1	1	-1	1	-1	1088
-1	-1	1	1	1	572
1	-1	1	1	-1	831
-1	1	1	1	-1	819
1	1	1	1	1	1104

Realice un análisis de varianza de los efectos principales y de interacciones de dos factores, suponga que todas las interacciones de tres factores y superiores son insignificantes. Utilice $\alpha = 0.05$.

15.11 Diseños de fracciones superiores y depurados

Algunas situaciones industriales requieren que el analista determine cuál de un gran número de factores controlables tiene un impacto sobre alguna respuesta importante. Los factores pueden ser cualitativos o variables de clase, variables de regresión o una mezcla de ambos. El procedimiento analítico puede incluir al análisis de varianza, el de regresión o ambos. A menudo el modelo de regresión que se utiliza implica sólo efectos principales lineales, aunque se pueden estimar algunas interacciones. La situación requiere la depuración de variables y los diseños experimentales que resultan se conocen como **diseños depurados**. Es claro que los diseños ortogonales de los niveles saturados o casi saturados son candidatos viables.

Resolución del diseño

Los diseños ortogonales de dos niveles a menudo se clasifican de acuerdo con su **resolución**, esta última se determina por medio de la siguiente definición.

Definición 15.2

La **resolución** de un diseño ortogonal de dos niveles es la longitud de la interacción más pequeña (menos compleja) entre el conjunto de contrastes de definición.

Si el diseño se construye como un factorial completo o fraccionado [es decir, un diseño 2^k o 2^{k-p} ($p = 1, 2, \dots, k-1$)], la noción de resolución del diseño es una ayuda para clasificar el impacto de la relación de alias. Por ejemplo, un diseño de resolución II tendría poco uso pues habría al menos un caso de alias de un efecto principal con otro. Un diseño de resolución III tendría todos los efectos principales (efectos lineales) ortogonales entre sí. Sin embargo, habrá algún alias entre los efectos lineales y las interacciones de dos factores. Es claro, entonces, que si el analista se interesa en estudiar los efectos principales (efectos lineales en el caso de la regresión) y no hay interacciones de dos factores, entonces se requiere un diseño de al menos resolución III.

Ejemplo 15.9 Suponga que un experimentador se interesa en estudiar cinco factores en un problema de regresión. Es importante que se estudien todos los efectos lineales. Ninguna interacción se considera importante. Construya un diseño de resolución III; utilice sólo ocho puntos de diseño.

SOLUCIÓN

Se requiere una fracción de un cuarto de un 2^5 y se deben utilizar tres letras o más como contrastes de definición. Elija ABC y $BCDE$ como contrastes de definición. Esto tiene como resultado a ADE como tercer contraste de definición. Claramente, el diseño es de resolución III y ningún efecto principal es alias de ningún otro efecto principal.

El uso de la resolución para clasificar diseños de dos niveles es muy importante. A menudo es deseable construir un diseño de primer orden que sea ortogonal y también compatible con los requerimientos del modelo. Por ejemplo, un analista podría desear probar efectos principales (o coeficientes de regresión lineal) como parte de una tarea de depuración de variables. Si las interacciones se consideran parte

del sistema, no se puede tolerar los alias entre los efectos principales y las interacciones de dos factores. En esta situación se necesita un diseño de resolución IV.

15.12 Construcción de diseños de resolución III y IV con 8, 16 y 32 puntos de diseño

Se pueden construir diseños de resoluciones III y IV útiles para 2 a 7 variables con 8 puntos de diseño. Simplemente comenzamos con un 2⁷ factorial que se satura de manera simbólica con interacciones.

x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$
-1	-1	-1	1	1	1	-1
1	-1	-1	-1	-1	1	1
-1	1	-1	-1	1	-1	1
-1	-1	1	1	-1	-1	1
1	1	-1	1	-1	-1	-1
1	-1	1	-1	1	-1	-1
-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1

Es claro que un diseño de resolución III se puede construir simplemente al reemplazar las columnas de interacción por nuevos efectos principales a través de siete variables. Por ejemplo, podemos definir

$$x_4 = x_1x_2 \quad (\text{contraste de definición } ABD)$$

$$x_5 = x_1x_3 \quad (\text{contraste de definición } ACE)$$

$$x_6 = x_2x_3 \quad (\text{contraste de definición } BCF)$$

$$x_7 = x_1x_2x_3 \quad (\text{contraste de definición } ABCG)$$

y obtener una fracción 2⁻⁴ de un 2⁷ factorial. Las expresiones anteriores identifican los contrastes de definición elegidos. Resultan once contrastes de definición adicionales y todos los contrastes de definición contienen al menos tres letras. De esta forma el diseño es de resolución III. Es claro que, si comenzamos con un subconjunto de las columnas aumentadas y concluimos con un diseño que incluye menos de siete variables de diseño, el resultado es un diseño de resolución III en menos de siete variables.

Se puede construir un conjunto similar de posibles diseños para 16 puntos de diseño al comenzar con un 2⁷ saturado con interacciones. Las definiciones de las variables que corresponden a estas interacciones producen diseños de resolución III a través de 15 variables. De manera similar los diseños que contienen 32 corridas se pueden construir al comenzar con un 2⁷.

Técnica del doblez

Se pueden complementar los diseños de resolución III que se describen antes para producir un diseño de resolución IV mediante el uso de la **técnica del doblez**. El do-

blez implica duplicar el tamaño del diseño al agregar la *negativa* de la matriz de diseño construida como se describe antes. La tabla 15.12 muestra un diseño de resolución IV de 16 corridas en siete variables mediante el uso de la técnica del doblez. Es obvio que podemos construir diseños de resolución IV que incluyen hasta 15 variables mediante el uso de la técnica del doblez sobre diseños desarrollados mediante el diseño 2^4 saturado.

Tabla 15.12 Diseño de resolución IV de dos niveles, con el uso de la técnica del doblez

	x_1	x_2	x_3	$x_4 = x_1x_2$	$x_5 = x_1x_3$	$x_6 = x_2x_3$	x_7
	-1	-1	-1	1	1	1	-1
	1	-1	-1	-1	-1	1	-1
	-1	1	-1	1	1	-1	-1
	-1	1	1	-1	-1	-1	-1
	1	-1	1	1	-1	1	-1
	1	1	1	-1	1	1	-1
	-1	-1	1	-1	-1	-1	1
	-1	1	1	1	1	1	1
	1	-1	-1	1	-1	-1	1
	1	1	-1	-1	1	1	1
	-1	-1	1	-1	-1	1	1
	-1	1	1	1	1	-1	1
	1	-1	1	1	-1	-1	1
	1	1	1	-1	1	1	1
	-1	-1	-1	-1	-1	-1	1
	-1	1	-1	1	1	-1	1
	-1	1	1	-1	-1	-1	1
	1	-1	1	1	-1	1	1
	1	1	1	-1	1	1	1
	-1	-1	1	-1	-1	-1	1
	-1	1	1	1	1	-1	1
	1	-1	1	1	-1	-1	1
	1	1	1	-1	1	1	1
	-1	-1	-1	-1	-1	-1	1
	-1	1	-1	1	1	-1	1
	-1	1	1	-1	-1	-1	1
	1	-1	1	1	-1	1	1
	1	1	1	-1	1	1	1

Este diseño se construye al "doblar" una fracción 1/8 de un 2^7 . La última columna se agrega como séptimo factor. En la práctica, la última columna a menudo juega el papel de variable de bloqueo. La técnica del doblez se utiliza con frecuencia en experimentación secuencial donde se analizan los datos del diseño inicial de resolución III. El experimentador puede considerar entonces, con base en el análisis, que se necesita un diseño de resolución IV. Como resultado, se puede necesitar una variable de bloqueo debido a que ocurre una separación temporal entre las dos partes del experimento. Además de la variable de bloqueo, el diseño final es una fracción de 1/4 de un experimento 2^6 .

15.13 Otros diseños de resolución III de dos niveles; diseños de Plackett-Burman

Una familia de diseños desarrollados por Plackett y Burman (véase la bibliografía) llenan los vacíos del tamaño de la muestra que existen en los factoriales fraccionados. Estos últimos son útiles con tamaños muestrales de 2^r (es decir, incluyen tamaños muestrales de 4, 8, 16, 32, 64, ...). Los diseños de Plackett-Burman incluyen $2r$ puntos de diseño y por ello se dispone de diseños de tamaño 12, 20, 24, 28, etcétera. Es-

tos diseños de Plackett-Burman de dos niveles son diseños de resolución III y son muy simples de construir. Se dan "líneas básicas" para cada tamaño muestral. Estas líneas de signos + y - son $n - 1$ en número. Para construir las columnas de la matriz de diseño se comienza con la línea básica y se realiza una permutación cíclica sobre las columnas hasta que se forman k columnas (número deseado de variables). Entonces se llena el último renglón con signos negativos. El resultado será un diseño de resolución III con k variables ($k = 1, 2, \dots, N$). Las líneas básicas son las siguientes:

$N = 12$	+	+	-	+	+	+	-	-	-	+	-									
$N = 16$	+	+	+	+	-	+	-	+	+	-	-	+	-	-	-					
$N = 20$	+	+	-	-	+	+	+	+	-	+	-	-	-	-	+	+	-			
$N = 24$	+	+	+	+	+	-	+	-	+	+	-	-	+	+	-	+	-	-	-	-

Ejemplo 15.10 Construya un diseño depurado de dos niveles con seis variables que contenga 12 puntos de diseño.

SOLUCIÓN

Comience con la línea básica en la columna inicial. La segunda columna se forma al traer la entrada de la parte inferior de la primera columna a la parte superior de la segunda columna y repetir la primera columna. La tercera columna se forma de la misma manera, con el uso de las entradas de la segunda columna. Cuando hay un número suficiente de columnas, simplemente se llena el último renglón con signos negativos. El diseño que resulta es el siguiente:

	x_1	x_2	x_3	x_4	x_5	x_6
	+	-	+	-	-	-
	+	+	-	+	-	-
	-	+	+	-	+	-
	+	-	+	+	-	+
	+	+	-	+	+	-
	+	+	+	-	+	+
	-	+	+	+	-	+
	-	-	+	+	+	-
	-	-	-	+	+	+
	+	-	-	+	+	+
	-	+	-	-	-	+
	-	-	-	-	-	-

Los diseños de Plackett-Burman son populares en la industria para depurar situaciones. Como los diseños de resolución III, todos los efectos lineales son ortogonales. Para cualquier tamaño muestral, el usuario dispone de un diseño para $k = 2, 3, \dots, N - 1$ variables. La estructura de alias para el diseño de Plackett-Burman es muy complicada y por ello el usuario no puede construir el diseño con un control completo sobre la estructura de alias como en el caso de diseños 2^k o 2^{k-p} . Sin embargo, en el caso de los

modelos de regresión el diseño de Plackett-Burman puede acomodar interacciones (aunque no sean ortogonales) cuando se dispone de suficientes grados de libertad.

15.14 Diseño robusto de parámetros de Taguchi

En este capítulo destacamos la noción del uso del diseño de experimentos (DOE) para aprender acerca de procesos de ingeniería y científicos. En el caso donde el proceso implica un producto, el DOE se puede utilizar para proporcionar el mejoramiento del producto o de la calidad. Como señalamos en el capítulo 1, se otorga mucha importancia al uso de métodos estadísticos en la mejora del producto. Un aspecto importante de este esfuerzo de mejora de la calidad de las décadas de 1980 y 1990 es el diseño de la calidad en procesos y productos en la etapa de investigación o en la etapa de diseño del proceso.

Gran parte del ímpetu de los métodos de mejora de la calidad lo motivó o, quizá, fue una reacción al aparente éxito que ingenieros y científicos en Japón tenían con el uso del diseño experimental. El diseño de experimentos es una importante herramienta de ingeniería en Japón. A principios de la década de 1980, el profesor Genichi Taguchi, ingeniero japonés, introdujo un método en su país para utilizar el diseño experimental en el desarrollo de productos y procesos que tiene las siguientes propiedades:

1. Insensible (robusto) a las condiciones del medio.
2. Insensible (robusto) a los factores que dificultan el control.
3. Proporciona variación mínima en su funcionamiento.

Esta filosofía y, en realidad, los métodos que la acompañan se denominan *diseño robusto de parámetros de Taguchi* (véase Taguchi, Taguchi y Wu y Kacker en la bibliografía). El término *diseño* en este contexto se refiere al diseño del proceso o sistema; *parámetro* se refiere a los parámetros del sistema. Éstos son los que llamamos *factores* o *variables*.

Es muy claro que los objetivos 1, 2 y 3 anteriores son bastante nobles. También es obvio que el personal de administración y los ingenieros en Estados Unidos no están particularmente atentos a estos temas. Por ejemplo, un ingeniero petrolero puede tener una mezcla fina de gasolina que funciona bastante bien mientras las condiciones son ideales o estables. Sin embargo, el funcionamiento se puede deteriorar debido a cambios en las condiciones del medio, como el tipo de conductor, condiciones climáticas, tipo de ingeniero, etcétera. Un científico de una compañía de alimentos puede tener una harina para pasteles bastante buena a menos que el usuario no siga exactamente las instrucciones de la caja que tienen que ver con la temperatura del horno, tiempo de horneado, etcétera. Un producto o proceso cuyo funcionamiento es consistente cuando se expone a estas condiciones cambiantes del medio se denomina **producto robusto** o **proceso robusto**. (Véase Myers y Montgomery (1995) en la bibliografía.)

Variables de control y ruido

Taguchi resalta la noción del uso de dos tipos de variables de diseño en un estudio. Estos factores son factores de control y factores de ruido.

Definición 15.3

Los **factores de control** son variables que se pueden controlar en el experimento y en el proceso. Los **factores de ruido** son variables que se pueden controlar o no en el experimento pero que no se pueden controlar en el proceso (o no se controlan bien en el proceso).

El enfoque de Taguchi (y el enfoque de muchos estadísticos prácticos en Estados Unidos) es el uso de variables de control y variables de ruido en el mismo experimento como efectos fijos. Los diseños ortogonales o arreglos ortogonales son diseños populares para su uso en este esfuerzo.

Objetivo del diseño robusto de parámetros

El objetivo del diseño robusto de parámetros es elegir los niveles de las variables de control (es decir, el diseño del proceso) que son más robustas (insensibles) a cambios en las variables de ruido.

Se debe notar que los cambios en las variables de ruido en realidad implican cambios durante el proceso, cambios en el campo, cambios en el medio, cambios en el manejo o uso del consumidor, etcétera.

Arreglo de producto

Un enfoque del diseño de experimentos que incluyen variables de control y de ruido es el uso de un plan experimental que requiere un diseño ortogonal para las variables de control y de ruido por separado. El experimento completo, entonces, es simplemente el producto o cruce de estos dos diseños ortogonales. El siguiente es un ejemplo sencillo de un arreglo de producto con dos variables de control y dos variables de ruido.

Ejemplo 15.11 En el artículo *The Taguchi Approach to Parameter Design* de D. M. Byrne y S. Taguchi, aparecido en *Quality Progress*, diciembre de 1987, los autores analizan un ejemplo interesante en el que se busca un método para ensamblar un conector electrométrico a un tubo de nailon que entrega el rendimiento de arranque requerido para que sea adecuado a una aplicación de motor automotriz. El objetivo es encontrar condiciones controlables que maximicen la fuerza de arranque. Entre las variables controlables están *A*, espesor de la pared del conector y *B*, profundidad de inserción. Durante la operación de rutina hay diversas variables que no se pueden controlar, aunque se controlarán durante el experimento. Entre ellas están *C*, tiempo de acondicionamiento y *D*, temperatura de acondicionamiento. Estos niveles se toman para cada variable de control y dos para cada variable de ruido. Como resultado, el arreglo cruzado es el siguiente. El arreglo de control es un arreglo de 3×3 y el arreglo de ruido es un 2^2 factorial ya familiar con 1), *c*, *d* y *cd* que representan las combinaciones de factores. El propósito del factor de ruido es crear la clase de variabilidad en la respuesta, fuerza de arranque, que se podría esperar en la operación diaria con el proceso. ■

	B (profundidad)		
	Superficial	Media	Profunda
Delgado	(1)	(1)	(1)
	c	c	c
	d	d	d
A (medio)	cd	cd	cd
	(1)	(1)	(1)
	c	c	c
Grueso	d	d	d
	cd	cd	cd
	(1)	(1)	(1)
	c	c	c
	d	d	d
	cd	cd	cd

Análisis

Hay varios procedimientos para el análisis del arreglo de producto. La aproximación que recomienda Taguchi y que ha sido adoptada por muchas compañías en Estados Unidos que tienen que ver con procesos de fabricación implica, inicialmente, la formación de estadísticas de resumen en cada combinación en el arreglo de control. Esta estadística de resumen se llama *razón señal-ruido* (SN). Suponga que llamamos y_1, y_2, \dots, y_n a un conjunto típico de corridas experimentales para el arreglo de ruido en una combinación fija del arreglo de control. La siguiente tabla describe algunas de las razones SN típicas:

Objetivo	Razón SN
Maximizar la respuesta	$SN_L = -10 \log \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2} \right)$
Lograr el objetivo	$SN_T = 10 \log \frac{\bar{y}^2}{s^2}$
Minimizar la respuesta	$SN_S = -10 \log \left(\frac{1}{n} \sum_{i=1}^n y_i^2 \right)$

Para cada uno de los casos anteriores se busca encontrar la combinación de las variables de control que *maximicen* SN.

Análisis de medias marginales

Una aproximación al análisis es simplemente tratar la razón SN como una respuesta estándar medida y realizar el análisis de varianza como se describe en este libro. Aunque a menudo se hace esto, se critica mucho (véase, por ejemplo, Box, 1986, en la bibliografía). A menudo el arreglo de control es un experimento factorial altamente fraccionado y el análisis implica el uso de *medias marginales* o un análisis típicamente fraccionado y el análisis implica el uso de *medias marginales* o un análisis típicamente fraccionado y el análisis implica el uso de *medias marginales* con un estudio de caso.

Ejemplo 15.12 Estudio de caso En un experimento que se describe en *Understanding Industrial Designed Experiments* de Schmidt y Launsby (véase la bibliografía), la optimización del proceso de soldadura se lleva a cabo en una planta de ensamblaje de tarjetas de circuitos. Las partes se insertan de forma manual o automática en una tarjeta sin revestir con un circuito impreso en ella. Después que se insertan las partes, la tarjeta se pone en una máquina soldadora de onda, que se usa para conectar todas las partes en el circuito. Las tarjetas se colocan en una banda transportadora y se toman a través de una serie de pasos. Se bañan en una mezcla fluida para eliminar óxido. Para minimizar la deformación, se precientan antes de aplicar la soldadura. La soldadura tiene lugar conforme las tarjetas se mueven por la onda de soldadura. El objeto del experimento es minimizar el número de defectos de soldadura por millón de unidades. El factor de control y los niveles son los siguientes:

Factor	(-1)	(+1)
A, temperatura del contenedor de la soldadura (°F)	480	510
B, velocidad de la banda transportadora (pie/min)	7.2	10
C, densidad de flujo	0.9 ^c	1.0 ^c
D, temperatura de precalentamiento (°F)	150	200
E, altura de la onda (pulgada)	0.5	0.6

Estos factores son fáciles de controlar en el nivel experimental pero son más terribles en el nivel de planta o de proceso. ■

Factores de ruido: tolerancias sobre los factores de control

A menudo en procesos como éste los factores naturales de ruido son tolerancias en los factores de control. Por ejemplo, en el proceso en línea real, la temperatura del contenedor de la soldadura y la velocidad de la banda transportadora son difíciles de controlar. Se sabe que el control de la temperatura está entre $\pm 5^\circ\text{F}$ y el control de la velocidad de la banda transportadora es ± 0.2 pie/min. En realidad es concebible que la variabilidad en la respuesta del producto (rendimiento de la soldadura) aumente debido a la incapacidad para controlar estos dos factores en algunos niveles nominales. El tercer factor de ruido es el tipo de ensamblaje involucrado. En la práctica, se utiliza uno de dos tipos de ensamblaje. De esta forma tenemos como factores de ruido:

Factor	(-1)	(+1)
A*, tolerancia de la temperatura del contenedor de la soldadura (°F) (desviación de la nominal)	-5 ^c	5 ^c
B*, tolerancia de la velocidad de la banda transportadora (pie/min) (desviación de la ideal)	-0.2	+0.2
C*, tipo de ensamblaje	1	2

Tanto el arreglo de control (arreglo interior) como el arreglo de ruido (arreglo exterior) se eligieron como factoriales fraccionados, el primero un 1/4 de un 2³ y el último en 1/2 de un 2³. El arreglo cruzado y los valores de respuesta son como sigue. Las primeras tres columnas del arreglo interior representan un 2³. Las columnas D = -AC y E = -BC. Así las interacciones de definición para el arreglo interior son ACD, BCE, y ADE. El

Arreglo interior					Arreglo exterior				
A	B	C	D	E	(i)	a*b*	a*c*	b*c*	(SN) _i
1	1	1	-1	-1	194	197	193	275	-46.75
1	1	-1	1	1	136	136	132	130	-42.61
1	-1	1	-1	1	185	261	264	264	-47.81
1	-1	-1	1	-1	47	125	127	42	-39.51
-1	1	1	-1	1	295	216	204	293	-48.15
-1	1	-1	1	1	234	159	231	157	-45.97
-1	-1	1	1	1	328	326	247	322	-45.76
-1	-1	-1	-1	-1	186	187	105	104	-43.59

arreglo exterior es una fracción estándar de resolución III de un 2^5 . Nótese que cada punto del arreglo interior contiene corridas del arreglo exterior. De esta manera se observan cuatro valores de respuesta en cada combinación del arreglo de control.

Razón señal-ruido y análisis

Como se indicó antes, Taguchi eligió "modelar" la razón señal-ruido en lugar de la respuesta natural. En este ejemplo la razón SN apropiada es

$$(SN)_i = -10 \log \left(\frac{1}{n} \sum_{j=1}^n y_j^2 \right)$$

La motivación es utilizar un criterio de funcionamiento que refleje una "pérdida de error cuadrático" o desviación cuadrática media del objetivo, en este caso el objetivo es cero (cero fallas en el proceso de soldadura). Con el signo negativo, se elige determinar valores de A, B, C, D y E que maximicen $(SN)_i$. El dispositivo analítico, siguiendo a Taguchi, es describir de forma gráfica un tipo de análisis de "sólo efectos principales" diseñado para maximizar (SN) . Las gráficas que indican las medias marginales con el uso de $(SN)_i$ como la respuesta se muestran en la figura 15.5.

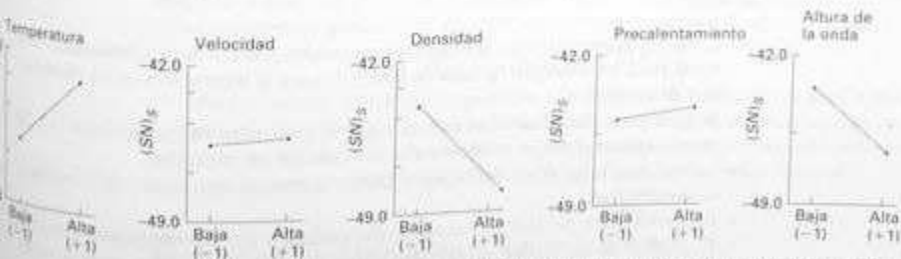


Figura 15.5 Gráfica que muestra la influencia de los factores en la razón señal-ruido.

La razón señal-ruido está dirigida a captar la media y la variabilidad de manera simultánea. Este análisis de las medias marginales en realidad sugeriría que la temperatura A y la densidad de flujo C son los más importantes entre los factores de control. Se pueden utilizar gráficas de medias marginales similares sobre la media \bar{y} (véase la figura 15.6).

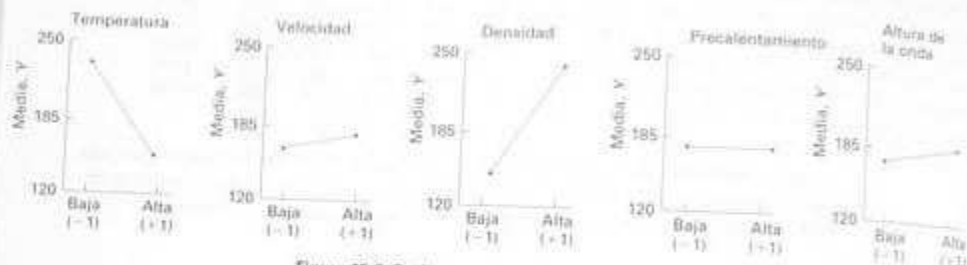


Figura 15.6 Gráfica que muestra la influencia de factores sobre la respuesta media.

Claramente, la temperatura y la densidad de flujo son los factores más importantes. Parecen influir en $(SN)_y$ y y . Por fortuna, se prefiere una temperatura alta y una densidad de flujo baja para $(SN)_y$ y la respuesta media. Por ello, las condiciones "óptimas" son

temperatura de soldadura = 510 °F, densidad de flujo = 0.9.

Si utilizamos el análisis que sugiere Taguchi, éste es el conjunto de condiciones más robusto.

Comentarios sobre el enfoque analítico de Taguchi

El enfoque de análisis que se presenta aquí es, en su mayor parte, el que propone Genichi Taguchi. Aunque sus contribuciones son en extremo importantes y bien documentadas, su aproximación específica al análisis es criticada. Entre las críticas están:

- Los diseños que sugiere a menudo son saturados o cerca de la saturación; éstos dejan poca información (grados de libertad) para la interacción entre las variables de control.
- Se hace poco modelado. Los métodos que se presentan en los capítulos 11 y 12 para construir el mejor modelo entre los factores no se utilizan.
- Las razones señal-ruido no siempre captan el sentido simultáneo de la media y la variabilidad.
- La aproximación del arreglo cruzado puede conducir a una experimentación excesiva.
- El análisis de las medias marginales utiliza una aproximación estricta de efectos principales, que ignora las interacciones entre las variables de control.

Para más comentarios bibliográficos con respecto al enfoque de Taguchi, consulte Kackar, Box y Box y Fung.

Aspectos positivos de la contribución de Taguchi

Las contribuciones de Taguchi a la ingeniería de diseño y a la optimización de procesos son muy importantes. No intentamos dar aquí una explicación detallada. Su trabajo ha impulsado a los ingenieros y estadísticos a replantear los métodos estadísticos en términos de sensibilidad a las variables del medio. Ha incitado a los estadísticos a resucitar viejas metodologías y desarrollar nuevas para utilizar la variabilidad de producto y proceso como una parte importante del criterio de rendimiento. Además de estas contribuciones técnicas ha guiado a muchos ingenieros hacia el uso de los métodos estadísticos y el diseño experimental y de manera indirecta influye en los estadísticos para trabajar en el área de la mejora de la calidad.

Enfoques alternativos al diseño robusto de parámetros

El trabajo de Taguchi influye en muchos estadísticos para buscar enfoques alternativos para resolver problemas en el diseño robusto de parámetros. Muchos de ellos siguen los principios de Taguchi del uso de variables de ruido y del énfasis en el trato con la variabilidad de la respuesta. Se debe consultar a Welch, Sachs y Kang; Vining y Myers; Myers, Khuri y Vining; y Shoemaker, Tsui y Wu. Ellos y otros intentan aliviar las críticas al enfoque de Taguchi mientras se mantienen en los principios del diseño robusto de parámetros.

Una aproximación sugerida por muchos es modelar la media muestral y la varianza muestral por separado en lugar de combinar los dos conceptos separados via una razón señal-ruido. El modelado separado a menudo ayuda al experimentador a obtener una mejor comprensión de los procesos involucrados. En el siguiente ejemplo, ilustramos este enfoque con el experimento del proceso de soldadura.

Ejemplo 15.13 Considere el conjunto de datos del ejemplo 15.12. Un análisis alternativo es plantear modelos separados para la media \bar{y} y la desviación estándar muestral. Suponga que utilizamos el código acostumbrado de +1 y -1 para los factores de control. Con base en la importancia evidente de la temperatura del contenedor de soldadura x_1 y de la densidad de flujo x_2 , el modelo de regresión lineal sobre la respuesta (número de errores por millón de uniones) produce el modelo

$$\hat{y} = 198.175 - 28.525x_1 + 57.975x_2$$

Para encontrar el nivel de temperatura y la densidad de flujo más robustos es importante producir un compromiso entre la respuesta media y la variabilidad, que requiere un modelado de la variabilidad. Una herramienta importante en este respecto es la transformación logarítmica (véase Bartlett y Kendall o Carroll y Ruppert):

$$\ln s^2 = \gamma_0 + \gamma_1(x_1) + \gamma_2(x_2)$$

Este proceso de modelado produce el siguiente resultado:

$$\ln s^2 = 7.4155 - 0.2067x_1 - 0.0309x_2$$

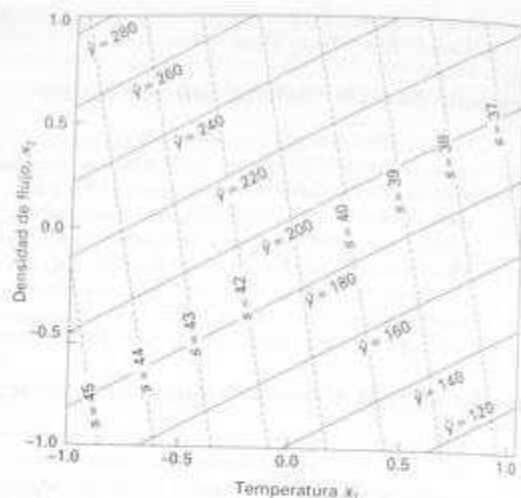


Figura 15.7 Media y desviación estándar para el ejemplo 15.13.

El análisis que es importante para el científico o ingeniero es hacer uso de los dos modelos de manera simultánea. Un enfoque gráfico puede ser muy útil. La figura 15.7 muestra un análisis simple de la media y la desviación estándar de manera simultánea. Como se podría esperar, la posición en temperatura y densidad de flujo que minimiza el número medio de errores es la misma que la que minimiza la variabilidad, a saber, temperatura alta y densidad de flujo baja. La aproximación gráfica de respuesta múltiple permite al usuario ver intercambios entre la media del proceso y la variabilidad del proceso. Para este ejemplo, el ingeniero puede estar poco satisfecho con las condiciones extremas en la temperatura de soldadura y densidad de flujo. La figura ofrece una estimación de la media y las condiciones de variabilidad que indican cuánto se pierde conforme se aleja de la condición óptima a cualesquiera condiciones intermedias. ■

Ejercicios

1. Utilice los datos de purificación del carbón del ejercicio 2, sección 15.3, para ajustar un modelo de tipo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3,$$

donde los niveles son

- x_1 : porcentaje de sólidos: 8, 12
- x_2 : tasa de flujo: 150, 250 gal/min
- x_3 : pH: 5, 6

Centre y escale las variables a las unidades de diseño. También realice una prueba para la falta de ajuste y comente lo adecuado del modelo de regresión lineal.

2. Se utiliza un plan 2^o factorial para construir un modelo de regresión que contiene coeficientes de primer orden y términos de modelo para todas las interacciones de dos factores. Se realizan corridas duplicadas para cada factor. Presente la tabla de análisis de varianza que muestre los grados de libertad para regresión, falta de ajuste y error puro.

3. Considere el 1/16 del 2^o factorial que se presenta en la sección 15.13. Liste los 11 contrastes de definición adicionales.

4. Construya un diseño de Plackett-Burman para 10 variables que contienen 24 corridas experimentales.

Ejercicios de repaso

1. Se utiliza un diseño de Plackett-Burman con el propósito de estudiar las propiedades reológicas de copolímeros de alto peso molecular. En el experimento se fijan dos niveles de cada una de seis variables. La viscosidad del polímero es la respuesta. Los datos se analizan en el Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia por personal del Departamento de Ingeniería Química de la universidad. Las variables son las siguientes: química de bloque duro x_1 , tasa de flujo de nitrógeno x_2 , tiempo de calentamiento x_3 , porcentaje de compresión x_4 , mediciones alta y baja x_5 , porcentaje de esfuerzo x_6 . Se obtienen los siguientes datos:

Observación	x_1	x_2	x_3	x_4	x_5	x_6	y
1	1	-1	1	-1	-1	-1	194,700
2	1	1	-1	1	-1	-1	588,400
3	-1	1	1	-1	1	-1	7,533
4	1	-1	1	1	-1	1	514,100
5	1	1	-1	1	1	-1	277,300
6	1	1	1	-1	1	1	493,500
7	-1	1	1	1	-1	1	8,969
8	-1	-1	1	1	1	-1	18,340
9	-1	-1	-1	1	1	1	6,793
10	1	-1	-1	-1	1	1	160,400
11	-1	1	-1	-1	-1	1	7,008
12	-1	-1	-1	-1	-1	-1	3,637

Construya una ecuación de regresión que relacione la viscosidad con los niveles de las seis variables. Realice pruebas t para todos los efectos principales. Recomendando factores que se deban conservar para estudios futuros y los que no. Utilice el cuadrado medio del residuo (5 grados de libertad) como medida del error experimental.

2. Una compañía grande de petróleo en el sudeste realiza experimentos en forma regular para probar los aditivos de los fluidos de perforación. La viscosidad plástica es una medición reológica que refleja el espesor del fluido. Se agregan varios polímeros al fluido para aumentar su viscosidad. El siguiente es un conjunto de datos en el que se utilizan dos polímeros en dos niveles cada uno y se mide la viscosidad. La concentración del polímero se indica como "baja" y "alta".

		Polímero 1	
		Baja	Alta
Polímero 2	Baja	3 3.5	11.3 12.0
	Alta	11.7 12.0	21.7 22.4

Realice un análisis del experimento 2^2 factorial. Pruebe los efectos de los dos polímeros y la interacción.

3. El Centro de Consulta Estadística del Instituto Politécnico y Universidad Estatal de Virginia analiza un experimento 2^2 factorial. El cliente es un miembro del Departamento de Vivienda, Diseño Interior y Administración de Recursos. El cliente se interesa en comparar hornos de arranque en frío y de precalentamiento en términos de la energía total que se entrega al producto. Además, las condiciones de consecución se comparan con el modo regular. Se realizan cuatro corridas experimentales en cada una de las cuatro combinaciones de los factores. A continuación se presentan los datos del experimento:

		Precautado	Arranque en frío
		618	575
Modo de consecución	619.3	373.7	
	629	574	
	611	572	
	581	558	
Modo regular	585.7	562	
	581	562	
	596	566	

Realice un análisis de varianza para estudiar los efectos principales y la interacción. Saque conclusiones.

4. Construya un diseño que incluya 12 corridas experimentales donde se hacen variar dos factores en dos niveles cada uno. Tiene la restricción adicional de que se deben utilizar bloques de tamaño 2 y debe ser capaz de realizar pruebas de significancia para los efectos principales y para el efecto de interacción.

5. En el estudio *The Use of Regression Analysis for Correcting Matrix Effects in the X-Ray Fluorescence Analysis of Pyrotechnic Compositions*, publicado en *Proceedings of the Tenth Conference on the Design of Experiments in Army Research Development and Testing*, ARD-D Report 65-3 (1965), se realiza un experimento en el que se permite variar las concentraciones de cuatro componentes de una mezcla de propulsor y los pesos de partículas finas y gruesas en la suspensión. Los factores *A*, *B*, *C* y *D*, cada uno en dos niveles, representan las concentraciones de los cuatro componentes y los factores *E* y *F*, también en dos niveles, representan los pesos de las partículas finas y gruesas presentes en la suspensión. El objetivo del análisis es determinar si las relaciones de intensidad de rayos X asociadas con el componente 1 del propulsor están influidas de manera significativa por la variación de las concentraciones de los diversos componentes y los pesos de los tamaños de las partículas en la mezcla. Se utiliza una fracción 1/8 de un experimento 2^o factorial con *ADE*, *BCE* y *ACF* como contrastes de definición. Los siguientes datos representan el total de un par de lecturas de intensidad:

El cuadrado medio del error combinado con 8 grados de libertad está dado por 0.02005. Analice los datos con el uso de un nivel de significancia de 0.05 para determinar si las concentraciones de los componentes y los pesos de las partículas finas y gruesas presentes en la suspensión tienen una influencia significativa en las relaciones de intensidad asociadas con el componente 1. Suponga que no existe interacción entre los seis factores.

6. Muestre el esquema de bloques para un experimento 2^o factorial en ocho bloques de tamaño 16 cada uno, utilice *ABCD*, *CDEFG* y *HDF* como contrastes de definición. Indique cuáles interacciones se sacrifican por completo en el experimento.

Lote:	Combinación de tratamientos:	Relación total de intensidad
1	<i>abef</i>	2.2489
2	<i>cdef</i>	1.8570
3	(1)	2.2428
4	<i>ace</i>	2.3270
5	<i>bde</i>	1.8830
6	<i>abcd</i>	1.8078
7	<i>adf</i>	2.1424
8	<i>bef</i>	1.9122

Estadística no paramétrica

16.1 Pruebas no paramétricas

La mayor parte de los procedimientos de prueba de hipótesis que se presentan en los capítulos anteriores se basan en la suposición de que las muestras aleatorias se seleccionan de poblaciones normales. Afortunadamente, la mayor parte de estas pruebas aún son confiables cuando experimentamos ligeras desviaciones de la normalidad, en particular cuando el tamaño de la muestra es grande. Tradicionalmente, estos procedimientos de prueba se denominan **métodos paramétricos**. En este capítulo consideramos varios procedimientos de prueba alternativos, llamados **no paramétricos** o **métodos de distribución libre**, que a menudo no suponen conocimiento de ninguna clase acerca de las distribuciones de las poblaciones fundamentales, excepto quizá que éstas son continuas.

Los procedimientos no paramétricos o de distribución libre se usan con mayor frecuencia por los analistas de datos. Hay muchas aplicaciones en la ciencia y la ingeniería donde los datos se reportan no como valores de un continuo sino más bien en una **escala ordinal** tal que es bastante natural asignar rangos a los datos. De hecho, el lector puede notar muy pronto en este capítulo que los métodos de distribución libre que se describen aquí implican un *análisis de rangos*. La mayoría de los analistas encuentran que los cálculos que se involucran en los métodos no paramétricos son muy atractivos e intuitivos.

Un ejemplo donde se aplica una prueba no paramétrica es el siguiente, dos jueces deben clasificar cinco marcas de cerveza de mucha demanda mediante la asignación de un grado de 1 a la marca que se considera que tiene la mejor calidad global, un grado 2 a la segunda mejor, etcétera. Se puede utilizar entonces

prueba no paramétrica para determinar dónde existe algún acuerdo entre los dos jueces.

Debemos señalar también que hay varias desventajas asociadas con las pruebas no paramétricas. En primer lugar, no utilizan la información que proporciona la muestra, y por ello una prueba no paramétrica será menos eficiente que el procedimiento paramétrico correspondiente, cuando se pueden aplicar ambos métodos. En consecuencia, para lograr la misma potencia, una prueba no paramétrica requerirá un tamaño muestral más grande que el que requerirá la correspondiente prueba no paramétrica.

Como indicamos antes, ligeras divergencias de la normalidad tienen como resultado desviaciones menores del ideal para las pruebas paramétricas estándar. Esto es cierto en particular para la prueba t y la prueba F . En el caso de la prueba t y la prueba F , el valor P citado puede ser ligeramente erróneo si existe una violación moderada de la suposición de normalidad.

En resumen, si se puede aplicar una prueba paramétrica y una no paramétrica al mismo conjunto de datos, debemos aplicar la técnica paramétrica más eficiente. Sin embargo, debemos reconocer que las suposiciones de normalidad a menudo no se pueden justificar, y que no siempre tenemos mediciones cuantitativas. Es una fortuna que los estadísticos nos proporcionen varios procedimientos no paramétricos útiles. Armado con las técnicas no paramétricas, el analista de datos tiene más munición para adaptar una variedad más amplia de situaciones experimentales. Se debe señalar que aun bajo las suposiciones de la teoría normal estándar, las eficiencias de las técnicas no paramétricas son notablemente cercanas a las del procedimiento paramétrico correspondiente. Por otro lado, las divergencias serias de la normalidad hacen al método no paramétrico más eficiente que el procedimiento paramétrico.

16.2 Prueba de signo

El lector debe recordar que los procedimientos que se presentan en la sección 10.7, para probar la hipótesis nula $\mu = \mu_0$, son válidos sólo si la población es aproximadamente normal o si la muestra es grande. Sin embargo, si $n < 30$ y la población es decididamente no normal, debemos recurrir a una prueba no paramétrica.

La prueba de signo se utiliza para probar hipótesis sobre una mediana poblacional. En el caso de muchos de los procedimientos no paramétricos, la media se reemplaza por la mediana como el **parámetro de ubicación** pertinente bajo prueba. Recuerde que en el capítulo 8 (definición 8.6) se define la mediana muestral. La contraparte poblacional, que se denota con $\tilde{\mu}$, tiene una definición análoga. Dada una variable aleatoria X , $\tilde{\mu}$ se define de modo que

$$P(X > \tilde{\mu}) = P(X < \tilde{\mu}) = 0.5.$$

Por supuesto, si la distribución es simétrica, la media y la mediana poblacionales son iguales. Al probar la hipótesis nula H_0 de que $\mu = \mu_0$ contra una alternativa apropiada, sobre la base de una muestra aleatoria de tamaño n , reemplazamos cada valor de la muestra que excede a μ_0 con un signo más y cada valor de la muestra menor que μ_0 con un signo menos. Si la hipótesis nula es verdadera y la población es simétrica, la suma de los signos más debe ser aproximadamente igual a la suma de los signos menos. Cuando un signo aparece con más frecuencia de lo que debería, con

base sólo en el azar, rechazamos la hipótesis de que la mediana poblacional μ es igual a μ_0 .

En teoría la prueba de signo se aplica sólo en situaciones donde μ_0 no puede ser igual al valor de cualquiera de las observaciones. Aunque hay una probabilidad cero de obtener una observación muestral exactamente igual a μ_0 cuando la población es continua, no obstante, en la práctica ocurrirá con frecuencia un valor muestral igual a μ_0 debido a una falta de precisión al registrar los datos. Cuando se observan valores muestrales iguales a μ_0 , se excluyen del análisis y el tamaño de la muestra se reduce en consecuencia.

La estadística de prueba apropiada para la prueba de signo es la variable aleatoria binomial X , que representa el número de signos más en nuestra muestra aleatoria. Si la hipótesis nula $\mu = \mu_0$ es verdadera, la probabilidad de que un valor muestral tenga como resultado un signo más o uno menos es igual a 1/2. Por tanto, para probar la hipótesis nula de que $\mu = \mu_0$, en realidad probamos la hipótesis nula de que el número de signos más es un valor de una variable aleatoria que tiene la distribución binomial con el parámetro $p = 1/2$. Los valores P para las alternativas unilateral y bilateral se pueden calcular entonces con el uso de esta distribución binomial. Por ejemplo, al probar

$$H_0: \mu = \mu_0,$$

$$H_1: \mu < \mu_0,$$

rechazaremos H_0 a favor de H_1 sólo si la proporción de signos más es bastante menor que 1/2; es decir, cuando el valor x de nuestra variable aleatoria es pequeño. De aquí, si el valor P que se calcula

$$P = P(X \leq x \text{ cuando } p = \frac{1}{2})$$

es menor que o igual a algún nivel de significancia α preestablecido, rechazamos H_0 a favor de H_1 . Por ejemplo, cuando $n = 15$ y $x = 3$, encontramos de la tabla A.1 que

$$P = P(X \leq 3 \text{ cuando } p = \frac{1}{2}) = \sum_{x=0}^3 b(x; 15, \frac{1}{2}) = 0.0176$$

por lo que la hipótesis nula $\mu = \mu_0$ se puede rechazar en realidad en el nivel de significancia de 0.05 pero no en el nivel 0.01.

Para probar la hipótesis

$$H_0: \mu = \mu_0,$$

$$H_1: \mu > \mu_0,$$

rechazamos H_0 a favor de H_1 sólo si la proporción de signos más es bastante mayor que 1/2; es decir, cuando x es grande. De aquí, si el valor P calculado

$$P = P(X \geq x \text{ cuando } p = \frac{1}{2})$$

es menor que α , rechazamos H_0 a favor de H_1 . Finalmente, para probar la hipótesis

$$H_0: \mu = \mu_0,$$

$$H_1: \mu \neq \mu_0,$$

rechazamos H_0 a favor de H_1 cuando la proporción de signos más es significativamente menor o mayor que 1/2. Esto, por supuesto, es equivalente a que x sea bastante pequeña o bastante grande. Por tanto, si $x < n/2$ y el valor P calculado

$$P = 2P(X \leq x \text{ cuando } p = \frac{1}{2})$$

es menor que o igual a α , o si $x > n/2$ y el valor P calculado

$$P = 2P(X \geq x \text{ cuando } p = \frac{1}{2})$$

es menor que o igual a α , rechazamos H_0 a favor de H_1 .

Siempre que $n > 10$, las probabilidades binomiales con $p = 1/2$ se pueden aproximar a partir de la curva normal, pues $np = nq > 5$. Suponga, por ejemplo, que deseamos probar la hipótesis

$$H_0: \mu = \mu_0,$$

$$H_1: \mu < \mu_0$$

en el nivel de significancia $\alpha = 0.05$ para una muestra aleatoria de tamaño $n = 20$ que produce $x = 6$ signos más. Con el uso de la aproximación de la curva normal con

$$\bar{\mu} = np = (20)(0.5) = 10$$

y

$$\sigma = \sqrt{npq} = \sqrt{(20)(0.5)(0.5)} = 2.236,$$

encontramos que

$$z = \frac{6.5 - 10}{2.236} = -1.57.$$

Por tanto,

$$P = P(X \leq 6) = P(Z < -1.57) = 0.0582,$$

que conduce a no rechazar la hipótesis nula.

Ejemplo 16.1 Los siguientes datos representan el número de horas que un compensador opera antes de requerir una recarga: 1.5, 2.2, 0.9, 1.3, 2.0, 1.6, 1.8, 1.5, 2.0, 1.2 y 1.7. Utilice la prueba de signo para probar la hipótesis en el nivel de significancia de 0.05 que este compensador particular opera con una mediana de 1.8 horas antes de requerir una recarga.

SOLUCIÓN

1. $H_0: \mu = 1.8$.
2. $H_1: \mu \neq 1.8$.
3. $\alpha = 0.05$.
4. Estadística de prueba: variable binomial X con $p = \frac{1}{2}$.
5. Cálculos: al reemplazar cada valor con el símbolo "+" si excede 1.8, con el símbolo "-" si es menor que 1.8 y descartar las mediciones que sean iguales a 1.8, obtenemos la siguiente secuencia

- + - - + - - + - -

para la que $n = 10$, $x = 3$, y $n/2 = 5$. Por tanto, de la tabla A.1 el valor P que se calcula es

$$P = 2P(X < 3 \text{ cuando } p = \frac{1}{2}) = 2 \sum_{x=0}^2 b(x; 10, \frac{1}{2}) = 0.3438 > 0.05.$$

6. Decisión: Decisión: no rechazar la hipótesis nula y concluir que el tiempo mediano de operación no es significativamente diferente de 1.8 horas. ■

También se puede utilizar la prueba de signo para probar la hipótesis nula $\mu_1 - \mu_2 = d_0$ para observaciones pareadas. Aquí reemplazamos cada diferencia, d_i , con un signo más o un signo menos dependiendo si la diferencia ajustada, $d_i - d_0$, es positiva o negativa. A lo largo de esta sección suponemos que las poblaciones son simétricas. Sin embargo, aun si las poblaciones son asimétricas podemos llevar a cabo el mismo procedimiento de prueba, pero las hipótesis se refieren a las medianas poblacionales en lugar de a las medias.

Ejemplo 16.2 Una compañía de taxis trata de decidir si el uso de llantas radiales en lugar de llantas regulares con cinturón mejora la economía de combustible. Se equipan 16 automóviles con llantas radiales y se manejan por un recorrido de prueba establecido. Sin cambiar de conductores, se equipan los mismos autos con las llantas regulares con cinturón y se manejan una vez más por el recorrido de prueba. Se registra el consumo de gasolina, en kilómetros por litro, de la siguiente manera:

Automóvil	Llantas radiales	Llantas con cinturón
1	4.2	4.1
2	4.7	4.9
3	6.6	6.2
4	7.0	6.9
5	6.7	6.8
6	4.5	4.4
7	5.7	5.7
8	6.0	5.8
9	7.4	6.9
10	4.9	4.9
11	6.1	6.0
12	5.2	4.9
13	5.7	5.3
14	6.9	6.5
15	6.8	7.1
16	4.9	4.8

¿Podemos concluir en el nivel de significancia de 0.05 que los autos equipados con llantas radiales obtienen mejores economías de combustible que los equipados con llantas regulares con cinturón?

SOLUCIÓN

Representemos con μ_1 y μ_2 los kilómetros por litro promedio para los autos equipados con llantas radiales y con cinturón, respectivamente

1. $H_0: \mu_1 - \mu_2 = 0$.
2. $H_1: \mu_1 - \mu_2 > 0$.
3. $\alpha = 0.05$.

- Estadística de prueba: variable binomial X con $p = 1/2$.
- Cálculo: después de reemplazar cada diferencia positiva con un símbolo "+", y cada diferencia negativa con un símbolo "-", y después descartar las dos diferencias cero, obtenemos la secuencia

+ - + + - + + + + + + - +

para la que $n = 14$ y $x = 11$. Con el uso de la aproximación de la curva normal, encontramos que

$$z = \frac{10.5 - 7}{\sqrt{14/2}} = 1.87,$$

y entonces

$$P = P(X \geq 11) = P(Z > 1.87) = 0.0307.$$

- Decisión: rechazar H_0 y concluir que, en promedio, las llantas radiales mejoran la economía de combustible. ■

La prueba de signo no sólo es uno de nuestros procedimientos no paramétricos más simples de aplicar, tiene la ventaja adicional de ser aplicable a datos dicotómicos que no se pueden registrar en una escala numérica pero que se pueden representar mediante respuestas positivas y negativas. Por ejemplo, la prueba de signo se aplica en experimentos donde se registra una respuesta cualitativa como "éxito" o "fracaso", y en experimentos de tipo sensorial donde se registra un signo más o un signo menos dependiendo de si el catador del sabor identifica de manera correcta o incorrecta el ingrediente que se desea.

Intentaremos realizar comparaciones entre varios de los procedimientos no paramétricos y las pruebas paramétricas correspondientes. En el caso de la prueba de signo la competencia es, por supuesto, la prueba t . Si se muestra de una distribución normal, el uso de la prueba t tendrá como resultado la potencia más grande de la prueba. Si la distribución sólo es simétrica, aunque no normal, se prefiere la prueba t en términos de potencia a menos que la distribución tenga "colas pesadas" en extremo en comparación con la distribución normal.

16.3 Prueba de rango con signo

El lector debe notar que la prueba de signo utiliza sólo los signos más y menos de las diferencias entre las observaciones y μ_0 en el caso de una muestra, o los signos más y menos de las diferencias entre los pares de observaciones en el caso de la muestra pareada, pero no toma en consideración la magnitud de estas diferencias. Una prueba que utiliza dirección y magnitud, propuesta en 1945 por Frank Wilcoxon, se llama ahora comúnmente **prueba de rango con signo de Wilcoxon**.

El analista puede extraer más información de los datos en una forma no paramétrica si resulta razonable aplicar una restricción adicional a la distribución de la que se toman los datos. La prueba de rango con signo de Wilcoxon se aplica en el caso de una **distribución continua simétrica**. Bajo esta condición podemos probar la hipótesis nula $\mu = \mu_0$. Primero restamos μ_0 de cada valor muestral y descartamos

todas las diferencias iguales a cero. Se clasifican entonces las diferencias restantes sin importar el signo. Se asigna un rango de 1 a la diferencia absoluta más pequeña (es decir, sin signo), un rango de 2 a la siguiente más pequeña, y así sucesivamente. Cuando el valor absoluto de dos o más diferencias es el mismo, se asigna a cada uno el promedio de los rangos que se asignarían si las diferencias se distinguieran. Por ejemplo, si las quinta y sexta diferencias son iguales en valor absoluto, a cada una se le asignaría un rango de 5.5. Si la hipótesis $\mu = \mu_0$ es verdadera, el total de los rangos que corresponden a las diferencias positivas debe ser casi igual al total de los rangos que corresponden a las diferencias negativas. Representemos estos totales con w_+ y w_- , respectivamente. Designamos el menor de w_+ y w_- con w .

Al seleccionar muestras repetidas esperaríamos que variarían w_+ y w_- , y por tanto w . De esta manera podemos considerar a w_+ , w_- , y w como valores de las correspondientes variables aleatorias W_+ , W_- , y W . La hipótesis nula $\mu = \mu_0$ se puede rechazar a favor de la alternativa $\mu < \mu_0$ sólo si w_+ es pequeña y w_- es grande. Del mismo modo, la alternativa $\mu > \mu_0$ se puede aceptar sólo si w_- es grande y w_+ es pequeña. Para una alternativa bilateral podemos rechazar H_0 a favor de H_1 si w_+ o w_- y por tanto w son suficientemente pequeñas. Por tanto, no importa cuál hipótesis alternativa pueda ser, rechazamos la hipótesis nula cuando el valor de la estadística apropiada W_+ , W_- , o W es suficientemente pequeño.

Dos muestras con observaciones pareadas

Para probar la hipótesis nula de que se muestrean dos poblaciones simétricas continuas con $\mu_1 = \mu_2$ para el caso de una muestra pareada, clasificamos las diferencias de las observaciones pareadas sin importar el signo y procedemos como en el caso de una sola muestra. Los diversos procedimientos de prueba para los casos de una sola muestra y de una muestra pareada se resumen en la tabla 16.1.

Tabla 16.1 Prueba de rango con signo

| Para probar H_0 | Contra H_1 | Calcular |
|-------------------|--------------------|----------|
| $\mu_1 = \mu_2$ | $\mu_1 < \mu_2$ | w_+ |
| | $\mu_1 > \mu_2$ | w_- |
| | $\mu_1 \neq \mu_2$ | w |
| $\mu_1 = \mu_2$ | $\mu_1 < \mu_2$ | w_+ |
| | $\mu_1 > \mu_2$ | w_- |
| | $\mu_1 \neq \mu_2$ | w |

No es difícil mostrar que siempre que $n < 5$ y el nivel de significancia no exceda 0.05 para una prueba de una cola o 0.10 para una prueba de dos colas, todos los valores posibles de w_+ , w_- , o w conducirán a la aceptación de la hipótesis nula. Sin embargo, cuando $5 \leq n \leq 30$, la tabla A.16 muestra valores críticos aproximados de W_+ y W_- para niveles de significancia iguales a 0.01, 0.025 y 0.05 para una prueba de una cola, y valores críticos de W para niveles de significancia iguales a 0.02, 0.05 y 0.10 para una prueba de dos colas. La hipótesis nula se rechaza

si el valor calculado $w_+ \leq w_{\alpha}$ o $w_- \geq w_{\alpha}$ es menor o igual que el valor tabulado apropiado. Por ejemplo, cuando $n = 12$, la tabla A.16 muestra que se requiere un valor de $w_{\alpha} \leq 17$ para que la alternativa unilateral $\mu < \mu_0$ sea significativa en el nivel 0.05.

Ejemplo 16.3 Repita el ejemplo 16.1 mediante el uso de la prueba de rango con signo.

SOLUCIÓN

1. $H_0: \mu = 1.8$.
2. $H_1: \mu \neq 1.8$.
3. $\alpha = 0.05$.
4. Región crítica: como $n = 10$, después de descartar la medición que es igual a 1.8, la tabla A.16 muestra que la región crítica es $w \leq 8$.
5. Cálculos: al restar 1.8 de cada medición y después clasificar las diferencias sin hacer caso del signo, tenemos

| | | | | | | | | | | |
|--------|------|-----|------|------|-----|------|------|-----|------|------|
| d_i | -0.3 | 0.4 | -0.9 | -0.5 | 0.2 | -0.2 | -0.3 | 0.2 | -0.6 | -0.1 |
| Rangos | 5.5 | 7 | 10 | 8 | 3 | 3 | 5.5 | 3 | 9 | 1 |

- Ahora bien, $w_+ = 13$ y $w_- = 42$ por lo que $w = 13$, el menor de w_+ y w_- .
6. Decisión: como antes, no rechazar H_0 y concluir que el tiempo promedio de operación no es significativamente diferente de 1.8 horas. ■

La prueba de rango con signo también se puede utilizar para probar la hipótesis nula $\mu_1 - \mu_2 = d_0$. En este caso las poblaciones no necesitan ser simétricas. Como con la prueba de signo, restamos d_0 de cada diferencia, clasificamos las diferencias ajustadas sin importar el signo y aplicamos el mismo procedimiento anterior.

Ejemplo 16.4 Se afirma que un estudiante universitario de último año puede aumentar su calificación en el área del campo de especialidad del examen de registro de graduados en al menos 50 puntos si de antemano se le proporcionan problemas de muestra. Para probar esta afirmación, se dividen 20 estudiantes del último año en 10 pares de modo que cada par tenga casi el mismo promedio de puntos de calidad general en sus tres primeros años en la universidad. Los problemas y respuestas de muestra se proporcionan al azar a un miembro de cada par una semana antes del examen. Se registran las siguientes calificaciones del examen:

| | Par: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Con problemas de muestra | | 531 | 621 | 603 | 579 | 451 | 660 | 591 | 719 | 543 | 575 |
| Sin problemas de muestra | | 509 | 540 | 688 | 502 | 424 | 683 | 568 | 748 | 530 | 524 |

Pruebe la hipótesis nula en el nivel de significancia de 0.05 de que los problemas de muestra aumentan las calificaciones en 50 puntos contra la hipótesis alternativa de que el aumento es menor a 50 puntos.

SOLUCIÓN

Representemos con μ_1 y μ_2 la calificación media de todos los estudiantes que resuelven el examen en cuestión con y sin problemas de muestra, respectivamente.

- $H_0: \mu_1 - \mu_2 = 50$.
- $H_1: \mu_1 - \mu_2 < 50$.
- $\alpha = 0.05$.
- Región crítica: como $n = 10$, la tabla A.16 muestra que la región crítica es $w_c \leq 11$.
- Cálculos:

| | Part. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------|-------|-----|----|-----|-----|-----|-----|-----|-----|-----|----|
| d_i | | 22 | 81 | -25 | 77 | 27 | -23 | 23 | -29 | 13 | 51 |
| $d_i - d_0$ | | -28 | 31 | -75 | 27 | -23 | -73 | -27 | -79 | -37 | 1 |
| Rangos | | 5 | 6 | 9 | 3.5 | 2 | 8 | 3.5 | 10 | 7 | 1 |

Encontramos ahora que $w_c = 6 + 3.5 + 1 = 10.5$.

- Decisión: rechazar H_0 y concluir que los problemas de muestra, en promedio, no aumentan las calificaciones de registro de graduados en 50 puntos. ■

Aproximación normal para muestras grandes

Cuando $n \geq 15$, la distribución muestral de W_n (o W_n^-) se aproxima a la distribución normal con media

$$\mu_{W_n} = \frac{n(n+1)}{4} \quad \text{y varianza} \quad \sigma_{W_n}^2 = \frac{n(n+1)(2n+1)}{24}$$

Por tanto, cuando n excede el valor más grande en la tabla A.16, se puede utilizar la estadística

$$Z = \frac{W_n - \mu_{W_n}}{\sigma_{W_n}}$$

para determinar la región crítica para nuestra prueba.

Ejercicios

1. Los siguientes datos representan el tiempo, en minutos, que un paciente tiene que esperar durante 12 visitas al consultorio de una doctora antes de ser atendido por ésta.

17 15 20 20
32 28 12 26
25 25 35 24

Utilice la prueba de signo al nivel de significancia de 0.05 para probar la afirmación de la doctora de que la media del tiempo de espera para sus pacientes no es mayor que 20 minutos antes de entrar al consultorio.

2. Los siguientes datos representan el número de horas de entrenamiento de vuelo que reciben 18 estudiantes para piloto de cierto instructor antes de su primer vuelo solos:

9 12 18 14 12 14
12 10 16 11 9 11
13 11 13 15 13 14

Con el uso de las probabilidades binomiales de la tabla A.1, realice una prueba de signo al nivel de significancia de 0.02 para probar la afirmación del instructor de que la mediana del tiempo que se requiere antes de que sus estudiantes vuelen solos es 12 horas de vuelo de entrenamiento.

3. Un inspector de alimentos examina 16 latas de cierta marca de jamón para determinar el porcentaje de impurezas externas. Se registraron los siguientes datos:

2.4 2.3 3.1 2.2
2.3 1.2 1.0 2.4
1.7 1.1 4.2 1.9
1.7 3.6 1.6 2.3

Con el uso de la aproximación normal a la distribución binomial, realice una prueba de signo al nivel de significancia de 0.05 para probar la hipótesis nula de que la mediana del porcentaje de impurezas en esta marca de jamón es 2.5% contra la alternativa de que la mediana del porcentaje de impurezas no es 2.5%.

4. Un proveedor de pintura afirma que un nuevo aditivo reducirá el tiempo de secado de su pintura acrílica. Para probar esta afirmación, se pintaron 12 paneles de madera, una mitad de cada panel con pintura que contiene el aditivo regular y la otra con pintura que contiene el nuevo aditivo. Los tiempos de secado, en horas, se registran a continuación:

| Panel | Tiempo de secado (horas) | |
|-------|--------------------------|-----------------|
| | Nuevo aditivo | Aditivo regular |
| 1 | 6.4 | 6.6 |
| 2 | 5.8 | 5.8 |
| 3 | 7.4 | 7.8 |
| 4 | 5.5 | 5.7 |
| 5 | 6.3 | 6.0 |
| 6 | 7.8 | 8.4 |
| 7 | 8.0 | 8.8 |

(continuación)

| Panel | Nuevo aditivo | Aditivo regular |
|-------|---------------|-----------------|
| 8 | 8.2 | 8.4 |
| 9 | 7.0 | 7.3 |
| 10 | 4.9 | 5.8 |
| 11 | 5.9 | 5.8 |
| 12 | 6.5 | 6.5 |

Utilice la prueba de signo en el nivel 0.05 para probar la hipótesis nula de que el nuevo aditivo no es mejor que el aditivo regular para reducir el tiempo de secado de este tipo de pintura.

5. Se afirma que una nueva dieta reducirá el peso de una persona 4.5 kilogramos, en promedio, en un periodo de dos semanas. Se registran los pesos de 10 mujeres que siguen esta dieta antes y después de un periodo de dos semanas, y se obtienen los siguientes datos:

| Mujer | Peso antes | Peso después |
|-------|------------|--------------|
| 1 | 58.5 | 60.0 |
| 2 | 60.3 | 54.9 |
| 3 | 61.7 | 58.1 |
| 4 | 69.0 | 62.1 |
| 5 | 64.0 | 58.5 |
| 6 | 62.6 | 59.9 |
| 7 | 56.7 | 54.4 |
| 8 | 63.6 | 60.2 |
| 9 | 68.2 | 62.3 |
| 10 | 59.4 | 58.7 |

Utilice la prueba de signo al nivel de significancia de 0.05 para probar la hipótesis de que la dieta reduce la mediana del peso en 4.5 kilogramos contra la hipótesis alternativa de que la mediana de la diferencia en pesos es menor que 4.5 kilogramos.

6. Se comparan dos tipos de instrumentos para medir la cantidad de monóxido de azufre en la atmósfera en un experimento de contaminación atmosférica. Se registraron las siguientes lecturas diarias en un periodo de dos semanas:

| Día | Monóxido de azufre | |
|-----|--------------------|---------------|
| | Instrumento A | Instrumento B |
| 1 | 0.96 | 0.87 |
| 2 | 0.82 | 0.74 |
| 3 | 0.75 | 0.63 |
| 4 | 0.61 | 0.55 |
| 5 | 0.89 | 0.76 |

(continuación)

| Día | Instrumento A | Instrumento B |
|-----|---------------|---------------|
| 6 | 0.64 | 0.70 |
| 7 | 0.81 | 0.69 |
| 8 | 0.68 | 0.57 |
| 9 | 0.65 | 0.53 |
| 10 | 0.84 | 0.88 |
| 11 | 0.59 | 0.51 |
| 12 | 0.94 | 0.79 |
| 13 | 0.91 | 0.84 |
| 14 | 0.77 | 0.63 |

Con el uso de la aproximación normal a la distribución binomial, realice una prueba de signo para determinar si los diferentes instrumentos conducen a diferentes resultados. Utilice un nivel de significancia de 0.05.

7. Las siguientes cifras dan la presión sanguínea sistólica de 16 corredores antes y después de una carrera de 8 kilómetros:

| Corredor | Antes | Después |
|----------|-------|---------|
| 1 | 158 | 164 |
| 2 | 149 | 158 |
| 3 | 160 | 163 |
| 4 | 155 | 160 |
| 5 | 164 | 172 |
| 6 | 138 | 147 |
| 7 | 163 | 167 |
| 8 | 159 | 169 |
| 9 | 165 | 173 |
| 10 | 145 | 147 |
| 11 | 150 | 156 |
| 12 | 161 | 164 |
| 13 | 132 | 133 |
| 14 | 155 | 161 |
| 15 | 146 | 154 |
| 16 | 159 | 170 |

Utilice una prueba de signo al nivel de significancia de 0.05 para probar la hipótesis nula de que correr 8 kilómetros aumenta la mediana de la presión sanguínea sistólica en ocho puntos contra la alternativa de que el aumento en la mediana es menor que ocho puntos.

8. Analice los datos del ejercicio 1 mediante el uso de la prueba de rango con signo.

9. Analice los datos del ejercicio 2 con el uso de la prueba de rango con signo.

10. Los pesos de cuatro personas antes de que dejen de fumar y cinco semanas después de dejar de fumar, en kilogramos, son los siguientes:

| Individuo: | 1 | 2 | 3 | 4 | 5 |
|------------|----|----|----|----|----|
| Antes | 66 | 80 | 69 | 52 | 75 |
| Después | 71 | 82 | 68 | 56 | 73 |

Utilice la prueba de rango con signo para observaciones pareadas para probar la hipótesis, en el nivel de significancia de 0.05, de que dejar de fumar no tiene efecto en el peso de una persona contra la alternativa de que el peso aumenta si se deja de fumar.

11. Repita el ejercicio 5 con el uso de la prueba de rango con signo.

12. Los siguientes son los números de recetas surtidas por dos farmacias en un periodo de 20 días:

| Día | Farmacia A | Farmacia B |
|-----|------------|------------|
| 1 | 19 | 17 |
| 2 | 21 | 15 |
| 3 | 15 | 12 |
| 4 | 17 | 12 |
| 5 | 24 | 16 |
| 6 | 12 | 15 |
| 7 | 19 | 11 |
| 8 | 14 | 13 |
| 9 | 20 | 14 |
| 10 | 18 | 21 |
| 11 | 23 | 19 |
| 12 | 21 | 15 |
| 13 | 17 | 11 |
| 14 | 12 | 10 |
| 15 | 16 | 20 |
| 16 | 15 | 12 |
| 17 | 20 | 13 |
| 18 | 18 | 17 |
| 19 | 14 | 16 |
| 20 | 22 | 18 |

Utilice la prueba de rango con signo al nivel de significancia de 0.01 para determinar si las dos farmacias, en promedio, surten el mismo número de recetas contra la alternativa de que la farmacia A surte más recetas que la farmacia B.

13. Repita el ejercicio 7 con el uso de la prueba de rango con signo.

16.4 Prueba de la suma de rangos

Como indicamos antes, el procedimiento no paramétrico por lo general es una alternativa apropiada para la prueba de la teoría normal cuando la suposición de normalidad no es válida. Cuando interesa probar la igualdad de las medias de dos distribuciones continuas que obviamente no son normales, y las muestras son independientes (es decir, no hay apareamiento de observaciones), la **prueba de la suma de rangos de Wilcoxon** o **prueba de dos muestras de Wilcoxon** es una alternativa apropiada a la prueba t de dos muestras que se describe en el capítulo 10.

Probaremos la hipótesis nula H_0 de que $\mu_1 = \mu_2$ contra alguna alternativa adecuada. Primero seleccionamos una muestra aleatoria de cada una de las poblaciones. Sea n_1 el número de observaciones en la muestra más pequeña y n_2 es el número de observaciones en la muestra más grande. Cuando las muestras son de igual tamaño, n_1 y n_2 se pueden asignar de manera aleatoria. Ordenar las $n_1 + n_2$ observaciones de las muestras combinadas en orden ascendente y sustituir un rango de 1, 2, ..., $n_1 + n_2$ para cada observación. En el caso de empates (observaciones idénticas), reemplazamos las observaciones por la media de los rangos que tendrían las observaciones si fueran distinguibles. Por ejemplo, si la séptima y octava observaciones son idénticas, asignaríamos un rango de 7.5 a cada una de las dos observaciones.

La suma de los rangos que corresponden a las n_1 observaciones en la muestra más pequeña se denota con w_1 . De manera similar, el valor w_2 representa la suma de los n_2 rangos que corresponden a la muestra más grande. El total $w_1 + w_2$ depende sólo del número de observaciones en las dos muestras y de ninguna manera resulta afectado por los resultados del experimento. De aquí, si $n_1 = 3$ y $n_2 = 4$, entonces $w_1 + w_2 = 1 + 2 + \dots + 7 = 28$, sin importar los valores numéricos de las observaciones. En general,

$$w_1 + w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2}$$

la suma aritmética de los enteros 1, 2, ..., $n_1 + n_2$. Una vez determinada w_1 , puede ser más fácil encontrar w_2 mediante la fórmula

$$w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - w_1$$

Al elegir muestras repetidas de tamaño n_1 y n_2 , esperaríamos que variara w_1 y por tanto w_2 . Podemos considerar entonces w_1 y w_2 como valores de las variables aleatorias W_1 y W_2 , respectivamente. La hipótesis nula $\mu_1 = \mu_2$ se rechazará a favor de la alternativa $\mu_1 < \mu_2$ sólo si w_1 es pequeña y w_2 es grande. Del mismo modo, la alternativa $\mu_1 > \mu_2$ se puede aceptar sólo si w_1 es grande y w_2 es pequeña. Para una prueba de dos colas, podemos rechazar H_0 a favor de H_1 si w_1 es pequeña y w_2 es grande o si w_1 es grande y w_2 es pequeña. En otras palabras se acepta la alternativa $\mu_1 < \mu_2$ si w_1 es suficientemente pequeña; la alternativa $\mu_1 > \mu_2$ se acepta si w_1 es suficientemente pequeña; y la alternativa $\mu_1 = \mu_2$ se acepta si el mínimo de w_1 y w_2 es suficientemente pequeño. En la práctica real por lo general basamos nuestra decisión en el valor

$$z_1 = w_1 - \frac{n_1(n_1 + 1)}{2} \quad \text{o} \quad z_2 = w_2 - \frac{n_2(n_2 + 1)}{2}$$

de la estadística relacionada U_1 o U_2 , o en el valor u de la estadística U , el mínimo de U_1 y U_2 . Estas estadísticas simplifican la construcción de tablas de valores críticos, pues U_1 y U_2 tienen distribuciones muestrales simétricas y toman valores en el intervalo de 0 a $n_1 n_2$, tales que $u_1 + u_2 = n_1 n_2$.

De las fórmulas para u_1 y u_2 , vemos que u_1 será pequeña cuando w_1 es pequeña y u_2 será pequeña cuando w_2 sea pequeña. En consecuencia, la hipótesis nula se rechazará siempre que las estadísticas apropiadas U_1 , U_2 , o U tomen un valor menor o igual que el valor crítico deseado dado en la tabla A.17. Los diversos procedimientos de prueba se resumen en la tabla 16.2.

Tabla 16.2 Prueba de la suma de rangos

| Para probar H_0 | Contra H_1 | Calcular |
|-------------------|--------------------|----------|
| $\mu_1 = \mu_2$ | $\mu_1 < \mu_2$ | u_1 |
| | $\mu_1 > \mu_2$ | u_2 |
| | $\mu_1 \neq \mu_2$ | u |

La tabla A.17 da valores críticos de U_1 y U_2 para niveles de significancia iguales a 0.001, 0.01, 0.025 y 0.05 para una prueba de una sola cola, y valores críticos de U para niveles de significancia iguales a 0.002, 0.02, 0.05 y 0.10 para una prueba de dos colas. Si el valor observado de u_1 , u_2 , o u es **menor o igual que** el valor crítico tabulado, se rechaza la hipótesis nula en el nivel de significancia que se indica en la tabla. Suponga, por ejemplo, que deseamos probar la hipótesis nula de que $\mu_1 = \mu_2$ contra la alternativa unilateral de que $\mu_1 < \mu_2$ en el nivel de significancia 0.05 para muestras aleatorias de tamaño $n_1 = 3$ y $n_2 = 5$ que dan el valor $w_1 = 8$. Se sigue que

$$u_1 = 8 - \frac{(3)(4)}{2} = 2$$

Nuestra prueba de una sola cola se basa en la estadística U_1 . Con el uso de la tabla A.17, rechazamos la hipótesis nula de medias iguales cuando $u_1 < 1$. Como $u_1 = 2$ cae en la región de aceptación, no se puede rechazar la hipótesis nula.

Ejemplo 16.5 Se encuentra que el contenido de nicotina de dos marcas de cigarrillos, medido en miligramos, es el siguiente:

| | | | | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Marca A | 2.1 | 4.0 | 6.3 | 5.4 | 4.8 | 3.7 | 6.1 | 3.3 | | |
| Marca B | 4.1 | 0.6 | 3.1 | 2.5 | 4.0 | 6.2 | 1.6 | 2.2 | 1.9 | 5.4 |

Pruebe la hipótesis, en el nivel de significancia de 0.05, de que el contenido promedio de nicotina de las dos marcas es igual contra la alternativa de que son diferentes.

SOLUCIÓN

1. $H_0: \mu_1 = \mu_2$.
2. $H_1: \mu_1 \neq \mu_2$.
3. $\alpha = 0.05$.
4. Región crítica: $u \leq 17$ (de la tabla A.17).
5. Cálculos: las observaciones se acomodan en orden ascendente y se les asignan rangos del 1 al 18.

| Datos originales | Rangos | Datos originales | Rangos |
|------------------|--------|------------------|--------|
| 0.6 | 1 | 4.0 | 10.5* |
| 1.6 | 2 | 4.0 | 10.5* |
| 1.9 | 3 | 4.1 | 12 |
| 2.1 | 4* | 4.8 | 13* |
| 2.2 | 5 | 5.4 | 14.5* |
| 2.5 | 6 | 5.4 | 14.5* |
| 3.1 | 7 | 6.1 | 16 |
| 3.3 | 8* | 6.2 | 17 |
| 3.7 | 9* | 6.3 | 18* |

*Los rangos con asterisco pertenecen a la muestra A.

Ahora

$$w_1 = 4 + 8 + 9 + 10.5 + 13 + 14.5 + 18 = 93$$

y

$$w_2 = \left[\frac{(18)(19)}{2} \right] - 93 = 78.$$

Por tanto,

$$u_1 = 93 - \left[\frac{(8)(9)}{2} \right] = 57, \quad u_2 = 78 - \left[\frac{(10)(11)}{2} \right] = 23.$$

6. Decisión: no rechazar H_0 y concluir que no hay diferencia significativa en el contenido promedio de nicotina en las dos marcas de cigarrillos.

Teoría normal de aproximación para dos muestras

Cuando n_1 y n_2 exceden 8, la distribución muestral de U_1 (o U_2) se aproxima a la distribución normal con media

$$\mu_{U_i} = \frac{n_1 n_2}{2} \quad \text{y varianza} \quad \sigma_{U_i}^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

En consecuencia, cuando n_i es mayor que 20, el valor máximo en la tabla A.17, y n_i es al menos 9, se puede utilizar la estadística

$$Z = \frac{U_i - \mu_{U_i}}{\sigma_{U_i}}$$

para nuestra prueba, con la región crítica que cae ya sea en alguna o en ambas colas de la distribución normal, dependiendo de la distribución de H_1 .

El uso de la prueba de suma de rangos de Wilcoxon no se restringe a poblaciones no normales. Se puede utilizar en lugar de la prueba t de dos muestras cuando las poblaciones son normales, aunque la potencia será menor. La prueba de suma de rangos de Wilcoxon siempre es superior a la prueba t para poblaciones decididamente no normales.

16.5 Prueba de Kruskal-Wallis

En los capítulos 13, 14 y 15, la técnica de análisis de varianza resalta como técnica analítica para probar la igualdad de $k \geq 2$ medias poblacionales. De nuevo, sin embargo, el lector debe recordar que se debe suponer la normalidad a fin de que la prueba F sea teóricamente correcta. En esta sección investigamos una alternativa no paramétrica al análisis de varianza.

La prueba de **Kruskal-Wallis**, también llamada **prueba H de Kruskal-Wallis**, es una generalización de la prueba de la suma de rangos para el caso de $k \geq 2$ muestras. Se utiliza para probar la hipótesis nula H_0 de que k muestras independientes son de poblaciones idénticas. Introducida en 1952 por W. H. Kruskal y W. A. Wallis, la prueba es un procedimiento no paramétrico para probar la igualdad de las medias en el análisis de varianza de un factor cuando el experimentador desea evitar la suposición de que las muestras se seleccionaron de poblaciones normales.

Sea n_i ($i = 1, 2, \dots, k$) el número de observaciones en la i -ésima muestra. Primero, combinamos todas las k muestras y acomodamos las $n = n_1 + n_2 + \dots + n_k$ observaciones en orden ascendente, y sustituimos el rango apropiado de $1, 2, \dots, n$ para cada observación. En el caso de empates (observaciones idénticas), seguimos el procedimiento acostumbrado de reemplazar las observaciones por las medias de los rangos que tendrían las observaciones si fueran distinguibles. La suma de los rangos que corresponde a las n_i observaciones en la i -ésima muestra se denota mediante la variable aleatoria R_i . Consideremos ahora la estadística

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1),$$

que se aproxima muy bien mediante una distribución χ^2 cuadrada con $k - 1$ grados de libertad cuando H_0 es verdadera y si cada muestra consiste en al menos 5 observaciones. Nótese que la estadística H toma el valor h , donde

$$h = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{r_i^2}{n_i} - 3(n+1),$$

cuando R_i toma el valor r_i , R_i toma el valor r_i , etcétera. El hecho de que h sea grande cuando las muestras independientes provienen de poblaciones que no son idénticas nos permite establecer el siguiente criterio de decisión para probar H_0 :

Prueba de Kruskal-Wallis

Para probar la hipótesis H_0 de que k muestras independientes son de poblaciones idénticas, calcular

$$h = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{r_i^2}{n_i} - 3(n+1).$$

Si h cae en la región crítica $H > \chi_{\alpha}^2$ con $v = k - 1$ grados de libertad, rechazar H_0 en el nivel de significancia α ; en cualquier otro caso, aceptar H_0 .

Ejemplo 16.6 En un experimento para determinar cuál de tres diferentes sistemas de misiles es preferible, se mide la tasa de utilización del propulsor. Los datos, des-

pués de codificarlos, se dan en la tabla 16.3. Utilice la prueba de Kruskal-Wallis y un nivel de significancia $\alpha = 0.05$ para probar la hipótesis de que las tasas de utilización del propulsor son las mismas para los tres sistemas de misiles.

SOLUCIÓN

1. H_0 : $\mu_1 = \mu_2 = \mu_3$.
2. H_1 : las tres medias son diferentes.
3. $\alpha = 0.05$.
4. Región crítica: $h > \chi_{0.05}^2 = 5.991$, para $v = 2$ grados de libertad.
5. Cálculos: en la tabla 16.4 convertimos las 19 observaciones a rangos y sumamos los rangos para cada sistema de misiles.

Tabla 16.3 Tasas de utilización del propulsor

| Sistemas de misiles: | 1 | 2 | 3 |
|----------------------|------|------|------|
| | 24.0 | 23.2 | 18.4 |
| | 16.7 | 19.8 | 19.1 |
| | 22.8 | 18.1 | 17.3 |
| | 19.8 | 17.6 | 17.3 |
| | 18.9 | 20.2 | 19.7 |
| | | 17.8 | 18.9 |
| | | | 18.8 |
| | | | 19.3 |

Tabla 16.4 Rangos para las tasas de utilización del propulsor

| Sistemas de misiles: | 1 | 2 | 3 |
|----------------------|--------------|--------------|--------------|
| | 19 | 18 | 7 |
| | 1 | 14.5 | 11 |
| | 17 | 6 | 2.5 |
| | 14.5 | 4 | 2.5 |
| | 9.5 | 16 | 15 |
| | $r_1 = 61.0$ | 5 | 9.5 |
| | | $r_2 = 63.5$ | 8 |
| | | | 12 |
| | | | $r_3 = 65.5$ |

Ahora, al sustituir $n_1 = 5$, $n_2 = 6$, $n_3 = 8$, $\bar{y}_1 = 61.0$, $\bar{y}_2 = 63.5$, $\bar{y}_3 = 65.5$, nuestra estadística de prueba H toma el valor

$$h = \frac{12}{(19)(20)} \left(\frac{61.0^2}{5} + \frac{63.5^2}{6} + \frac{65.5^2}{8} \right) - (3)(20) = 1.66$$

6. Decisión: como $h = 1.66$ no cae en la región crítica $h > 5.991$, tenemos insuficiente evidencia para rechazar la hipótesis de que las tasas de utilización del propulsor son las mismas para los tres sistemas de misiles. ■

Ejercicios

1. Un fabricante de cigarrillos afirma que el contenido de alquitrán de la marca de cigarrillos B es menor que la de la marca A. Para probar esta afirmación, se registran las siguientes determinaciones de contenido de alquitrán, en miligramos:

| | | | | | | |
|---------|---|----|---|----|----|----|
| Marca A | 1 | 12 | 9 | 13 | 11 | 14 |
| Marca B | 8 | 10 | 7 | | | |

Utilice la prueba de suma de rangos con $\alpha = 0.05$ para probar si la afirmación es válida.

2. Para determinar si un nuevo suero detendrá la leucemia, se seleccionan nueve pacientes que ya alcanzaron una etapa avanzada de la enfermedad. Cinco pacientes reciben el tratamiento y cuatro no. Los tiempos de sobrevivencia, en años, a partir del momento en que comienza el experimento son

| | | | | | |
|-----------------|-----|-----|-----|-----|-----|
| Con tratamiento | 2.1 | 5.3 | 1.4 | 4.6 | 0.9 |
| Sin tratamiento | 1.9 | 0.5 | 2.8 | 3.1 | |

Utilice la prueba de suma de rangos, en el nivel de significancia de 0.05, para determinar si el suero es efectivo.

3. Los siguientes datos representan el número de horas que operan dos diferentes tipos de calculadoras científicas de bolsillo antes de que requieran recarga.

| | | | | | | | | | |
|---------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Calculadora A | 5.5 | 5.6 | 6.3 | 4.6 | 5.3 | 5.0 | 6.2 | 5.8 | 5.1 |
| Calculadora B | 3.8 | 4.8 | 4.3 | 4.2 | 4.0 | 4.9 | 4.5 | 3.2 | 4.5 |

Utilice la prueba de la suma de rangos con $\alpha = 0.01$ para determinar si la calculadora A opera más tiempo que la calculadora B con una carga completa de la batería.

4. Se fabrica un hilo para pescar por medio de dos procesos. Para determinar si hay una diferencia en la resistencia media de ruptura de los hilos, se seleccionan 10 piezas de cada proceso y después se prueba su resistencia. Los resultados son los siguientes:

| Proceso 1 | Proceso 2 |
|-----------|-----------|
| 10.4 | 8.7 |
| 9.8 | 11.2 |
| 11.5 | 9.8 |
| 10.0 | 10.1 |
| 9.9 | 10.8 |
| 9.6 | 9.5 |
| 10.9 | 11.0 |
| 11.8 | 9.8 |
| 9.3 | 10.5 |
| 10.7 | 9.9 |

Utilice la prueba de suma de rangos con $\alpha = 0.1$ para determinar si hay una diferencia entre las resistencias medias de ruptura de los hilos fabricados por los dos procesos.

5. De una clase de matemáticas de 12 estudiantes con capacidades iguales que utilizan material programado, se seleccionan cinco al azar y el maestro les da instrucción adicional. Los resultados del examen final son los siguientes:

| | Calificación | | | | | |
|---------------------------|--------------|----|----|----|----|----|
| Con instrucción adicional | 87 | 69 | 78 | 91 | 80 | |
| Sin instrucción adicional | 75 | 88 | 64 | 82 | 93 | 79 |

Utilice la prueba de la suma de rangos con $\alpha = 0.05$ para determinar si la instrucción adicional afecta la calificación promedio.

6. Los siguientes datos representan los pesos, en kilogramos, del equipaje personal que llevan en varios vuelos un miembro de un equipo de beisbol y un jugador de un equipo de baloncesto.

Peso del equipaje (kilogramos)

| Jugador de beisbol | Jugador de baloncesto |
|--------------------|-----------------------|
| 16.3 | 20.0 |
| 18.1 | 15.0 |
| 15.9 | 18.6 |
| 14.1 | 14.5 |
| 17.7 | 19.1 |
| 16.3 | 13.8 |
| 13.2 | 17.2 |
| 18.6 | 15.4 |
| 15.6 | 17.7 |
| 18.6 | 16.8 |
| 18.3 | 12.7 |
| 17.4 | 15.0 |
| 14.8 | 15.9 |
| 16.5 | 16.3 |

Utilice la prueba de la suma de rangos con $\alpha = 0.05$ para probar la hipótesis nula de que los dos atletas llevan la misma cantidad de equipaje en promedio contra la hipótesis alternativa de que los pesos promedio del equipaje para los dos atletas son diferentes.

7. Los siguientes datos representan los tiempos de operación en horas para tres tipos de calculadoras científicas de bolsillo antes de que requieran recarga:

| Calculadora | | |
|-------------|-----|-----|
| A | B | C |
| 4.9 | 5.5 | 6.4 |
| 6.1 | 5.4 | 6.8 |
| 4.3 | 6.2 | 5.6 |
| 4.6 | 5.8 | 6.5 |
| 5.3 | 5.5 | 6.3 |
| | 5.2 | 6.6 |
| | 4.8 | |

Utilice la prueba de Kruskal-Wallis, en el nivel de significancia 0.01, para probar la hipótesis de que los tiempos de operación para las tres calculadoras son iguales.

8. En el ejercicio 8 de la página 474 utilice la prueba de Kruskal-Wallis en el nivel de significancia de 0.05 para determinar si los solventes químicos orgánicos difieren de manera significativa en la tasa de absorción.

16.6 Pruebas de corridas

Al aplicar los diversos conceptos estadísticos que se presentan a lo largo de este libro, siempre se supone que nuestros datos muestrales se reúnen mediante algún

procedimiento aleatorio. Las **pruebas de corridas**, que se basan en el orden en el que se obtienen las observaciones muestrales, es una técnica útil para probar la hipótesis nula H_0 de que las observaciones en realidad se extraen al azar.

Para ilustrar las pruebas de corridas, supongamos que se encuestan 12 personas para saber si utilizan cierto producto. Se cuestionaría seriamente la supuesta aleatoriedad de la muestra si las 12 personas fueran del mismo sexo. Designátemos un hombre y una mujer con los símbolos M y F , respectivamente, y registremos los resultados de acuerdo con su sexo en el orden en que ocurren. Una secuencia típica para el experimento podría ser

$$\underline{M} \underline{M} \underline{F} \underline{F} \underline{F} \underline{M} \underline{F} \underline{F} \underline{M} \underline{M} \underline{M} \underline{M}$$

donde agrupamos las subsecuencias de símbolos similares. Tales agrupamientos se llaman **corridas**.

Definición 16.1

Una **corrida** es una subsecuencia de uno o más símbolos idénticos que representan una propiedad común de los datos.

Sin importar si las mediciones de nuestra muestra representan datos cualitativos o cuantitativos, la prueba de corridas divide los datos en dos categorías mutuamente excluyentes: masculino o femenino; defectuoso o no defectuoso; caras o cruces; arriba o abajo de la mediana; etcétera. En consecuencia, una secuencia siempre estará limitada a dos símbolos distintos. Sea n_1 el número de símbolos asociado con la categoría que ocurre menos y n_2 el número de símbolos que pertenecen a la otra categoría. Entonces el tamaño de la muestra $n = n_1 + n_2$.

Para los $n = 12$ símbolos en nuestra encuesta tenemos cinco corridas, con la primera que contiene dos M , la segunda tres F , etcétera. Si el número de corridas es mayor o menor que el que esperaríamos al azar, se debe rechazar la hipótesis de que la muestra se extrajo al azar. Ciertamente, una muestra que tiene como resultado sólo dos corridas,

$$M M M M M M M F F F F F F$$

o la inversa, es más improbable que ocurra a partir de un proceso de selección aleatoria. Tal resultado indica que las primeras siete personas entrevistadas fueron todas hombres seguidas de cinco mujeres. De la misma manera, si la muestra tiene como resultado el número máximo de 12 corridas, como en la secuencia alternante

$$M F M F M F M F M F M F$$

de nuevo sospecharíamos del orden en que se seleccionaron los individuos para la encuesta.

La prueba de corridas para la aleatoriedad se basa en la variable aleatoria V , el número total de corridas que ocurren en la secuencia completa de nuestro experimento. En la tabla A.18, se dan valores de $P(V \leq v^*$ cuando H_0 es verdadera) para $v^* = 2, 3, \dots, 20$ corridas, y valores de n_1 y n_2 menores que o iguales a 10. Los valores P para pruebas de una cola y de dos colas se pueden obtener con el uso de estos valores tabulados.

En la encuesta anterior exhibimos un total de 5 *F* y 7 *M*. De aquí, con $n_1 = 5$, $n_2 = 7$, y $v = 5$, notamos de la tabla A.18 para una prueba de dos colas que el valor P es:

$$P = 2P(V \leq 5 \text{ cuando } H_0 \text{ es real}) = 0.394 > 0.05.$$

Es decir, el valor $v = 5$ es razonable en el nivel de significancia de 0.05 cuando H_0 es verdadera, y por tanto no tenemos suficiente evidencia para rechazar la hipótesis de aleatoriedad en nuestra muestra.

Cuando el número de corridas es grande, por ejemplo si $v = 11$ y $n_1 = 5$ y $n_2 = 7$, entonces el valor P en una prueba de dos colas es

$$\begin{aligned} P &= 2P(V \geq 11 \text{ cuando } H_0 \text{ es real}) = 2[1 - P(V \leq 10) \text{ cuando } H_0 \text{ es real}] \\ &= 2(1 - 0.992) = 0.016 < 0.05, \end{aligned}$$

que nos lleva a rechazar la hipótesis de que los valores de la muestra ocurren al azar.

La prueba de corridas también se puede utilizar para detectar desviaciones en la aleatoriedad de una secuencia de mediciones cuantitativas en el tiempo, ocasionadas por tendencias o periodicidades. Al reemplazar cada medición en el orden en el que se obtienen por un símbolo *más* si caen por arriba de la mediana, por un símbolo *menos* si caen por debajo de la mediana, y omitimos todas las mediciones que son exactamente iguales a la mediana, generamos una secuencia de símbolos más y menos que se prueban por su aleatoriedad como se ilustra en el siguiente ejemplo.

Ejemplo 16.7 Se ajusta una máquina para servir adelgazador de pintura acrílica en un contenedor. ¿Diría que la cantidad de adelgazador de pintura que despacha la máquina varía de forma aleatoria si se mide el contenido de los siguientes 15 contenedores y se encuentra que es 3.6, 3.9, 4.1, 3.6, 3.8, 3.7, 3.4, 4.0, 3.8, 4.1, 3.9, 4.0, 3.8, 4.2 y 4.1 litros? Utilice un nivel de significancia de 0.1.

SOLUCIÓN

1. H_0 : la secuencia es aleatoria.
2. H_1 : la secuencia no es aleatoria.
3. $\alpha = 0.1$.
4. Estadística de prueba: V , número total de corridas.
5. Cálculos: para la muestra dada encontramos $\bar{x} = 3.9$. Al reemplazar cada medición por el símbolo "+" si cae por arriba de 3.9, por el símbolo "-" si cae por debajo de 3.9, y omitimos las dos mediciones que son iguales a 3.9, obtenemos la secuencia:

- + - - - + + + + - + +

para la que $n_1 = 6$, $n_2 = 7$, y $v = 6$. Por tanto, de la tabla A.18, el valor P calculado es:

$$P = 2P(V \leq 6 \text{ cuando } H_0 \text{ es real}) = 0.592 > 0.1,$$

6. Decisión: no rechazar la hipótesis de que la secuencia de mediciones varía de forma aleatoria. ■

La prueba de corridas, aunque menos poderosa, también se puede utilizar como una alternativa para la prueba de dos muestras de Wilcoxon para probar la afirmación de que dos muestras aleatorias provienen de poblaciones que tienen la misma distribución y por tanto medias iguales. Si las poblaciones son simétricas, el rechazo de la afirmación de distribuciones iguales es equivalente a aceptar la hipótesis alternativa de que las medias no son iguales. Al realizar la prueba, primero combinamos las observaciones de ambas muestras y las acomodamos en orden ascendente. Asignamos ahora la letra *A* a cada observación tomada de una de las poblaciones y la letra *B* a cada observación de la segunda población, con lo que se genera una secuencia que consiste en los símbolos *A* y *B*. Si las observaciones de una población empatan con las observaciones de la otra población, la secuencia de símbolos *A* y *B* que se genera no será única y en consecuencia es poco probable que el número de corridas sea único. Los procedimientos para quitar los empates por lo general tienen como resultado tediosos cálculos adicionales, y por esta razón se puede preferir la aplicación de la prueba de la suma de rangos de Wilcoxon siempre que ocurran estas situaciones.

Para ilustrar el uso de las corridas en la prueba de medias iguales, considere los tiempos de sobrevivencia de los pacientes de leucemia del ejercicio 2 de la página 624 para los que tenemos

| | | | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0.5 | 0.9 | 1.4 | 1.9 | 2.1 | 2.8 | 3.1 | 4.6 | 5.3 |
| <i>B</i> | <i>A</i> | <i>A</i> | <i>B</i> | <i>A</i> | <i>B</i> | <i>B</i> | <i>A</i> | <i>A</i> |

de donde resultan $v = 6$ corridas. Si las dos poblaciones simétricas tienen medias iguales, las observaciones de las dos muestras estarán entremezcladas, lo que resultará en muchas corridas. Sin embargo, si las medias poblacionales son significativamente diferentes, esperaríamos que la mayoría de las observaciones para una de las dos muestras fueran más pequeñas que las de la otra muestra. En el caso extremo donde las poblaciones no se traslapan, obtendríamos una secuencia de la forma

A A A A A B B B B o *B B B B A A A A A*

y en cualquier caso sólo hay dos corridas. En consecuencia, la hipótesis de medias poblacionales iguales se rechazará en el nivel de significancia α sólo cuando v es suficientemente pequeña de modo que

$$P = P(V \leq v \text{ cuando } H_0 \text{ es real}) \leq \alpha,$$

lo que implica una prueba de una cola.

De regreso a los datos del ejercicio 2 de la página 624 para los que $n_1 = 4$, $n_2 = 5$, y $v = 6$, encontramos de la tabla A.18 que

$$P = P(V \leq 6 \text{ cuando } H_0 \text{ es real}) = 0.786 > 0.05$$

y por tanto se falla al rechazar la hipótesis nula de medias iguales. De aquí concluimos que el nuevo suero no prolonga la vida al no detener la leucemia.

Cuando n_1 y n_2 aumentan en tamaño, la distribución de muestreo de V se aproxima a la distribución normal con media

$$\mu_V = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad \text{y varianza} \quad \sigma_V^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

En consecuencia, cuando n_1 y n_2 son ambos mayores que 10, se puede utilizar la estadística

$$Z = \frac{V - \mu_2}{\sigma_2}$$

para establecer la región crítica para la prueba de corridas.

16.7 Límites de tolerancia

Los límites de tolerancia para una distribución normal de mediciones se presenta en el capítulo 9. En esta sección consideramos un método para construir intervalos de tolerancia independientes de la forma de la distribución subyacente. Como se podría sospechar, para un grado de confianza razonable serán considerablemente más grandes que los que se construyen cuando se supone la normalidad, y el tamaño de la muestra que se requiere es por lo general muy grande. Los límites de tolerancia no paramétricos se establecen en función de la observación más grande y de la más pequeña en nuestra muestra.

Límites de tolerancia bilaterales

Para cualquier distribución de mediciones, los límites de tolerancia bilaterales se indican mediante las observaciones más pequeña y más grande en una muestra de tamaño n , donde n se determina de modo que se pueda afirmar con una confianza de $100(1 - \gamma)\%$ que **al menos** la proporción $1 - \alpha$ de la distribución está incluida entre los extremos de la muestra.

La tabla A.19 da los tamaños muestrales que se requieren para valores seleccionados de γ y $1 - \alpha$. Por ejemplo, cuando $\gamma = 0.01$ y $1 - \alpha = 0.95$, debemos elegir una muestra aleatoria de tamaño $n = 130$ a fin de tener 99% de confianza de que al menos 95% de la distribución de mediciones está incluida entre los extremos de la muestra.

En lugar de determinar el tamaño muestral n de modo que una proporción específica de mediciones esté contenida entre los extremos de la muestra, es deseable en muchos procesos industriales determinar el tamaño de la muestra de forma que una proporción fija de la población caiga por debajo de la observación más grande (o por arriba de la más pequeña) de la muestra. Tales límites se llaman límites de tolerancia unilaterales.

Límites de tolerancia unilaterales

Para cualquier distribución de mediciones, un límite de tolerancia unilateral se determina mediante la observación más pequeña (más grande) en una muestra de tamaño n , donde n se determina de modo que se pueda asegurar con $100(1 - \gamma)\%$ de confianza que **al menos** la proporción $1 - \alpha$ de la distribución excederá la más pequeña (será menor que la mayor) observación de la muestra.

La tabla A.20 muestra los tamaños muestrales requeridos correspondientes a valores seleccionados de γ y $1 - \alpha$. De aquí, cuando $\gamma = 0.05$ y $1 - \alpha = 0.70$, debemos elegir una muestra de tamaño $n = 9$ a fin de tener una confianza de 95% de que 70% de nuestra distribución de mediciones excederá la observación más pequeña de la muestra.

16.8 Coeficiente de correlación de rango

En el capítulo 11 utilizamos el coeficiente de correlación muestral r para medir la relación lineal entre dos variables continuas X y Y . Si los rangos $1, 2, \dots, n$ se asignan a las observaciones x en orden de magnitud y de manera similar a las observaciones y , y si estos rangos se sustituyen después por los valores numéricos reales en la fórmula para el coeficiente de correlación del capítulo 11, obtenemos la contraparte no paramétrica del coeficiente de correlación convencional. Un coeficiente de correlación calculado de esta forma se conoce como **coeficiente de correlación de rangos de Spearman** y se denota por r_s . Cuando no hay empates entre ambos conjuntos de mediciones, la fórmula para r_s se reduce a una expresión mucho más simple que incluye las diferencias d_i entre los rangos asignados a los n pares de x y y , que establecemos ahora.

Coeficiente de correlación de rangos

Una medición no paramétrica de la asociación entre dos variables X y Y está dada por el **coeficiente de correlación de rangos**

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

donde d_i es la diferencia entre los rangos asignados x_i y y_i , y n es el número de pares de datos.

En la práctica la fórmula anterior también se usa cuando hay empates entre las observaciones x o y . Los rangos para observaciones empatadas se asignan como en la prueba del rango con signo al promediar los rangos que se habrían asignado si las observaciones fuesen distinguibles.

El valor de r_s por lo general estará cercano al valor que se obtiene al encontrar r con base en mediciones numéricas y se interpreta casi en la misma forma. Como antes, el valor de r_s irá de -1 a $+1$. Un valor de $+1$ o -1 indica una asociación perfecta entre X y Y , el signo más ocurre para rangos idénticos y el signo menos para rangos inversos. Cuando r_s es cercano a cero, concluiremos que las variables no están correlacionadas.

Ejemplo 16.8 Las cifras que se listan en la tabla 16.5, publicadas por la Comisión Federal de Comercio, muestran los miligramos de alquitrán y nicotina que se encuentra en 10 marcas de cigarrillos. Calcule el coeficiente de correlación de rangos para medir el grado de relación entre el contenido de alquitrán y nicotina en cigarrillos.

SOLUCIÓN

Representemos con X y Y los contenidos de alquitrán y nicotina, respectivamente. Primero asignamos rangos a cada conjunto de medidas, con el rango de 1 asignado al número más bajo en cada conjunto, el rango de 2 al segundo número más bajo en cada conjunto, y así sucesivamente, hasta que se asigna el rango 10 al número más grande. La tabla 16.6 muestra los rangos individuales de las mediciones y las diferencias en rangos para los 10 pares de observaciones.

Tabla 16.5 Contenidos de alquitrán y nicotina

| Marca de cigarrillos | Contenido de alquitrán | Contenido de nicotina |
|----------------------|------------------------|-----------------------|
| Viceroy | 14 | 0.9 |
| Marlboro | 17 | 1.1 |
| Chesterfield | 28 | 1.6 |
| Kool | 17 | 1.3 |
| Kent | 16 | 1.0 |
| Raleigh | 13 | 0.8 |
| Old Gold | 24 | 1.5 |
| Philip Morris | 25 | 1.4 |
| Oasis | 18 | 1.2 |
| Players | 31 | 2.0 |

Tabla 16.6 Rangos para los contenidos de alquitrán y nicotina

| Marca de cigarrillos | x_i | y_i | d_i |
|----------------------|-------|-------|-------|
| Viceroy | 2 | 2 | 0 |
| Marlboro | 4.5 | 4 | 0.5 |
| Chesterfield | 9 | 9 | 0 |
| Kool | 4.5 | 6 | -1.5 |
| Kent | 3 | 3 | 0 |
| Raleigh | 1 | 1 | 0 |
| Old Gold | 7 | 8 | -1 |
| Philip Morris | 8 | 7 | 1 |
| Oasis | 6 | 5 | 1 |
| Players | 10 | 10 | 0 |

Al sustituir en la fórmula para r_s , encontramos que

$$r_s = 1 - \frac{(6)(5.75)}{(10)(100 - 1)} = 0.965,$$

lo que indica una correlación positiva alta entre la cantidad de alquitrán y de nicotina que se encuentra en los cigarrillos.

Existen algunas ventajas en el uso de r_s en lugar de r . Por ejemplo, ya no suponemos que la relación fundamental entre X y Y es lineal y por tanto, cuando los datos poseen una relación curvilínea distinta, el coeficiente de correlación de rangos probablemente será más confiable que la medición convencional. Una segunda ventaja del uso del coeficiente de correlación de rangos es el hecho de que no se hacen suposiciones de normalidad con respecto a las distribuciones de X y Y . Quizá la mayor ventaja ocurre cuando se es incapaz de hacer mediciones numéricas significativas sin embargo se pueden establecer rangos. Tal es el caso, por ejemplo, cuando diferentes jueces clasifican un grupo de individuos de acuerdo con algún atributo. El coeficiente de correlación de rangos se puede utilizar en esta situación como una medida de la consistencia de los dos jueces.

Para probar la hipótesis de que $\rho = 0$ con el uso de un coeficiente de correlación de rangos, se necesita considerar la distribución muestral de los valores r_s bajo

la suposición de no correlación. En la tabla A.21 aparecen valores críticos calculados para $\alpha = 0.05, 0.025, 0.01,$ y 0.05 . La elaboración de esta tabla es similar a la tabla de valores críticos para la distribución t excepto para la columna izquierda, que ahora da el número de pares de observaciones en lugar de los grados de libertad. Como la distribución de los valores r_s es simétrica alrededor de cero cuando $\rho = 0$, el valor r_s que deja un área de α a la izquierda es igual al negativo del valor r_s que deja un área α a la derecha. Para una hipótesis alternativa bilateral, la región crítica de tamaño α cae igualmente en las dos colas de la distribución. Para una prueba en la que la hipótesis alternativa es negativa, la región crítica está completamente en la cola izquierda de la distribución y, cuando la alternativa es positiva, la región crítica se coloca por completo en la cola derecha. ■

Ejemplo 16.9 Refiérase al ejemplo 16.8 y pruebe la hipótesis de que la correlación entre la cantidad de alquitrán y nicotina en los cigarrillos es cero contra la alternativa de que es mayor que cero. Utilice un nivel de significancia de 0.01.

SOLUCIÓN

1. $H_0: \rho = 0$.
2. $H_1: \rho > 0$.
3. $\alpha = 0.01$.
4. Región crítica: $r_s > 0.745$, de la tabla A.21.
5. Cálculos: del ejemplo 16.8, $r_s = 0.965$.
6. Decisión: rechazar H_0 y concluir que hay una correlación o relación significativa entre la cantidad de alquitrán y nicotina que se encuentra en los cigarrillos. ■

Bajo la suposición de no correlación, se puede mostrar que la distribución de los valores r_s se aproxima a una distribución normal con una media 0 y desviación estándar de $1/\sqrt{n-1}$ conforme n aumenta. En consecuencia, cuando n excede los valores dados en la tabla A.21, se podría probar la correlación de significancia mediante el cálculo de

$$z = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s \sqrt{n-1}$$

y la comparación con los valores críticos de la distribución normal estándar que se muestran en la tabla A.3.

Ejercicios

1. Se selecciona una muestra aleatoria de 15 adultos que viven en una pequeña ciudad para estimar la proporción de votantes que favorecen a cierto candidato para alcalde. También se le preguntó a cada individuo si era graduado universitario. Se obtiene la siguiente secuencia, al hacer que Y y N designen las respuestas de "sí" y "no" a la pregunta sobre instrucción:

A N N N Y Y N Y Y N Y N N N

Utilice la prueba de corridas en el nivel de significancia de 0.1 para determinar si la secuencia apoya la afirmación de que la muestra se seleccionó al azar.

2. Se utiliza un proceso de ploteado para cubrir cierto tipo de charola de servicio. Cuando el proceso es-

12 bajo control, el espesor de la plata sobre la charola variará de forma aleatoria siguiendo una distribución normal con una media de 0.02 milímetros y una desviación estándar de 0.005 milímetros. Suponga que las siguientes 12 charolas examinadas muestran los siguientes espesores de plata: 0.019, 0.021, 0.020, 0.019, 0.020, 0.018, 0.023, 0.021, 0.024, 0.022, 0.023, 0.022. Utilice la prueba de corridas para determinar si las fluctuaciones en el espesor de una charola a otra son aleatorias. Sea $\alpha = 0.05$.

3. Use la prueba de corridas para probar si hay una diferencia en el tiempo promedio de operación para las dos calculadoras del ejercicio 3 en la página 625.

4. En una línea de producción industrial, los artículos se inspeccionan de forma periódica en busca de defectuosos. Lo siguiente es una secuencia de artículos defectuosos, D , y no defectuosos, N , producidos por esta línea:

D D N N N D N N D D N N N N
N D D D N N D N N N N D N D

Utilice la teoría de muestras grandes para la prueba de corridas, con un nivel de significancia de 0.05, para determinar si los defectuosos ocurren o no al azar.

5. Suponga que las mediciones del ejercicio 2 de la página 68 se registran en renglones sucesivos de izquierda a derecha conforme se reúnen. Utilice la prueba de corridas, con $\alpha = 0.05$, para probar la hipótesis de que los datos representan una secuencia aleatoria.

6. ¿Qué tan grande se requiere que sea una muestra para tener 95% de confianza de que al menos 85% de la distribución de medidas se incluye entre los extremos de la muestra?

7. ¿Cuál es la probabilidad de que el rango de una muestra aleatoria de tamaño 24 incluya al menos 90% de la población?

8. ¿Qué tan grande se requiere que sea una muestra para tener 99% de confianza de que al menos 80% de la población sea menor que la observación más grande de la muestra?

9. ¿Cuál es la probabilidad de que al menos 98% de una población exceda el valor más pequeño en una muestra aleatoria de tamaño $n = 135$?

10. La siguiente tabla da las calificaciones registradas de 10 estudiantes en un examen de mitad del trimestre y la del examen final en un curso de cálculo.

(a) Calcule el coeficiente de correlación de rangos.

| Estudiante | Examen de mitad del trimestre | Examen final |
|------------|-------------------------------|--------------|
| L.S.A. | 84 | 73 |
| W.P.B. | 98 | 63 |
| R.W.K. | 91 | 87 |
| J.R.L. | 72 | 66 |
| J.K.L. | 86 | 78 |
| D.L.P. | 93 | 78 |
| B.L.P. | 80 | 91 |
| D.W.M. | 0 | 0 |
| M.N.M. | 92 | 88 |
| R.H.S. | 87 | 77 |

(b) Pruebe la hipótesis nula de que $\rho = 0$ contra la alternativa de que $\rho > 0$. Utilice $\alpha = 0.025$.

11. Con referencia a los datos del ejercicio 1 de la página 363,

(a) calcule el coeficiente de correlación de rangos,
 (b) pruebe la hipótesis nula en el nivel de significancia 0.05 de que $\rho = 0$ contra la alternativa de que $\rho \neq 0$. Compare sus resultados con los que se obtienen en el ejercicio 5 de la página 400.

12. Calcule el coeficiente de correlación de rangos para la precipitación pluvial diaria y la cantidad de partículas eliminadas en el ejercicio 9 de la página 364.

13. Con referencia a los pesos y tamaños del tórax de los infantes del ejercicio 4 en la página 400,

(a) calcule el coeficiente de correlación de rangos;
 (b) pruebe la hipótesis en el nivel de significancia 0.025 de que $\rho = 0$ contra la alternativa de que $\rho > 0$.

14. Un grupo de consumidores prueba la calidad general de nueve marcas de hornos de microondas. Los rangos asignados por el grupo y los precios de venta sugeridos son los siguientes:

| Fabricante | Clasificación del grupo | Precio sugerido |
|------------|-------------------------|-----------------|
| A | 6 | \$480 |
| B | 9 | 395 |
| C | 2 | 575 |
| D | 8 | 350 |
| E | 5 | 510 |
| F | 1 | 545 |
| G | 7 | 400 |
| H | 4 | 465 |
| I | 3 | 420 |

¿Existe una relación significativa entre la calidad y el precio de un horno de microondas? Utilice un nivel de significancia de 0.05.

15. Dos jueces en un desfile de regreso a clases califican ocho carros alegóricos en el siguiente orden:

| | Carro alegórico: | | | | | | | |
|--------|------------------|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Juez A | 5 | 8 | 4 | 3 | 6 | 2 | 7 | 1 |
| Juez B | 7 | 5 | 4 | 2 | 8 | 1 | 6 | 3 |

- (a) Calcule la correlación de rangos.
 (b) Pruebe la hipótesis nula de que $\rho = 0$ contra la alternativa de que $\rho > 0$. Use $\alpha = 0.05$.

16. En el artículo titulado *Risky Assumptions* de Paul Slovic, Baruch Fischhoff y Sarah Lichtenstein, publicado en *Psychology Today* (junio de 1980), miembros de la Liga de Mujeres Votantes y expertos profesionalmente involucrados en la evaluación de riesgos clasifican el riesgo de muerte en Estados Unidos para 30 actividades y tecnologías.

- (a) Calcule el coeficiente de correlación de rangos.
 (b) Pruebe la hipótesis nula de cero correlación entre las clasificaciones de la Liga de Mujeres Votantes y de los expertos contra la alternativa de que la correlación no es cero. Utilice un nivel de significancia de 0.05.

Las clasificaciones se muestran en la tabla siguiente.

| Riesgo de la actividad o tecnología | Votantes | Expertos |
|-------------------------------------|----------|----------|
| Energía nuclear | 1 | 20 |
| Vehículos de motor | 2 | 1 |
| Pistolas | 3 | 4 |
| Tabaquismo | 4 | 2 |
| Motocicletas | 5 | 6 |
| Bebidas alcohólicas | 6 | 3 |
| Aviación privada | 7 | 12 |
| Traabajo policiaco | 8 | 17 |
| Pesticidas | 9 | 8 |
| Cirugía | 10 | 5 |
| Bombero | 11 | 18 |
| Construcción | 12 | 13 |
| Cacería | 13 | 23 |
| Latas de aerosol | 14 | 26 |
| Montañismo | 15 | 29 |
| Bicicletas | 16 | 15 |
| Aviación comercial | 17 | 16 |
| Energía eléctrica | 18 | 9 |
| Natación | 19 | 10 |
| Anticonceptivos | 20 | 11 |
| Esquí | 21 | 30 |
| Rayos X | 22 | 7 |
| Fútbol | 23 | 27 |
| Ferrocarriles | 24 | 19 |
| Conservadores de alimentos | 25 | 14 |
| Colorantes de alimentos | 26 | 21 |
| Podadoras | 27 | 28 |
| Antibióticos | 28 | 24 |
| Electrodomésticos | 29 | 22 |
| Vacunas | 30 | 25 |

Ejercicios de repaso

1. Un estudio de una compañía química compara las propiedades de desecación de dos diferentes polímeros. Se utilizan 10 lodos diferentes y se permite que ambos polímeros secan cada lodo. El secado libre se mide en ml/min.

| Tipo de lodo | Polímero A | Polímero B |
|--------------|------------|------------|
| 1 | 12.7 | 12.0 |
| 2 | 14.6 | 15.0 |
| 3 | 18.6 | 19.2 |
| 4 | 17.5 | 17.3 |
| 5 | 11.8 | 12.2 |
| 6 | 16.9 | 16.6 |
| 7 | 19.9 | 20.1 |
| 8 | 17.6 | 17.6 |
| 9 | 15.6 | 16.0 |
| 10 | 16.0 | 16.1 |

(a) Utilice la prueba de signos en el nivel 0.05 para probar la hipótesis nula de que el polímero A tiene la misma mediana de secado que el polímero B.

(b) Utilice la prueba de rangos con signo para probar la hipótesis de la parte (a).

2. En el ejercicio 5 de la página 524, use la prueba de Kruskal-Wallis, en el nivel de significancia de 0.05, para determinar si los análisis químicos realizados por los cuatro laboratorios dan, en promedio, los mismos resultados.

3. Use los datos del ejercicio 2 de la página 485 para ver si la cantidad mediana de pérdida de nitrógeno en la transpiración es diferente para los tres niveles de proteína dietética.

Control estadístico de calidad

17.1 Introducción

La noción del uso de las técnicas de muestreo y de análisis estadístico en un escenario de producción tiene sus comienzos en la década de 1920. El objetivo de este concepto altamente exitoso es la reducción sistemática de la variabilidad y el aislamiento asociado de las fuentes de dificultades *durante la producción*. En 1924, Walter A. Shewhart de la empresa Bell Telephone Laboratories desarrolló el concepto de una gráfica de control. Sin embargo, no fue sino hasta la Segunda Guerra Mundial que se generalizó el uso de gráficas de control. Esto se debió a la importancia de mantener la calidad en los procesos de producción durante dicho periodo. En las décadas de 1950 y 1960, el desarrollo del control de calidad y el área general de seguridad de la calidad crecieron de manera rápida, en particular con el surgimiento del programa espacial en Estados Unidos. En Japón hay un amplio y exitoso uso del control de calidad gracias a los esfuerzos de W. Edwards Deming, quien trabajó como consultor en Japón después de la Segunda Guerra Mundial. El control de calidad ha sido, y es, un ingrediente importante en el desarrollo de la industria y la economía de Japón.

El control de calidad recibe una creciente atención como herramienta de administración en la que importantes características de un producto se observan, evalúan y comparan con algún tipo de estándar. Los diversos procedimientos en el control de calidad implican un uso considerable de los procedimientos de muestreo y principios estadísticos ya presentados en capítulos anteriores. Los usuarios principales del control de calidad son, por supuesto, las corporaciones industriales. Es claro que un programa efectivo de control de calidad aumenta la calidad del artículo que se produce

y las utilidades. Esto es en particular cierto hoy en día pues los productos se fabrican en volúmenes altos. Antes del movimiento hacia los métodos de control, la calidad a menudo sufrió debido a la falta de eficiencia que, por supuesto, aumenta el costo.

Gráfica de control

El propósito de una gráfica de control es determinar si el comportamiento de un proceso se mantiene en un nivel aceptable de calidad. Se espera, por supuesto, que cualquier proceso experimente una variabilidad natural, es decir, variabilidad debido a fuentes de variación poco importantes e incontrolables. Por otro lado, un proceso puede experimentar tipos más serios de variabilidad en mediciones clave de comportamiento. Estas fuentes de variabilidad pueden surgir de uno de varios tipos de "causas asignables" no aleatorias, como errores del operador o indicadores mal ajustados en una máquina. Un proceso que opera en este estado se denomina **fuera de control**. Se dice que un proceso que experimenta sólo variaciones aleatorias está en **control estadístico**. Por supuesto, un proceso de producción exitoso puede operar en un estado de control durante un período grande. Se supone que durante este período el proceso elabora un producto aceptable. Sin embargo, puede haber un "corrimiento" gradual o súbito que requiere detección.

Una gráfica de control tiene la intención de ser un dispositivo para detectar el estado no aleatorio o fuera de control de un proceso. Por lo general, la gráfica de control toma la forma que se indica en la figura 17.1. Es importante que el corrimiento se detecte de forma rápida de modo que se pueda corregir el problema. Obviamente, si la detección es lenta se producen muchos artículos defectuosos o fuera de las especificaciones, lo que tiene como resultado un considerable desperdicio y un costo aumentado.

Se deben considerar algunos tipos de características de la calidad y las unidades del proceso se deben muestrear conforme pasa el tiempo. Por ejemplo, la característica puede ser la circunferencia de un cojinete de motor. La línea central representa el valor promedio de la característica cuando el proceso está controlado. Los puntos que se indican en la figura pueden representar los resultados de, digá-

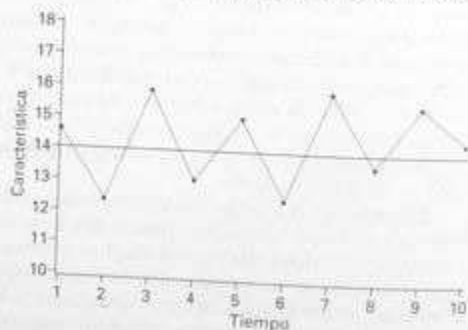


Figura 17.1 Gráfica de control típica.

mos, promedios muestrales de esta característica, con las muestras tomadas respecto al tiempo. Los límites de control superior e inferior se eligen de modo que se esperaría que todos los puntos muestrales queden cubiertos por estos límites si el proceso está controlado. Como resultado, la forma general de los puntos graficados respecto al tiempo determina si se concluye que el proceso está dentro de control. La evidencia para estar "dentro de control" la produce un patrón aleatorio de puntos; con todos los valores graficados dentro de los límites de control. Cuando un punto cae fuera de los límites de control, esto se toma como evidencia de un proceso que está fuera de control, y se sugiere una búsqueda de la causa. Además, un patrón no aleatorio de puntos se puede considerar sospechoso y en realidad una indicación de que se necesita una investigación de la acción correctiva apropiada.

17.2 Naturaleza de los límites de control

Las ideas fundamentales en las que se basan las gráficas de control son similares en estructura a la prueba de hipótesis. Los límites de control se establecen para controlar la probabilidad de cometer el error de concluir que el proceso está fuera de control cuando de hecho no lo está. Esto corresponde a la probabilidad de cometer un error tipo I si probáramos la hipótesis nula de que el proceso está bajo control. Por otro lado, debemos estar atentos al error del segundo tipo, a saber, no encontrar el proceso fuera de control cuando de hecho lo está (error tipo II). De esta manera la elección de los límites de control es similar a la elección de una región crítica.

Como en el caso de la prueba de hipótesis, el tamaño de la muestra en cada punto es importante. La consideración del tamaño de la muestra depende en gran medida de la sensibilidad o potencia de detección del estado fuera de control. En esta aplicación, la noción de *potencia* es muy similar a la situación de la prueba de hipótesis. Es claro que entre más grande sea la muestra en cada periodo, más rápida es la detección de un proceso fuera de control. En un sentido, los límites de control en realidad definen lo que el usuario considera como *bajo control*. En otras palabras, la amplitud dada por los límites de control obviamente debe depender en cierto sentido de la variabilidad del proceso. Como resultado, el cálculo de los límites de control dependerá por completo de manera natural de los datos que se tomen de los resultados del proceso. De esta forma cualquier control de calidad debe tener su comienzo en el cálculo a partir de una muestra preliminar o conjunto de muestras que establecerán la línea central y los límites de control de calidad.

17.3 Propósitos de la gráfica de control

Un propósito obvio de la gráfica de control es la mera vigilancia del proceso; es decir, determinar si se necesitan realizar cambios. Además, la constante obtención sistemática de datos a menudo permite a la administración evaluar la capacidad del proceso. Claramente, si una sola característica de comportamiento es importante, el muestreo continuo y la estimación de la media y la desviación estándar de la característica de comportamiento ofrece la actualización de lo que el proceso puede hacer en términos de comportamiento medio y variación aleatoria. Esto es valioso aun si el proceso permanece bajo control durante periodos largos. La estructura siste-

mática y formal de la gráfica a menudo puede prevenir una reacción desmesurada a cambios que representen sólo fluctuaciones aleatorias. Obviamente, en muchas situaciones, los cambios realizados por una reacción desmesurada pueden crear problemas serios difíciles de resolver.

Las características de calidad de las gráficas de control caen por lo general en dos categorías, **variables** y **atributos**. Como resultado, los tipos de gráficas de control con frecuencia tienen las mismas clasificaciones. En el caso de las gráficas de tipo variables, la característica por lo general es una medición sobre un continuo, como el diámetro, peso, etcétera. Para la gráfica de atributos, la característica refleja el producto individual *concreto* (es o no defectuoso). Las aplicaciones de estas dos situaciones distintas son obvias.

En el caso de la gráfica de variables, se debe ejercer control sobre la tendencia central y la variabilidad. Un analista de control de calidad se debe preocupar de si existe *en promedio* un corrimiento de los valores de la característica de comportamiento. Además, siempre habrá una preocupación acerca de si algún cambio en las condiciones del proceso tiene como resultado una disminución en la precisión (es decir, un aumento en la variabilidad). Para tratar con estos dos conceptos son esenciales gráficas de control separadas. La tendencia central está controlada por la gráfica \bar{X} , donde las medias de muestras relativamente pequeñas se grafican en la gráfica de control. La variabilidad alrededor de la media se controla por el *rango* en la muestra, o la *desviación estándar* muestral. En el caso de muestreo de atributos, la *proporción de defectuosos* de una muestra a menudo es la cantidad que se grafica. En la siguiente sección presentamos el desarrollo de gráficas de control para la característica de comportamiento tipo *variable*.

17.4 Gráficas de control para variables

Un ejemplo es una forma relativamente fácil para comprender los rudimentos de la gráfica \bar{X} para variables. Suponga que las gráficas de control de calidad se deben utilizar en un proceso de fabricación de cierta parte de un motor. Suponga que la media del proceso es $\mu = 50$ mm y que la desviación estándar es $\sigma = 0.01$ mm. Suponga que se muestrean grupos de cinco cada hora y que los valores de la media

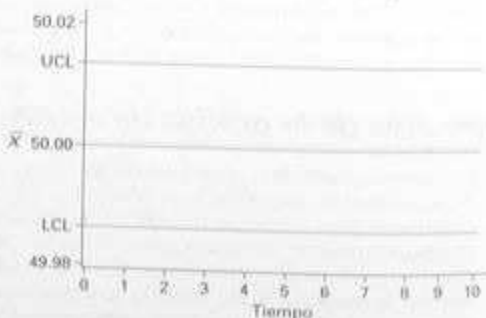


Figura 17.2. Los límites de control 3σ para el ejemplo de la parte del motor.

muestra \bar{X} se registran y grafican como en la figura 17.2. Los límites para las gráficas \bar{X} se basan en la desviación estándar de la variable aleatoria \bar{X} . Sabemos del material del capítulo 8 que para el promedio de observaciones independientes en una muestra de tamaño n ,

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

donde σ es la desviación estándar de una observación individual. Los límites de control se determinan de modo que tengan como resultado una pequeña probabilidad de que un valor dado de \bar{X} esté fuera de los límites dado que, en realidad, el proceso está bajo control (es decir, $\mu = 50$). Si recurrimos al teorema del límite central, tenemos que bajo la condición de que el proceso está bajo control,

$$\bar{X} \sim N\left(50, \frac{0.01}{\sqrt{5}}\right).$$

Como resultado, $100(1 - \alpha)\%$ de los valores \bar{X} cae dentro de los límites cuando el proceso está bajo control si utilizamos los límites

$$LCL = \mu - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 50 - z_{0.0045}(0.0045), \quad UCL = \mu + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 50 + z_{0.0045}(0.0045).$$

Con frecuencia las gráficas \bar{X} se basan en límites que se denominan como límites "tres-sigma", con referencia, por supuesto, a $z_{\alpha/2} = 3$ y a límites que se convierten en

$$\mu \pm 3 \frac{\sigma}{\sqrt{n}}.$$

En nuestro ejemplo los límites superior e inferior son

$$UCL = 50 + 3(0.0045) = 50.0135, \quad LCL = 50 - 3(0.0045) = 49.9865.$$

De esta forma, si vemos la estructura de los límites 3σ desde el punto de vista de la prueba de hipótesis, para un punto muestral dado, la probabilidad de que el valor \bar{X} caiga fuera de los límites de control es 0.0026, dado que el proceso está bajo control. Esta es la probabilidad de que el analista determine de manera errónea que el proceso está fuera de control (véase la tabla A.3).

El ejemplo anterior no sólo ilustra la gráfica \bar{X} para variables, sino también debe proporcionar al lector una idea de la naturaleza de las gráficas de control en general. La línea central por lo general refleja el valor ideal de un parámetro importante. Los límites de control se establecen a partir del conocimiento de las propiedades de muestreo de la estadística que estima el parámetro en cuestión. Muy a menudo éstos implican un múltiplo de la desviación estándar de la estadística. Se ha vuelto una práctica general utilizar límites 3σ . En el caso de la gráfica \bar{X} que se proporciona aquí, el teorema del límite central proporciona al usuario una buena aproximación de la probabilidad de determinar en falso que el proceso está fuera de control. En general, sin embargo, el usuario puede no ser capaz de confiar en la normalidad de la estadística de la línea central. Como resultado podría no conocerse la probabilidad exacta de un "error tipo I". A pesar de esto, es casi estándar el uso de límites $k\sigma$. Mientras el uso de los límites 3σ es amplio, a veces el usuario puede desear desviarse de esta aproximación. Un múltiplo más pequeño de σ puede ser apro-

piado cuando es importante detectar de forma rápida una situación fuera de control. Debido a consideraciones económicas, puede resultar costoso permitir que un proceso continúe funcionando fuera de control incluso por periodos cortos, mientras que el costo de la búsqueda y corrección de las causas imputables puede ser relativamente pequeño. Es claro, en este caso, que los límites de control que son más estrictos que los límites 3σ son apropiados.

Subgrupos racionales

Los valores muestrales a ser usados en un esfuerzo de control de calidad se dividen en subgrupos con una *muestra* que representa un subgrupo. Como indicamos antes, el orden en el tiempo de producción es en realidad una base natural para la selección de los subgrupos. Se puede ver el esfuerzo de control de calidad de manera muy simple como 1) muestreo, 2) detección de un estado fuera de control y 3) búsqueda de las causas imputables que puedan ocurrir con el tiempo. La selección de la base para estos grupos muestrales parece ser bastante directa. La elección de estos subgrupos de información muestral puede tener un efecto importante en el éxito del programa de control de calidad. Estos subgrupos con frecuencia se denominan **subgrupos racionales**. Generalmente, si el analista se interesa en la detección de un *corrimiento de la ubicación*, se considera que los subgrupos se deben elegir de modo que la variabilidad dentro del subgrupo sea pequeña y que las causas asignables, si se presentan, puedan tener la mayor posibilidad de detección. De esta manera deseamos elegir los subgrupos de modo que se maximice la variabilidad entre subgrupos. La elección de unidades en un subgrupo que se producen en tiempos muy cercanos, por ejemplo, es una aproximación razonable. Por otro lado, las gráficas de control a menudo se utilizan para controlar la variabilidad, en cuyo caso la estadística de comportamiento es la *variabilidad dentro de la muestra*. Por ello, es más importante elegir los subgrupos racionales para maximizar la variabilidad dentro de la muestra. En este caso, las observaciones en los subgrupos se deben comportar más como una muestra aleatoria y esta variabilidad dentro de las muestras necesita ser una descripción de la variabilidad del proceso.

Es importante notar que las gráficas de control sobre la variabilidad se deben establecer antes del desarrollo de las gráficas sobre el centro de ubicación (digamos, gráficas \bar{X}). Cualquier gráfica de control sobre el centro de ubicación en realidad dependerá de la variabilidad. Por ejemplo, vimos un ejemplo de la gráfica de tendencia central y ésta depende de σ . En las secciones que siguen, se presentará una estimación de σ a partir de los datos.

Gráficas \bar{X} con parámetros estimados

Ilustramos con anterioridad las nociones de la gráfica \bar{X} que hacen uso del teorema del límite central y empleamos valores *conocidos* de la media y desviación estándar del proceso. Como se indicó, se utilizan los límites de control

$$LCL = \mu - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad UCL = \mu + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

y un valor \bar{X} que cae fuera de estos límites se ve como evidencia de que la media μ cambió y por ello el proceso puede estar fuera de control.

En muchas situaciones prácticas no es razonable suponer que conocemos μ y σ . Como resultado, se deben proporcionar estimaciones de los datos que se obtienen cuando el proceso está bajo control. Típicamente, las estimaciones se determinan durante un periodo en el que se reúne información de origen o de inicio. Se elige una base para subgrupos racionales y los datos se reúnen con muestras de tamaño n en cada subgrupo. Los tamaños muestrales por lo general son pequeños, digamos, 4, 5 o 6, y se toman k muestras, con k al menos igual a 20. Durante este periodo en el que se supone que el proceso está bajo control, el usuario establece estimaciones de μ y σ sobre las que se basa la gráfica de control. La información importante reunida durante este periodo incluye las medias muestrales en el subgrupo, la media general y el rango de la muestra en cada subgrupo. En los siguientes párrafos señalamos cómo se utiliza esta información para producir la gráfica de control.

Una parte de la información muestral de estas k muestras toma la forma $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$, donde la variable aleatoria \bar{X}_i es el promedio de los valores en la i -ésima muestra. Obviamente, el promedio global es la variable aleatoria

$$\bar{\bar{X}} = \frac{\sum_{i=1}^k \bar{X}_i}{k}$$

Éste es el estimador apropiado de la media del proceso y, como resultado, es la línea central en la gráfica de control \bar{X} . En aplicaciones de control de calidad a menudo es conveniente estimar σ a partir de la información relacionada con los rangos en las muestras en lugar de las desviaciones estándar de las muestras. Definamos para la i -ésima muestra a

$$R_i = X_{\max,i} - X_{\min,i}$$

como el rango para los datos en la i -ésima muestra. Aquí $X_{\max,i}$ y $X_{\min,i}$ son, respectivamente, la observación más grande y la más pequeña en la muestra. La estimación apropiada de σ es una función del rango promedio

$$\bar{R} = \frac{\sum_{i=1}^k R_i}{k}$$

Una estimación de σ , digamos $\hat{\sigma}$, se obtiene mediante

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

donde d_2 es una constante que depende del tamaño de la muestra. Los valores de d_2 se muestran en la tabla A.23.

El uso del rango para producir una estimación de σ tiene sus raíces en aplicaciones del tipo de control de calidad, en particular debido a que el rango era muy fácil de calcular en la época anterior al periodo en que el tiempo de cálculo no se considera una dificultad. La suposición de normalidad de las observaciones individuales está implícita en la gráfica \bar{X} . Por supuesto, la existencia del teorema del límite central es ciertamente útil a este respecto. Bajo la suposición de normalidad, hacemos uso de una variable aleatoria que se denomina rango relativo, dada por

$$W = \frac{R}{\sigma}$$

Resulta que los momentos de W son funciones simples del tamaño muestral n (véase la referencia a Montgomery en la bibliografía). El valor esperado de W a menudo se denomina d_2 . Así, al tomar el valor esperado de W anterior,

$$\frac{E(R)}{\sigma} = d_2$$

Como resultado, se comprende con facilidad el fundamento para la estimación de $\hat{\sigma} = R/d_2$. Se sabe bien que el método del rango produce un estimador eficiente de σ en muestras relativamente pequeñas. Esto hace que el estimador sea en particular atractivo en aplicaciones de control de calidad dado que los tamaños muestrales en los subgrupos por lo general son pequeños. El uso del método del rango para la estimación de σ tiene como resultado gráficas de control con los siguientes parámetros:

$$\text{UCL} = \bar{X} + \frac{3R}{d_2 \sqrt{n}}, \quad \text{línea central} = \bar{X}, \quad \text{LCL} = \bar{X} - \frac{3R}{d_2 \sqrt{n}}$$

Al definir la cantidad

$$A_2 = \frac{3}{d_2 \sqrt{n}},$$

tenemos que

$$\text{UCL} = \bar{X} + A_2 R, \quad \text{LCL} = \bar{X} - A_2 R.$$

Para simplificar la estructura, el usuario de las gráficas \bar{X} a menudo encuentra valores tabulados de A_2 . En la tabla A.23 se dan tabulaciones de los valores de A_2 para varios tamaños muestrales.

Gráficas R para control de variación

Hasta este punto todos los ejemplos y detalles trataron con el intento del analista de control de calidad de detectar condiciones fuera de control producidas por un *corrimiento en la media*. Los límites de control se basan en la distribución de la variable aleatoria \bar{X} y dependen de la suposición de normalidad de las observaciones individuales. Es importante para el control que se aplique a la variabilidad así como al centro de ubicación. De hecho, muchos expertos consideran que el control de variabilidad de la característica de comportamiento es más importante y que se debe establecer antes de considerar el centro de ubicación. La variabilidad del proceso se puede controlar a través del uso de *gráficas del rango muestral*. Una gráfica de los rangos muestrales respecto al tiempo se denomina **gráfica R** . Se puede utilizar la misma estructura general como en el caso de la gráfica \bar{X} , con R como *línea central* y los límites de control dependen de una estimación de la desviación estándar de la variable aleatoria R . Así, como en el caso de la gráfica \bar{X} , establecen límites 3σ donde " 3σ " implica $3\sigma_R$. La cantidad σ_R se debe estimar a partir de los datos justos como se estima $\sigma_{\bar{X}}$.

La estimación de σ_R , la desviación estándar, también se basa en la distribución del rango relativo

$$W = \frac{R}{\sigma}$$

La desviación estándar de W es una función conocida del tamaño muestral y por lo general se denota como d_2 . Como resultado,

$$\sigma_R = \sigma d_2$$

Podemos reemplazar ahora σ por $\bar{\sigma} = R/d_2$, y de esta forma el estimador de σ_R es

$$\hat{\sigma}_R = \frac{R d_2}{d_2}$$

Así las cantidades que definen a la gráfica- R son

$$UCL = \bar{R} D_4, \quad \text{línea central} = \bar{R}, \quad LCL = \bar{R} D_3,$$

donde las constantes D_4 y D_3 (dependiendo sólo en n) son

$$D_4 = 1 + 3 \frac{d_2}{d_2}, \quad D_3 = 1 - 3 \frac{d_2}{d_2}$$

Las constantes D_4 y D_3 se tabulan en la tabla A.23.

Gráficas \bar{X} y R para variables

Se controla un proceso de fabricación de partes componentes para misiles, con la resistencia a la tensión, en libras por pulgada cuadrada, como característica de comportamiento. Se toman muestras de tamaño 5 cada hora y se reportan 25 muestras. Los datos se muestran en la tabla 17.1.

Como indicamos antes, es importante en un principio establecer condiciones de la variabilidad "bajo control". La línea central calculada de la gráfica R es

$$\bar{R} = \frac{\sum_{i=1}^{25} R_i}{25} = 10.72$$

De la tabla A.23 encontramos que para $n = 5$, $D_3 = 0$ y $D_4 = 2.115$. Como resultado, los límites de control para la gráfica- R son

$$LCL = \bar{R} D_3 = (10.72)(0) = 0,$$

$$UCL = \bar{R} D_4 = (10.72)(2.115) = 22.6728.$$

En la figura 17.3 se muestra la gráfica- R . Ninguno de los rangos graficados cae fuera de los límites de control. Como resultado, no hay indicación de una situación fuera de control.

Ahora se puede construir la gráfica \bar{X} para las lecturas de la resistencia a la tensión. La línea central es

$$\bar{\bar{X}} = \frac{\sum_{i=1}^{25} \bar{X}_i}{25} = 1507.328.$$

Tabla 17.1 Información muestral sobre datos de resistencia a la tensión

| Número de muestra | Observaciones | | | | | \bar{X}_i | R_i |
|-------------------|---------------|------|------|------|------|-------------|-------|
| 1 | 1515 | 1518 | 1512 | 1498 | 1511 | 1510.8 | 20 |
| 2 | 1504 | 1511 | 1507 | 1499 | 1502 | 1504.6 | 12 |
| 3 | 1517 | 1513 | 1504 | 1521 | 1520 | 1515.0 | 17 |
| 4 | 1497 | 1503 | 1510 | 1508 | 1502 | 1504.0 | 15 |
| 5 | 1507 | 1502 | 1497 | 1509 | 1512 | 1505.4 | 15 |
| 6 | 1519 | 1522 | 1523 | 1517 | 1511 | 1518.4 | 12 |
| 7 | 1498 | 1497 | 1507 | 1511 | 1508 | 1504.2 | 14 |
| 8 | 1511 | 1518 | 1507 | 1503 | 1509 | 1509.6 | 15 |
| 9 | 1506 | 1503 | 1498 | 1508 | 1506 | 1504.2 | 10 |
| 10 | 1503 | 1506 | 1511 | 1501 | 1500 | 1504.2 | 11 |
| 11 | 1499 | 1503 | 1507 | 1503 | 1501 | 1502.6 | 8 |
| 12 | 1507 | 1503 | 1502 | 1500 | 1501 | 1502.6 | 7 |
| 13 | 1500 | 1506 | 1501 | 1498 | 1507 | 1502.4 | 9 |
| 14 | 1501 | 1509 | 1503 | 1508 | 1503 | 1504.8 | 8 |
| 15 | 1507 | 1508 | 1502 | 1509 | 1501 | 1505.4 | 8 |
| 16 | 1511 | 1509 | 1503 | 1510 | 1507 | 1508.0 | 8 |
| 17 | 1508 | 1511 | 1513 | 1509 | 1506 | 1509.4 | 7 |
| 18 | 1508 | 1509 | 1512 | 1515 | 1519 | 1512.6 | 11 |
| 19 | 1520 | 1517 | 1519 | 1522 | 1516 | 1518.8 | 6 |
| 20 | 1506 | 1511 | 1517 | 1516 | 1508 | 1511.6 | 11 |
| 21 | 1500 | 1498 | 1503 | 1504 | 1508 | 1502.6 | 10 |
| 22 | 1511 | 1514 | 1509 | 1508 | 1506 | 1509.6 | 8 |
| 23 | 1505 | 1508 | 1500 | 1509 | 1503 | 1505.0 | 9 |
| 24 | 1501 | 1498 | 1505 | 1502 | 1505 | 1502.2 | 7 |
| 25 | 1509 | 1511 | 1507 | 1500 | 1499 | 1505.2 | 12 |

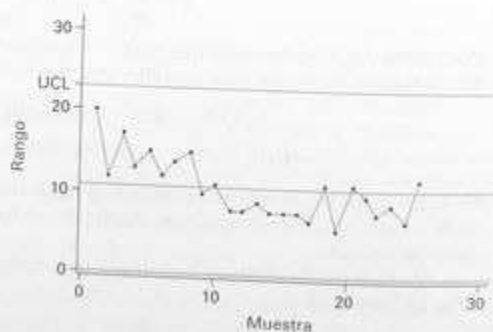


Figura 17.3 Gráfica R para el ejemplo de resistencia a la tensión.

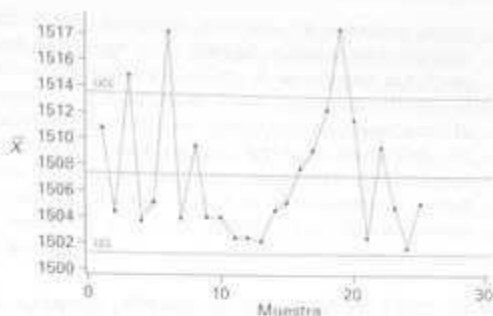


Figura 17.4 Gráfica \bar{X} para el ejemplo de resistencia a la tensión.

Para muestras de tamaño 5, encontramos de la tabla A.23 que $A_2 = 0.577$. Por ello los límites de control son

$$UCL = \bar{X} + A_2\bar{R} = 1507.328 + (0.577)(10.72) = 1513.5134.$$

$$LCL = \bar{X} - A_2\bar{R} = 1507.328 - (0.577)(10.72) = 1501.1426.$$

En la figura 17.4 se muestra la gráfica \bar{X} . Como el lector puede observar, tres valores caen fuera de los límites de control. Como resultado, los límites de control para \bar{X} no se deben usar para la línea de control de calidad.

Comentarios adicionales acerca de las gráficas de control para variables

Un proceso puede parecer que está bajo control y, de hecho, puede permanecer bajo control durante un periodo largo. ¿Esto necesariamente significa que el proceso opera exitosamente? Un proceso que opera *bajo control* es simplemente uno en el que la media y la variabilidad del proceso son estables. De manera aparente, no ocurren cambios serios. “Bajo control” implica que el proceso permanece consistente con *variabilidad natural*. Las gráficas de control de calidad se pueden ver como un método en el que la variabilidad natural inherente gobierna la amplitud de los límites de control. No hay implicación, sin embargo, de hasta qué punto un proceso bajo control satisface las *especificaciones* predeterminadas que el proceso requiere. Las especificaciones son límites que establece el consumidor. Si la variabilidad natural actual del proceso es mayor que la dictada por la especificación, el proceso no producirá artículos que cumplan las especificaciones con una frecuencia alta, aunque el proceso sea estable y esté bajo control.

Aludimos a la suposición de normalidad sobre las observaciones individuales en una gráfica de control de variables. Para la gráfica \bar{X} , si las observaciones individuales son normales, la estadística \bar{X} es normal. Como resultado, el analista de control de calidad tiene en este caso control sobre la probabilidad de un error tipo I. Si las \bar{X} individuales no son normales \bar{X} es aproximadamente normal y por ello existe un control aproximado sobre la probabilidad de un error tipo I para el caso en el

que se conoce α . Sin embargo, el uso del método del rango para estimar la desviación estándar también depende de la suposición de normalidad. Estudios con respecto a la robustez de la gráfica \bar{X} para desviaciones de la normalidad indican que la muestras de tamaño $k > 4$, la gráfica \bar{X} tiene como resultado un riesgo α cercano al anunciado (véase el trabajo de Montgomery, Schilling y Nelson en la bibliografía). Indicamos antes que la aproximación $\pm k\sigma_R$ a la gráfica R es una cuestión de conveniencia y tradición. Aun si la distribución de observaciones individuales es normal, la distribución de R no es normal. De hecho, la distribución de R no es ni siquiera simétrica. Los límites de control simétricos $\pm k\sigma_R$ sólo dan una aproximación al riesgo α , y en algunos casos la aproximación no es particularmente buena.

Elección del tamaño de la muestra (función característica de operación) en el caso de la gráfica \bar{X}

Los científicos e ingenieros que tratan con el control de calidad a menudo se refieren a los factores que afectan el *diseño de la gráfica de control*. Los componentes que determinan el diseño de la gráfica incluyen el tamaño de la muestra que se toma en cada subgrupo, la amplitud de los límites de control y la frecuencia del muestreo. Todos estos factores dependen en gran medida de consideraciones económicas y prácticas. La frecuencia de muestreo obviamente depende del costo del muestreo y del costo que ocurre si el proceso continúa fuera de control durante un periodo largo. Estos mismos factores afectan la amplitud de la región "bajo control". El costo asociado con la investigación y la búsqueda de causas imputables tiene un impacto sobre la amplitud de la región y sobre la frecuencia de muestreo. Se dedica una cantidad considerable de atención al diseño óptimo de gráficas de control y no se darán aquí detalles más extensos. Se remite al lector al trabajo de Montgomery y colaboradores, que se cita en la bibliografía para un excelente recuento histórico de gran parte de esta investigación.

La elección del tamaño muestral y la frecuencia de muestreo implica balancear los recursos disponibles para estos dos esfuerzos. En muchos casos, el analista puede necesitar hacer cambios en la estrategia hasta que se logra el balance apropiado. El analista siempre debe estar consciente de que si el costo de producción de artículos no adecuados es grande, una alta frecuencia de muestreo con tamaño muestral relativamente pequeño es una estrategia apropiada.

Se deben tomar en cuenta muchos factores en la elección de un tamaño muestral. En la ilustración y discusión enfatizamos el uso de $n = 4, 5$ o 6 . Estos valores se consideran relativamente pequeños para problemas generales en inferencia estadística pero quizá tamaños muestrales apropiados para control de calidad. Una justificación, por supuesto, es que el control de calidad es un proceso continuo y los resultados producidos por una muestra o conjunto de unidades estarán seguidos por resultados de muchas más. Así, el tamaño muestral "efectivo" de todo el esfuerzo de control de calidad es muchas veces mayor que el que se utiliza en un subgrupo. Por lo general se considera más efectivo *muestrear con frecuencia* con un tamaño muestral pequeño.

El analista puede hacer uso de la noción de *poder* de una muestra para obtener alguna idea de la efectividad del tamaño muestral que se elige. Esto es importante en particular debido a que los tamaños muestrales pequeños por lo general se

utilizan en cada subgrupo. Refiérase a los capítulos 10 y 13 para un análisis de la potencia de pruebas formales sobre medias y del análisis de varianzas. Aunque las pruebas formales de hipótesis en realidad no se realizan en el control de calidad, se puede tratar la información muestral como si la estrategia en cada subgrupo fuera probar una hipótesis, ya sea sobre la media poblacional μ o sobre la desviación estándar σ . Es de interés la *probabilidad de detección* de una condición fuera de control para una muestra dada y, quizá más importante, el número esperado de corridas que se requieren para la detección. La probabilidad de detección de una condición fuera de control específica corresponde al poder de una prueba. No es nuestra intención mostrar el desarrollo de la potencia para todos los tipos de gráficas de control que aquí se presentan, sino más bien mostrar el desarrollo de la gráfica \bar{X} y presentar los resultados de potencia para la gráfica R .

Considere la gráfica \bar{X} para σ conocida. Suponga que el estado bajo control tiene $\mu = \mu_0$. Un estudio del papel del tamaño muestral del subgrupo es equivalente a investigar el riesgo β , es decir, la probabilidad de que un valor \bar{X} permanezca dentro de los límites de control dado que, en realidad, ocurre un corrimiento en la media. Suponga que la forma que toma el corrimiento es

$$\mu = \mu_0 + r\sigma.$$

De nuevo, al hacer uso de la normalidad de \bar{X} , tenemos

$$\beta = P\{\text{LCL} < \bar{X} < \text{UCL} \mid \mu = \mu_0 + r\sigma\}.$$

Para el caso de límites $k\sigma$,

$$\text{LCL} = \mu_0 - \frac{k\sigma}{\sqrt{n}} \quad \text{y} \quad \text{UCL} = \mu_0 + \frac{k\sigma}{\sqrt{n}}.$$

Como resultado, si denotamos con Z la variable aleatoria normal estándar

$$\begin{aligned} \beta &= P\left\{Z < \left[\frac{\mu_0 + k\sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}}\right]\right\} - P\left\{Z < \left[\frac{\mu_0 - k\sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}}\right]\right\} \\ &= P\left\{Z < \left[\frac{\mu_0 + k\sigma/(\mu_0 + r\sigma)}{\sigma/\sqrt{n}}\right]\right\} - P\left\{Z < \left[\frac{\mu_0 - k\sigma/\sqrt{n} - (\mu_0 + r\sigma)}{\sigma/\sqrt{n}}\right]\right\} \\ &= P\{Z < (k - r\sqrt{n})\} - P\{Z < (-k - r\sqrt{n})\}. \end{aligned}$$

Nótese el papel de n , r y k en la expresión para el riesgo β . La probabilidad de no detectar un corrimiento específico con claridad aumenta con un aumento de k , como se esperaba. β disminuye con un aumento de r , la magnitud del corrimiento y disminuye con un aumento en el tamaño muestral n .

Se debe enfatizar que la expresión anterior tiene como resultado el riesgo β (probabilidad de un error tipo II) para el caso de una *sola muestra*. Por ejemplo, suponga que en el caso de una muestra de tamaño 4, ocurre un corrimiento de σ en la media. La probabilidad de detectar el corrimiento (poder) en la *primera muestra a continuación del corrimiento* es (suponga límites 3σ):

$$1 - \beta = 1 - [P(Z < 1) - P(Z < -5)] = 0.159.$$

Por otro lado, la probabilidad de detectar un corrimiento de 2σ es

$$1 - \beta = 1 - [P(Z < -1) - P(Z < -7)] = 0.8413.$$

Los resultados anteriores ilustran una probabilidad bastante modesta de detectar un corrimiento de magnitud σ y una probabilidad bastante alta de detectar un corrimiento de magnitud 2σ . La presentación completa de cómo se comportan los límites de control, digamos, 3σ para la gráfica \bar{X} que aquí se describe se muestra en la figura 17.5. En lugar de graficar el poder, se da una gráfica de β contra r , donde el corrimiento en la media tiene magnitud $r\sigma$. Por supuesto, los tamaños muestrales de $n = 4, 5, 6$ tienen como resultado una probabilidad pequeña de detectar un corrimiento de 1.0σ o incluso 1.5σ sobre la primera muestra después del corrimiento.

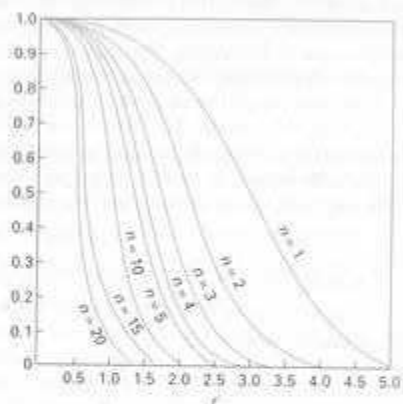


Figura 17.5 Curvas características de operación para la gráfica \bar{X} con límites 3σ . Aquí β es el error de probabilidad tipo II sobre la primera muestra después de un corrimiento en la media de $r\sigma$.

Pero si el muestreo se realiza con frecuencia, la probabilidad puede no ser tan importante como el número promedio o esperado de corridas que se requiere antes de la detección del corrimiento. Una detección rápida es importante y en realidad es posible aunque la probabilidad de detección sobre la primera muestra no sea alta. Resulta que las gráficas \bar{X} con estas pequeñas muestras tendrá como resultado una detección relativamente rápida. Si β es la probabilidad de no detectar un corrimiento sobre la primera muestra que sigue al corrimiento, entonces la probabilidad de detectar el corrimiento sobre la muestra i -ésima después de éste es (si suponemos muestras independientes):

$$P_i = (1 - \beta)\beta^{i-1}$$

El lector debe reconocer ésta como una aplicación de la distribución geométrica. El valor promedio o esperado del número de muestras que se requieren para la detección es

$$\sum_{i=1}^{\infty} i\beta^{i-1}(1 - \beta) = \frac{1}{1 - \beta}$$

Así el número esperado de muestras que se requieren para detectar el corrimiento en la media es el *recíproco del poder* (es decir, la probabilidad de detección de la primera muestra después del corrimiento).

Ejemplo 17.1 En cierto esfuerzo de control de calidad es importante para el analista de control de calidad ser capaz de detectar rápidamente corrimientos en la media de $\pm\sigma$ mientras utiliza una gráfica de control 3σ con un tamaño muestral $n = 4$. El número esperado de muestras que se requieren después del corrimiento para la detección del estado fuera de control puede ser una ayuda en la evaluación del procedimiento de control de calidad.

De la figura 17.5, para $n = 4$ y $r = 1$, se puede ver que $\beta = 0.82$. Si denotamos con s el número de muestras que se requieren para detectar el corrimiento, la media de s es

$$E(s) = \frac{1}{1 - \beta} = \frac{1}{0.18} = 5.5.$$

De esta manera, en promedio, se requieren seis subgrupos antes de la detección de un corrimiento de $\pm\sigma$. ■

Elección del tamaño muestral para la gráfica R

La curva CO de la gráfica R se muestra en la figura 17.6. Como la gráfica R se utiliza para control de la desviación estándar del proceso, el riesgo β se grafica como función de la desviación estándar bajo control, σ_0 , y la desviación estándar después de que el proceso queda fuera de control. La última desviación estándar se denotará con σ_1 . Sea

$$A = \frac{\sigma_1}{\sigma_0}.$$

Se grafica β contra A para varios tamaños muestrales.

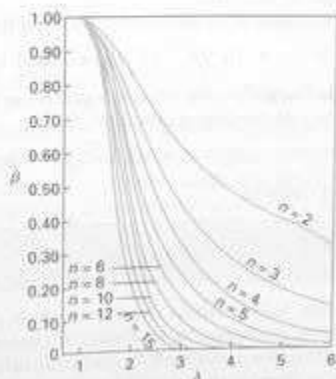


Figura 17.6 Curva característica operativa para las gráficas R con límites de 3σ .

Gráficas \bar{X} y S para variables

Es natural para el estudiante de estadística anticipar el uso de la varianza muestral en la gráfica \bar{X} y en una gráfica para control de la variabilidad. El rango es eficiente como estimador para σ , pero esta eficiencia disminuye conforme el tamaño de la muestra se hace más grande. Para n tan grande como 10, se debe utilizar la estadística familiar

$$S = \sqrt{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}}$$

en la gráfica de control para la media y la variabilidad. El lector debe recordar del capítulo 9 que S^2 es un estimador insesgado para σ^2 pero que S no es insesgado para σ . Es una costumbre corregir S para sesgos en aplicaciones de la gráfica de control. Sabemos, en general, que

$$E(S) \neq \sigma.$$

En el caso en que las X_i son independientes, distribuidas de forma normal con media μ y varianza σ^2 ,

$$E(S) = c_4 \sigma, \quad \text{donde} \quad c_4 = \left(\frac{2}{n-1} \right)^{1/2} \frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}$$

y $\Gamma(\cdot)$ se refiere a la función gamma (véase el capítulo 6). Por ejemplo, para $n = 50$, $c_4 = 3/8 \sqrt{2\pi}$. Además, la varianza del estimador S es

$$\text{Var}(S) = \sigma^2(1 - c_4^2).$$

Establezcamos las propiedades de S que nos permitirán escribir límites de control para \bar{X} y S . Para construir una estructura apropiada, comenzamos con la suposición de que se conoce σ . Después presentamos la estimación de σ a partir de un conjunto preliminar de muestras.

Si se gráfica la estadística S , los parámetros obvios de la gráfica de control son

$$\text{UCL} = c_4 \sigma + 3\sigma \sqrt{1 - c_4^2}, \quad \text{línea central} = c_4 \sigma, \quad \text{LCL} = c_4 \sigma - 3\sigma \sqrt{1 - c_4^2}.$$

Como de costumbre, los límites de control se definen de manera más sucinta a través del uso de constantes tabuladas. Sean

$$B_5 = c_4 - 3\sqrt{1 - c_4^2}, \quad B_6 = c_4 + 3\sqrt{1 - c_4^2},$$

entonces tenemos

$$\text{UCL} = B_6 \sigma, \quad \text{línea central} = c_4 \sigma, \quad \text{LCL} = B_5 \sigma$$

En la tabla A.23 se tabulan los valores de B_5 y B_6 para varios tamaños muestrales.

Ahora, por supuesto, los límites de control anteriores sirven como base para el desarrollo de los parámetros de control de calidad para la situación que con más frecuencia se observa en la práctica; a saber, en la que se desconoce σ . Debemos suponer una vez más que se toma un conjunto de *muestras base* o *muestras preliminares*

para producir una estimación de σ durante lo que se supone como período "bajo control". Las desviaciones estándar muestrales S_1, S_2, \dots, S_m se obtienen a partir de muestras que son cada una de tamaño n . A menudo se utiliza un estimador insesgado del tipo

$$\frac{\bar{S}}{c_4} = \left(\frac{1}{m} \sum_{i=1}^m S_i \right) / c_4$$

para σ . Aquí, por supuesto, \bar{S} , el valor promedio de la desviación estándar muestral en la muestra preliminar, es la línea central lógica en la gráfica de control para el control de la variabilidad. Los límites de control superior e inferior son estimadores insesgados de los límites de control apropiados para el caso donde se conoce σ . Como

$$E\left(\frac{\bar{S}}{c_4}\right) = \sigma,$$

la estadística \bar{S} es una línea central apropiada (como estimador insesgado de $c_4\sigma$) y las cantidades

$$\bar{S} - 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2} \quad \text{y} \quad \bar{S} + 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2}$$

son los límites de control 3σ inferior y superior apropiados, respectivamente. Como resultado, la línea central y los límites para la gráfica \bar{S} para control de variabilidad son

$$\text{UCL} = B_4 \bar{S}, \quad \text{línea central} = \bar{S}, \quad \text{LCL} = B_3 \bar{S},$$

donde

$$B_3 = 1 - \frac{3}{c_4} \sqrt{1 - c_4^2}, \quad B_4 = 1 + \frac{3}{c_4} \sqrt{1 - c_4^2}.$$

Las constantes B_3 y B_4 aparecen en la tabla A.23.

Podemos escribir ahora los parámetros de la gráfica \bar{X} correspondiente que implican el uso de la desviación estándar muestral. Supongamos que \bar{S} y \bar{X} están disponibles de la muestra preliminar base. La línea central sigue siendo \bar{X} y los límites 3σ son simplemente de la forma $\bar{X} \pm 3\hat{\sigma}/\sqrt{n}$, donde $\hat{\sigma}$ es un estimador insesgado. Simplemente proporcionamos \bar{S}/c_4 como un estimador de σ , y de esta forma tenemos

$$\text{UCL} = \bar{X} + A_1 \bar{S}, \quad \text{línea central} = \bar{X}, \quad \text{LCL} = \bar{X} - A_1 \bar{S},$$

donde

$$A_1 = \frac{3}{c_4 \sqrt{n}}.$$

En la tabla A.23 aparece la constante A_1 para varios tamaños muestrales.

Ejemplo 17.2 Se producen contenedores mediante un proceso donde el volumen de los contenedores se sujeta a un control de calidad. Se utilizan 25 muestras de tamaño 10 cada una para establecer los parámetros de control de calidad. En la tabla 17.2 se documenta la información de estas muestras.

Tabla 17.2 Volumen de muestras de contenedores para 25 muestras en una muestra preliminar (en centímetros cúbicos)

| Muestra | Observaciones | | | | | \bar{X} | S_x |
|---------|---------------|--------|--------|--------|--------|---------------------|--------|
| 1 | 62.255 | 62.301 | 62.289 | 62.289 | 62.311 | 62.269 | 0.0498 |
| 2 | 62.187 | 62.225 | 62.337 | 62.297 | 62.307 | 62.271 | 0.0622 |
| 3 | 62.421 | 62.377 | 62.257 | 62.295 | 62.222 | 62.314 | 0.0829 |
| 4 | 62.301 | 62.315 | 62.293 | 62.317 | 62.409 | 62.327 | 0.0469 |
| 5 | 62.400 | 62.375 | 62.295 | 62.272 | 62.372 | 62.343 | 0.0558 |
| 6 | 62.372 | 62.275 | 62.315 | 62.372 | 62.302 | 62.327 | 0.0434 |
| 7 | 62.297 | 62.303 | 62.337 | 62.392 | 62.344 | 62.335 | 0.0361 |
| 8 | 62.325 | 62.362 | 62.351 | 62.371 | 62.397 | 62.361 | 0.0264 |
| 9 | 62.327 | 62.297 | 62.318 | 62.342 | 62.318 | 62.320 | 0.0163 |
| 10 | 62.297 | 62.325 | 62.303 | 62.307 | 62.333 | 62.313 | 0.0151 |
| 11 | 62.315 | 62.366 | 62.308 | 62.318 | 62.319 | 62.325 | 0.0232 |
| 12 | 62.297 | 62.322 | 62.344 | 62.342 | 62.313 | 62.324 | 0.0198 |
| 13 | 62.375 | 62.287 | 62.362 | 62.319 | 62.382 | 62.345 | 0.0406 |
| 14 | 62.317 | 62.321 | 62.297 | 62.372 | 62.319 | 62.325 | 0.0279 |
| 15 | 62.299 | 62.307 | 62.383 | 62.341 | 62.394 | 62.345 | 0.0431 |
| 16 | 62.308 | 62.319 | 62.344 | 62.319 | 62.378 | 62.334 | 0.0281 |
| 17 | 62.319 | 62.357 | 62.277 | 62.315 | 62.295 | 62.313 | 0.0300 |
| 18 | 62.333 | 62.362 | 62.292 | 62.327 | 62.314 | 62.326 | 0.0257 |
| 19 | 62.313 | 62.387 | 62.315 | 62.318 | 62.341 | 62.335 | 0.0313 |
| 20 | 62.375 | 62.321 | 62.354 | 62.342 | 62.375 | 62.353 | 0.0239 |
| 21 | 62.399 | 62.308 | 62.292 | 62.372 | 62.299 | 62.334 | 0.0483 |
| 22 | 62.309 | 62.403 | 62.318 | 62.295 | 62.317 | 62.328 | 0.0427 |
| 23 | 62.293 | 62.293 | 62.342 | 62.315 | 62.349 | 62.318 | 0.0264 |
| 24 | 62.388 | 62.308 | 62.315 | 62.392 | 62.303 | 62.341 | 0.0448 |
| 25 | 62.328 | 62.318 | 62.317 | 62.295 | 62.319 | 62.314 | 0.0111 |
| | | | | | | $\bar{X} = 62.3266$ | |
| | | | | | | $S = 0.0361$ | |

Del apéndice: $B_1 = 0$, $B_2 = 2.089$, $A_1 = 1.427$. Como resultado, los límites de control para \bar{X} están dados por

$$\bar{X} + A_1 S = 62.3771, \quad \bar{X} - A_1 S = 62.2740$$

y los límites de control para la gráfica S son

$$UCL = B_2 S = 0.0754, \quad LCL = B_3 S = 0$$

Las figuras 17.7 y 17.8 muestran las gráficas de control \bar{X} y S , respectivamente, para el ejemplo 17.2. En las gráficas se representa la información muestral para las 25 muestras en el conjunto de datos preliminar. Parece que el control se establece después de las primeras muestras.

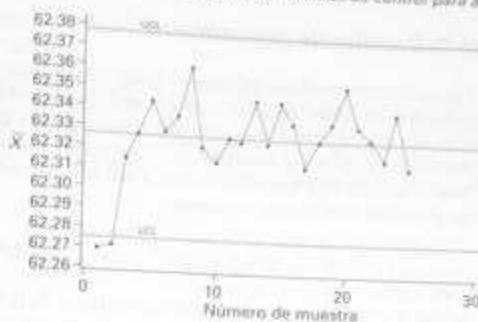


Figura 17.7 Gráfica \bar{X} con límites de control establecidos por los datos del ejemplo 17.2.

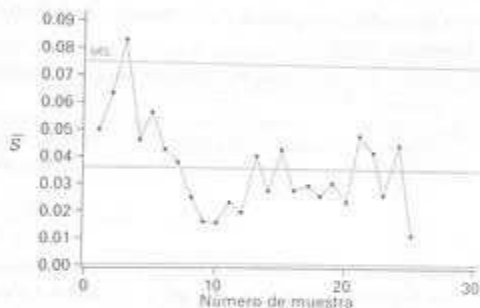


Figura 17.8 Gráfica S con límites de control establecidos por los datos del ejemplo 17.2.

17.5 Gráficas de control para atributos

Como indicamos antes en este capítulo, muchas aplicaciones industriales de control de calidad requieren que la característica de calidad indique sólo la afirmación de que el artículo "se adapta". En otras palabras, no hay medición continua que es crucial para el comportamiento del artículo. Una ilustración obvia de este tipo de muestreo, que se llama **muestreo por atributos**, es el comportamiento de un foco —que funciona o no de manera satisfactoria—. El artículo es *defectuoso* o *no defectuoso*. Las piezas metálicas fabricadas pueden contener deformaciones. Los contenedores de una línea de producción pueden tener fugas. En ambos casos un artículo defectuoso impide su uso por el consumidor. La gráfica de control estándar para esta situación es la gráfica- p , o gráfica para la fracción de defectuosos. Como se podría esperar, la distribución de probabilidad involucrada es la distribución binomial. Se remite al lector al capítulo 4 para información básica de la distribución binomial.

Gráfica p para la fracción de defectuosos

Cualquier artículo fabricado puede tener varias características que son importantes y lo debe examinar un inspector. Sin embargo, todo el desarrollo se enfoca aquí a una sola característica. Suponga que para todos los artículos la probabilidad de un defectuoso es p , y que todos los artículos se producen de forma independiente. Entonces, en una muestra aleatoria de n artículos producidos, con X como el número de artículos defectuosos, tenemos

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

Como se podría sospechar, la media y varianza de la variable aleatoria binomial jugarán un papel importante en el desarrollo de la gráfica de control. El lector debe recordar que

$$E(X) = np, \quad \text{y} \quad \text{Var}(X) = np(1-p).$$

Un estimador insesgado de p es la **fracción de defectuosos** o la **proporción de defectuosos**, \hat{p} , donde

$$\hat{p} = \frac{\text{número de defectuosos en la muestra de tamaño } n}{n}$$

Como en el caso de las gráficas de control de variables, las propiedades de distribución de \hat{p} son importantes en el desarrollo de la gráfica de control. Sabemos que

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}.$$

Aquí aplicamos los mismos principios 3σ que utilizamos para las gráficas de variables. Supongamos inicialmente que se conoce p . La estructura, entonces, de las gráficas de control implica el uso de límites 3σ con

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}.$$

De esta manera los límites son

$$\text{UCL} = p + 3\sqrt{\frac{p(1-p)}{n}}, \quad \text{LCL} = p - 3\sqrt{\frac{p(1-p)}{n}},$$

con el proceso considerado bajo control cuando los valores \hat{p} de la muestra yacen dentro de los límites de control.

En general, por supuesto, no se conoce el valor de p y se debe estimar a partir de un conjunto base de muestras de forma muy similar al caso de μ y σ en las gráficas de variables. Suponga que hay m muestras preliminares de tamaño n . Para una muestra dada, cada una de las n observaciones se reporta como "defectuosa" o "no defectuosa". El estimador insesgado obvio para p a utilizar en la gráfica de control es

$$\hat{p} = \frac{\sum_{i=1}^m \hat{p}_i}{m}.$$

donde \hat{p}_i es la proporción de defectuosos en la i -ésima muestra. Como resultado, los límites de control son

$$UCL = \bar{p} + 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}, \text{ línea central} = \bar{p}, \text{ LCL} = \bar{p} - 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Ejemplo 17.3 Considere los datos que se muestran en la tabla 17.3 sobre el número de componentes electrónicos defectuosos en muestras de tamaño 50. Se tomaron 20 muestras a fin de establecer valores preliminares de la gráfica de control. Las gráficas de control determinadas por este periodo preliminar tendrán una línea central $\bar{p} = 0.088$ y límites de control

$$UCL = \bar{p} + 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{50}} = 0.2082, \text{ LCL} = \bar{p} - 3 \sqrt{\frac{\bar{p}(1-\bar{p})}{50}} = -0.0322$$

Tabla 17.3 Datos para el ejemplo 17.3 para establecer límites de control para gráficas p , muestras de tamaño 50

| Muestra | Número de componentes defectuosos | Fración defectuoso, \hat{p}_i |
|---------|-----------------------------------|---------------------------------|
| 1 | 8 | 0.16 |
| 2 | 6 | 0.12 |
| 3 | 5 | 0.10 |
| 4 | 7 | 0.14 |
| 5 | 2 | 0.04 |
| 6 | 5 | 0.10 |
| 7 | 3 | 0.06 |
| 8 | 8 | 0.16 |
| 9 | 4 | 0.08 |
| 10 | 4 | 0.08 |
| 11 | 3 | 0.06 |
| 12 | 1 | 0.02 |
| 13 | 5 | 0.10 |
| 14 | 4 | 0.08 |
| 15 | 4 | 0.08 |
| 16 | 2 | 0.04 |
| 17 | 3 | 0.06 |
| 18 | 5 | 0.10 |
| 19 | 6 | 0.12 |
| 20 | 3 | 0.06 |
| | | <hr/> |
| | | $\bar{p} = 0.088$ |

Obviamente, con un valor calculado negativo, el LCL se ajusta a cero. Se hace evidente a partir de los valores de los límites de control que el proceso está bajo control durante este periodo preliminar. ■

Selección del tamaño muestral para la gráfica p

La elección del tamaño muestral para la gráfica p para atributos incluye los mismos tipos generales de consideraciones que los de la gráfica para variables. Se requiere un tamaño muestral que sea suficientemente grande para tener una probabilidad alta de detección de una condición fuera de control cuando, de hecho, ocurre un cambio específico en p . No hay un *mejor método* para la elección del tamaño de la muestra. Sin embargo, una aproximación razonable, sugerida por Duncan (véase la bibliografía), es elegir n de modo que haya una probabilidad de 0.5 de detectar un corrimiento de un monto particular en p . La solución que resulta para n es bastante simple. Suponga que se aplica la aproximación normal a la distribución binomial. Deseamos, bajo la condición de que p tiene un corrimiento Δ , digamos, $p_1 > p_0$, que

$$P(\hat{p} > \text{UCL}) = 0.5 = P\left[Z \geq \frac{\text{UCL} - p_1}{\sqrt{p_1(1-p_1)/n}}\right] = 0.5.$$

Como $P(Z > 0) = 0.5$, hacemos

$$\frac{\text{UCL} - p_1}{\sqrt{p_1(1-p_1)/n}} = 0.$$

Al sustituir

$$\bar{p} + 3\sqrt{\frac{p(1-p)}{n}} = \text{UCL}, \quad \text{tenemos} \quad (p - p_1) + 3\sqrt{\frac{p(1-p)}{n}} = 0.$$

Podemos resolver ahora para n , el tamaño de cada muestra:

$$n = \frac{9}{\Delta^2} p(1-p),$$

donde, por supuesto, Δ es el "corrimiento" en el valor de p , y p es la probabilidad de un defectuoso sobre la que se basan los límites de control. Sin embargo, si las gráficas de control se basan en límites $k\sigma$ entonces

$$n = \frac{k^2}{\Delta^2} p(1-p).$$

Ejemplo 17.4 Suponga que se diseña una gráfica de control de calidad de atributos con un valor de $p = 0.01$ para la probabilidad bajo control de un defectuoso. ¿Cuál es el tamaño de la muestra por subgrupo que produce una probabilidad de 0.5 de que se detecte un proceso que se corre a $p = p_1 = 0.05$? La gráfica resultante incluirá límites 3σ .

Aquí tenemos $\Delta = 0.04$. El tamaño apropiado de la muestra es:

$$n = \frac{9}{(0.04)^2} (0.01)(0.99) = 56.$$

Gráficas de control para defectuosos (uso del modelo de Poisson)

En el desarrollo anterior supusimos que el artículo bajo consideración es uno que es defectuoso (es decir no funcional) o no defectuoso. En el último caso es funcional y por ello aceptable para el consumidor. En muchas situaciones este enfoque de "defectuoso o no" es demasiado simplista. Las unidades pueden contener defectos o no cumplimiento de la norma pero aún funcionar bastante bien para el consumidor. En realidad, en este caso, puede ser importante ejercer control sobre el *número de defectos* o *número de diferencias*. Este tipo de esfuerzo de control de calidad encuentra aplicación cuando las unidades son no simplistas o quizá grandes. Por ejemplo, el número de defectos puede ser bastante útil como el objeto de control cuando el artículo o unidad es, digamos, una computadora personal. Otro ejemplo es una unidad defrita por 50 pies de tubería fabricada, donde el número de soldaduras defectuosas es el objeto del control de calidad, el número de defectos de 50 pies de tejido para alfombras fabricado o el número de "burbujas" en una hoja grande de vidrio fabricado.

Es claro a partir de lo que describimos aquí que la distribución binomial no es apropiada. El número total de diferencias en una unidad o el número promedio por unidad se puede usar como la medida para la gráfica de control. Bastante a menudo se supone que el número de diferencias en una muestra de artículos sigue la distribución de Poisson. Este tipo de gráfica con frecuencia se llama **gráfica C**.

Suponga que el número de defectos X en una unidad de producto sigue la distribución de Poisson con parámetro λ . (Aquí $r = 1$ para el modelo de Poisson.) Recuerde que para la distribución de Poisson,

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Aquí, la variable aleatoria X es el número de diferencias. En el capítulo 6 aprendimos que la media y la varianza de la variable aleatoria de Poisson son ambas λ . De esta forma si la gráfica de control de calidad se estructurara de acuerdo con los límites 3σ acostumbrados, podríamos tener, para λ conocida,

$$UCL = \lambda + 3\sqrt{\lambda}, \quad \text{línea central} = \lambda, \quad LCL = \lambda - 3\sqrt{\lambda}.$$

Como de costumbre, λ a menudo debe provenir de un estimador de los datos. Una estimación insesgada de λ es el número *promedio* de diferencias por muestra. Denote esta estimación con $\hat{\lambda}$. Así la gráfica de control tiene los límites

$$UCL = \hat{\lambda} + 3\sqrt{\hat{\lambda}}, \quad \text{línea central} = \hat{\lambda}, \quad LCL = \hat{\lambda} - 3\sqrt{\hat{\lambda}}.$$

Ejemplo 17.5 La tabla 17.4 representa el número de defectos en 20 muestras sucesivas de rollos de hoja metálica cada una de 100 pies de longitud. Se debe desarrollar una gráfica de control a partir de estos datos preliminares con el propósito de controlar el número de defectos en tales muestras. La estimación del parámetro de Poisson λ está dada por $\hat{\lambda} = 5.95$. Como resultado, los límites de control sugeridos por estos datos preliminares son (límites 3σ):

$$UCL = \hat{\lambda} + 3\sqrt{\hat{\lambda}} = 13.2678, \quad LCL = \hat{\lambda} - 3\sqrt{\hat{\lambda}} = -1.3678 \quad (\text{LCL igualada a cero}).$$

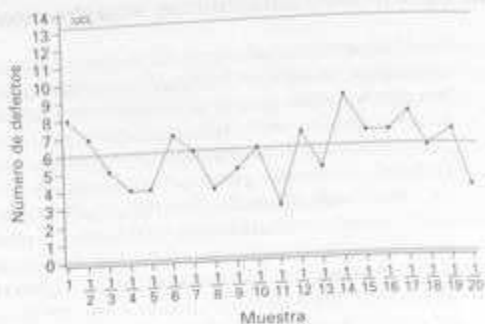


Figura 17.9 Datos preliminares representados en la gráfica de control para el ejemplo 17.5.

La figura 17.9 muestra una gráfica de los datos preliminares con los límites de control.

Tabla 17.4 Datos para el ejemplo 17.5; el control incluye el número de defectos en un rollo de hoja metálica

| Número de muestra | Número de defectos | Número de muestra | Número de defectos |
|-------------------|--------------------|-------------------|--------------------|
| 1 | 8 | 11 | 3 |
| 2 | 7 | 12 | 2 |
| 3 | 5 | 13 | 5 |
| 4 | 4 | 14 | 9 |
| 5 | 4 | 15 | 7 |
| 6 | 7 | 16 | 7 |
| 7 | 6 | 17 | 8 |
| 8 | 4 | 18 | 6 |
| 9 | 5 | 19 | 7 |
| 10 | 6 | 20 | 4 |
| Promedio: 5.98 | | | |

La tabla 17.5 muestra datos adicionales tomados del proceso de producción. Para cada muestra, se investiga la unidad en la que se basa la gráfica, a saber, 100 pies del metal. Se revela la información de 20 muestras, la figura 17.10 muestra una gráfica de los datos adicionales de producción. Es claro que el proceso está bajo control, al menos a lo largo del periodo en el que se tomaron los datos.

En el ejemplo 17.5, dejamos muy claro que la unidad de muestreo o de inspección es, a saber, 100 pies de metal. En muchos casos donde el artículo es específico (por ejemplo, una computadora personal o un tipo específico de dispositivo electrónico), la unidad de inspección puede ser un conjunto de artículos. Por ejemplo, el analista puede decidir utilizar 10 computadoras en cada subgrupo y de esta forma observar un conteo del número total de defectos que se encuentran. De esta forma la muestra preliminar para la construcción de la gráfica de control incluiría el uso de varias muestras, cada una de 10 computadoras. La elección del tamaño muestral

Tabla 17.5 Datos adicionales del proceso de producción del ejemplo 17.5

| Número de muestra | Número de defectos | Número de muestra | Número de defectos |
|-------------------|--------------------|-------------------|--------------------|
| 1 | 3 | 11 | 7 |
| 2 | 5 | 12 | 5 |
| 3 | 8 | 13 | 9 |
| 4 | 5 | 14 | 4 |
| 5 | 8 | 15 | 6 |
| 6 | 4 | 16 | 5 |
| 7 | 3 | 17 | 3 |
| 8 | 6 | 18 | 2 |
| 9 | 5 | 19 | 1 |
| 10 | 2 | 20 | 6 |

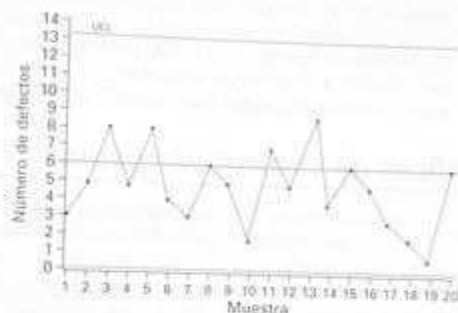


Figura 17.10 Datos adicionales de producción para el ejemplo 17.5.

puede depender de muchos factores. A menudo, se puede querer un tamaño muestral que asegure un LCL positivo.

El analista puede desear utilizar el número promedio de defectos por unidad de muestreo como la medida básica en la gráfica de control. Por ejemplo, para el caso de la computadora personal, midamos la variable aleatoria

$$U = \frac{\text{número total de defectos}}{n}$$

para cada muestra de, digamos, $n = 10$. Se puede utilizar el método de las funciones generadoras de momento para mostrar que U es una variable aleatoria de Poisson (véase el ejercicio 1) si suponemos que el número de defectos por unidad de muestreo es de Poisson con parámetro λ . De esta manera la gráfica de control para esta situación se caracteriza por lo siguiente:

$$UCL = \bar{U} + 3\sqrt{\frac{\bar{U}}{n}}, \quad \text{línea central} = \bar{U}, \quad LCL = \bar{U} - 3\sqrt{\frac{\bar{U}}{n}}$$

Aquí, por supuesto, \bar{U} es el promedio de los valores U en el conjunto de datos preliminares o base. El término \bar{U}/n se deriva del resultado que

$$E(U) = \lambda, \quad \text{Var}(U) = \frac{\lambda}{n}$$

y por ello \bar{U} es un estimador incesgado de $E(U) = \lambda$ y \bar{U}/n es un estimador incesgado de $\text{Var}(U) = \lambda/n$. Este tipo de gráfica de control a menudo se llama **gráfica U** .

En todo el desarrollo de esta sección basamos nuestra producción de las gráficas de control en el modelo de probabilidad de Poisson. Este modelo se utiliza en combinación con el concepto 3σ . Como explicamos antes en este capítulo, la noción de límites 3σ tiene sus raíces en la aproximación normal, aunque muchos usuarios consideran que el concepto trabaja bien como herramienta pragmática incluso si la normalidad no es siquiera aproximadamente correcta. La dificultad, por supuesto, es que en ausencia de normalidad no se puede controlar la probabilidad de una especificación incorrecta de un estado fuera de control. En el caso del modelo de Poisson, cuando λ es pequeña la distribución es bastante asimétrica, condición que puede producir resultados indeseables si se conserva la aproximación 3σ . ■

17.6 Gráficas de control de cusum

La desventaja con las gráficas de control tipo Shewhart, que se desarrollan e ilustran en las secciones anteriores, yace en su incapacidad para detectar pequeños cambios en la media. Un mecanismo de control de calidad que recibe una considerable atención en la literatura estadística y uso en la industria es la **gráfica de suma acumulada (cusum)**. El método para la gráfica de suma acumulada es simple y su atractivo es que es intuitivo. Debe ser obvio para el lector por qué responde mejor a pequeños cambios en la media. Considere una gráfica de control para la media con un nivel de referencia establecido en el valor W . Considere las observaciones particulares X_1, X_2, \dots, X_n .

Las primeras r sumas acumuladas son

$$\begin{aligned} S_1 &= X_1 - W \\ S_2 &= S_1 + (X_2 - W) \\ S_3 &= S_2 + (X_3 - W) \\ &\vdots \\ S_r &= S_{r-1} + (X_r - W) \end{aligned}$$

Es claro que la suma acumulada es simplemente la acumulación de las diferencias del nivel de referencia. Es decir,

$$S_k = \sum_{i=1}^k (X_i - W), \quad k = 1, 2, \dots$$

La gráfica de suma acumulada es, entonces, una gráfica de S_k contra el tiempo.

Suponga que consideramos que el nivel de referencia w es un valor aceptable de la media μ . Claramente, si no hay corrimiento en μ , la gráfica de suma acumulada debe ser

aproximadamente horizontal, con algunas fluctuaciones menores balanceadas alrededor de cero. Ahora, si sólo hay un cambio moderado en la media, debe resultar un cambio relativamente grande en la *pendiente* de la gráfica de suma acumulada, pues cada nueva observación tiene una oportunidad de contribuir con un corrimiento y la medida que se grafica se acumula a estos corrimientos. Por supuesto, la señal de que la media está recorrida yace en la naturaleza de la pendiente de la gráfica de suma acumulada. El propósito de la gráfica es detectar cambios que se alejan del nivel de referencia. Una pendiente diferente de cero (en cualquier dirección) representa un cambio a partir del nivel de referencia. Una pendiente positiva indica un aumento en la media por arriba del nivel de referencia, mientras que una pendiente negativa señala una disminución.

Las gráficas de suma acumulada a menudo se diseñan con un nivel de *calidad aceptable* definido (AQL) y un nivel de *calidad rechazable* (RQL) preestablecido por el usuario. Ambos representan valores de la media. Éstos se pueden ver como si jugaran papeles similares a los de las medias nula y alternativa en la prueba de hipótesis. Considere una situación donde el analista desea detectar un aumento en el valor de la media del proceso. Usaremos la notación μ_0 para AQL y μ_1 para RQL y $\mu_1 > \mu_0$. El nivel de referencia se fija ahora en

$$W = \frac{\mu_0 + \mu_1}{2}$$

Los valores de S_r ($r = 1, 2, \dots$) tendrán una pendiente negativa si la media del proceso está en μ_0 y una pendiente positiva si la media del proceso está en μ_1 .

Regla de decisión para las gráficas cusum

Como indicamos antes, la pendiente de la gráfica de suma acumulada proporciona la señal de acción para el analista de control de calidad. La regla de decisión requiere la acción si, en el r -ésimo periodo de muestreo,

$$d_r > h,$$

donde h es un valor preestablecido que se llama **longitud del intervalo de decisión** y

$$d_r = S_r - \min_{1 \leq i \leq r-1} S_i$$

En otras palabras, se toma la acción si los datos revelan que el valor de la suma acumulada real excede en una cantidad específica al valor previo de la suma acumulada más pequeña.

Una modificación en la mecánica que se describió antes permite la facilidad en el empleo del método. Describimos un procedimiento que grafica las sumas acumuladas y calcula las diferencias. Una modificación simple implica graficar las diferencias de manera directa y permitir la verificación contra el intervalo de decisión. La expresión general para d_r es bastante simple. Para el procedimiento de la suma acumulada donde se detectan aumentos en la media,

$$d_r = \max\{0, d_{r-1} + (X_r - W)\}.$$

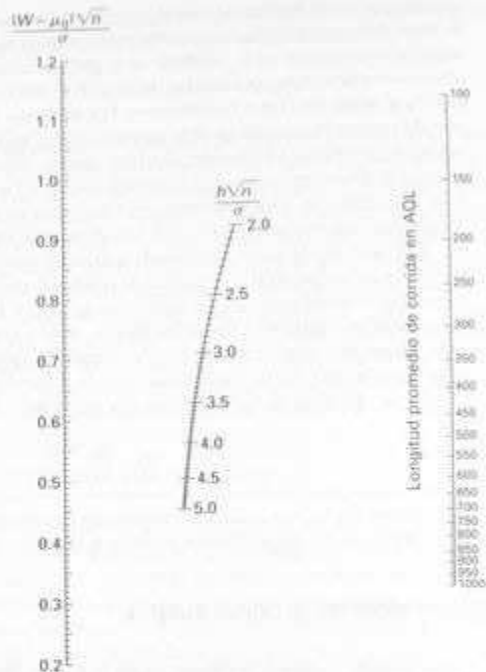


Figura 17.11 Nomograma para la gráfica de control de la suma acumulada en términos de la longitud promedio de prueba en $AQL = \mu_1$.

La elección del valor de h es, por supuesto, muy importante. En este libro dejamos no proporcionar los detalles que aparecen en la literatura que trata de esta elección. Se remite al lector a Ewan y Kemp y Montgomery (véase la bibliografía) para una exposición más completa. Una consideración importante es la longitud esperada de la corrida. Idealmente, la longitud esperada de la corrida es bastante grande bajo $\mu = \mu_0$ y bastante pequeña cuando $\mu = \mu_1$.

Ewan y Kemp produjeron gráficas que ayudan en la elección del intervalo de decisión h . Estas gráficas se reproducen en las figuras 17.11 y 17.12. Se llaman **nomogramas** y permiten la elección de h para controlar la longitud promedio de corrida bajo la condición $AQL = \mu_0$.

El siguiente ejemplo ilustra el uso de los nomogramas para elegir un valor de h . El valor de n en las expresiones de la gráfica indica el tamaño de las muestras en casos donde las X_1, X_2, \dots observables son *promedios*.

Ejemplo 17.6 Suponga que un analista de control de calidad desea producir una gráfica de suma acumulada donde la longitud promedio de corrida en un nivel de

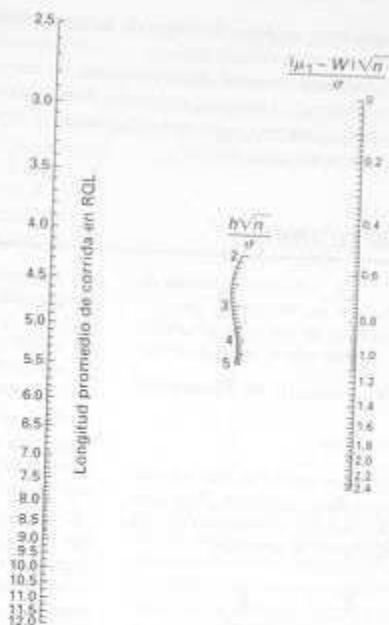


Figura 17.12 Normograma para la gráfica de control de suma acumulada en términos de la longitud promedio de corrida en RQL = μ_1 .

calidad aceptable es 500, cuando de hecho, es de interés detectar un aumento en la media de a lo más σ . ¿Cuál es el valor del intervalo de decisión, h ? Se utilizarán muestras de tamaño 4 con promedios calculados con el procedimiento de suma acumulada.

SOLUCIÓN

Como

$$W = \frac{\mu_0 + \mu_1}{2}, \quad \frac{(W - \mu_0) \sqrt{n}}{\sigma} = 1.0.$$

Si conectamos el punto en el eje vertical izquierdo de la figura 17.11 en 1.0 con el valor de 500 en el eje vertical derecho, observamos que

$$\frac{h \sqrt{n}}{\sigma} = 2.35.$$

Entonces usamos

$$h = \frac{2.35 \sigma}{2} = 1.175 \sigma.$$

Obviamente, se debe disponer de una estimación de σ a fin de que el intervalo de decisión se elija en la práctica.

Podemos observar ahora mediante un uso similar de la figura 17.12 que para estas condiciones la longitud promedio de corrida en ROL es aproximadamente 3.2. Esto proporciona una respuesta relativamente rápida cuando, en realidad, la media se corre en una cantidad σ .

Ejercicios de repaso

1. Considere x_1, x_2, \dots, x_n variables aleatorias de Poisson independientes con parámetros $\mu_1, \mu_2, \dots, \mu_n$. Utilice las propiedades de las funciones generadoras de momento para mostrar que la variable aleatoria $\sum_{i=1}^n x_i$ es una variable aleatoria de Poisson con media $\sum_{i=1}^n \mu_i$ y varianza $\sum_{i=1}^n \mu_i$.

2. Considere los siguientes datos tomados en subgrupos de tamaño 5. Los datos contienen 20 promedios, y rangos del diámetro (en milímetros) de una parte componente importante de un motor.

| Muestra | \bar{X} | R |
|---------|-----------|--------|
| 1 | 2.3972 | 0.0052 |
| 2 | 2.4191 | 0.0117 |
| 3 | 2.4215 | 0.0062 |
| 4 | 2.3917 | 0.0089 |
| 5 | 2.4151 | 0.0095 |
| 6 | 2.4027 | 0.0101 |
| 7 | 2.3921 | 0.0091 |
| 8 | 2.4171 | 0.0059 |
| 9 | 2.3951 | 0.0068 |
| 10 | 2.4215 | 0.0048 |
| 11 | 2.3687 | 0.0082 |
| 12 | 2.4107 | 0.0032 |
| 13 | 2.4009 | 0.0077 |
| 14 | 2.3992 | 0.0107 |
| 15 | 2.3889 | 0.0025 |
| 16 | 2.4107 | 0.0138 |
| 17 | 2.4109 | 0.0037 |
| 18 | 2.3944 | 0.0052 |
| 19 | 2.3951 | 0.0038 |
| 20 | 2.4015 | 0.0017 |

Elabore gráficas \bar{X} y R . ¿El proceso parece estar bajo control?

3. Suponga para el ejercicio 2 que el comprador fija especificaciones para la parte. Las especificaciones requieren que el diámetro caiga en el rango cubierto por 2.4000 ± 0.0100 mm. ¿Qué proporción de unidades producidas por este proceso no cumplirán con las especificaciones?

4. Para la situación del ejercicio 2, proporcione estimaciones numéricas de la media y de la desviación estándar del diámetro para la parte que se fabrica en el proceso.

5. Considere los datos de la tabla 17.1. Suponga que se toman muestras adicionales de tamaño 5 y se registra la resistencia de ruptura. El muestreo produce los siguientes resultados (en libras por pulgada cuadrada).

(a) Grafique los datos, utilice las gráficas \bar{X} y R para los datos preliminares de la tabla 17.1.

(b) ¿El proceso parece estar bajo control? Si no, explique por qué.

| Muestra | \bar{X}_i | R_i |
|---------|-------------|-------|
| 1 | 1511 | 22 |
| 2 | 1508 | 14 |
| 3 | 1522 | 11 |
| 4 | 1488 | 18 |
| 5 | 1519 | 6 |
| 6 | 1524 | 11 |
| 7 | 1519 | 8 |
| 8 | 1504 | 7 |
| 9 | 1500 | 8 |
| 10 | 1519 | 14 |

6. Considere un proceso bajo control con media $\mu = 25$ y $\sigma = 1.0$. Suponga que se usan subgrupos de tamaño 5 con límites de control, $\mu \pm 3\sigma/\sqrt{n}$, y línea central en μ . Suponga que ocurre un corrimiento en la media, y por ello la nueva media es $\mu = 26.5$.

(a) ¿Cuál es el número promedio de muestras que se requiere (después del corrimiento) para detectar la situación fuera de control?

(b) ¿Cuál es la desviación estándar del número de corridas que se requiere?

7. Considere la situación del ejemplo 17.2. Se toman los siguientes datos de muestras adicionales de tamaño 5. Grafique los valores \bar{X} y S sobre las gráficas \bar{X} y S que producen los datos en la muestra preliminar. ¿El proceso parece estar bajo control? Explique por qué sí o por qué no.

| Muestra | \bar{X}_i | S_i |
|---------|-------------|-------|
| 1 | 62.280 | 0.062 |
| 2 | 62.319 | 0.049 |
| 3 | 62.297 | 0.077 |
| 4 | 62.318 | 0.042 |
| 5 | 62.315 | 0.038 |
| 6 | 62.389 | 0.052 |
| 7 | 62.401 | 0.059 |
| 8 | 62.315 | 0.042 |
| 9 | 62.298 | 0.036 |
| 10 | 62.337 | 0.068 |

8. Se toman muestras de tamaño 50 cada hora de un proceso que produce cierto tipo de artículo que se considera defectuoso o no defectuoso. Se toman 20 muestras.

| Muestra | Número de artículos defectuosos | Muestra | Número de artículos defectuosos |
|---------|---------------------------------|---------|---------------------------------|
| 1 | 4 | 11 | 2 |
| 2 | 3 | 12 | 4 |
| 3 | 5 | 13 | 1 |
| 4 | 3 | 14 | 2 |
| 5 | 2 | 15 | 3 |
| 6 | 2 | 16 | 1 |
| 7 | 2 | 17 | 1 |
| 8 | 1 | 18 | 2 |
| 9 | 4 | 19 | 3 |
| 10 | 3 | 20 | 1 |

(a) Construya una gráfica de control para controlar la proporción de defectuosos.

(b) ¿El proceso parece estar bajo control? Explique.

9. Para la situación del ejercicio 8, suponga que se colectan datos adicionales como se muestra a continuación (muestras de tamaño 50):

| Muestra | Número de artículos defectuosos |
|---------|---------------------------------|
| 1 | 3 |
| 2 | 4 |
| 3 | 2 |
| 4 | 2 |
| 5 | 3 |
| 6 | 1 |
| 7 | 3 |
| 8 | 5 |
| 9 | 7 |
| 10 | 7 |

¿El proceso parece estar bajo control? Explique.

10. Se intenta un esfuerzo de control de calidad para un proceso donde se fabrican grandes placas de acero e interesan los defectos superficiales. El objetivo es establecer una gráfica de control de calidad para el número de defectos por placa. Los datos son los siguientes:

| Gráfica muestral | Número de defectos | Gráfica muestral | Número de defectos |
|------------------|--------------------|------------------|--------------------|
| 1 | 4 | 11 | 1 |
| 2 | 2 | 12 | 2 |
| 3 | 1 | 13 | 2 |
| 4 | 3 | 14 | 3 |
| 5 | 0 | 15 | 1 |
| 6 | 4 | 16 | 4 |
| 7 | 5 | 17 | 3 |
| 8 | 3 | 18 | 2 |
| 9 | 2 | 19 | 1 |
| 10 | 2 | 20 | 3 |

Establezca la gráfica de control apropiada; utilice esta información muestral. ¿El proceso parece estar bajo control?

Bibliografía

- BARTLETT, M. S., and D. G. KENDALL. "The Statistical Analysis of Variance Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society*, Ser. B, 8, 128-150, 1946.
- BICKEL, P. J., and DOKSUM, K. A. *Mathematical Statistics*. Oakland, Calif.: Holden-Day, Inc., 1977.
- BOWKER, A. H., and G. J. LIEBERMAN. *Engineering Statistics*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, Inc., 1972.
- BOX, G. E. P. "Signal to Noise Ratios, Performance Criteria and Transformations (with discussion)," *Technometrics*, 307 1-40, 1988.
- BOX, G. E. P., and C. A. FUNG. "Studies in Quality Improvement: Minimizing Transmitted Variation by Parameter Design," Report 8, University of Wisconsin-Madison, Center for Quality and Productivity Improvement.
- BOX, G. E. P., W. G. HUNTER, and J. S. HUNTER. *Statistics for Experimenters*. New York: John Wiley & Sons, Inc., 1978.
- BROWNLIE, K. A. *Statistical Theory and Methodology in Science and Engineering*, 2nd ed. New York: John Wiley & Sons, Inc., 1965.
- CARROLL, R. J., and D. RUPPERT. "Transformation and Weighting in Regression," New York: Chapman and Hall, 1988, p. 348.
- CHATTERJEE, S., and B. PRICE. *Regression Analysis by Example*. New York: John Wiley & Sons, Inc., 1977.
- COOK, R. D., and S. WEISBERG. *Residuals and Influence in Regression*. New York: Chapman and Hall, 1982.
- DANIEL, C., and F. WOOD. *Fitting Equations to Data*, 2nd ed. New York: John Wiley & Sons, Inc., 1980.
- DANIEL, W. W. *Applied Nonparametric Statistics*. Boston: Houghton Mifflin Company, 1978.
- DERMAN, C., L. GLASER, and I. OLKIN. *Probability Models and Applications*. New York: Macmillan Publishing Company, 1980.
- DEVORE, J. L. *Probability and Statistics for Engineering and the Sciences*, 4th ed. Monterey, Calif.: Brooks/Cole Publishing Co., 1995.
- DIXON, W. J., and F. J. MASSEY, JR. *Introduction to Statistical Analysis*, 3rd ed. New York: McGraw-Hill Book Company, 1969.
- DRAPER, N., and H. SMITH. *Applied Regression Analysis*, 2nd ed. New York: John Wiley & Sons, Inc., 1981.

- DYER, D. D., and J. P. KEATING, "On the Determination of Critical Values for Bartlett's Test," *J. Am. Stat. Assoc.*, Vol. 75, 1980.
- EWAN, W. D., and K. W. KEMP, "Sampling Inspection of Continuous Processes with No Autocorrelation Between Successive Results," *Biometrika*, Vol. 41, 363-380, 1960.
- GUNST, R. F., and R. L. MASON, *Regression Analysis and Its Application: A Data-Oriented Approach*, New York: Marcel Dekker, Inc., 1980.
- GUTTMAN, I., and S. S. WILKS, *Introductory Engineering Statistics*, New York: John Wiley & Sons, Inc., 1965.
- HICKS, C. R., *Fundamental Concepts in the Design of Experiments*, 2nd ed. New York: Holt, Rinehart and Winston, 1973.
- HOCKING, R. R., "The Analysis and Selection of Variables in Linear Regression," *Biometrics*, Vol. 32, 1976.
- HOERL, A. E., and R. W. KENNARD, "Ridge Regression: Applications to Nonorthogonal Problems," *Technometrics*, Vol. 12, No. 1, 1970.
- HOGG, R. V., and A. T. CRAIG, *Introduction to Mathematical Statistics*, 4th ed. New York: Macmillan Publishing Company, 1978.
- HOGG, R. V., and J. LEDOLTER, *Engineering Statistics*, 2nd ed. Upper Saddle River, N.J.: Prentice-Hall, Inc., 1992.
- HOLLANDER, M., and D. WOLFE, *Nonparametric Statistical Methods*, New York: John Wiley & Sons, Inc., 1973.
- JOHNSON, N. L., and F. C. LEONE, *Statistics and Experimental Design: In Engineering and the Physical Sciences*, Vols. I and II, 2nd ed. New York: John Wiley & Sons, Inc., 1977.
- KACKAR, R., "Off-Line Quality Control, Parameter Design, and the Taguchi Methods," *Journal of Quality Technology*, 17, 176-188, 1985.
- KOOPMANS, L. H., *An Introduction to Contemporary Statistics*, Boston: Duxbury Press, 1981.
- LARSEN, R. J., and M. L. MORRIS, *Introduction to Mathematical Statistics*, Upper Saddle River, N.J.: Prentice-Hall, Inc., 1981.
- LEHMANN, E. L., *Nonparametrics: Statistical Methods Based on Ranks*, San Francisco: Holden-Day, Inc., 1975.
- LENTNER, M., and T. BISHOP, *Design and Analysis of Experiments*, 2nd ed. Blacksburg, Va.: Valley Book Co., 1986.
- McLAVE, J., and F. DIETRICH, *Statistics*, San Francisco: Dellen Publishing Co., 1978.
- MONTGOMERY, D. C., *Introduction to Statistical Quality Control*, 3rd ed. New York: John Wiley & Sons, Inc., 1996.
- MOSTELLER, F., and J. TUKEY, *Data Analysis and Regression*, Reading, Mass.: Addison-Wesley Publishing Co., Inc., 1977.
- MYERS, R. H., *Classical and Modern Regression with Applications*, 2nd ed. Boston: Duxbury Press, 1990.
- MYERS, R. H., A. I. KHURI, and G. YINING, "Response Surface Alternatives

- to the Taguchi Robust Parameter Design Approach," *The American Statistician*, Vol. 46, No. 2, 131-139, 1992.
- MYERS, R. H. and D. C. MONTGOMERY, *Response Surface Methodology: Product and Process Optimization Using Designed Experiments*. New York: John Wiley and Sons, Inc., 1995.
- NOETHER, G. E. *Introduction to Statistics: A Nonparametric Approach*, 2nd ed. Boston: Houghton Mifflin Company, 1976.
- OTT, L. *An Introduction to Statistical Methods and Data Analysis*, 4th ed. Boston: Duxbury Press, 1994.
- PLACKETT, R. L., and J. P. BURMAN, "The Design of Multifactor Experiments," *Biometrika*, Vol. 33, 305-325, 1946.
- ROSS, S. *Introduction to Applied Probability Models*, 2nd ed. New York: Academic Press, Inc., 1980.
- SHILLING, E. G., and P. R. NELSON, "The Effect of Nonnormality on the Control Limits of \bar{X} Charts," *J. Quality Tech.*, Vol. 8, 1976.
- SCHMIDT, S. R., and R. G. LAUNSBY, "Understanding Industrial Designed Experiments," Colorado Springs, Colo., 1991.
- SHOEMAKER, A. C., K. L. TSUI, and C. F. J. WU, "Economical Experimentation Methods for Robust Parameter Design," Paper presented at the Fall Technical Conference, American Society of Quality Control, Houston, Tex., 1989.
- SNEDECOR, G., and W. G. COCHRAN, *Statistical Methods*, 7th ed. Ames, Iowa: The Iowa State University Press, 1980.
- STEEL, R. G. D., and J. H. TORRIE, *Principles and Procedures of Statistics*, 2nd ed. New York: McGraw-Hill Book Company, 1979.
- TAGUCHI, G. *Introduction to Quality Engineering*. White Plains, N.Y.: UNIPUB/Kraus International, 1991.
- TAGUCHI, G., and Y. WU, "Introduction to Off-Line Quality Control," Central Japan Quality Control Association (available from American Supplier Institute, 32100 Detroit Industrial Expressway, Romulus, Mich. 48174), 1980.
- THOMPSON, W. O., and F. B. CADY, *Proceedings of the University of Kentucky Conference on Regression with a Large Number of Predictor Variables*, Lexington, 1973.
- TUKEY, T. W. *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley Publishing Co., Inc., 1977.
- VINING, G. G., and R. H. MYERS, "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach," *Journal of Quality Technology*, 22, 38-45, 1990.
- WELCH, W. J., T. K. YU, S. M. KANG, and J. SACKS, "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology*, 22, 15-22, 1990.
- WINER, B. J. *Statistical Principles in Experimental Design*, 2nd ed. New York: McGraw-Hill Book Company, 1971.

Apéndice: tablas estadísticas

- A.1 Sumas de probabilidad binomial $\sum_{x=0}^n b(x; n, p)$
- A.2 Sumas de probabilidad de Poisson $\sum_{x=0}^{\infty} p(x; \mu)$
- A.3 Áreas bajo la curva normal
- A.4 Valores críticos de la distribución t
- A.5 Valores críticos de la distribución χ^2 cuadrada
- A.6 Valores críticos de la distribución F
- A.7 Factores de tolerancia para distribuciones normales
- A.8 Tamaño muestral para la prueba t de la media
- A.9 Tamaño muestral para la prueba t de la diferencia entre dos medias
- A.10 Valores críticos para la prueba de Bartlett
- A.11 Valores críticos para la prueba de Cochran
- A.12 Rangos studentizados significativos mínimos r_p
- A.13 Valores de $d_{n,r}(k, v)$ para comparaciones bilaterales entre k tratamientos y un control
- A.14 Valores de $d_{n,r}(k, v)$ para comparaciones unilaterales entre k tratamientos y un control
- A.15 Potencia de la prueba de análisis de varianza
- A.16 Valores críticos para la prueba de rangos con signo
- A.17 Valores críticos para la prueba de suma de rangos
- A.18 $P(V \leq v^*$ cuando H_0 es verdadera) en la prueba de corridas
- A.19 Tamaño muestral para límites de tolerancia no paramétricos bilaterales
- A.20 Tamaño muestral para límites de tolerancia no paramétricos unilaterales
- A.21 Valores críticos del coeficiente de correlación de rangos de Spearman
- A.22 Puntos porcentuales superiores de la distribución de rango studentizado
- A.23 Factores para la elaboración de gráficas de control
- A.24 Tabla de la función gamma incompleta

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=0}^r b(x; n, p)$

| n | r | p | | | | | | | | | |
|----|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
| 8 | 0 | 0.4305 | 0.1678 | 0.1001 | 0.0576 | 0.0168 | 0.0039 | 0.0007 | 0.0001 | 0.0000 | |
| | 1 | 0.8131 | 0.5033 | 0.3671 | 0.2553 | 0.1064 | 0.0352 | 0.0085 | 0.0013 | 0.0001 | |
| | 2 | 0.9619 | 0.7869 | 0.6785 | 0.5518 | 0.3154 | 0.1445 | 0.0498 | 0.0113 | 0.0012 | 0.0000 |
| | 3 | 0.9950 | 0.9437 | 0.8862 | 0.8059 | 0.6941 | 0.5603 | 0.4337 | 0.0560 | 0.0104 | 0.0004 |
| | 4 | 0.9996 | 0.9896 | 0.9727 | 0.9420 | 0.8263 | 0.6367 | 0.4059 | 0.1941 | 0.0563 | 0.0050 |
| | 5 | 1.0000 | 0.9988 | 0.9958 | 0.9882 | 0.9502 | 0.8555 | 0.6846 | 0.4482 | 0.2031 | 0.0381 |
| | 6 | | 0.9991 | 0.9996 | 0.9987 | 0.9915 | 0.9648 | 0.8956 | 0.7447 | 0.4967 | 0.1869 |
| | 7 | | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9961 | 0.9832 | 0.9424 | 0.8322 | 0.5695 |
| | 8 | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 9 | 0 | 0.3874 | 0.1342 | 0.0751 | 0.0404 | 0.0101 | 0.0020 | 0.0003 | 0.0000 | | |
| | 1 | 0.7748 | 0.4362 | 0.3003 | 0.1960 | 0.0795 | 0.0195 | 0.0038 | 0.0004 | 0.0000 | |
| | 2 | 0.9470 | 0.7382 | 0.6007 | 0.4628 | 0.2318 | 0.0898 | 0.0250 | 0.0043 | 0.0003 | 0.0000 |
| | 3 | 0.9917 | 0.9144 | 0.8343 | 0.7297 | 0.4926 | 0.2359 | 0.0994 | 0.0253 | 0.0051 | 0.0001 |
| | 4 | 0.9991 | 0.9804 | 0.9511 | 0.9012 | 0.7334 | 0.5000 | 0.2666 | 0.0988 | 0.0196 | 0.0009 |
| | 5 | 0.9999 | 0.9969 | 0.9900 | 0.9747 | 0.8006 | 0.7461 | 0.5174 | 0.2703 | 0.0856 | 0.0063 |
| | 6 | 1.0000 | 0.9997 | 0.9987 | 0.9957 | 0.9750 | 0.9102 | 0.7682 | 0.5372 | 0.2618 | 0.0530 |
| | 7 | | 1.0000 | 0.9999 | 0.9996 | 0.9962 | 0.9805 | 0.9295 | 0.8040 | 0.5638 | 0.2252 |
| | 8 | | | 1.0000 | 1.0000 | 0.9997 | 0.9980 | 0.9899 | 0.9596 | 0.8658 | 0.6126 |
| 9 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |
| 10 | 0 | 0.3487 | 0.1074 | 0.0563 | 0.0282 | 0.0060 | 0.0010 | 0.0001 | 0.0000 | | |
| | 1 | 0.7361 | 0.3758 | 0.2440 | 0.1493 | 0.0464 | 0.0107 | 0.0017 | 0.0001 | 0.0000 | |
| | 2 | 0.9298 | 0.6778 | 0.5256 | 0.3828 | 0.1873 | 0.0547 | 0.0123 | 0.0016 | 0.0001 | |
| | 3 | 0.9872 | 0.8791 | 0.7759 | 0.6496 | 0.3823 | 0.1719 | 0.0548 | 0.0106 | 0.0009 | 0.0000 |
| | 4 | 0.9984 | 0.9672 | 0.9219 | 0.8497 | 0.6331 | 0.3770 | 0.1662 | 0.0474 | 0.0064 | 0.0002 |
| | 5 | 0.9999 | 0.9936 | 0.9803 | 0.9527 | 0.8338 | 0.6230 | 0.3669 | 0.1503 | 0.0328 | 0.0016 |
| | 6 | 1.0000 | 0.9991 | 0.9965 | 0.9894 | 0.9452 | 0.8281 | 0.6177 | 0.3364 | 0.1299 | 0.0128 |
| | 7 | | 0.9999 | 0.9996 | 0.9984 | 0.9877 | 0.9453 | 0.8327 | 0.6172 | 0.3222 | 0.0702 |
| | 8 | | 1.0000 | 1.0000 | 0.9999 | 0.9983 | 0.9893 | 0.9536 | 0.8507 | 0.6242 | 0.2639 |
| 9 | | | | 1.0000 | 0.9999 | 0.9990 | 0.9940 | 0.9718 | 0.8926 | 0.6512 | |
| 10 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |
| 11 | 0 | 0.3138 | 0.0859 | 0.0422 | 0.0198 | 0.0036 | 0.0005 | 0.0000 | | | |
| | 1 | 0.6974 | 0.3221 | 0.1971 | 0.1130 | 0.0302 | 0.0059 | 0.0007 | 0.0000 | | |
| | 2 | 0.9104 | 0.6174 | 0.4852 | 0.3127 | 0.1189 | 0.0327 | 0.0059 | 0.0006 | 0.0000 | |
| | 3 | 0.9815 | 0.8369 | 0.7133 | 0.5696 | 0.2865 | 0.1133 | 0.0293 | 0.0043 | 0.0002 | |
| | 4 | 0.9972 | 0.9496 | 0.8854 | 0.7897 | 0.5328 | 0.2744 | 0.0994 | 0.0216 | 0.0020 | 0.0000 |
| | 5 | 0.9997 | 0.9883 | 0.9657 | 0.9218 | 0.7535 | 0.5000 | 0.2465 | 0.0782 | 0.0117 | 0.0003 |
| | 6 | 1.0000 | 0.9980 | 0.9924 | 0.9784 | 0.9006 | 0.7256 | 0.4672 | 0.2103 | 0.0504 | 0.0028 |
| | 7 | | 0.9998 | 0.9988 | 0.9957 | 0.9707 | 0.8867 | 0.7037 | 0.4304 | 0.1611 | 0.0185 |
| | 8 | | 1.0000 | 0.9999 | 0.9994 | 0.9941 | 0.9673 | 0.8811 | 0.6873 | 0.3828 | 0.0896 |
| 9 | | | 1.0000 | 1.0000 | 0.9993 | 0.9941 | 0.9698 | 0.8870 | 0.6779 | 0.3026 | |
| 10 | | | | | 1.0000 | 0.9995 | 0.9964 | 0.9802 | 0.9141 | 0.6882 | |
| 11 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=0}^r b(x; n, p)$

| n | r | p | | | | | | | | | | |
|----|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | |
| 12 | 0 | 0.2824 | 0.0687 | 0.0317 | 0.0138 | 0.0022 | 0.0002 | 0.0000 | | | | |
| | 1 | 0.6590 | 0.2749 | 0.1584 | 0.0850 | 0.0196 | 0.0032 | 0.0003 | 0.0000 | | | |
| | 2 | 0.8891 | 0.5583 | 0.3907 | 0.2528 | 0.0834 | 0.0193 | 0.0028 | 0.0002 | 0.0000 | | |
| | 3 | 0.9744 | 0.7946 | 0.6488 | 0.4925 | 0.2253 | 0.0730 | 0.0153 | 0.0017 | 0.0001 | | |
| | 4 | 0.9957 | 0.9274 | 0.8424 | 0.7237 | 0.4382 | 0.1938 | 0.0573 | 0.0095 | 0.0006 | 0.0000 | |
| | 5 | 0.9995 | 0.9806 | 0.9456 | 0.8821 | 0.6652 | 0.3872 | 0.1582 | 0.0386 | 0.0039 | 0.0001 | |
| | 6 | 0.9999 | 0.9961 | 0.9857 | 0.9614 | 0.8418 | 0.6128 | 0.3348 | 0.1178 | 0.0194 | 0.0005 | |
| | 7 | 1.0000 | 0.9994 | 0.9972 | 0.9905 | 0.9427 | 0.8062 | 0.5618 | 0.2763 | 0.0726 | 0.0043 | |
| | 8 | | 0.9999 | 0.9996 | 0.9983 | 0.9847 | 0.9270 | 0.7747 | 0.5075 | 0.2054 | 0.0256 | |
| | 9 | | | 1.0000 | 0.9998 | 0.9972 | 0.9807 | 0.9166 | 0.7472 | 0.4417 | 0.1109 | |
| | 10 | | | | 1.0000 | 0.9997 | 0.9968 | 0.9804 | 0.9150 | 0.7251 | 0.3410 | |
| | 11 | | | | | 1.0000 | 0.9998 | 0.9978 | 0.9862 | 0.9517 | 0.7176 | |
| 12 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | |
| 13 | 0 | 0.2542 | 0.0550 | 0.0238 | 0.0097 | 0.0013 | 0.0001 | 0.0000 | | | | |
| | 1 | 0.6213 | 0.2356 | 0.1267 | 0.0637 | 0.0126 | 0.0017 | 0.0001 | 0.0000 | | | |
| | 2 | 0.8661 | 0.5017 | 0.3326 | 0.2025 | 0.0579 | 0.0112 | 0.0013 | 0.0001 | | | |
| | 3 | 0.9658 | 0.7473 | 0.5843 | 0.4206 | 0.1686 | 0.0461 | 0.0078 | 0.0007 | 0.0000 | | |
| | 4 | 0.9935 | 0.9009 | 0.7940 | 0.6543 | 0.3530 | 0.1334 | 0.0321 | 0.0040 | 0.0002 | 0.0000 | |
| | 5 | 0.9991 | 0.9700 | 0.9198 | 0.8346 | 0.5744 | 0.2905 | 0.0977 | 0.0182 | 0.0012 | 0.0000 | |
| | 6 | 0.9999 | 0.9920 | 0.9757 | 0.9376 | 0.7712 | 0.5000 | 0.2288 | 0.0624 | 0.0079 | 0.0001 | |
| | 7 | 1.0000 | 0.9980 | 0.9944 | 0.9818 | 0.9023 | 0.7095 | 0.4256 | 0.1654 | 0.0300 | 0.0009 | |
| | 8 | | 0.9998 | 0.9990 | 0.9960 | 0.9679 | 0.8666 | 0.6470 | 0.3457 | 0.0991 | 0.0065 | |
| | 9 | | 1.0000 | 0.9999 | 0.9993 | 0.9922 | 0.9539 | 0.8314 | 0.5794 | 0.2527 | 0.0342 | |
| | 10 | | | 1.0000 | 0.9999 | 0.9987 | 0.9888 | 0.9421 | 0.7975 | 0.4903 | 0.1379 | |
| | 11 | | | | 1.0000 | 0.9999 | 0.9983 | 0.9874 | 0.9363 | 0.7664 | 0.3787 | |
| | 12 | | | | | 1.0000 | 0.9999 | 0.9987 | 0.9903 | 0.9450 | 0.7438 | |
| 13 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | |
| 14 | 0 | 0.2288 | 0.0440 | 0.0178 | 0.0068 | 0.0008 | 0.0001 | 0.0000 | | | | |
| | 1 | 0.5846 | 0.1979 | 0.1010 | 0.0475 | 0.0081 | 0.0009 | 0.0001 | | | | |
| | 2 | 0.8416 | 0.4481 | 0.2811 | 0.1608 | 0.0398 | 0.0065 | 0.0006 | 0.0000 | | | |
| | 3 | 0.9559 | 0.6982 | 0.5213 | 0.3552 | 0.1243 | 0.0287 | 0.0039 | 0.0002 | | | |
| | 4 | 0.9908 | 0.8702 | 0.7415 | 0.5842 | 0.2793 | 0.0898 | 0.0175 | 0.0017 | 0.0000 | | |
| | 5 | 0.9985 | 0.9561 | 0.8883 | 0.7805 | 0.4859 | 0.2120 | 0.0583 | 0.0083 | 0.0004 | 0.0000 | |
| | 6 | 0.9998 | 0.9884 | 0.9617 | 0.9067 | 0.6925 | 0.3953 | 0.1501 | 0.0315 | 0.0024 | 0.0000 | |
| | 7 | 1.0000 | 0.9976 | 0.9897 | 0.9685 | 0.8499 | 0.6047 | 0.3075 | 0.0933 | 0.0116 | 0.0002 | |
| | 8 | | 0.9996 | 0.9978 | 0.9917 | 0.9417 | 0.7880 | 0.5141 | 0.2195 | 0.0439 | 0.0015 | |
| | 9 | | 1.0000 | 0.9997 | 0.9983 | 0.9825 | 0.9102 | 0.7207 | 0.4158 | 0.1298 | 0.0062 | |
| | 10 | | | 1.0000 | 0.9998 | 0.9961 | 0.9713 | 0.8757 | 0.6448 | 0.3018 | 0.0441 | |
| | 11 | | | | 1.0000 | 0.9994 | 0.9935 | 0.9602 | 0.8392 | 0.5519 | 0.1584 | |
| | 12 | | | | | 0.9999 | 0.9991 | 0.9919 | 0.9525 | 0.8021 | 0.4154 | |
| | 13 | | | | | | 1.0000 | 0.9999 | 0.9992 | 0.9932 | 0.9560 | 0.7712 |
| 14 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | |

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=r}^n b(x; n, p)$

| n | r | p | | | | | | | | | | | | | |
|----|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--|--|--|--|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | | | | |
| 15 | 0 | 0.2059 | 0.0352 | 0.0134 | 0.0047 | 0.0005 | 0.0000 | | | | | | | | |
| | 1 | 0.5490 | 0.1671 | 0.0802 | 0.0353 | 0.0052 | 0.0005 | 0.0000 | | | | | | | |
| | 2 | 0.8159 | 0.3980 | 0.2361 | 0.1268 | 0.0271 | 0.0037 | 0.0003 | 0.0000 | | | | | | |
| | 3 | 0.9444 | 0.6482 | 0.4613 | 0.2969 | 0.0905 | 0.0176 | 0.0019 | 0.0001 | | | | | | |
| | 4 | 0.9873 | 0.8358 | 0.6865 | 0.5155 | 0.2173 | 0.0592 | 0.0094 | 0.0007 | 0.0000 | | | | | |
| | 5 | 0.9978 | 0.9389 | 0.8516 | 0.7216 | 0.4032 | 0.1509 | 0.0338 | 0.0037 | 0.0001 | | | | | |
| | 6 | 0.9997 | 0.9819 | 0.9434 | 0.8689 | 0.6098 | 0.3036 | 0.0951 | 0.0152 | 0.0008 | | | | | |
| | 7 | 1.0000 | 0.9958 | 0.9827 | 0.9500 | 0.7969 | 0.5000 | 0.2131 | 0.0900 | 0.0042 | 0.0000 | | | | |
| | 8 | | 0.9992 | 0.9958 | 0.9848 | 0.9050 | 0.6984 | 0.2902 | 0.1331 | 0.0181 | 0.0002 | | | | |
| | 9 | | 0.9999 | 0.9992 | 0.9963 | 0.9602 | 0.8491 | 0.5968 | 0.2784 | 0.0611 | 0.0023 | | | | |
| | 10 | | 1.0000 | 0.9999 | 0.9993 | 0.9907 | 0.9408 | 0.7827 | 0.4845 | 0.1642 | 0.0127 | | | | |
| | 11 | | | 1.0000 | 0.9999 | 0.9981 | 0.9824 | 0.9095 | 0.7031 | 0.3518 | 0.0556 | | | | |
| | 12 | | | | 1.0000 | 0.9997 | 0.9963 | 0.9729 | 0.8732 | 0.6620 | 0.1841 | | | | |
| | 13 | | | | | 1.0000 | 0.9995 | 0.9948 | 0.9647 | 0.8329 | 0.4510 | | | | |
| | 14 | | | | | | 1.0000 | 0.9995 | 0.9953 | 0.9648 | 0.7941 | | | | |
| 15 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | |
| 16 | 0 | 0.1853 | 0.0281 | 0.0100 | 0.0033 | 0.0003 | 0.0000 | | | | | | | | |
| | 1 | 0.5147 | 0.1407 | 0.0635 | 0.0261 | 0.0033 | 0.0003 | 0.0000 | | | | | | | |
| | 2 | 0.7892 | 0.3518 | 0.1971 | 0.0994 | 0.0183 | 0.0021 | 0.0001 | | | | | | | |
| | 3 | 0.9316 | 0.5981 | 0.4050 | 0.2459 | 0.0651 | 0.0106 | 0.0009 | 0.0000 | | | | | | |
| | 4 | 0.9830 | 0.7982 | 0.6302 | 0.4499 | 0.1669 | 0.0384 | 0.0049 | 0.0003 | | | | | | |
| | 5 | 0.9967 | 0.9183 | 0.8103 | 0.6598 | 0.3288 | 0.1051 | 0.0191 | 0.0016 | 0.0000 | | | | | |
| | 6 | 0.9995 | 0.9723 | 0.9204 | 0.8247 | 0.5272 | 0.2272 | 0.0583 | 0.0071 | 0.0002 | | | | | |
| | 7 | 0.9999 | 0.9930 | 0.9729 | 0.9256 | 0.7161 | 0.4018 | 0.1423 | 0.0237 | 0.0015 | 0.0000 | | | | |
| | 8 | 1.0000 | 0.9985 | 0.9925 | 0.9743 | 0.8577 | 0.5982 | 0.2839 | 0.0744 | 0.0070 | 0.0001 | | | | |
| | 9 | | 0.9998 | 0.9984 | 0.9929 | 0.9417 | 0.7728 | 0.4728 | 0.1753 | 0.0267 | 0.0005 | | | | |
| | 10 | | 1.0000 | 0.9997 | 0.9984 | 0.9809 | 0.8949 | 0.6712 | 0.3402 | 0.0817 | 0.0033 | | | | |
| | 11 | | | 1.0000 | 0.9997 | 0.9951 | 0.9616 | 0.8334 | 0.5501 | 0.2018 | 0.0170 | | | | |
| | 12 | | | | 1.0000 | 0.9991 | 0.9894 | 0.9349 | 0.7541 | 0.4019 | 0.0684 | | | | |
| | 13 | | | | | 0.9999 | 0.9979 | 0.9817 | 0.9006 | 0.6482 | 0.2108 | | | | |
| | 14 | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9739 | 0.8593 | 0.4853 | | | | |
| | 15 | | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9719 | 0.8147 | | | | |
| 16 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | |

Tabla A.1 (continuación) Sumas de probabilidad binomial $\sum_{x=0}^r b(x; n, p)$

| n | r | p | | | | | | | | | | | | | | | | | | |
|----|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--|--|--|--|--|--|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 | | | | | | | | | |
| 19 | 0 | 0.1251 | 0.0144 | 0.0042 | 0.0011 | 0.0001 | | | | | | | | | | | | | | |
| | 1 | 0.4283 | 0.0829 | 0.0310 | 0.0104 | 0.0008 | 0.0000 | | | | | | | | | | | | | |
| | 2 | 0.7054 | 0.2369 | 0.1113 | 0.0462 | 0.0055 | 0.0004 | 0.0000 | | | | | | | | | | | | |
| | 3 | 0.8830 | 0.4551 | 0.2631 | 0.1332 | 0.0230 | 0.0022 | 0.0001 | | | | | | | | | | | | |
| | 4 | 0.9646 | 0.6733 | 0.4654 | 0.2822 | 0.0696 | 0.0096 | 0.0006 | 0.0000 | | | | | | | | | | | |
| | 5 | 0.9914 | 0.8369 | 0.6678 | 0.4739 | 0.1629 | 0.0318 | 0.0031 | 0.0001 | | | | | | | | | | | |
| | 6 | 0.9983 | 0.9324 | 0.8251 | 0.6655 | 0.3081 | 0.0835 | 0.0116 | 0.0006 | | | | | | | | | | | |
| | 7 | 0.9997 | 0.9767 | 0.9225 | 0.8180 | 0.4878 | 0.1796 | 0.0352 | 0.0028 | 0.0000 | | | | | | | | | | |
| | 8 | 1.0000 | 0.9923 | 0.9713 | 0.9161 | 0.6675 | 0.3238 | 0.0885 | 0.0105 | 0.0003 | | | | | | | | | | |
| | 9 | | 0.9984 | 0.9911 | 0.9674 | 0.8139 | 0.5000 | 0.1861 | 0.0326 | 0.0016 | | | | | | | | | | |
| | 10 | | 0.9997 | 0.9977 | 0.9895 | 0.9115 | 0.6762 | 0.3325 | 0.0839 | 0.0067 | 0.0000 | | | | | | | | | |
| | 11 | | 0.9999 | 0.9995 | 0.9972 | 0.9648 | 0.8204 | 0.5122 | 0.1820 | 0.0233 | 0.0003 | | | | | | | | | |
| | 12 | | 1.0000 | 0.9999 | 0.9994 | 0.9884 | 0.9165 | 0.6919 | 0.3345 | 0.0676 | 0.0017 | | | | | | | | | |
| | 13 | | | 1.0000 | 0.9999 | 0.9969 | 0.9682 | 0.8371 | 0.5261 | 0.1631 | 0.0086 | | | | | | | | | |
| | 14 | | | | 1.0000 | 0.9994 | 0.9964 | 0.9364 | 0.7178 | 0.3267 | 0.0352 | | | | | | | | | |
| | 15 | | | | | 0.9999 | 0.9978 | 0.9770 | 0.8668 | 0.5449 | 0.1150 | | | | | | | | | |
| | 16 | | | | | | 1.0000 | 0.9996 | 0.9945 | 0.9338 | 0.7631 | 0.2946 | | | | | | | | |
| | 17 | | | | | | | 1.0000 | 0.9992 | 0.9886 | 0.9171 | 0.5797 | | | | | | | | |
| | 18 | | | | | | | | 0.9999 | 0.9989 | 0.9856 | 0.8649 | | | | | | | | |
| | 19 | | | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | | | |
| 20 | 0 | 0.1216 | 0.0115 | 0.0032 | 0.0008 | 0.0000 | | | | | | | | | | | | | | |
| | 1 | 0.3917 | 0.0682 | 0.0243 | 0.0076 | 0.0005 | 0.0000 | | | | | | | | | | | | | |
| | 2 | 0.6769 | 0.2061 | 0.0913 | 0.0355 | 0.0036 | 0.0002 | 0.0000 | | | | | | | | | | | | |
| | 3 | 0.8670 | 0.4114 | 0.2252 | 0.1071 | 0.0100 | 0.0013 | 0.0001 | | | | | | | | | | | | |
| | 4 | 0.9568 | 0.6296 | 0.4148 | 0.2375 | 0.0510 | 0.0059 | 0.0003 | | | | | | | | | | | | |
| | 5 | 0.9867 | 0.8042 | 0.6172 | 0.4164 | 0.1256 | 0.0207 | 0.0016 | 0.0000 | | | | | | | | | | | |
| | 6 | 0.9976 | 0.9133 | 0.7858 | 0.6080 | 0.2500 | 0.0577 | 0.0065 | 0.0003 | | | | | | | | | | | |
| | 7 | 0.9996 | 0.9679 | 0.8982 | 0.7723 | 0.4159 | 0.1326 | 0.0210 | 0.0013 | 0.0000 | | | | | | | | | | |
| | 8 | 0.9999 | 0.9900 | 0.9591 | 0.8667 | 0.5956 | 0.2517 | 0.0565 | 0.0051 | 0.0001 | | | | | | | | | | |
| | 9 | 1.0000 | 0.9974 | 0.9861 | 0.9520 | 0.7553 | 0.4119 | 0.1275 | 0.0171 | 0.0006 | | | | | | | | | | |
| | 10 | | 0.9994 | 0.9961 | 0.9629 | 0.8725 | 0.5881 | 0.2447 | 0.0480 | 0.0026 | 0.0000 | | | | | | | | | |
| | 11 | | 0.9999 | 0.9991 | 0.9949 | 0.9435 | 0.7485 | 0.4044 | 0.1133 | 0.0100 | 0.0001 | | | | | | | | | |
| | 12 | | 1.0000 | 0.9998 | 0.9987 | 0.9790 | 0.8684 | 0.5841 | 0.2277 | 0.0321 | 0.0004 | | | | | | | | | |
| | 13 | | | 1.0000 | 0.9997 | 0.9935 | 0.9423 | 0.7500 | 0.3920 | 0.0867 | 0.0024 | | | | | | | | | |
| | 14 | | | | 1.0000 | 0.9984 | 0.9791 | 0.8744 | 0.5836 | 0.1958 | 0.0113 | | | | | | | | | |
| | 15 | | | | | 0.9997 | 0.9941 | 0.9490 | 0.7625 | 0.3204 | 0.0432 | | | | | | | | | |
| | 16 | | | | | | 1.0000 | 0.9987 | 0.9840 | 0.8929 | 0.5886 | 0.1230 | | | | | | | | |
| | 17 | | | | | | | 0.9998 | 0.9964 | 0.9645 | 0.7959 | 0.3231 | | | | | | | | |
| | 18 | | | | | | | | 1.0000 | 0.9995 | 0.9924 | 0.9308 | 0.6083 | | | | | | | |
| | 19 | | | | | | | | | 1.0000 | 0.9992 | 0.9885 | 0.8784 | | | | | | | |
| | 20 | | | | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | | | | | |

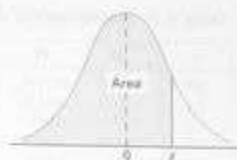
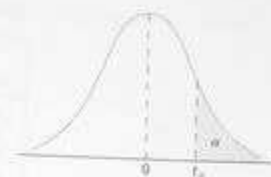


Tabla A.3 Áreas bajo la curva normal

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -3.4 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 |
| -3.3 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0003 |
| -3.2 | 0.0007 | 0.0007 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0005 | 0.0005 | 0.0005 |
| -3.1 | 0.0010 | 0.0009 | 0.0009 | 0.0009 | 0.0008 | 0.0008 | 0.0008 | 0.0008 | 0.0007 | 0.0007 |
| -3.0 | 0.0013 | 0.0012 | 0.0012 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |
| -2.9 | 0.0019 | 0.0018 | 0.0017 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| -2.8 | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| -2.7 | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| -2.6 | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0043 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| -2.5 | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| -2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| -2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| -2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| -2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| -2.0 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| -1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| -1.8 | 0.0359 | 0.0352 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| -1.7 | 0.0446 | 0.0438 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| -1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| -1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| -1.4 | 0.0808 | 0.0792 | 0.0776 | 0.0764 | 0.0749 | 0.0735 | 0.0722 | 0.0708 | 0.0694 | 0.0681 |
| -1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0839 | 0.0823 |
| -1.2 | 0.1151 | 0.1133 | 0.1112 | 0.1095 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| -1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| -1.0 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| -0.9 | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| -0.8 | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| -0.7 | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2296 | 0.2266 | 0.2236 | 0.2206 | 0.2177 | 0.2148 |
| -0.6 | 0.2743 | 0.2709 | 0.2676 | 0.2643 | 0.2611 | 0.2578 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| -0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| -0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| -0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| -0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3974 | 0.3936 | 0.3897 | 0.3859 |
| -0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4481 | 0.4441 | 0.4401 | 0.4361 | 0.4321 | 0.4280 | 0.4241 |
| -0.0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |


 Tabla A.4 Valores críticos de la distribución t

| ν | α | | | | | | |
|----------|----------|-------|-------|-------|-------|-------|--------|
| | 0.40 | 0.30 | 0.20 | 0.15 | 0.10 | 0.05 | 0.025 |
| 1 | 0.325 | 0.727 | 1.376 | 1.963 | 3.078 | 6.314 | 12.706 |
| 2 | 0.289 | 0.617 | 1.061 | 1.386 | 1.886 | 2.920 | 4.303 |
| 3 | 0.277 | 0.584 | 0.978 | 1.250 | 1.638 | 2.353 | 3.182 |
| 4 | 0.271 | 0.569 | 0.941 | 1.190 | 1.533 | 2.132 | 2.776 |
| 5 | 0.267 | 0.559 | 0.920 | 1.156 | 1.476 | 2.015 | 2.571 |
| 6 | 0.265 | 0.553 | 0.906 | 1.134 | 1.440 | 1.943 | 2.447 |
| 7 | 0.263 | 0.549 | 0.896 | 1.119 | 1.415 | 1.895 | 2.365 |
| 8 | 0.262 | 0.546 | 0.889 | 1.108 | 1.397 | 1.860 | 2.306 |
| 9 | 0.261 | 0.543 | 0.883 | 1.100 | 1.383 | 1.833 | 2.262 |
| 10 | 0.260 | 0.542 | 0.879 | 1.093 | 1.372 | 1.812 | 2.228 |
| 11 | 0.260 | 0.540 | 0.876 | 1.088 | 1.363 | 1.796 | 2.201 |
| 12 | 0.259 | 0.539 | 0.873 | 1.083 | 1.356 | 1.782 | 2.179 |
| 13 | 0.259 | 0.537 | 0.870 | 1.079 | 1.350 | 1.771 | 2.160 |
| 14 | 0.258 | 0.537 | 0.868 | 1.076 | 1.345 | 1.761 | 2.145 |
| 15 | 0.258 | 0.536 | 0.866 | 1.074 | 1.341 | 1.753 | 2.131 |
| 16 | 0.258 | 0.535 | 0.865 | 1.071 | 1.337 | 1.746 | 2.120 |
| 17 | 0.257 | 0.534 | 0.863 | 1.069 | 1.333 | 1.740 | 2.110 |
| 18 | 0.257 | 0.534 | 0.862 | 1.067 | 1.330 | 1.734 | 2.101 |
| 19 | 0.257 | 0.533 | 0.861 | 1.066 | 1.328 | 1.729 | 2.093 |
| 20 | 0.257 | 0.533 | 0.860 | 1.064 | 1.325 | 1.725 | 2.086 |
| 21 | 0.257 | 0.532 | 0.859 | 1.063 | 1.323 | 1.721 | 2.080 |
| 22 | 0.256 | 0.532 | 0.858 | 1.061 | 1.321 | 1.717 | 2.074 |
| 23 | 0.256 | 0.532 | 0.858 | 1.060 | 1.319 | 1.714 | 2.069 |
| 24 | 0.256 | 0.531 | 0.857 | 1.059 | 1.318 | 1.711 | 2.064 |
| 25 | 0.256 | 0.531 | 0.856 | 1.058 | 1.316 | 1.708 | 2.060 |
| 26 | 0.256 | 0.531 | 0.856 | 1.058 | 1.315 | 1.706 | 2.056 |
| 27 | 0.256 | 0.531 | 0.855 | 1.057 | 1.314 | 1.703 | 2.052 |
| 28 | 0.256 | 0.530 | 0.855 | 1.056 | 1.313 | 1.701 | 2.048 |
| 29 | 0.256 | 0.530 | 0.854 | 1.055 | 1.311 | 1.699 | 2.045 |
| 30 | 0.256 | 0.530 | 0.854 | 1.055 | 1.310 | 1.697 | 2.042 |
| 40 | 0.255 | 0.529 | 0.851 | 1.050 | 1.303 | 1.684 | 2.021 |
| 60 | 0.254 | 0.527 | 0.848 | 1.045 | 1.296 | 1.671 | 2.000 |
| 120 | 0.254 | 0.526 | 0.845 | 1.041 | 1.289 | 1.658 | 1.980 |
| ∞ | 0.253 | 0.524 | 0.842 | 1.036 | 1.282 | 1.645 | 1.960 |

Tabla A.4 (continuación) Valores críticos de la distribución t

| v | α | | | | | | |
|----------|----------|--------|--------|--------|--------|---------|---------|
| | 0.02 | 0.015 | 0.01 | 0.0075 | 0.005 | 0.0025 | 0.0005 |
| 1 | 15.895 | 21.205 | 31.821 | 42.434 | 60.657 | 127.322 | 636.500 |
| 2 | 4.849 | 5.643 | 6.965 | 8.073 | 9.925 | 14.089 | 31.598 |
| 3 | 3.482 | 3.896 | 4.541 | 5.047 | 5.841 | 7.453 | 12.924 |
| 4 | 2.999 | 3.298 | 3.747 | 4.088 | 4.604 | 5.598 | 8.610 |
| 5 | 2.757 | 3.003 | 3.365 | 3.634 | 4.032 | 4.773 | 6.859 |
| 6 | 2.612 | 2.829 | 3.143 | 3.372 | 3.707 | 4.317 | 5.959 |
| 7 | 2.517 | 2.715 | 2.998 | 3.203 | 3.499 | 4.029 | 5.408 |
| 8 | 2.449 | 2.634 | 2.896 | 3.085 | 3.355 | 3.833 | 5.041 |
| 9 | 2.398 | 2.574 | 2.821 | 2.998 | 3.250 | 3.690 | 4.781 |
| 10 | 2.359 | 2.527 | 2.764 | 2.932 | 3.169 | 3.581 | 4.587 |
| 11 | 2.328 | 2.491 | 2.738 | 2.879 | 3.106 | 3.497 | 4.437 |
| 12 | 2.303 | 2.461 | 2.681 | 2.836 | 3.055 | 3.428 | 4.318 |
| 13 | 2.282 | 2.436 | 2.650 | 2.801 | 3.012 | 3.372 | 4.221 |
| 14 | 2.264 | 2.415 | 2.624 | 2.771 | 2.977 | 3.326 | 4.140 |
| 15 | 2.249 | 2.397 | 2.602 | 2.746 | 2.947 | 3.286 | 4.073 |
| 16 | 2.235 | 2.382 | 2.583 | 2.724 | 2.921 | 3.252 | 4.015 |
| 17 | 2.224 | 2.368 | 2.567 | 2.706 | 2.898 | 3.222 | 3.965 |
| 18 | 2.214 | 2.356 | 2.552 | 2.689 | 2.878 | 3.197 | 3.922 |
| 19 | 2.205 | 2.346 | 2.539 | 2.674 | 2.861 | 3.174 | 3.881 |
| 20 | 2.197 | 2.336 | 2.528 | 2.661 | 2.845 | 3.153 | 3.844 |
| 21 | 2.189 | 2.328 | 2.518 | 2.649 | 2.831 | 3.135 | 3.810 |
| 22 | 2.183 | 2.320 | 2.508 | 2.639 | 2.819 | 3.119 | 3.782 |
| 23 | 2.177 | 2.313 | 2.500 | 2.629 | 2.807 | 3.104 | 3.758 |
| 24 | 2.172 | 2.307 | 2.492 | 2.620 | 2.797 | 3.091 | 3.745 |
| 25 | 2.167 | 2.301 | 2.485 | 2.612 | 2.787 | 3.078 | 3.725 |
| 26 | 2.162 | 2.296 | 2.479 | 2.605 | 2.779 | 3.067 | 3.707 |
| 27 | 2.158 | 2.291 | 2.473 | 2.598 | 2.771 | 3.057 | 3.690 |
| 28 | 2.154 | 2.286 | 2.467 | 2.592 | 2.763 | 3.047 | 3.674 |
| 29 | 2.150 | 2.282 | 2.462 | 2.586 | 2.756 | 3.038 | 3.659 |
| 30 | 2.147 | 2.278 | 2.457 | 2.581 | 2.750 | 3.030 | 3.646 |
| 40 | 2.125 | 2.250 | 2.423 | 2.542 | 2.704 | 2.971 | 3.551 |
| 60 | 2.099 | 2.223 | 2.390 | 2.504 | 2.660 | 2.915 | 3.480 |
| 120 | 2.076 | 2.196 | 2.358 | 2.468 | 2.617 | 2.860 | 3.373 |
| ∞ | 2.054 | 2.170 | 2.326 | 2.432 | 2.576 | 2.807 | 3.291 |

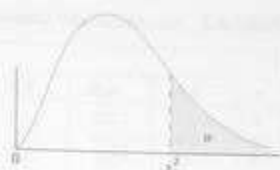


Tabla A.5 Valores críticos de la distribución ji cuadrada

| v | α | | | | | | | | | |
|----|---------|--------|---------|---------|---------|--------|--------|--------|--------|--------|
| | 0.995 | 0.99 | 0.98 | 0.975 | 0.95 | 0.90 | 0.80 | 0.75 | 0.70 | 0.50 |
| 1 | 0.00393 | 0.0157 | 0.01628 | 0.01682 | 0.00393 | 0.0158 | 0.0642 | 0.102 | 0.148 | 0.435 |
| 2 | 0.0100 | 0.0201 | 0.0404 | 0.0506 | 0.103 | 0.211 | 0.446 | 0.575 | 0.713 | 1.386 |
| 3 | 0.0717 | 0.115 | 0.185 | 0.216 | 0.352 | 0.584 | 1.005 | 1.213 | 1.424 | 2.366 |
| 4 | 0.207 | 0.297 | 0.429 | 0.484 | 0.711 | 1.064 | 1.649 | 1.923 | 2.195 | 3.357 |
| 5 | 0.412 | 0.554 | 0.752 | 0.831 | 1.145 | 1.610 | 2.343 | 2.675 | 3.000 | 4.351 |
| 6 | 0.676 | 0.872 | 1.134 | 1.237 | 1.635 | 2.204 | 3.070 | 3.455 | 3.828 | 5.348 |
| 7 | 0.989 | 1.239 | 1.564 | 1.690 | 2.167 | 2.833 | 3.822 | 4.255 | 4.671 | 6.346 |
| 8 | 1.344 | 1.646 | 2.032 | 2.180 | 2.733 | 3.490 | 4.594 | 5.071 | 5.527 | 7.344 |
| 9 | 1.735 | 2.088 | 2.532 | 2.700 | 3.325 | 4.168 | 5.380 | 5.899 | 6.393 | 8.343 |
| 10 | 2.156 | 2.558 | 3.059 | 3.247 | 3.940 | 4.865 | 6.179 | 6.737 | 7.267 | 9.342 |
| 11 | 2.603 | 3.053 | 3.609 | 3.816 | 4.575 | 5.578 | 6.980 | 7.584 | 8.148 | 10.341 |
| 12 | 3.074 | 3.571 | 4.178 | 4.404 | 5.226 | 6.304 | 7.807 | 8.438 | 9.034 | 11.340 |
| 13 | 3.565 | 4.107 | 4.765 | 5.009 | 5.892 | 7.042 | 8.634 | 9.299 | 9.926 | 12.340 |
| 14 | 4.075 | 4.660 | 5.368 | 5.629 | 6.571 | 7.790 | 9.467 | 10.165 | 10.821 | 13.339 |
| 15 | 4.601 | 5.229 | 5.985 | 6.262 | 7.261 | 8.547 | 10.307 | 11.036 | 11.721 | 14.339 |
| 16 | 5.142 | 5.812 | 6.614 | 6.908 | 7.962 | 9.312 | 11.152 | 11.912 | 12.624 | 15.338 |
| 17 | 5.697 | 6.408 | 7.255 | 7.564 | 8.672 | 10.085 | 12.002 | 12.792 | 13.531 | 16.338 |
| 18 | 6.265 | 7.015 | 7.906 | 8.231 | 9.390 | 10.865 | 12.857 | 13.675 | 14.440 | 17.338 |
| 19 | 6.844 | 7.633 | 8.567 | 8.907 | 10.117 | 11.651 | 13.710 | 14.562 | 15.352 | 18.338 |
| 20 | 7.434 | 8.260 | 9.237 | 9.591 | 10.851 | 12.443 | 14.578 | 15.452 | 16.260 | 19.337 |
| 21 | 8.034 | 8.897 | 9.915 | 10.283 | 11.591 | 13.240 | 15.445 | 16.344 | 17.182 | 20.337 |
| 22 | 8.643 | 9.542 | 10.600 | 10.982 | 12.338 | 14.041 | 16.314 | 17.240 | 18.101 | 21.337 |
| 23 | 9.260 | 10.196 | 11.293 | 11.688 | 13.091 | 14.848 | 17.187 | 18.137 | 19.021 | 22.337 |
| 24 | 9.886 | 10.856 | 11.992 | 12.401 | 13.848 | 15.659 | 18.062 | 19.037 | 19.943 | 23.337 |
| 25 | 10.520 | 11.524 | 12.697 | 13.120 | 14.611 | 16.473 | 18.940 | 19.939 | 20.867 | 24.337 |
| 26 | 11.160 | 12.198 | 13.409 | 13.844 | 15.379 | 17.292 | 19.820 | 20.843 | 21.792 | 25.336 |
| 27 | 11.808 | 12.879 | 14.125 | 14.573 | 16.151 | 18.114 | 20.703 | 21.749 | 22.719 | 26.336 |
| 28 | 12.461 | 13.565 | 14.847 | 15.308 | 16.928 | 18.939 | 21.588 | 22.657 | 23.647 | 27.336 |
| 29 | 13.121 | 14.256 | 15.574 | 16.047 | 17.708 | 19.768 | 22.475 | 23.567 | 24.577 | 28.336 |
| 30 | 13.787 | 14.953 | 16.306 | 16.791 | 18.493 | 20.599 | 23.364 | 24.478 | 25.508 | 29.336 |

Tabla A.5 (continuación) *Valores críticos de la distribución ji cuadrada*

| v | α | | | | | | | | | |
|----|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 0.30 | 0.25 | 0.20 | 0.10 | 0.05 | 0.02 | 0.01 | 0.005 | 0.001 | |
| 1 | 1.074 | 1.323 | 1.642 | 2.706 | 3.841 | 5.024 | 5.412 | 6.635 | 7.879 | 10.827 |
| 2 | 2.408 | 2.773 | 3.219 | 4.605 | 5.991 | 7.378 | 7.824 | 9.210 | 10.597 | 13.815 |
| 3 | 3.665 | 4.108 | 4.642 | 6.251 | 7.815 | 9.348 | 9.837 | 11.345 | 12.838 | 16.268 |
| 4 | 4.878 | 5.385 | 5.989 | 7.779 | 9.488 | 11.143 | 11.668 | 13.277 | 14.860 | 18.465 |
| 5 | 6.064 | 6.626 | 7.289 | 9.236 | 11.070 | 12.832 | 13.388 | 15.089 | 16.750 | 20.517 |
| 6 | 7.231 | 7.841 | 8.558 | 10.645 | 12.592 | 14.449 | 15.033 | 16.812 | 18.548 | 22.457 |
| 7 | 8.383 | 9.037 | 9.803 | 12.017 | 14.067 | 16.013 | 16.622 | 18.475 | 20.278 | 24.332 |
| 8 | 9.524 | 10.219 | 11.030 | 13.362 | 15.507 | 17.535 | 18.168 | 20.090 | 21.955 | 26.125 |
| 9 | 10.656 | 11.389 | 12.242 | 14.684 | 16.919 | 19.023 | 19.679 | 21.666 | 23.589 | 27.877 |
| 10 | 11.781 | 12.549 | 13.442 | 15.987 | 18.307 | 20.483 | 21.161 | 23.209 | 25.188 | 29.588 |
| 11 | 12.899 | 13.701 | 14.631 | 17.275 | 19.675 | 21.920 | 22.618 | 24.725 | 26.757 | 31.264 |
| 12 | 14.011 | 14.845 | 15.812 | 18.549 | 21.026 | 23.337 | 24.054 | 26.217 | 28.300 | 32.909 |
| 13 | 15.119 | 15.984 | 16.985 | 19.812 | 22.362 | 24.736 | 25.472 | 27.688 | 29.819 | 34.538 |
| 14 | 16.222 | 17.117 | 18.151 | 21.064 | 23.685 | 26.119 | 26.873 | 29.141 | 31.319 | 36.123 |
| 15 | 17.322 | 18.245 | 19.311 | 22.307 | 24.996 | 27.488 | 28.259 | 30.578 | 32.801 | 37.697 |
| 16 | 18.418 | 19.369 | 20.465 | 23.542 | 26.296 | 28.845 | 29.633 | 32.000 | 34.267 | 39.252 |
| 17 | 19.513 | 20.489 | 21.615 | 24.769 | 27.587 | 30.191 | 30.995 | 33.409 | 35.718 | 40.790 |
| 18 | 20.601 | 21.605 | 22.760 | 25.989 | 28.869 | 31.526 | 32.346 | 34.805 | 37.156 | 42.332 |
| 19 | 21.689 | 22.718 | 23.900 | 27.204 | 30.144 | 32.852 | 33.687 | 36.191 | 38.582 | 43.820 |
| 20 | 22.775 | 23.828 | 25.038 | 28.412 | 31.410 | 34.170 | 35.020 | 37.566 | 39.997 | 45.315 |
| 21 | 23.858 | 24.935 | 26.171 | 29.615 | 32.671 | 35.479 | 36.343 | 38.932 | 41.401 | 46.797 |
| 22 | 24.939 | 26.039 | 27.301 | 30.813 | 33.924 | 36.781 | 37.639 | 40.289 | 42.796 | 48.266 |
| 23 | 26.018 | 27.141 | 28.429 | 32.007 | 35.172 | 38.076 | 38.968 | 41.638 | 44.181 | 49.728 |
| 24 | 27.096 | 28.241 | 29.553 | 33.196 | 36.415 | 39.364 | 40.270 | 42.980 | 45.558 | 51.179 |
| 25 | 28.172 | 29.339 | 30.675 | 34.382 | 37.652 | 40.646 | 41.566 | 44.314 | 46.938 | 52.620 |
| 26 | 29.246 | 30.434 | 31.795 | 35.563 | 38.885 | 41.923 | 42.856 | 45.642 | 48.290 | 54.052 |
| 27 | 30.319 | 31.528 | 32.912 | 36.741 | 40.113 | 43.194 | 44.140 | 46.963 | 49.645 | 55.476 |
| 28 | 31.391 | 32.620 | 34.027 | 37.916 | 41.337 | 44.461 | 45.419 | 48.278 | 50.993 | 56.893 |
| 29 | 32.461 | 33.711 | 35.139 | 39.087 | 42.557 | 45.722 | 46.693 | 49.588 | 52.336 | 58.302 |
| 30 | 33.530 | 34.800 | 36.250 | 40.256 | 43.773 | 46.979 | 47.962 | 50.892 | 53.672 | 59.703 |



Tabla A.6*: Valores críticos de la distribución F

$$f_{\alpha}(v_1, v_2)$$

| v_2 | v_1 | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 236.8 | 238.9 | 240.5 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 |
| 25 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 |
| 26 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.24 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.22 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.25 | 2.18 | 2.12 |
| 60 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.17 | 2.09 | 2.02 | 1.96 |
| ∞ | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 |

*Reproducida de la tabla 18 de *Biometrika Tables for Statisticians*, vol. 1, con permiso de E. S. Pearson y de los administradores de Biometrika.

Tabla A.6 (continuación) Valores críticos de la distribución F
 $f_{0.05}(v_1, v_2)$

| v_2 | v_1 | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 1 | 241.9 | 243.9 | 245.9 | 248.0 | 249.1 | 250.1 | 251.1 | 252.2 | 253.3 | 254.3 |
| 2 | 19.40 | 19.41 | 19.43 | 19.45 | 19.45 | 19.46 | 19.47 | 19.48 | 19.49 | 19.50 |
| 3 | 8.79 | 8.74 | 8.70 | 8.66 | 8.64 | 8.62 | 8.59 | 8.57 | 8.55 | 8.53 |
| 4 | 5.06 | 5.01 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 |
| 5 | 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.46 | 4.43 | 4.40 | 4.36 |
| 6 | 4.06 | 4.00 | 3.94 | 3.87 | 3.84 | 3.81 | 3.77 | 3.74 | 3.70 | 3.67 |
| 7 | 3.64 | 3.57 | 3.51 | 3.44 | 3.41 | 3.38 | 3.34 | 3.30 | 3.27 | 3.23 |
| 8 | 3.35 | 3.28 | 3.22 | 3.15 | 3.12 | 3.08 | 3.04 | 3.01 | 2.97 | 2.93 |
| 9 | 3.14 | 3.07 | 3.01 | 2.94 | 2.90 | 2.86 | 2.83 | 2.79 | 2.75 | 2.71 |
| 10 | 2.98 | 2.91 | 2.85 | 2.77 | 2.74 | 2.70 | 2.66 | 2.62 | 2.58 | 2.54 |
| 11 | 2.85 | 2.79 | 2.72 | 2.65 | 2.61 | 2.57 | 2.53 | 2.49 | 2.45 | 2.40 |
| 12 | 2.75 | 2.69 | 2.62 | 2.54 | 2.51 | 2.47 | 2.43 | 2.38 | 2.34 | 2.30 |
| 13 | 2.67 | 2.60 | 2.53 | 2.46 | 2.42 | 2.38 | 2.34 | 2.30 | 2.25 | 2.21 |
| 14 | 2.60 | 2.53 | 2.46 | 2.39 | 2.35 | 2.31 | 2.27 | 2.22 | 2.18 | 2.13 |
| 15 | 2.54 | 2.48 | 2.40 | 2.33 | 2.29 | 2.25 | 2.20 | 2.16 | 2.11 | 2.07 |
| 16 | 2.49 | 2.42 | 2.35 | 2.28 | 2.24 | 2.19 | 2.15 | 2.11 | 2.06 | 2.01 |
| 17 | 2.45 | 2.38 | 2.31 | 2.23 | 2.19 | 2.15 | 2.10 | 2.06 | 2.01 | 1.96 |
| 18 | 2.41 | 2.34 | 2.27 | 2.19 | 2.15 | 2.11 | 2.06 | 2.02 | 1.97 | 1.92 |
| 19 | 2.38 | 2.31 | 2.23 | 2.16 | 2.11 | 2.07 | 2.03 | 1.98 | 1.93 | 1.88 |
| 20 | 2.35 | 2.28 | 2.20 | 2.12 | 2.08 | 2.04 | 1.99 | 1.95 | 1.90 | 1.84 |
| 21 | 2.32 | 2.25 | 2.18 | 2.10 | 2.05 | 2.01 | 1.96 | 1.92 | 1.87 | 1.81 |
| 22 | 2.30 | 2.23 | 2.15 | 2.07 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.78 |
| 23 | 2.27 | 2.20 | 2.13 | 2.05 | 2.01 | 1.96 | 1.91 | 1.86 | 1.81 | 1.76 |
| 24 | 2.25 | 2.18 | 2.11 | 2.03 | 1.98 | 1.94 | 1.89 | 1.84 | 1.79 | 1.73 |
| 25 | 2.24 | 2.16 | 2.09 | 2.01 | 1.96 | 1.92 | 1.87 | 1.82 | 1.77 | 1.71 |
| 26 | 2.22 | 2.15 | 2.07 | 1.99 | 1.95 | 1.90 | 1.85 | 1.80 | 1.75 | 1.69 |
| 27 | 2.20 | 2.13 | 2.06 | 1.97 | 1.93 | 1.88 | 1.84 | 1.79 | 1.73 | 1.67 |
| 28 | 2.19 | 2.12 | 2.04 | 1.96 | 1.91 | 1.87 | 1.82 | 1.77 | 1.71 | 1.65 |
| 29 | 2.18 | 2.10 | 2.03 | 1.94 | 1.90 | 1.85 | 1.81 | 1.75 | 1.70 | 1.64 |
| 30 | 2.16 | 2.09 | 2.01 | 1.93 | 1.89 | 1.84 | 1.79 | 1.74 | 1.68 | 1.62 |
| 40 | 2.08 | 2.00 | 1.92 | 1.84 | 1.79 | 1.74 | 1.69 | 1.64 | 1.58 | 1.51 |
| 60 | 1.99 | 1.92 | 1.84 | 1.75 | 1.70 | 1.65 | 1.59 | 1.53 | 1.47 | 1.39 |
| 120 | 1.91 | 1.83 | 1.75 | 1.66 | 1.61 | 1.55 | 1.50 | 1.43 | 1.35 | 1.25 |
| ∞ | 1.83 | 1.75 | 1.67 | 1.57 | 1.52 | 1.46 | 1.39 | 1.32 | 1.22 | 1.10 |

Tabla A.6: (continuación) Valores críticos de la distribución F

$$f_{\alpha, \nu_1, \nu_2}$$

| ν_1 | ν_2 | | | | | | | | |
|----------|---------|--------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 4052 | 4999.5 | 5403 | 5625 | 5764 | 5859 | 5928 | 5981 | 6022 |
| 2 | 98.50 | 99.00 | 99.17 | 99.25 | 99.30 | 99.33 | 99.36 | 99.37 | 99.39 |
| 3 | 34.12 | 30.82 | 29.46 | 28.71 | 28.34 | 27.91 | 27.67 | 27.49 | 27.35 |
| 4 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | 15.21 | 14.98 | 14.80 | 14.66 |
| 5 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 |
| 6 | 13.75 | 10.92 | 9.78 | 9.15 | 8.75 | 8.47 | 8.26 | 8.10 | 7.98 |
| 7 | 12.25 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.99 | 6.84 | 6.72 |
| 8 | 11.26 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.91 |
| 9 | 10.56 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.61 | 5.47 | 5.35 |
| 10 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 |
| 11 | 9.65 | 7.21 | 6.22 | 5.67 | 5.32 | 5.07 | 4.89 | 4.74 | 4.63 |
| 12 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 |
| 13 | 9.07 | 6.70 | 5.74 | 5.21 | 4.86 | 4.62 | 4.44 | 4.30 | 4.19 |
| 14 | 8.86 | 6.51 | 5.56 | 5.04 | 4.69 | 4.46 | 4.28 | 4.14 | 4.03 |
| 15 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 |
| 16 | 8.53 | 6.23 | 5.29 | 4.77 | 4.44 | 4.20 | 4.03 | 3.89 | 3.78 |
| 17 | 8.40 | 6.11 | 5.18 | 4.67 | 4.34 | 4.10 | 3.93 | 3.79 | 3.68 |
| 18 | 8.29 | 6.01 | 5.09 | 4.58 | 4.25 | 4.01 | 3.84 | 3.71 | 3.60 |
| 19 | 8.18 | 5.93 | 5.01 | 4.50 | 4.17 | 3.94 | 3.77 | 3.63 | 3.52 |
| 20 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 |
| 21 | 8.02 | 5.78 | 4.87 | 4.37 | 4.04 | 3.81 | 3.64 | 3.51 | 3.40 |
| 22 | 7.95 | 5.72 | 4.82 | 4.31 | 3.99 | 3.76 | 3.59 | 3.45 | 3.35 |
| 23 | 7.88 | 5.66 | 4.76 | 4.26 | 3.94 | 3.71 | 3.54 | 3.41 | 3.30 |
| 24 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.26 |
| 25 | 7.77 | 5.57 | 4.68 | 4.18 | 3.85 | 3.63 | 3.46 | 3.32 | 3.22 |
| 26 | 7.72 | 5.53 | 4.64 | 4.14 | 3.82 | 3.59 | 3.42 | 3.29 | 3.18 |
| 27 | 7.68 | 5.49 | 4.60 | 4.11 | 3.78 | 3.56 | 3.39 | 3.26 | 3.15 |
| 28 | 7.64 | 5.45 | 4.57 | 4.07 | 3.75 | 3.53 | 3.36 | 3.23 | 3.12 |
| 29 | 7.60 | 5.42 | 4.54 | 4.04 | 3.73 | 3.50 | 3.33 | 3.20 | 3.09 |
| 30 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 3.07 |
| 40 | 7.31 | 5.18 | 4.31 | 3.83 | 3.51 | 3.29 | 3.12 | 2.99 | 2.89 |
| 60 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.72 |
| 120 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.56 |
| ∞ | 6.63 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.41 |

Tabla A.6 (continuación) Valores críticos de la distribución F

| v_2 | v_1 | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 1 | 6056 | 6106 | 6157 | 6209 | 6235 | 6261 | 6287 | 6313 | 6339 | 6366 |
| 2 | 99.40 | 99.42 | 99.43 | 99.45 | 99.46 | 99.47 | 99.47 | 99.48 | 99.49 | 99.50 |
| 3 | 27.23 | 27.05 | 26.87 | 26.69 | 26.60 | 26.50 | 26.41 | 26.32 | 26.22 | 26.13 |
| 4 | 14.35 | 14.37 | 14.20 | 14.02 | 13.95 | 13.84 | 13.75 | 13.65 | 13.56 | 13.46 |
| 5 | 10.05 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.29 | 9.20 | 9.11 | 9.02 |
| 6 | 7.87 | 7.72 | 7.56 | 7.40 | 7.31 | 7.23 | 7.14 | 7.06 | 6.97 | 6.88 |
| 7 | 6.62 | 6.47 | 6.31 | 6.16 | 6.07 | 5.99 | 5.91 | 5.82 | 5.74 | 5.65 |
| 8 | 5.81 | 5.67 | 5.52 | 5.36 | 5.28 | 5.20 | 5.12 | 5.03 | 4.95 | 4.86 |
| 9 | 5.26 | 5.11 | 4.96 | 4.81 | 4.73 | 4.65 | 4.57 | 4.48 | 4.40 | 4.31 |
| 10 | 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.17 | 4.08 | 4.00 | 3.91 |
| 11 | 4.54 | 4.40 | 4.25 | 4.10 | 4.02 | 3.94 | 3.86 | 3.78 | 3.69 | 3.60 |
| 12 | 4.30 | 4.16 | 4.01 | 3.86 | 3.78 | 3.70 | 3.62 | 3.54 | 3.45 | 3.36 |
| 13 | 4.10 | 3.96 | 3.82 | 3.66 | 3.59 | 3.51 | 3.43 | 3.34 | 3.25 | 3.17 |
| 14 | 3.94 | 3.80 | 3.66 | 3.51 | 3.43 | 3.35 | 3.27 | 3.18 | 3.09 | 3.00 |
| 15 | 3.80 | 3.67 | 3.52 | 3.37 | 3.29 | 3.21 | 3.13 | 3.05 | 2.96 | 2.87 |
| 16 | 3.69 | 3.55 | 3.41 | 3.26 | 3.18 | 3.10 | 3.02 | 2.93 | 2.84 | 2.75 |
| 17 | 3.59 | 3.46 | 3.31 | 3.16 | 3.08 | 3.00 | 2.92 | 2.83 | 2.75 | 2.65 |
| 18 | 3.51 | 3.37 | 3.23 | 3.08 | 3.00 | 2.92 | 2.84 | 2.75 | 2.66 | 2.57 |
| 19 | 3.43 | 3.30 | 3.15 | 3.00 | 2.92 | 2.84 | 2.76 | 2.67 | 2.58 | 2.49 |
| 20 | 3.37 | 3.23 | 3.09 | 2.94 | 2.86 | 2.78 | 2.69 | 2.61 | 2.52 | 2.42 |
| 21 | 3.31 | 3.17 | 3.03 | 2.88 | 2.80 | 2.72 | 2.64 | 2.55 | 2.46 | 2.36 |
| 22 | 3.26 | 3.12 | 2.98 | 2.83 | 2.75 | 2.67 | 2.58 | 2.50 | 2.40 | 2.31 |
| 23 | 3.21 | 3.07 | 2.93 | 2.78 | 2.70 | 2.62 | 2.54 | 2.45 | 2.35 | 2.26 |
| 24 | 3.17 | 3.03 | 2.89 | 2.74 | 2.66 | 2.58 | 2.49 | 2.40 | 2.31 | 2.21 |
| 25 | 3.13 | 2.99 | 2.85 | 2.70 | 2.62 | 2.54 | 2.45 | 2.36 | 2.27 | 2.17 |
| 26 | 3.09 | 2.96 | 2.81 | 2.66 | 2.58 | 2.50 | 2.42 | 2.33 | 2.23 | 2.13 |
| 27 | 3.06 | 2.93 | 2.78 | 2.63 | 2.55 | 2.47 | 2.38 | 2.29 | 2.20 | 2.10 |
| 28 | 3.03 | 2.90 | 2.75 | 2.60 | 2.52 | 2.44 | 2.35 | 2.26 | 2.17 | 2.06 |
| 29 | 3.00 | 2.87 | 2.73 | 2.57 | 2.49 | 2.41 | 2.33 | 2.23 | 2.14 | 2.03 |
| 30 | 2.98 | 2.84 | 2.70 | 2.55 | 2.47 | 2.39 | 2.30 | 2.21 | 2.11 | 2.01 |
| 40 | 2.80 | 2.66 | 2.52 | 2.37 | 2.29 | 2.20 | 2.11 | 2.02 | 1.92 | 1.80 |
| 60 | 2.65 | 2.50 | 2.35 | 2.20 | 2.12 | 2.03 | 1.94 | 1.84 | 1.73 | 1.60 |
| 120 | 2.47 | 2.34 | 2.19 | 2.03 | 1.95 | 1.86 | 1.76 | 1.66 | 1.53 | 1.38 |
| ∞ | 2.32 | 2.18 | 2.04 | 1.88 | 1.79 | 1.70 | 1.59 | 1.47 | 1.32 | 1.00 |

Tabla A.7* Factores de tolerancia para distribuciones normales
 $\gamma = 0.05$

| n | 1 - α | | |
|----|--------------|--------|--------|
| | 0.90 | 0.95 | 0.99 |
| 2 | 32.019 | 37.674 | 48.430 |
| 3 | 8.380 | 9.916 | 12.861 |
| 4 | 5.369 | 6.370 | 8.299 |
| 5 | 4.275 | 5.079 | 6.634 |
| 6 | 3.712 | 4.414 | 5.775 |
| 7 | 3.369 | 4.007 | 5.248 |
| 8 | 3.136 | 3.732 | 4.891 |
| 9 | 2.967 | 3.532 | 4.631 |
| 10 | 2.839 | 3.379 | 4.433 |
| 11 | 2.737 | 3.259 | 4.277 |
| 12 | 2.655 | 3.162 | 4.150 |
| 13 | 2.587 | 3.081 | 4.044 |
| 14 | 2.529 | 3.012 | 3.955 |
| 15 | 2.480 | 2.954 | 3.878 |
| 16 | 2.437 | 2.903 | 3.812 |
| 17 | 2.400 | 2.858 | 3.754 |
| 18 | 2.366 | 2.819 | 3.702 |
| 19 | 2.337 | 2.784 | 3.656 |
| 20 | 2.310 | 2.752 | 3.615 |
| 25 | 2.208 | 2.631 | 3.457 |
| 30 | 2.140 | 2.549 | 3.350 |
| 35 | 2.090 | 2.490 | 3.272 |
| 40 | 2.052 | 2.445 | 3.213 |

$\gamma = 0.01$

| n | 1 - α | | |
|----|--------------|---------|---------|
| | 0.90 | 0.95 | 0.99 |
| 2 | 160.193 | 188.491 | 242.300 |
| 3 | 18.930 | 22.401 | 29.055 |
| 4 | 9.398 | 11.150 | 14.527 |
| 5 | 6.612 | 7.855 | 10.260 |
| 6 | 5.337 | 6.345 | 8.301 |
| 7 | 4.613 | 5.488 | 7.187 |
| 8 | 4.147 | 4.936 | 6.468 |
| 9 | 3.822 | 4.550 | 5.966 |
| 10 | 3.582 | 4.265 | 5.594 |
| 11 | 3.397 | 4.045 | 5.308 |
| 12 | 3.250 | 3.870 | 5.079 |
| 13 | 3.130 | 3.727 | 4.893 |
| 14 | 3.029 | 3.608 | 4.737 |
| 15 | 2.945 | 3.507 | 4.605 |
| 16 | 2.872 | 3.421 | 4.492 |
| 17 | 2.808 | 3.345 | 4.393 |
| 18 | 2.753 | 3.279 | 4.307 |
| 19 | 2.703 | 3.221 | 4.230 |
| 20 | 2.659 | 3.168 | 4.161 |
| 25 | 2.494 | 2.972 | 3.904 |
| 30 | 2.385 | 2.841 | 3.733 |
| 35 | 2.306 | 2.748 | 3.611 |
| 40 | 2.247 | 2.677 | 3.518 |

*Adaptada de C. Eisenhart, M. W. Hastay y W. A. Wallis, *Techniques of Statistical Analysis*, capítulo 2, McGraw-Hill Book Company, Nueva York, 1947. Utilizada con permiso de McGraw-Hill Book Company.

Tabla A.7 (continuación) Factores de tolerancia para distribuciones normales

| $\gamma = 0.05$ | | | | $\gamma = 0.01$ | | | |
|-----------------|--------------|-------|-------|-----------------|--------------|-------|-------|
| n | $1 - \alpha$ | | | n | $1 - \alpha$ | | |
| | 0.90 | 0.95 | 0.99 | | 0.90 | 0.95 | 0.99 |
| 45 | 2.021 | 2.408 | 3.165 | 45 | 2.200 | 2.621 | 3.444 |
| 50 | 1.996 | 2.379 | 3.126 | 50 | 2.162 | 2.576 | 3.385 |
| 55 | 1.976 | 2.354 | 3.094 | 55 | 2.130 | 2.538 | 3.335 |
| 60 | 1.958 | 2.333 | 3.066 | 60 | 2.103 | 2.506 | 3.293 |
| 65 | 1.943 | 2.315 | 3.042 | 65 | 2.080 | 2.478 | 3.257 |
| 70 | 1.929 | 2.299 | 3.021 | 70 | 2.060 | 2.454 | 3.225 |
| 75 | 1.917 | 2.285 | 3.002 | 75 | 2.042 | 2.433 | 3.197 |
| 80 | 1.907 | 2.272 | 2.986 | 80 | 2.026 | 2.414 | 3.173 |
| 85 | 1.897 | 2.261 | 2.971 | 85 | 2.012 | 2.397 | 3.150 |
| 90 | 1.889 | 2.251 | 2.958 | 90 | 1.999 | 2.382 | 3.130 |
| 95 | 1.881 | 2.241 | 2.945 | 95 | 1.987 | 2.368 | 3.112 |
| 100 | 1.874 | 2.233 | 2.934 | 100 | 1.977 | 2.355 | 3.096 |
| 150 | 1.825 | 2.175 | 2.859 | 150 | 1.905 | 2.270 | 2.983 |
| 200 | 1.798 | 2.143 | 2.816 | 200 | 1.865 | 2.222 | 2.921 |
| 250 | 1.780 | 2.121 | 2.788 | 250 | 1.839 | 2.191 | 2.880 |
| 300 | 1.767 | 2.106 | 2.767 | 300 | 1.820 | 2.169 | 2.850 |
| 400 | 1.749 | 2.084 | 2.739 | 400 | 1.794 | 2.138 | 2.809 |
| 500 | 1.737 | 2.070 | 2.721 | 500 | 1.777 | 2.117 | 2.783 |
| 600 | 1.729 | 2.060 | 2.707 | 600 | 1.764 | 2.102 | 2.763 |
| 700 | 1.722 | 2.052 | 2.697 | 700 | 1.755 | 2.091 | 2.748 |
| 800 | 1.717 | 2.046 | 2.688 | 800 | 1.747 | 2.082 | 2.736 |
| 900 | 1.712 | 2.040 | 2.682 | 900 | 1.741 | 2.075 | 2.728 |
| 1000 | 1.709 | 2.036 | 2.676 | 1000 | 1.736 | 2.068 | 2.718 |
| ∞ | 1.645 | 1.960 | 2.576 | ∞ | 1.645 | 1.960 | 2.576 |

Tabla A.8* Tamaño muestral para la prueba t de la media

| Prueba unilateral
Prueba bilateral | Nivel de la prueba t | | | | | | | | | | | | | | | | | | | | |
|---------------------------------------|-------------------------------------|------|-----|-----|-----|------------------------------------|------|-----|-----|-----|-------------------------------------|------|-----|-----|-----|-----------------------------------|------|-----|-----|-----|---|
| | $\alpha = 0.005$
$\alpha = 0.01$ | | | | | $\alpha = 0.01$
$\alpha = 0.02$ | | | | | $\alpha = 0.025$
$\alpha = 0.05$ | | | | | $\alpha = 0.05$
$\alpha = 0.1$ | | | | | |
| | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | |
| 0.05 | | | | | | | | | | | | | | | | | | | | | |
| 0.10 | | | | | | | | | | | | | | | | | | | | 122 | |
| 0.15 | | | | | | | | | | | | | | | | | | | 70 | | |
| 0.20 | | | | | | | | | | 139 | | | | | | | | | | | |
| 0.25 | | | | | 110 | | | | | 90 | | | | | | | | | | | |
| 0.30 | | | | 134 | 78 | | | | | | | | | | 128 | 64 | | | | | |
| 0.35 | | | 125 | 99 | 58 | | | | | 115 | 65 | | | | 119 | 90 | 45 | | | 122 | |
| 0.40 | | 115 | 97 | 77 | 45 | | | | | 109 | 85 | 47 | | | 109 | 88 | 67 | 34 | | 90 | |
| 0.45 | | 92 | 77 | 62 | 37 | | | | | 101 | 85 | 66 | 37 | | 117 | 84 | 68 | 51 | 20 | 101 | |
| 0.50 | 100 | 75 | 63 | 51 | 30 | | | | | 90 | 86 | 55 | 43 | 25 | | | | | | 80 | |
| 0.55 | 83 | 63 | 53 | 42 | 26 | | | | | 75 | 55 | 46 | 36 | 21 | | | | | | 65 | |
| 0.60 | 71 | 53 | 45 | 36 | 22 | | | | | 63 | 47 | 39 | 31 | 18 | | | | | | 54 | |
| 0.65 | 61 | 46 | 39 | 31 | 20 | | | | | 53 | 38 | 32 | 24 | 13 | | | | | | 46 | |
| 0.70 | 53 | 40 | 34 | 28 | 17 | | | | | 46 | 33 | 27 | 21 | 13 | | | | | | 39 | |
| 0.75 | 47 | 36 | 30 | 25 | 16 | | | | | 40 | 29 | 24 | 19 | 10 | | | | | | 34 | |
| 0.80 | 41 | 32 | 27 | 22 | 14 | | | | | 35 | 26 | 21 | 16 | 9 | | | | | | 30 | |
| 0.85 | 37 | 29 | 24 | 20 | 13 | | | | | 31 | 22 | 19 | 15 | 9 | | | | | | 27 | |
| 0.90 | 34 | 26 | 22 | 18 | 12 | | | | | 28 | 21 | 17 | 13 | 8 | | | | | | 24 | |
| 0.95 | 31 | 24 | 20 | 17 | 11 | | | | | 25 | 19 | 16 | 12 | 7 | | | | | | 21 | |
| 1.00 | 28 | 22 | 19 | 16 | 10 | | | | | 23 | 17 | 14 | 11 | 7 | | | | | | 19 | |
| 1.1 | 24 | 19 | 16 | 14 | 9 | | | | | 21 | 16 | 13 | 10 | 6 | | | | | | 18 | |
| 1.2 | 21 | 16 | 14 | 12 | 8 | | | | | 18 | 13 | 11 | 9 | 6 | | | | | | 15 | |
| 1.3 | 18 | 15 | 13 | 11 | 8 | | | | | 15 | 12 | 10 | 8 | 5 | | | | | | 13 | |
| 1.4 | 16 | 13 | 12 | 10 | 7 | | | | | 14 | 10 | 9 | 7 | 5 | | | | | | 11 | |
| 1.5 | 15 | 12 | 11 | 9 | 7 | | | | | 13 | 9 | 8 | 7 | 6 | | | | | | 10 | |
| 1.6 | 13 | 11 | 10 | 8 | 6 | | | | | 12 | 8 | 7 | 6 | 5 | | | | | | 9 | |
| 1.7 | 12 | 10 | 9 | 8 | 6 | | | | | 11 | 8 | 7 | 6 | 5 | | | | | | 8 | |
| 1.8 | 12 | 10 | 9 | 8 | 6 | | | | | 10 | 8 | 7 | 6 | 5 | | | | | | 8 | |
| 1.9 | 11 | 9 | 8 | 7 | 6 | | | | | 10 | 8 | 7 | 6 | 5 | | | | | | 7 | |
| 2.0 | 10 | 8 | 8 | 7 | 5 | | | | | 9 | 7 | 7 | 6 | 5 | | | | | | 7 | |
| 2.1 | | 10 | 8 | 7 | 7 | | | | | 8 | 7 | 6 | 6 | 5 | | | | | | 7 | |
| 2.2 | | 9 | 8 | 7 | 6 | | | | | 8 | 7 | 6 | 5 | 5 | | | | | | 7 | |
| 2.3 | | 9 | 7 | 7 | 6 | | | | | 8 | 6 | 6 | 6 | 5 | | | | | | 6 | |
| 2.4 | | 8 | 7 | 7 | 6 | | | | | 7 | 6 | 6 | 6 | 5 | | | | | | 6 | |
| 2.5 | | 8 | 7 | 6 | 6 | | | | | 7 | 6 | 6 | 6 | 5 | | | | | | 6 | |
| 3.0 | | 7 | 6 | 6 | 5 | | | | | 6 | 5 | 5 | 5 | 5 | | | | | | 5 | |
| 3.5 | | 6 | 5 | 5 | 5 | | | | | 6 | 5 | 5 | 5 | 5 | | | | | | 5 | |
| 4.0 | | | | | 6 | | | | | | | | | | | | | | | | 6 |

*Reproducida con permiso de O. L. Davies, ed., *Design and Analysis of Industrial Experiments*, Oliver & Boyd, Edimburgo, 1956.

Tabla A.9* *Tamaño muestral para la prueba t de la diferencia entre dos medias*

| Prueba unilateral
Prueba bilateral | Nivel de la prueba t | | | | | | | | | | | | | | | | | | | |
|---------------------------------------|------------------------|------|-----|-----|-----|-----------------|------|-----|-----|-----|------------------|------|-----|-----|-----|-----------------|------|-----|-----|-----|
| | $\alpha = 0.005$ | | | | | $\alpha = 0.01$ | | | | | $\alpha = 0.025$ | | | | | $\alpha = 0.05$ | | | | |
| | $\alpha = 0.01$ | | | | | $\alpha = 0.02$ | | | | | $\alpha = 0.05$ | | | | | $\alpha = 0.1$ | | | | |
| $\beta =$ | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 | 0.01 | 0.05 | 0.1 | 0.2 | 0.5 |
| 0.05 | | | | | | | | | | | | | | | | | | | | |
| 0.10 | | | | | | | | | | | | | | | | | | | | |
| 0.15 | | | | | | | | | | | | | | | | | | | | |
| 0.20 | | | | | | | | | | | | | | | | | | | | 137 |
| 0.25 | | | | | | | | | | | | | | | | 124 | | | | 88 |
| 0.30 | | | | | | | | | | 123 | | | | | | | | | | 61 |
| 0.35 | | | | | | | | | | 90 | | | | | | | | | | 102 |
| 0.40 | | | | | 110 | | | | | 70 | | | | | | | | | | 45 |
| 0.45 | | | | | 85 | | | | | 70 | | | | | | 100 | 50 | | | 108 |
| 0.50 | | | | | 118 | 68 | | | | 101 | 55 | | | | | 105 | 79 | 39 | | 108 |
| 0.55 | | | | | 96 | 55 | | | | 106 | 82 | 45 | | | | 106 | 86 | 64 | 32 | 88 |
| 0.60 | | | | | 101 | 79 | 46 | | | 106 | 88 | 68 | 38 | | | 87 | 71 | 53 | 27 | 112 |
| 0.65 | | | | | 101 | 85 | 67 | 39 | | 90 | 74 | 58 | 32 | | | 104 | 74 | 60 | 45 | 23 |
| 0.70 | | | | | 87 | 73 | 57 | 34 | | 104 | 77 | 64 | 49 | 27 | | 88 | 63 | 51 | 39 | 20 |
| 0.75 | | | | | 100 | 75 | 63 | 50 | 29 | 90 | 66 | 55 | 43 | 24 | | 76 | 55 | 44 | 34 | 17 |
| 0.80 | | | | | 88 | 66 | 55 | 44 | 26 | 79 | 58 | 48 | 38 | 21 | | 67 | 48 | 39 | 29 | 15 |
| 0.85 | | | | | 77 | 58 | 49 | 39 | 23 | 70 | 51 | 43 | 33 | 19 | | 59 | 42 | 34 | 26 | 14 |
| 0.90 | | | | | 69 | 51 | 43 | 35 | 21 | 62 | 46 | 38 | 30 | 17 | | 52 | 37 | 31 | 23 | 12 |
| 0.95 | | | | | 62 | 46 | 39 | 31 | 19 | 55 | 41 | 34 | 27 | 15 | | 47 | 34 | 27 | 21 | 11 |
| 1.00 | | | | | 50 | 38 | 32 | 26 | 15 | 45 | 33 | 28 | 22 | 13 | | 42 | 30 | 25 | 19 | 10 |
| 1.1 | | | | | 42 | 32 | 27 | 22 | 13 | 38 | 28 | 23 | 19 | 11 | | 38 | 27 | 23 | 17 | 9 |
| 1.2 | | | | | 36 | 27 | 23 | 18 | 11 | 32 | 24 | 20 | 16 | 9 | | 32 | 23 | 19 | 14 | 8 |
| 1.3 | | | | | 31 | 23 | 20 | 16 | 10 | 28 | 21 | 17 | 14 | 8 | | 27 | 20 | 16 | 12 | 7 |
| 1.4 | | | | | 27 | 20 | 17 | 14 | 9 | 24 | 18 | 15 | 12 | 8 | | 23 | 17 | 14 | 11 | 6 |
| 1.5 | | | | | 24 | 18 | 15 | 13 | 8 | 21 | 16 | 14 | 11 | 7 | | 20 | 15 | 12 | 10 | 6 |
| 1.6 | | | | | 21 | 16 | 14 | 11 | 7 | 19 | 14 | 12 | 10 | 6 | | 18 | 13 | 11 | 9 | 5 |
| 1.7 | | | | | 19 | 15 | 13 | 10 | 7 | 17 | 13 | 11 | 9 | 6 | | 16 | 12 | 10 | 8 | 5 |
| 1.8 | | | | | 17 | 13 | 11 | 10 | 6 | 15 | 12 | 10 | 8 | 5 | | 14 | 11 | 9 | 7 | 4 |
| 1.9 | | | | | 16 | 12 | 11 | 9 | 6 | 14 | 11 | 9 | 8 | 5 | | 13 | 10 | 8 | 6 | 4 |
| 2.0 | | | | | 14 | 11 | 10 | 8 | 6 | 13 | 10 | 9 | 7 | 5 | | 12 | 9 | 7 | 6 | 4 |
| 2.1 | | | | | 13 | 10 | 9 | 8 | 5 | 12 | 9 | 8 | 7 | 5 | | 11 | 8 | 7 | 6 | 4 |
| 2.2 | | | | | 12 | 10 | 8 | 7 | 5 | 11 | 9 | 7 | 6 | 4 | | 10 | 8 | 6 | 5 | 3 |
| 2.3 | | | | | 11 | 9 | 8 | 7 | 5 | 10 | 8 | 7 | 6 | 4 | | 9 | 7 | 6 | 5 | 3 |
| 2.4 | | | | | 11 | 9 | 8 | 6 | 5 | 10 | 8 | 7 | 6 | 4 | | 9 | 7 | 6 | 5 | 3 |
| 2.5 | | | | | 10 | 8 | 7 | 6 | 4 | 9 | 7 | 6 | 5 | 4 | | 8 | 6 | 5 | 4 | 3 |
| 3.0 | | | | | 8 | 6 | 6 | 5 | 4 | 7 | 6 | 5 | 4 | 3 | | 7 | 5 | 4 | 3 | 2 |
| 3.5 | | | | | 6 | 5 | 5 | 4 | 3 | 6 | 5 | 4 | 4 | 3 | | 6 | 5 | 4 | 3 | 2 |
| 4.0 | | | | | 6 | 5 | 4 | 4 | 3 | 5 | 4 | 4 | 3 | 3 | | 5 | 4 | 4 | 3 | 2 |

*Reproducida con permiso de O. L. Davies, ed., *Design and Analysis of Industrial Experiments*, Oliver & Boyd, Edimburgo, 1956.

Tabla A.10* Valores críticos para la prueba de Bartlett

| n | $b_{\alpha} (0,01; n)$ | | | | | | | | | |
|-----|--------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--|
| | Numero de poblaciones, k | | | | | | | | | |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 3 | 0.1411 | 0.1672 | * | * | * | * | * | * | * | |
| 4 | 0.2843 | 0.3165 | 0.3475 | 0.3729 | 0.3937 | 0.4110 | * | * | * | |
| 5 | 0.3984 | 0.4304 | 0.4607 | 0.4850 | 0.5046 | 0.5207 | 0.5343 | 0.5458 | 0.5558 | |
| 6 | 0.4850 | 0.5149 | 0.5430 | 0.5653 | 0.5832 | 0.5978 | 0.6100 | 0.6204 | 0.6293 | |
| 7 | 0.5512 | 0.5787 | 0.6045 | 0.6248 | 0.6410 | 0.6542 | 0.6652 | 0.6744 | 0.6824 | |
| 8 | 0.6031 | 0.6282 | 0.6518 | 0.6704 | 0.6851 | 0.6970 | 0.7069 | 0.7153 | 0.7225 | |
| 9 | 0.6445 | 0.6676 | 0.6892 | 0.7062 | 0.7197 | 0.7305 | 0.7395 | 0.7471 | 0.7536 | |
| 10 | 0.6783 | 0.6996 | 0.7195 | 0.7352 | 0.7475 | 0.7575 | 0.7657 | 0.7726 | 0.7786 | |
| 11 | 0.7063 | 0.7260 | 0.7445 | 0.7590 | 0.7703 | 0.7795 | 0.7871 | 0.7935 | 0.7990 | |
| 12 | 0.7299 | 0.7483 | 0.7654 | 0.7789 | 0.7894 | 0.7980 | 0.8050 | 0.8109 | 0.8160 | |
| 13 | 0.7501 | 0.7672 | 0.7832 | 0.7958 | 0.8056 | 0.8135 | 0.8201 | 0.8256 | 0.8303 | |
| 14 | 0.7674 | 0.7835 | 0.7985 | 0.8103 | 0.8195 | 0.8269 | 0.8330 | 0.8382 | 0.8426 | |
| 15 | 0.7825 | 0.7977 | 0.8118 | 0.8229 | 0.8315 | 0.8385 | 0.8443 | 0.8491 | 0.8532 | |
| 16 | 0.7958 | 0.8101 | 0.8235 | 0.8339 | 0.8421 | 0.8486 | 0.8541 | 0.8586 | 0.8625 | |
| 17 | 0.8076 | 0.8211 | 0.8338 | 0.8436 | 0.8514 | 0.8576 | 0.8627 | 0.8670 | 0.8707 | |
| 18 | 0.8181 | 0.8300 | 0.8429 | 0.8523 | 0.8596 | 0.8655 | 0.8704 | 0.8745 | 0.8780 | |
| 19 | 0.8275 | 0.8397 | 0.8512 | 0.8601 | 0.8670 | 0.8727 | 0.8773 | 0.8811 | 0.8845 | |
| 20 | 0.8360 | 0.8476 | 0.8586 | 0.8671 | 0.8737 | 0.8791 | 0.8833 | 0.8871 | 0.8903 | |
| 21 | 0.8437 | 0.8548 | 0.8653 | 0.8734 | 0.8797 | 0.8848 | 0.8890 | 0.8926 | 0.8956 | |
| 22 | 0.8507 | 0.8614 | 0.8714 | 0.8791 | 0.8852 | 0.8901 | 0.8941 | 0.8975 | 0.9004 | |
| 23 | 0.8571 | 0.8673 | 0.8769 | 0.8844 | 0.8902 | 0.8949 | 0.8988 | 0.9020 | 0.9047 | |
| 24 | 0.8630 | 0.8728 | 0.8820 | 0.8892 | 0.8948 | 0.8993 | 0.9030 | 0.9061 | 0.9087 | |
| 25 | 0.8684 | 0.8779 | 0.8867 | 0.8936 | 0.8990 | 0.9034 | 0.9069 | 0.9099 | 0.9124 | |
| 26 | 0.8734 | 0.8825 | 0.8911 | 0.8977 | 0.9029 | 0.9071 | 0.9105 | 0.9134 | 0.9158 | |
| 27 | 0.8781 | 0.8869 | 0.8951 | 0.9015 | 0.9065 | 0.9105 | 0.9138 | 0.9166 | 0.9190 | |
| 28 | 0.8824 | 0.8909 | 0.8988 | 0.9050 | 0.9099 | 0.9138 | 0.9169 | 0.9196 | 0.9219 | |
| 29 | 0.8864 | 0.8946 | 0.9023 | 0.9083 | 0.9130 | 0.9167 | 0.9198 | 0.9224 | 0.9246 | |
| 30 | 0.8902 | 0.8981 | 0.9056 | 0.9114 | 0.9159 | 0.9195 | 0.9225 | 0.9250 | 0.9271 | |
| 40 | 0.9175 | 0.9235 | 0.9291 | 0.9335 | 0.9370 | 0.9397 | 0.9420 | 0.9439 | 0.9455 | |
| 50 | 0.9339 | 0.9387 | 0.9433 | 0.9468 | 0.9496 | 0.9518 | 0.9536 | 0.9551 | 0.9564 | |
| 60 | 0.9449 | 0.9489 | 0.9527 | 0.9557 | 0.9580 | 0.9599 | 0.9614 | 0.9626 | 0.9637 | |
| 80 | 0.9586 | 0.9617 | 0.9646 | 0.9668 | 0.9685 | 0.9699 | 0.9711 | 0.9720 | 0.9728 | |
| 100 | 0.9669 | 0.9693 | 0.9716 | 0.9734 | 0.9748 | 0.9759 | 0.9769 | 0.9776 | 0.9783 | |

*Reproducido de D. D. Dyer y J. P. Keating: "On the determination of critical values for Bartlett's test", *J. Am. Stat. Assoc.*, vol. 75, 1980, con permiso del consejo de directores.

Tabla A.10 (continuación) Valores críticos para la prueba de Bartlett
 $b_1(0.05; n)$

| n | Número de poblaciones, k | | | | | | | | |
|-----|--------------------------|--------|--------|--------|--------|--------|--------|--------|--------|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 3 | 0.3123 | 0.3058 | 0.3173 | 0.3299 | * | * | * | * | * |
| 4 | 0.4780 | 0.4699 | 0.4803 | 0.4921 | 0.5028 | 0.5122 | 0.5204 | 0.5277 | 0.5341 |
| 5 | 0.5845 | 0.5762 | 0.5850 | 0.5952 | 0.6045 | 0.6126 | 0.6197 | 0.6260 | 0.6315 |
| 6 | 0.6563 | 0.6483 | 0.6559 | 0.6646 | 0.6727 | 0.6798 | 0.6860 | 0.6914 | 0.6961 |
| 7 | 0.7075 | 0.7000 | 0.7065 | 0.7142 | 0.7213 | 0.7275 | 0.7329 | 0.7376 | 0.7418 |
| 8 | 0.7456 | 0.7387 | 0.7444 | 0.7512 | 0.7574 | 0.7629 | 0.7677 | 0.7719 | 0.7757 |
| 9 | 0.7751 | 0.7686 | 0.7737 | 0.7798 | 0.7854 | 0.7903 | 0.7946 | 0.7984 | 0.8017 |
| 10 | 0.7984 | 0.7924 | 0.7970 | 0.8025 | 0.8076 | 0.8121 | 0.8160 | 0.8194 | 0.8224 |
| 11 | 0.8175 | 0.8118 | 0.8160 | 0.8210 | 0.8257 | 0.8298 | 0.8333 | 0.8365 | 0.8392 |
| 12 | 0.8332 | 0.8280 | 0.8317 | 0.8364 | 0.8407 | 0.8444 | 0.8477 | 0.8506 | 0.8531 |
| 13 | 0.8465 | 0.8415 | 0.8450 | 0.8493 | 0.8533 | 0.8568 | 0.8598 | 0.8625 | 0.8648 |
| 14 | 0.8578 | 0.8532 | 0.8564 | 0.8604 | 0.8641 | 0.8673 | 0.8701 | 0.8726 | 0.8748 |
| 15 | 0.8676 | 0.8632 | 0.8662 | 0.8699 | 0.8734 | 0.8764 | 0.8790 | 0.8814 | 0.8834 |
| 16 | 0.8761 | 0.8719 | 0.8747 | 0.8782 | 0.8815 | 0.8843 | 0.8868 | 0.8890 | 0.8909 |
| 17 | 0.8836 | 0.8796 | 0.8823 | 0.8856 | 0.8886 | 0.8913 | 0.8936 | 0.8957 | 0.8975 |
| 18 | 0.8902 | 0.8865 | 0.8890 | 0.8921 | 0.8949 | 0.8975 | 0.8997 | 0.9016 | 0.9033 |
| 19 | 0.8961 | 0.8926 | 0.8949 | 0.8979 | 0.9006 | 0.9030 | 0.9051 | 0.9069 | 0.9086 |
| 20 | 0.9015 | 0.8980 | 0.9003 | 0.9031 | 0.9057 | 0.9080 | 0.9100 | 0.9117 | 0.9132 |
| 21 | 0.9063 | 0.9030 | 0.9051 | 0.9078 | 0.9103 | 0.9124 | 0.9143 | 0.9160 | 0.9175 |
| 22 | 0.9106 | 0.9075 | 0.9095 | 0.9120 | 0.9144 | 0.9165 | 0.9183 | 0.9199 | 0.9213 |
| 23 | 0.9146 | 0.9116 | 0.9135 | 0.9159 | 0.9182 | 0.9202 | 0.9219 | 0.9235 | 0.9248 |
| 24 | 0.9182 | 0.9153 | 0.9172 | 0.9195 | 0.9217 | 0.9236 | 0.9253 | 0.9267 | 0.9280 |
| 25 | 0.9216 | 0.9187 | 0.9205 | 0.9228 | 0.9249 | 0.9267 | 0.9283 | 0.9297 | 0.9309 |
| 26 | 0.9246 | 0.9219 | 0.9236 | 0.9258 | 0.9278 | 0.9296 | 0.9311 | 0.9325 | 0.9338 |
| 27 | 0.9275 | 0.9249 | 0.9265 | 0.9286 | 0.9305 | 0.9322 | 0.9337 | 0.9350 | 0.9361 |
| 28 | 0.9301 | 0.9276 | 0.9292 | 0.9312 | 0.9330 | 0.9347 | 0.9361 | 0.9374 | 0.9385 |
| 29 | 0.9326 | 0.9301 | 0.9316 | 0.9336 | 0.9354 | 0.9370 | 0.9383 | 0.9396 | 0.9406 |
| 30 | 0.9348 | 0.9325 | 0.9340 | 0.9358 | 0.9376 | 0.9391 | 0.9404 | 0.9416 | 0.9426 |
| 40 | 0.9513 | 0.9495 | 0.9506 | 0.9520 | 0.9533 | 0.9545 | 0.9555 | 0.9564 | 0.9572 |
| 50 | 0.9612 | 0.9597 | 0.9606 | 0.9617 | 0.9628 | 0.9637 | 0.9645 | 0.9652 | 0.9658 |
| 60 | 0.9677 | 0.9665 | 0.9672 | 0.9681 | 0.9690 | 0.9698 | 0.9705 | 0.9710 | 0.9716 |
| 80 | 0.9758 | 0.9749 | 0.9754 | 0.9761 | 0.9768 | 0.9774 | 0.9779 | 0.9783 | 0.9787 |
| 100 | 0.9807 | 0.9799 | 0.9804 | 0.9809 | 0.9815 | 0.9819 | 0.9823 | 0.9827 | 0.9830 |

Tabla A.11* Valores críticos para la prueba de Cochran

$\alpha = 0.01$

| k \ n | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 17 | 37 | 145 | ∞ |
|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| 2 | 0.9999 | 0.9950 | 0.9794 | 0.9586 | 0.9373 | 0.9172 | 0.8988 | 0.8823 | 0.8674 | 0.8539 | 0.7949 | 0.7007 | 0.6002 | 0.5000 |
| 3 | 0.9923 | 0.9423 | 0.8831 | 0.8335 | 0.7933 | 0.7606 | 0.7335 | 0.7107 | 0.6912 | 0.6743 | 0.6059 | 0.5153 | 0.4240 | 0.3333 |
| 4 | 0.9676 | 0.8643 | 0.7814 | 0.7212 | 0.6761 | 0.6410 | 0.6129 | 0.5897 | 0.5702 | 0.5536 | 0.4884 | 0.4057 | 0.3251 | 0.2500 |
| 5 | 0.9279 | 0.7885 | 0.6957 | 0.6329 | 0.5875 | 0.5531 | 0.5289 | 0.5037 | 0.4854 | 0.4697 | 0.4094 | 0.3351 | 0.2644 | 0.2000 |
| 6 | 0.8828 | 0.7218 | 0.6258 | 0.5635 | 0.5195 | 0.4866 | 0.4608 | 0.4401 | 0.4229 | 0.4084 | 0.3529 | 0.2858 | 0.2229 | 0.1667 |
| 7 | 0.8376 | 0.6644 | 0.5685 | 0.5080 | 0.4659 | 0.4347 | 0.4105 | 0.3911 | 0.3751 | 0.3616 | 0.3105 | 0.2494 | 0.1929 | 0.1429 |
| 8 | 0.7945 | 0.6132 | 0.5209 | 0.4627 | 0.4226 | 0.3932 | 0.3704 | 0.3522 | 0.3373 | 0.3248 | 0.2779 | 0.2214 | 0.1700 | 0.1250 |
| 9 | 0.7544 | 0.5727 | 0.4810 | 0.4251 | 0.3870 | 0.3592 | 0.3378 | 0.3207 | 0.3067 | 0.2950 | 0.2514 | 0.1992 | 0.1521 | 0.1111 |
| 10 | 0.7175 | 0.5358 | 0.4469 | 0.3934 | 0.3572 | 0.3308 | 0.3106 | 0.2945 | 0.2813 | 0.2704 | 0.2297 | 0.1811 | 0.1376 | 0.1000 |
| 12 | 0.6528 | 0.4751 | 0.3919 | 0.3428 | 0.3099 | 0.2861 | 0.2680 | 0.2535 | 0.2419 | 0.2320 | 0.1961 | 0.1535 | 0.1157 | 0.0833 |
| 15 | 0.5747 | 0.4069 | 0.3317 | 0.2882 | 0.2593 | 0.2386 | 0.2228 | 0.2104 | 0.2002 | 0.1918 | 0.1612 | 0.1251 | 0.0934 | 0.0667 |
| 20 | 0.4799 | 0.3297 | 0.2654 | 0.2308 | 0.2048 | 0.1877 | 0.1748 | 0.1646 | 0.1567 | 0.1501 | 0.1248 | 0.0980 | 0.0769 | 0.0500 |
| 24 | 0.4247 | 0.2871 | 0.2295 | 0.1970 | 0.1759 | 0.1608 | 0.1495 | 0.1406 | 0.1338 | 0.1283 | 0.1060 | 0.0810 | 0.0595 | 0.0417 |
| 30 | 0.3652 | 0.2412 | 0.1913 | 0.1635 | 0.1454 | 0.1327 | 0.1232 | 0.1157 | 0.1100 | 0.1054 | 0.0867 | 0.0658 | 0.0480 | 0.0333 |
| 40 | 0.2940 | 0.1915 | 0.1508 | 0.1281 | 0.1135 | 0.1033 | 0.0957 | 0.0898 | 0.0853 | 0.0816 | 0.0668 | 0.0503 | 0.0363 | 0.0250 |
| 60 | 0.2151 | 0.1371 | 0.1069 | 0.0902 | 0.0796 | 0.0722 | 0.0668 | 0.0625 | 0.0594 | 0.0567 | 0.0461 | 0.0344 | 0.0245 | 0.0167 |
| 120 | 0.1225 | 0.0759 | 0.0585 | 0.0489 | 0.0429 | 0.0387 | 0.0357 | 0.0334 | 0.0316 | 0.0302 | 0.0242 | 0.0178 | 0.0125 | 0.0083 |
| ∞ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Reproducida de C. Eisenhart, M. W. Hartley, y W. A. Wallis, *Techniques of Statistical Analysis*, capítulo 15, McGraw-Hill Book Company, Nueva York, 1947. Utilizada con permiso de McGraw-Hill Book Company.

Tabla A.12* Rangos studentizados significativos mínimos r_p

$\alpha = 0.05$

| p | P | | | | | | | | | |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 1 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 | 17.97 |
| 2 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 | 6.085 |
| 3 | 4.501 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 | 4.516 |
| 4 | 3.927 | 4.013 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 | 4.033 |
| 5 | 3.635 | 3.749 | 3.797 | 3.814 | 3.814 | 3.814 | 3.814 | 3.814 | 3.814 | 3.814 |
| 6 | 3.461 | 3.587 | 3.649 | 3.680 | 3.694 | 3.697 | 3.697 | 3.697 | 3.697 | 3.697 |
| 7 | 3.344 | 3.477 | 3.548 | 3.588 | 3.611 | 3.622 | 3.626 | 3.626 | 3.626 | 3.626 |
| 8 | 3.261 | 3.399 | 3.475 | 3.521 | 3.549 | 3.566 | 3.575 | 3.579 | 3.579 | 3.579 |
| 9 | 3.199 | 3.339 | 3.420 | 3.470 | 3.502 | 3.523 | 3.536 | 3.544 | 3.547 | 3.547 |
| 10 | 3.151 | 3.293 | 3.376 | 3.430 | 3.465 | 3.489 | 3.505 | 3.516 | 3.522 | 3.522 |
| 11 | 3.113 | 3.256 | 3.342 | 3.397 | 3.435 | 3.462 | 3.480 | 3.493 | 3.501 | 3.501 |
| 12 | 3.082 | 3.225 | 3.313 | 3.370 | 3.410 | 3.439 | 3.459 | 3.474 | 3.484 | 3.484 |
| 13 | 3.055 | 3.200 | 3.289 | 3.348 | 3.389 | 3.419 | 3.442 | 3.458 | 3.470 | 3.470 |
| 14 | 3.033 | 3.178 | 3.268 | 3.329 | 3.372 | 3.403 | 3.426 | 3.444 | 3.457 | 3.457 |
| 15 | 3.014 | 3.160 | 3.250 | 3.312 | 3.356 | 3.389 | 3.413 | 3.432 | 3.446 | 3.446 |
| 16 | 2.998 | 3.144 | 3.235 | 3.298 | 3.343 | 3.376 | 3.402 | 3.422 | 3.437 | 3.437 |
| 17 | 2.984 | 3.130 | 3.222 | 3.285 | 3.331 | 3.366 | 3.392 | 3.412 | 3.429 | 3.429 |
| 18 | 2.971 | 3.118 | 3.210 | 3.274 | 3.321 | 3.356 | 3.383 | 3.405 | 3.421 | 3.421 |
| 19 | 2.960 | 3.107 | 3.199 | 3.264 | 3.311 | 3.347 | 3.375 | 3.397 | 3.415 | 3.415 |
| 20 | 2.950 | 3.097 | 3.190 | 3.255 | 3.303 | 3.339 | 3.368 | 3.391 | 3.409 | 3.409 |
| 24 | 2.919 | 3.066 | 3.160 | 3.226 | 3.276 | 3.315 | 3.345 | 3.370 | 3.390 | 3.390 |
| 30 | 2.888 | 3.035 | 3.131 | 3.199 | 3.250 | 3.290 | 3.322 | 3.349 | 3.371 | 3.371 |
| 40 | 2.858 | 3.006 | 3.102 | 3.171 | 3.224 | 3.266 | 3.300 | 3.328 | 3.352 | 3.352 |
| 60 | 2.829 | 2.976 | 3.073 | 3.143 | 3.198 | 3.241 | 3.277 | 3.307 | 3.333 | 3.333 |
| 120 | 2.800 | 2.947 | 3.045 | 3.116 | 3.172 | 3.217 | 3.254 | 3.287 | 3.314 | 3.314 |
| ∞ | 2.772 | 2.918 | 3.017 | 3.089 | 3.146 | 3.193 | 3.232 | 3.265 | 3.294 | 3.294 |

* Condensada de H. Leon Hartel, "Critical Values for Duncan's New Multiple Range Test", *Biometrics*, vol. 16, Núm. 4, 1960, con permiso del autor y del editor.

Tabla A.12 (continuación) Rangos studentizados significativos mínimos r_p
 $\alpha = 0.01$

| r | p | | | | | | | | | |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| 1 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 | 90.03 |
| 2 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 | 14.04 |
| 3 | 8.261 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 | 8.321 |
| 4 | 6.512 | 6.677 | 6.740 | 6.756 | 6.756 | 6.756 | 6.756 | 6.756 | 6.756 | 6.756 |
| 5 | 5.702 | 5.893 | 5.989 | 6.040 | 6.065 | 6.074 | 6.074 | 6.074 | 6.074 | 6.074 |
| 6 | 5.243 | 5.439 | 5.549 | 5.614 | 5.655 | 5.680 | 5.694 | 5.701 | 5.703 | 5.703 |
| 7 | 4.949 | 5.145 | 5.260 | 5.334 | 5.383 | 5.416 | 5.439 | 5.454 | 5.464 | 5.464 |
| 8 | 4.746 | 4.939 | 5.057 | 5.135 | 5.189 | 5.227 | 5.256 | 5.276 | 5.291 | 5.291 |
| 9 | 4.596 | 4.787 | 4.906 | 4.986 | 5.043 | 5.086 | 5.118 | 5.142 | 5.160 | 5.160 |
| 10 | 4.482 | 4.671 | 4.790 | 4.871 | 4.931 | 4.975 | 5.010 | 5.037 | 5.058 | 5.058 |
| 11 | 4.392 | 4.579 | 4.697 | 4.780 | 4.841 | 4.887 | 4.924 | 4.952 | 4.975 | 4.975 |
| 12 | 4.320 | 4.504 | 4.622 | 4.706 | 4.767 | 4.815 | 4.852 | 4.883 | 4.907 | 4.907 |
| 13 | 4.260 | 4.442 | 4.560 | 4.644 | 4.706 | 4.755 | 4.793 | 4.824 | 4.850 | 4.850 |
| 14 | 4.210 | 4.391 | 4.508 | 4.591 | 4.654 | 4.704 | 4.743 | 4.775 | 4.802 | 4.802 |
| 15 | 4.168 | 4.347 | 4.463 | 4.547 | 4.610 | 4.660 | 4.700 | 4.733 | 4.760 | 4.760 |
| 16 | 4.131 | 4.309 | 4.425 | 4.509 | 4.572 | 4.622 | 4.663 | 4.699 | 4.724 | 4.724 |
| 17 | 4.099 | 4.275 | 4.391 | 4.475 | 4.539 | 4.589 | 4.630 | 4.664 | 4.693 | 4.693 |
| 18 | 4.071 | 4.246 | 4.362 | 4.445 | 4.509 | 4.560 | 4.601 | 4.635 | 4.664 | 4.664 |
| 19 | 4.046 | 4.220 | 4.335 | 4.419 | 4.483 | 4.534 | 4.575 | 4.610 | 4.639 | 4.639 |
| 20 | 4.024 | 4.197 | 4.312 | 4.395 | 4.459 | 4.510 | 4.552 | 4.587 | 4.617 | 4.617 |
| 24 | 3.956 | 4.126 | 4.239 | 4.322 | 4.386 | 4.437 | 4.480 | 4.516 | 4.546 | 4.546 |
| 30 | 3.889 | 4.056 | 4.168 | 4.250 | 4.314 | 4.366 | 4.409 | 4.445 | 4.477 | 4.477 |
| 40 | 3.825 | 3.988 | 4.098 | 4.180 | 4.244 | 4.296 | 4.339 | 4.376 | 4.408 | 4.408 |
| 60 | 3.762 | 3.922 | 4.031 | 4.111 | 4.174 | 4.226 | 4.270 | 4.307 | 4.340 | 4.340 |
| 120 | 3.702 | 3.858 | 3.965 | 4.044 | 4.107 | 4.158 | 4.202 | 4.239 | 4.272 | 4.272 |
| ∞ | 3.643 | 3.796 | 3.900 | 3.978 | 4.040 | 4.091 | 4.135 | 4.172 | 4.205 | 4.205 |

Tabla A.13 Valores de $d_{\alpha/2}(k, v)$ para comparaciones bilaterales entre k tratamientos y un control

| | | $\alpha = 0.05$ | | | | | | | | |
|----------|------|---|------|------|------|------|------|------|------|--|
| | | $k = \text{número de medias de tratamiento (excluye el control)}$ | | | | | | | | |
| v | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 5 | 2.57 | 3.03 | 3.29 | 3.48 | 3.62 | 3.73 | 3.82 | 3.90 | 3.97 | |
| 6 | 2.45 | 2.86 | 3.10 | 3.26 | 3.39 | 3.49 | 3.57 | 3.64 | 3.71 | |
| 7 | 2.36 | 2.75 | 2.97 | 3.12 | 3.24 | 3.33 | 3.41 | 3.47 | 3.53 | |
| 8 | 2.31 | 2.67 | 2.88 | 3.02 | 3.13 | 3.22 | 3.29 | 3.35 | 3.41 | |
| 9 | 2.26 | 2.61 | 2.81 | 2.95 | 3.05 | 3.14 | 3.20 | 3.26 | 3.32 | |
| 10 | 2.23 | 2.57 | 2.76 | 2.89 | 2.99 | 3.07 | 3.14 | 3.19 | 3.24 | |
| 11 | 2.20 | 2.53 | 2.72 | 2.84 | 2.94 | 3.02 | 3.08 | 3.14 | 3.19 | |
| 12 | 2.18 | 2.50 | 2.68 | 2.81 | 2.90 | 2.98 | 3.04 | 3.09 | 3.14 | |
| 13 | 2.16 | 2.48 | 2.65 | 2.78 | 2.87 | 2.94 | 3.00 | 3.06 | 3.10 | |
| 14 | 2.14 | 2.46 | 2.63 | 2.75 | 2.84 | 2.91 | 2.97 | 3.02 | 3.07 | |
| 15 | 2.13 | 2.44 | 2.61 | 2.73 | 2.82 | 2.89 | 2.95 | 3.00 | 3.04 | |
| 16 | 2.12 | 2.42 | 2.59 | 2.71 | 2.80 | 2.87 | 2.92 | 2.97 | 3.02 | |
| 17 | 2.11 | 2.41 | 2.58 | 2.69 | 2.78 | 2.85 | 2.90 | 2.95 | 3.00 | |
| 18 | 2.10 | 2.40 | 2.56 | 2.68 | 2.76 | 2.83 | 2.89 | 2.94 | 2.98 | |
| 19 | 2.09 | 2.39 | 2.55 | 2.66 | 2.75 | 2.81 | 2.87 | 2.92 | 2.96 | |
| 20 | 2.09 | 2.38 | 2.54 | 2.65 | 2.73 | 2.80 | 2.86 | 2.90 | 2.95 | |
| 24 | 2.06 | 2.35 | 2.51 | 2.61 | 2.70 | 2.76 | 2.81 | 2.86 | 2.90 | |
| 30 | 2.04 | 2.32 | 2.47 | 2.58 | 2.66 | 2.72 | 2.77 | 2.82 | 2.86 | |
| 40 | 2.02 | 2.29 | 2.44 | 2.54 | 2.62 | 2.68 | 2.73 | 2.77 | 2.81 | |
| 60 | 2.00 | 2.27 | 2.41 | 2.51 | 2.58 | 2.64 | 2.69 | 2.73 | 2.77 | |
| 120 | 1.98 | 2.24 | 2.38 | 2.47 | 2.55 | 2.60 | 2.65 | 2.69 | 2.73 | |
| ∞ | 1.96 | 2.21 | 2.35 | 2.44 | 2.51 | 2.57 | 2.61 | 2.65 | 2.69 | |

*Reproducida de Charles W. Dunnett, "New Tables for Multiple Comparison with a Control", *Biometrics*, vol. 20, Núm. 3, 1964, con permiso del autor y del editor.

Tabla A.13* (continuación) Valores de $d_{\alpha/2}(k, v)$ para comparaciones bilaterales entre k tratamientos y un control

$\alpha = 0.01$

| v | $k = \text{número de medias de tratamiento (excluye el control)}$ | | | | | | | | |
|----------|---|------|------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 4.03 | 4.63 | 4.98 | 5.22 | 5.41 | 5.56 | 5.69 | 5.80 | 5.89 |
| 6 | 3.71 | 4.21 | 4.51 | 4.71 | 4.87 | 5.00 | 5.10 | 5.20 | 5.28 |
| 7 | 3.50 | 3.95 | 4.21 | 4.39 | 4.53 | 4.64 | 4.74 | 4.82 | 4.89 |
| 8 | 3.36 | 3.77 | 4.00 | 4.17 | 4.29 | 4.40 | 4.48 | 4.56 | 4.62 |
| 9 | 3.25 | 3.63 | 3.85 | 4.01 | 4.12 | 4.22 | 4.30 | 4.37 | 4.43 |
| 10 | 3.17 | 3.53 | 3.74 | 3.88 | 3.99 | 4.08 | 4.16 | 4.22 | 4.28 |
| 11 | 3.11 | 3.45 | 3.65 | 3.79 | 3.89 | 3.98 | 4.05 | 4.11 | 4.16 |
| 12 | 3.05 | 3.39 | 3.58 | 3.71 | 3.81 | 3.89 | 3.96 | 4.02 | 4.07 |
| 13 | 3.01 | 3.33 | 3.52 | 3.65 | 3.74 | 3.82 | 3.89 | 3.94 | 3.99 |
| 14 | 2.98 | 3.29 | 3.47 | 3.59 | 3.69 | 3.76 | 3.83 | 3.88 | 3.93 |
| 15 | 2.95 | 3.25 | 3.43 | 3.55 | 3.64 | 3.71 | 3.78 | 3.83 | 3.88 |
| 16 | 2.92 | 3.22 | 3.39 | 3.51 | 3.60 | 3.67 | 3.73 | 3.78 | 3.83 |
| 17 | 2.90 | 3.19 | 3.36 | 3.47 | 3.56 | 3.63 | 3.69 | 3.74 | 3.79 |
| 18 | 2.88 | 3.17 | 3.33 | 3.44 | 3.53 | 3.60 | 3.66 | 3.71 | 3.75 |
| 19 | 2.86 | 3.15 | 3.31 | 3.42 | 3.50 | 3.57 | 3.63 | 3.68 | 3.72 |
| 20 | 2.85 | 3.13 | 3.29 | 3.40 | 3.48 | 3.55 | 3.60 | 3.65 | 3.69 |
| 24 | 2.80 | 3.07 | 3.22 | 3.32 | 3.40 | 3.47 | 3.52 | 3.57 | 3.61 |
| 30 | 2.75 | 3.01 | 3.15 | 3.25 | 3.33 | 3.39 | 3.44 | 3.49 | 3.52 |
| 40 | 2.70 | 2.95 | 3.09 | 3.19 | 3.26 | 3.32 | 3.37 | 3.41 | 3.44 |
| 60 | 2.66 | 2.90 | 3.03 | 3.12 | 3.19 | 3.25 | 3.29 | 3.33 | 3.37 |
| 120 | 2.62 | 2.85 | 2.97 | 3.06 | 3.12 | 3.18 | 3.22 | 3.26 | 3.29 |
| ∞ | 2.58 | 2.79 | 2.92 | 3.00 | 3.06 | 3.11 | 3.15 | 3.19 | 3.22 |

Tabla A.14* Valores de $d_{\alpha}(k, v)$ para comparaciones unilaterales entre k tratamientos y un control

$\alpha = 0.05$

| v | $k = \text{número de medias de tratamiento (excluye el control)}$ | | | | | | | | |
|----------|---|------|------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 2.02 | 2.44 | 2.68 | 2.85 | 2.98 | 3.08 | 3.16 | 3.24 | 3.30 |
| 6 | 1.94 | 2.34 | 2.56 | 2.71 | 2.83 | 2.92 | 3.00 | 3.07 | 3.12 |
| 7 | 1.89 | 2.27 | 2.48 | 2.62 | 2.73 | 2.82 | 2.89 | 2.95 | 3.01 |
| 8 | 1.86 | 2.22 | 2.42 | 2.55 | 2.66 | 2.74 | 2.81 | 2.87 | 2.92 |
| 9 | 1.83 | 2.18 | 2.37 | 2.50 | 2.60 | 2.68 | 2.75 | 2.81 | 2.86 |
| 10 | 1.81 | 2.15 | 2.34 | 2.47 | 2.56 | 2.64 | 2.70 | 2.76 | 2.81 |
| 11 | 1.80 | 2.13 | 2.31 | 2.44 | 2.53 | 2.60 | 2.67 | 2.72 | 2.77 |
| 12 | 1.78 | 2.11 | 2.29 | 2.41 | 2.50 | 2.58 | 2.64 | 2.69 | 2.74 |
| 13 | 1.77 | 2.09 | 2.27 | 2.39 | 2.48 | 2.55 | 2.61 | 2.66 | 2.71 |
| 14 | 1.76 | 2.08 | 2.25 | 2.37 | 2.46 | 2.53 | 2.59 | 2.64 | 2.69 |
| 15 | 1.75 | 2.07 | 2.24 | 2.36 | 2.44 | 2.51 | 2.57 | 2.62 | 2.67 |
| 16 | 1.75 | 2.06 | 2.23 | 2.34 | 2.43 | 2.50 | 2.56 | 2.61 | 2.65 |
| 17 | 1.74 | 2.05 | 2.22 | 2.33 | 2.42 | 2.49 | 2.54 | 2.59 | 2.64 |
| 18 | 1.73 | 2.04 | 2.21 | 2.32 | 2.41 | 2.48 | 2.53 | 2.58 | 2.62 |
| 19 | 1.73 | 2.03 | 2.20 | 2.31 | 2.40 | 2.47 | 2.52 | 2.57 | 2.61 |
| 20 | 1.72 | 2.03 | 2.19 | 2.30 | 2.39 | 2.46 | 2.51 | 2.56 | 2.60 |
| 24 | 1.71 | 2.01 | 2.17 | 2.28 | 2.36 | 2.43 | 2.48 | 2.53 | 2.57 |
| 30 | 1.70 | 1.99 | 2.15 | 2.25 | 2.33 | 2.40 | 2.45 | 2.50 | 2.54 |
| 40 | 1.68 | 1.97 | 2.13 | 2.23 | 2.31 | 2.37 | 2.42 | 2.47 | 2.51 |
| 60 | 1.67 | 1.95 | 2.10 | 2.21 | 2.28 | 2.35 | 2.39 | 2.44 | 2.48 |
| 120 | 1.66 | 1.93 | 2.08 | 2.18 | 2.26 | 2.32 | 2.37 | 2.41 | 2.45 |
| ∞ | 1.64 | 1.92 | 2.06 | 2.16 | 2.23 | 2.29 | 2.34 | 2.38 | 2.42 |

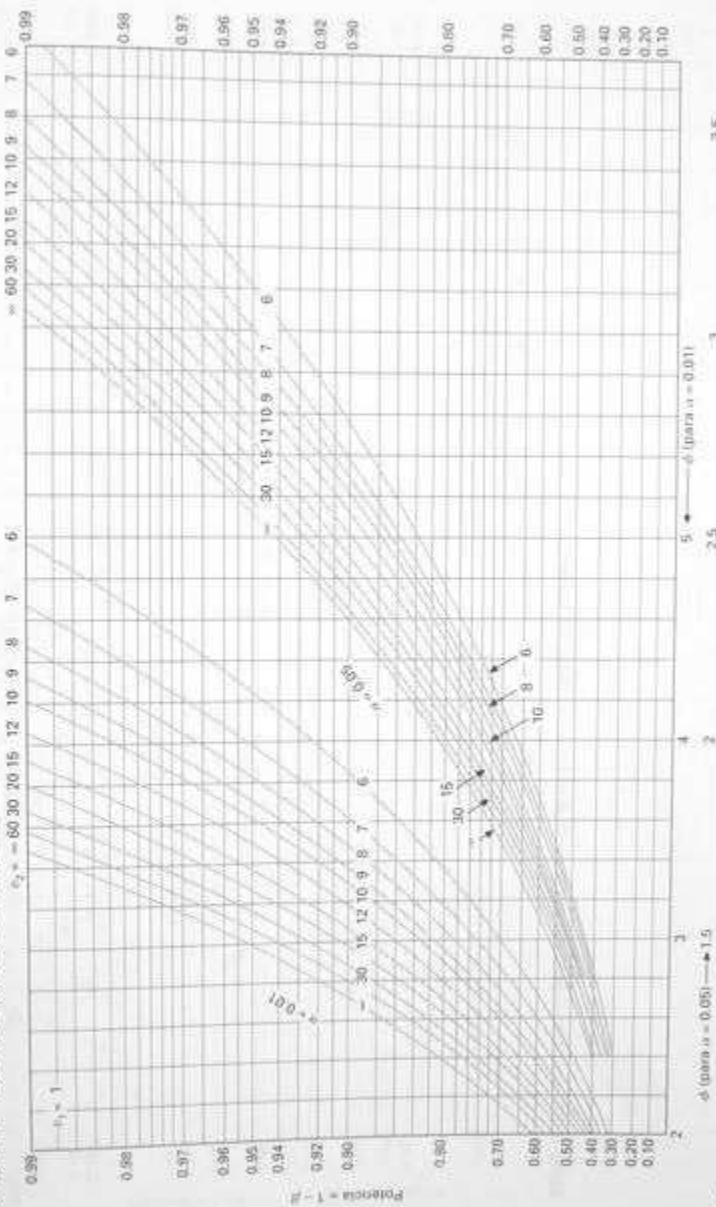
*Reproducida de Charles W. Dunnett, "A Multiple Comparison Procedure for Comparing Several Treatments with a Control", *J. Am. Stat. Assoc.*, vol. 50, 1955, 1096-1121, con permiso del autor y del editor.

Tabla A.14* (continuación) Valores de $d_{\alpha/2}(k, v)$ para comparaciones unilaterales entre k tratamientos y un control

$\alpha = 0.01$

| v | $k = \text{número de medias de tratamiento (excluye el control)}$ | | | | | | | | |
|----------|---|------|------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5 | 3.37 | 3.90 | 4.21 | 4.43 | 4.60 | 4.73 | 4.85 | 4.94 | 5.03 |
| 6 | 3.14 | 3.61 | 3.88 | 4.07 | 4.21 | 4.33 | 4.43 | 4.51 | 4.59 |
| 7 | 3.00 | 3.42 | 3.66 | 3.83 | 3.96 | 4.07 | 4.15 | 4.23 | 4.30 |
| 8 | 2.90 | 3.29 | 3.51 | 3.67 | 3.79 | 3.88 | 3.96 | 4.03 | 4.09 |
| 9 | 2.82 | 3.19 | 3.40 | 3.55 | 3.66 | 3.75 | 3.82 | 3.89 | 3.94 |
| 10 | 2.76 | 3.11 | 3.31 | 3.45 | 3.56 | 3.64 | 3.71 | 3.78 | 3.83 |
| 11 | 2.72 | 3.06 | 3.25 | 3.38 | 3.48 | 3.56 | 3.63 | 3.69 | 3.74 |
| 12 | 2.68 | 3.01 | 3.19 | 3.32 | 3.42 | 3.50 | 3.56 | 3.62 | 3.67 |
| 13 | 2.65 | 2.97 | 3.15 | 3.27 | 3.37 | 3.44 | 3.51 | 3.56 | 3.61 |
| 14 | 2.62 | 2.94 | 3.11 | 3.23 | 3.32 | 3.40 | 3.46 | 3.51 | 3.56 |
| 15 | 2.60 | 2.91 | 3.08 | 3.20 | 3.29 | 3.36 | 3.42 | 3.47 | 3.52 |
| 16 | 2.58 | 2.88 | 3.05 | 3.17 | 3.26 | 3.33 | 3.39 | 3.44 | 3.48 |
| 17 | 2.57 | 2.86 | 3.03 | 3.14 | 3.23 | 3.30 | 3.36 | 3.41 | 3.45 |
| 18 | 2.55 | 2.84 | 3.01 | 3.12 | 3.21 | 3.27 | 3.33 | 3.38 | 3.42 |
| 19 | 2.54 | 2.83 | 2.99 | 3.10 | 3.18 | 3.25 | 3.31 | 3.36 | 3.40 |
| 20 | 2.53 | 2.81 | 2.97 | 3.08 | 3.17 | 3.23 | 3.29 | 3.34 | 3.38 |
| 24 | 2.49 | 2.77 | 2.92 | 3.03 | 3.11 | 3.17 | 3.22 | 3.27 | 3.31 |
| 30 | 2.46 | 2.72 | 2.87 | 2.97 | 3.05 | 3.11 | 3.16 | 3.21 | 3.24 |
| 40 | 2.42 | 2.68 | 2.82 | 2.92 | 2.99 | 3.05 | 3.10 | 3.14 | 3.18 |
| 60 | 2.39 | 2.64 | 2.78 | 2.87 | 2.94 | 3.00 | 3.04 | 3.08 | 3.12 |
| 120 | 2.36 | 2.60 | 2.73 | 2.82 | 2.89 | 2.94 | 2.99 | 3.03 | 3.06 |
| ∞ | 2.33 | 2.56 | 2.68 | 2.77 | 2.84 | 2.89 | 2.93 | 2.97 | 3.00 |

Tabla A.15* Potencia de la prueba de análisis de varianza



*Reproducida de E. S. Pearson y H. O. Hartley, "Charts of the Power Function for Analysis of Variance Tests, Derived from the Non-central F Distribution", *Biometrika*, vol. 38, 1951, 112, con permiso del editor.

Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza

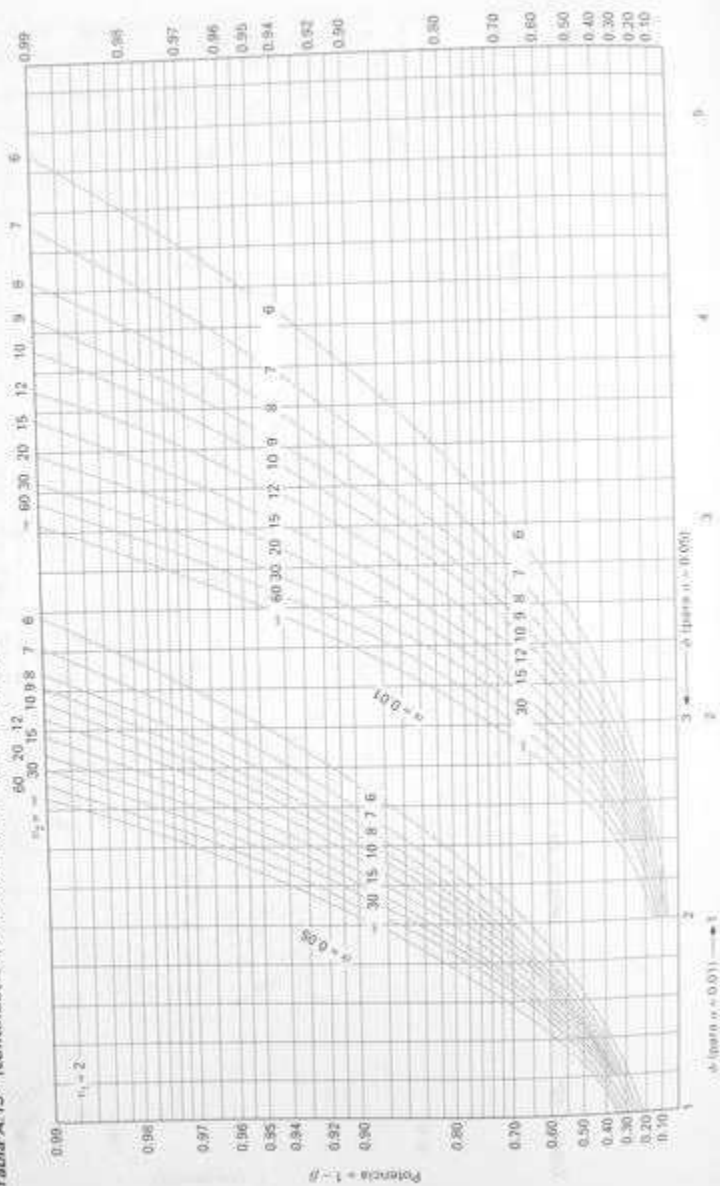


Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza

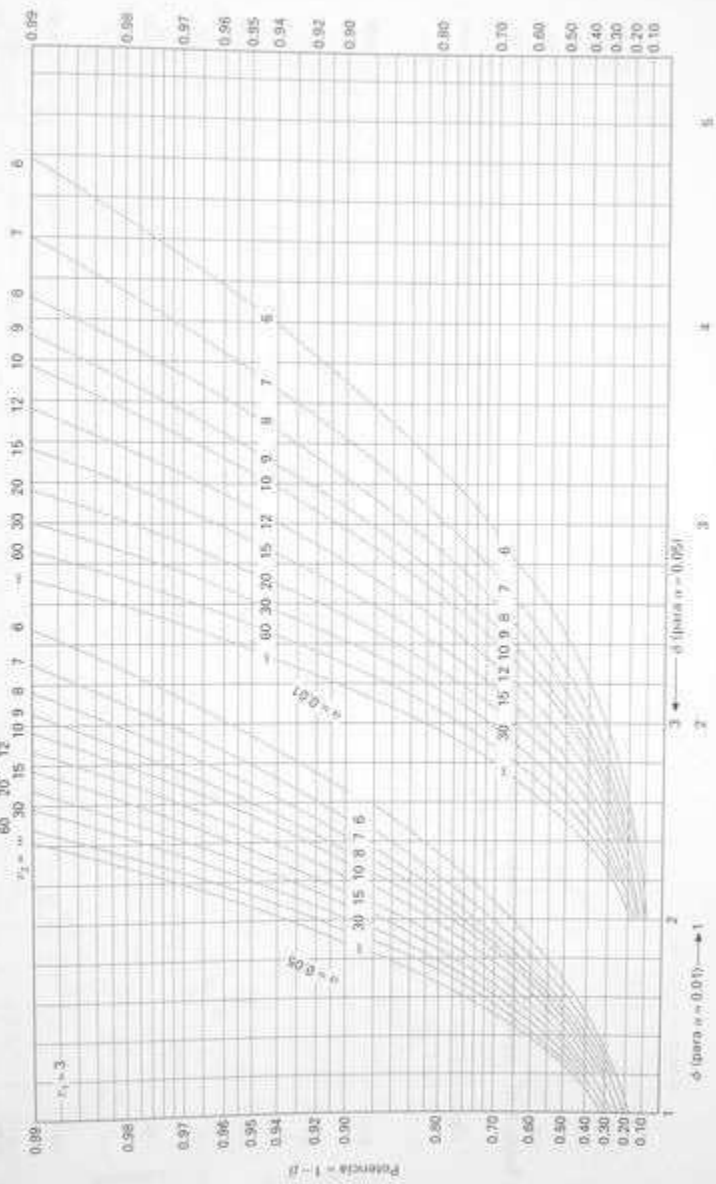


Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza

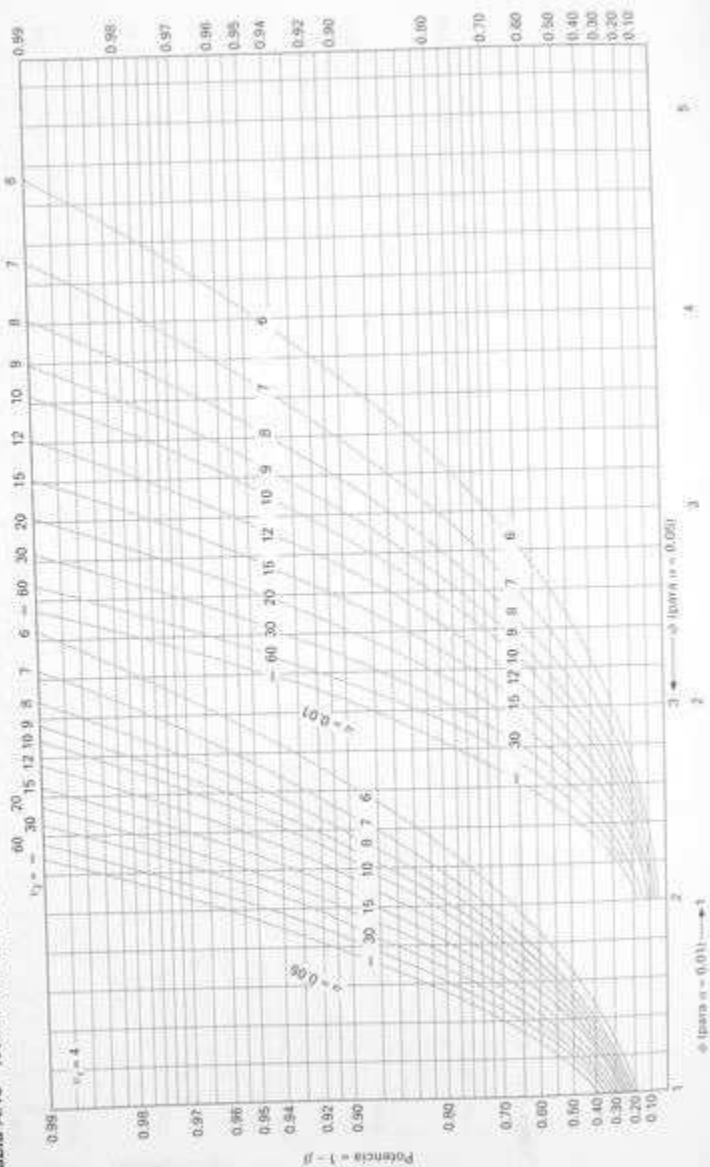
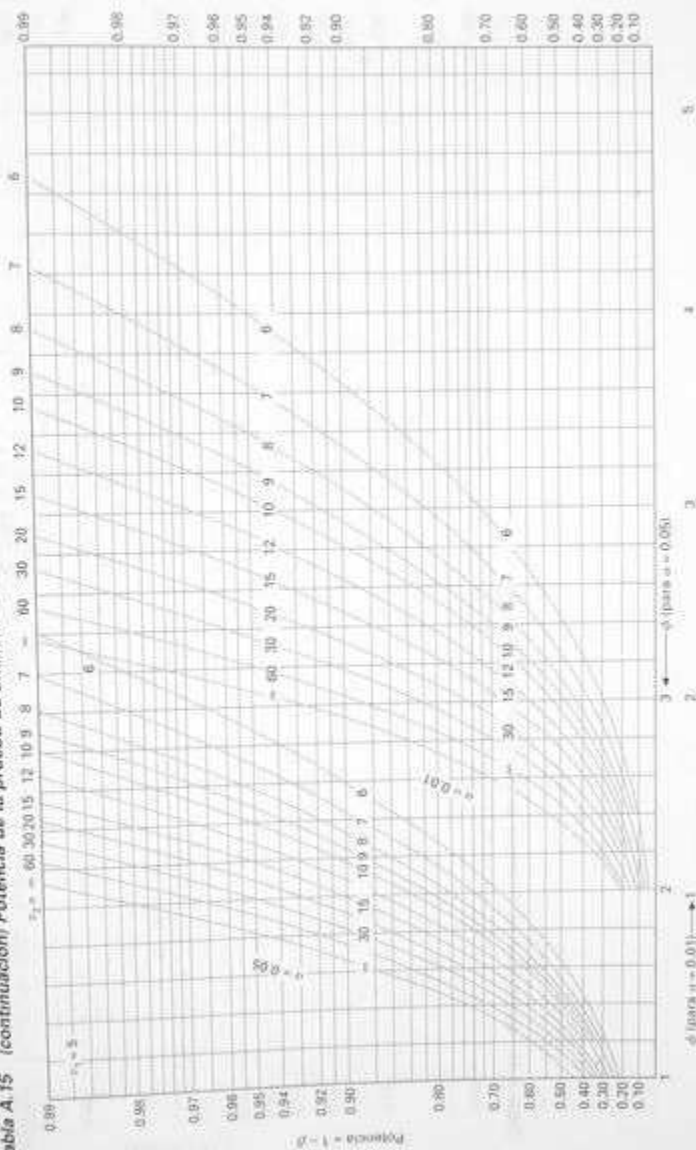


Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza



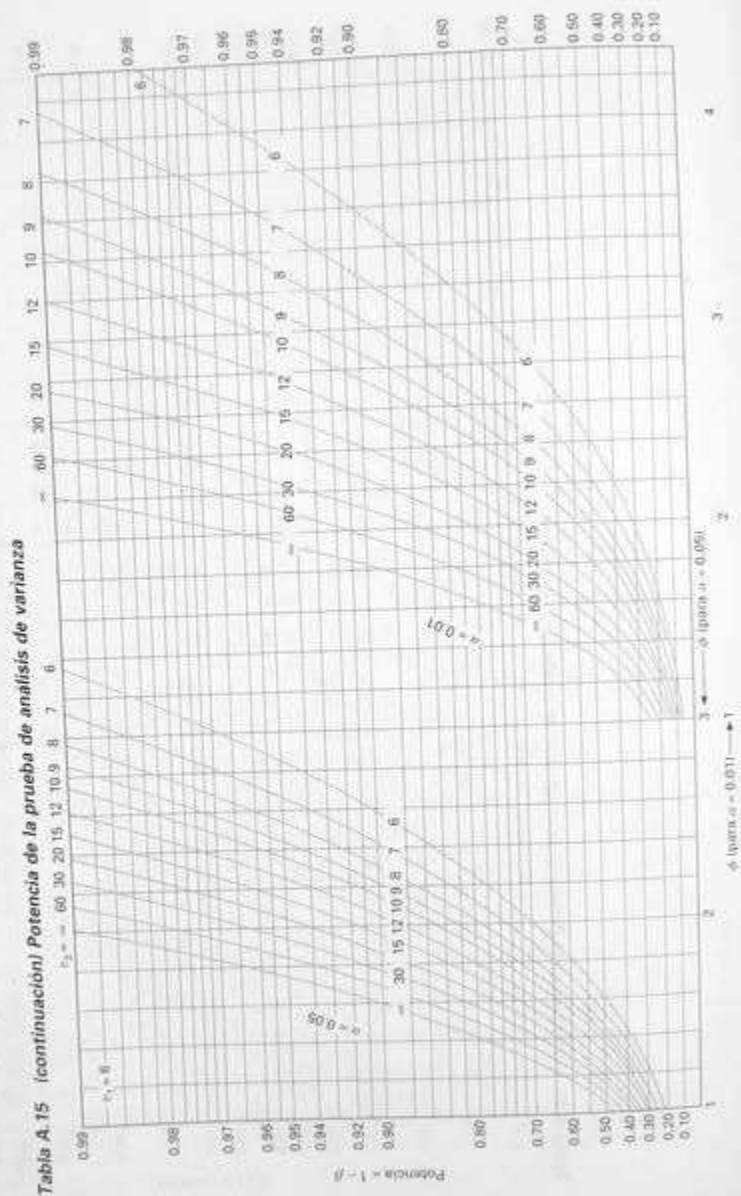


Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza

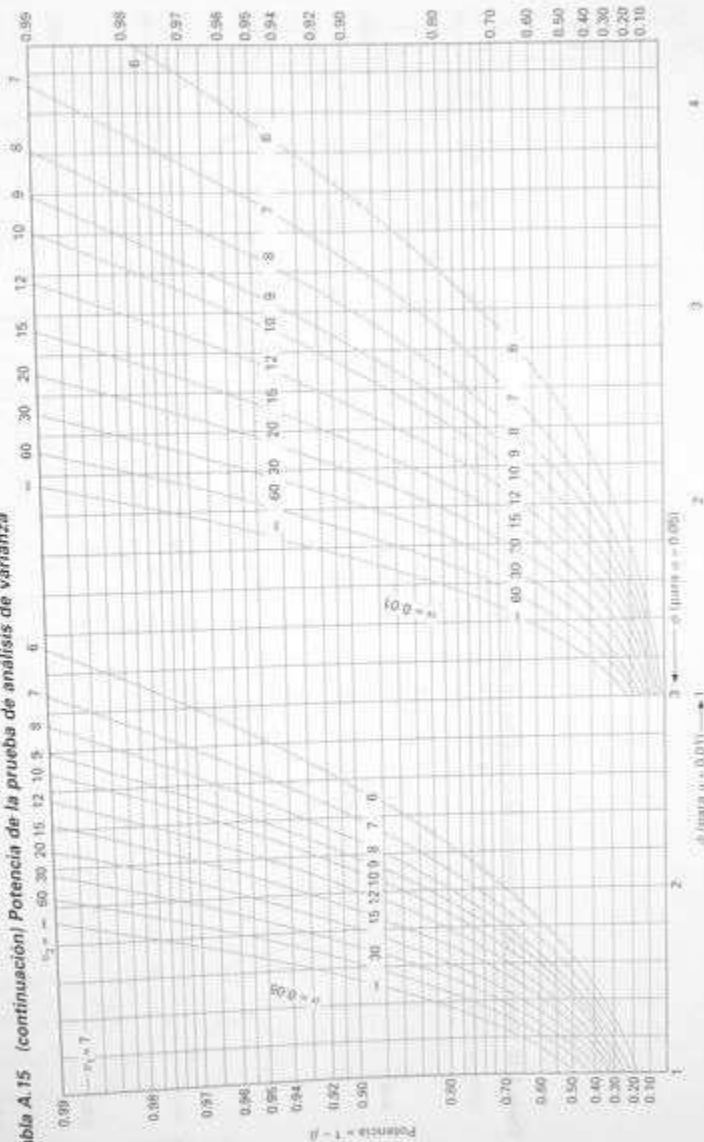


Tabla A.15 (continuación) Potencia de la prueba de análisis de varianza

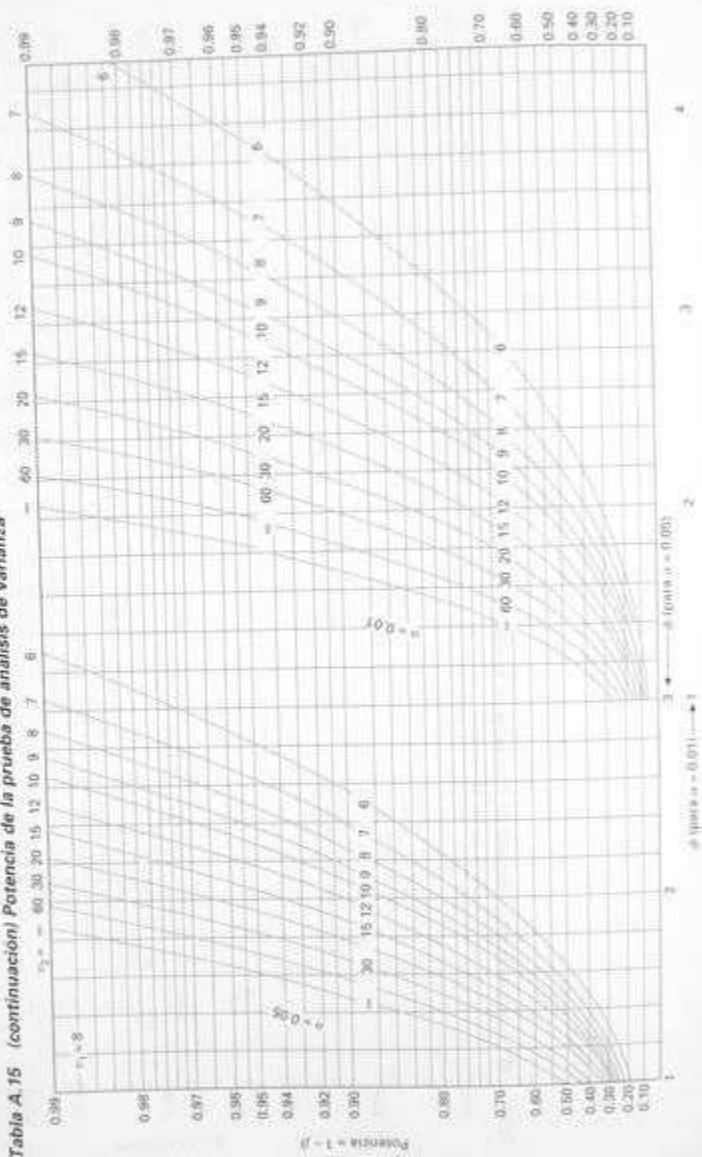


Tabla A.16* Valores críticos para la prueba de rangos con signo

| n | Unilateral $\alpha = 0.01$
Bilateral $\alpha = 0.02$ | Unilateral $\alpha = 0.025$
Bilateral $\alpha = 0.05$ | Unilateral $\alpha = 0.05$
Bilateral $\alpha = 0.10$ |
|-----|---|--|---|
| 5 | | | 1 |
| 6 | | 1 | 2 |
| 7 | 0 | 2 | 4 |
| 8 | 2 | 4 | 6 |
| 9 | 3 | 6 | 8 |
| 10 | 5 | 8 | 11 |
| 11 | 7 | 11 | 14 |
| 12 | 10 | 14 | 17 |
| 13 | 13 | 17 | 21 |
| 14 | 16 | 21 | 26 |
| 15 | 20 | 25 | 30 |
| 16 | 24 | 30 | 36 |
| 17 | 28 | 35 | 41 |
| 18 | 33 | 40 | 47 |
| 19 | 38 | 46 | 54 |
| 20 | 43 | 52 | 60 |
| 21 | 49 | 59 | 68 |
| 22 | 56 | 66 | 75 |
| 23 | 62 | 73 | 83 |
| 24 | 69 | 81 | 92 |
| 25 | 77 | 90 | 101 |
| 26 | 85 | 98 | 110 |
| 27 | 93 | 107 | 120 |
| 28 | 102 | 117 | 130 |
| 29 | 111 | 127 | 141 |
| 30 | 120 | 137 | 152 |

*Reproducida de F. Wilcoxon y R. A. Wilcox. *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, Pearl River, N. Y., 1964, con permiso de la American Cyanamid Company.

Tabla A.17* Valores críticos para la prueba de suma de rangos

Prueba de una cola con $\alpha = 0.001$ o prueba de dos colas con $\alpha = 0.002$

| n_1 | n_2 | | | | | | | | | | | | | | | | | | | |
|-------|-------|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|-----|-----|--|--|
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | | | | | |
| 1 | | | | | | | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | 0 | 0 | 0 | 0 | | | | | |
| 3 | | | | | | | | 1 | 1 | 1 | 2 | 2 | 3 | 3 | 3 | | | | | |
| 4 | | | | | 0 | 0 | 0 | 2 | 2 | 3 | 3 | 4 | 5 | 5 | 6 | 7 | 7 | | | |
| 5 | | 0 | 0 | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 11 | 12 | | | |
| 6 | 0 | 1 | 2 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 13 | 14 | 15 | 16 | | |
| 7 | | 2 | 3 | 3 | 5 | 6 | 6 | 7 | 8 | 9 | 10 | 11 | 15 | 17 | 18 | 20 | 21 | 21 | | |
| 8 | | | 5 | 5 | 6 | 8 | 9 | 11 | 12 | 14 | 15 | 17 | 19 | 21 | 23 | 25 | 26 | 26 | | |
| 9 | | | | 7 | 8 | 10 | 12 | 14 | 17 | 19 | 21 | 23 | 25 | 27 | 29 | 32 | 32 | 32 | | |
| 10 | | | | | 10 | 12 | 14 | 17 | 19 | 21 | 23 | 25 | 27 | 29 | 32 | 34 | 37 | 37 | | |
| 11 | | | | | | 15 | 17 | 20 | 22 | 24 | 27 | 29 | 32 | 34 | 37 | 40 | 42 | 42 | | |
| 12 | | | | | | | 20 | 23 | 25 | 28 | 31 | 34 | 37 | 40 | 45 | 48 | 48 | 48 | | |
| 13 | | | | | | | | 26 | 29 | 32 | 35 | 38 | 42 | 45 | 50 | 54 | 54 | 54 | | |
| 14 | | | | | | | | | 32 | 36 | 39 | 43 | 46 | 50 | 55 | 59 | 59 | 59 | | |
| 15 | | | | | | | | | | 40 | 43 | 47 | 51 | 55 | 60 | 65 | 65 | 65 | | |
| 16 | | | | | | | | | | | 48 | 52 | 56 | 60 | 66 | 70 | 70 | 70 | | |
| 17 | | | | | | | | | | | | 57 | 61 | 66 | 71 | 76 | 76 | 76 | | |
| 18 | | | | | | | | | | | | | 66 | 71 | 77 | 82 | 82 | 82 | | |
| 19 | | | | | | | | | | | | | | 77 | 82 | 88 | 88 | 88 | | |
| 20 | | | | | | | | | | | | | | | 88 | 94 | 100 | 100 | | |

Prueba de una cola con $\alpha = 0.01$ o prueba de dos colas con $\alpha = 0.02$

| n_1 | n_2 | | | | | | | | | | | | | | | | | | | |
|-------|-------|---|---|---|----|----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|
| | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | | | | |
| 1 | | | | | | | | | | | | | | | | | | | | |
| 2 | | | | | | | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 4 | | | | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 5 | 5 | 5 | 5 |
| 5 | 0 | 1 | 1 | 2 | 3 | 3 | 4 | 5 | 5 | 6 | 7 | 7 | 8 | 9 | 9 | 10 | 10 | 10 | 10 | 10 |
| 6 | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 15 | 16 | 18 | 19 | 20 | 22 | 22 |
| 7 | | | 3 | 4 | 6 | 7 | 8 | 9 | 11 | 12 | 13 | 15 | 16 | 18 | 19 | 21 | 23 | 24 | 26 | 28 |
| 8 | | | | 6 | 8 | 9 | 11 | 12 | 14 | 16 | 17 | 19 | 21 | 23 | 24 | 26 | 28 | 30 | 32 | 34 |
| 9 | | | | | 10 | 11 | 13 | 15 | 17 | 20 | 22 | 24 | 26 | 28 | 30 | 33 | 36 | 38 | 40 | 40 |
| 10 | | | | | | 14 | 16 | 18 | 21 | 23 | 26 | 28 | 31 | 33 | 36 | 38 | 41 | 44 | 47 | 47 |
| 11 | | | | | | | 19 | 22 | 24 | 27 | 30 | 33 | 36 | 38 | 41 | 44 | 47 | 50 | 53 | 53 |
| 12 | | | | | | | | 25 | 28 | 31 | 34 | 37 | 41 | 44 | 47 | 50 | 53 | 56 | 60 | 60 |
| 13 | | | | | | | | | 31 | 35 | 38 | 42 | 46 | 49 | 53 | 56 | 60 | 63 | 67 | 67 |
| 14 | | | | | | | | | | 39 | 43 | 47 | 51 | 55 | 59 | 63 | 67 | 73 | 73 | 73 |
| 15 | | | | | | | | | | | 47 | 51 | 56 | 60 | 65 | 69 | 75 | 80 | 80 | 80 |
| 16 | | | | | | | | | | | | 56 | 61 | 66 | 70 | 75 | 80 | 85 | 87 | 87 |
| 17 | | | | | | | | | | | | | 66 | 71 | 76 | 82 | 88 | 93 | 93 | 93 |
| 18 | | | | | | | | | | | | | | 77 | 82 | 88 | 94 | 100 | 100 | 100 |
| 19 | | | | | | | | | | | | | | | 88 | 94 | 101 | 107 | 107 | 107 |
| 20 | | | | | | | | | | | | | | | | 107 | 114 | 114 | 114 | 114 |

*Basada en parte de las tablas 1, 3, 5 y 7 de D. Aulbe, "Extended tables for the Mann-Whitney statistic", *Bulletin of the Institute of Educational Research at Indiana University*, vol. 1, Núm. 2, 1953, con permiso del director.

Tabla A.18* $P\{V \leq v\}$ cuando H_0 es verdadera en la prueba de corridas

| (n_1, n_2) | v^* | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| (2, 3) | 0.200 | 0.500 | 0.900 | 1.000 | | | | | | |
| (2, 4) | 0.133 | 0.400 | 0.800 | 1.000 | | | | | | |
| (2, 5) | 0.095 | 0.333 | 0.714 | 1.000 | | | | | | |
| (2, 6) | 0.071 | 0.286 | 0.643 | 1.000 | | | | | | |
| (2, 7) | 0.056 | 0.250 | 0.583 | 1.000 | | | | | | |
| (2, 8) | 0.044 | 0.222 | 0.533 | 1.000 | | | | | | |
| (2, 9) | 0.036 | 0.200 | 0.491 | 1.000 | | | | | | |
| (2, 10) | 0.030 | 0.182 | 0.455 | 1.000 | | | | | | |
| (3, 3) | 0.100 | 0.300 | 0.700 | 0.900 | 1.000 | | | | | |
| (3, 4) | 0.057 | 0.200 | 0.543 | 0.800 | 0.971 | 1.000 | | | | |
| (3, 5) | 0.036 | 0.143 | 0.429 | 0.714 | 0.929 | 1.000 | | | | |
| (3, 6) | 0.024 | 0.107 | 0.345 | 0.643 | 0.881 | 1.000 | | | | |
| (3, 7) | 0.017 | 0.083 | 0.283 | 0.583 | 0.835 | 1.000 | | | | |
| (3, 8) | 0.012 | 0.067 | 0.236 | 0.533 | 0.788 | 1.000 | | | | |
| (3, 9) | 0.009 | 0.055 | 0.200 | 0.491 | 0.745 | 1.000 | | | | |
| (3, 10) | 0.007 | 0.045 | 0.171 | 0.455 | 0.706 | 1.000 | | | | |
| (4, 4) | 0.029 | 0.114 | 0.371 | 0.629 | 0.886 | 0.971 | 1.000 | | | |
| (4, 5) | 0.016 | 0.071 | 0.262 | 0.500 | 0.786 | 0.929 | 0.992 | 1.000 | | |
| (4, 6) | 0.010 | 0.048 | 0.190 | 0.405 | 0.690 | 0.881 | 0.976 | 1.000 | | |
| (4, 7) | 0.006 | 0.033 | 0.142 | 0.333 | 0.606 | 0.833 | 0.954 | 1.000 | | |
| (4, 8) | 0.004 | 0.024 | 0.109 | 0.279 | 0.533 | 0.788 | 0.929 | 1.000 | | |
| (4, 9) | 0.003 | 0.018 | 0.085 | 0.236 | 0.471 | 0.745 | 0.902 | 1.000 | | |
| (4, 10) | 0.002 | 0.014 | 0.068 | 0.203 | 0.419 | 0.706 | 0.874 | 1.000 | | |
| (5, 5) | 0.008 | 0.040 | 0.167 | 0.357 | 0.643 | 0.833 | 0.960 | 0.992 | 1.000 | |
| (5, 6) | 0.004 | 0.024 | 0.110 | 0.262 | 0.522 | 0.738 | 0.911 | 0.976 | 0.998 | |
| (5, 7) | 0.003 | 0.015 | 0.076 | 0.197 | 0.424 | 0.632 | 0.854 | 0.955 | 0.992 | |
| (5, 8) | 0.002 | 0.010 | 0.054 | 0.152 | 0.347 | 0.576 | 0.793 | 0.929 | 0.984 | |
| (5, 9) | 0.001 | 0.007 | 0.039 | 0.119 | 0.287 | 0.510 | 0.734 | 0.902 | 0.972 | |
| (5, 10) | 0.001 | 0.005 | 0.029 | 0.095 | 0.239 | 0.455 | 0.678 | 0.874 | 0.958 | |
| (6, 6) | 0.002 | 0.013 | 0.067 | 0.175 | 0.392 | 0.608 | 0.825 | 0.933 | 0.987 | |
| (6, 7) | 0.001 | 0.008 | 0.043 | 0.121 | 0.296 | 0.500 | 0.733 | 0.879 | 0.966 | |
| (6, 8) | 0.001 | 0.005 | 0.028 | 0.086 | 0.226 | 0.413 | 0.646 | 0.821 | 0.937 | |
| (6, 9) | 0.000 | 0.003 | 0.019 | 0.063 | 0.175 | 0.343 | 0.566 | 0.762 | 0.902 | |
| (6, 10) | 0.000 | 0.002 | 0.013 | 0.047 | 0.137 | 0.288 | 0.497 | 0.706 | 0.864 | |
| (7, 7) | 0.001 | 0.004 | 0.025 | 0.078 | 0.209 | 0.383 | 0.617 | 0.791 | 0.922 | |
| (7, 8) | 0.000 | 0.002 | 0.015 | 0.051 | 0.149 | 0.296 | 0.514 | 0.704 | 0.867 | |
| (7, 9) | 0.000 | 0.001 | 0.010 | 0.035 | 0.108 | 0.231 | 0.427 | 0.622 | 0.806 | |
| (7, 10) | 0.000 | 0.001 | 0.006 | 0.024 | 0.080 | 0.182 | 0.355 | 0.549 | 0.743 | |
| (8, 8) | 0.000 | 0.001 | 0.009 | 0.032 | 0.100 | 0.214 | 0.405 | 0.595 | 0.786 | |
| (8, 9) | 0.000 | 0.001 | 0.005 | 0.020 | 0.069 | 0.157 | 0.319 | 0.500 | 0.702 | |
| (8, 10) | 0.000 | 0.000 | 0.003 | 0.013 | 0.048 | 0.117 | 0.251 | 0.419 | 0.621 | |
| (9, 9) | 0.000 | 0.000 | 0.003 | 0.012 | 0.044 | 0.109 | 0.238 | 0.399 | 0.603 | |
| (9, 10) | 0.000 | 0.000 | 0.002 | 0.008 | 0.029 | 0.077 | 0.179 | 0.319 | 0.510 | |
| (10, 10) | 0.000 | 0.000 | 0.001 | 0.004 | 0.019 | 0.051 | 0.128 | 0.242 | 0.414 | |

*Reproducida de C. Eisenhart y F. Sved, "Tables for testing Randomness of Grouping in a Sequence of Alternatives", *Ann. Math. Stat.*, vol. 14, 1943, con permiso del editor.

Tabla A.18 (continuación) $P(V \leq v^*)$, cuando H_0 es verdadera) en la prueba de corridas

| (n_1, n_2) | v^* | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| (2, 3) | | | | | | | | | | |
| (2, 4) | | | | | | | | | | |
| (2, 5) | | | | | | | | | | |
| (2, 6) | | | | | | | | | | |
| (2, 7) | | | | | | | | | | |
| (2, 8) | | | | | | | | | | |
| (2, 9) | | | | | | | | | | |
| (2, 10) | | | | | | | | | | |
| (3, 3) | | | | | | | | | | |
| (3, 4) | | | | | | | | | | |
| (3, 5) | | | | | | | | | | |
| (3, 6) | | | | | | | | | | |
| (3, 7) | | | | | | | | | | |
| (3, 8) | | | | | | | | | | |
| (3, 9) | | | | | | | | | | |
| (3, 10) | | | | | | | | | | |
| (4, 4) | | | | | | | | | | |
| (4, 5) | | | | | | | | | | |
| (4, 6) | | | | | | | | | | |
| (4, 7) | | | | | | | | | | |
| (4, 8) | | | | | | | | | | |
| (4, 9) | | | | | | | | | | |
| (4, 10) | | | | | | | | | | |
| (5, 5) | | | | | | | | | | |
| (5, 6) | 1.000 | | | | | | | | | |
| (5, 7) | 1.000 | | | | | | | | | |
| (5, 8) | 1.000 | | | | | | | | | |
| (5, 9) | 1.000 | | | | | | | | | |
| (5, 10) | 1.000 | | | | | | | | | |
| (6, 6) | 0.998 | 1.000 | | | | | | | | |
| (6, 7) | 0.992 | 0.999 | 1.000 | | | | | | | |
| (6, 8) | 0.984 | 0.998 | 1.000 | | | | | | | |
| (6, 9) | 0.972 | 0.994 | 1.000 | | | | | | | |
| (6, 10) | 0.958 | 0.990 | 1.000 | | | | | | | |
| (7, 7) | 0.975 | 0.996 | 0.999 | 1.000 | | | | | | |
| (7, 8) | 0.949 | 0.988 | 0.998 | 1.000 | 1.000 | | | | | |
| (7, 9) | 0.916 | 0.975 | 0.994 | 0.999 | 1.000 | | | | | |
| (7, 10) | 0.879 | 0.957 | 0.990 | 0.998 | 1.000 | | | | | |
| (8, 8) | 0.900 | 0.968 | 0.991 | 0.999 | 1.000 | 1.000 | | | | |
| (8, 9) | 0.843 | 0.939 | 0.980 | 0.996 | 0.999 | 1.000 | 1.000 | | | |
| (8, 10) | 0.782 | 0.903 | 0.964 | 0.990 | 0.998 | 1.000 | 1.000 | | | |
| (9, 9) | 0.762 | 0.891 | 0.956 | 0.985 | 0.997 | 1.000 | 1.000 | 1.000 | | |
| (9, 10) | 0.681 | 0.834 | 0.923 | 0.974 | 0.992 | 0.999 | 1.000 | 1.000 | 1.000 | |
| (10, 10) | 0.586 | 0.758 | 0.872 | 0.949 | 0.981 | 0.996 | 0.999 | 1.000 | 1.000 | 1.000 |

Tabla A.19* Tamaño muestral para límites de tolerancia no paramétricos bilaterales

| $1 - \alpha$ | $1 - \gamma$ | | | | | |
|--------------|--------------|------|------|------|-------|-------|
| | 0.50 | 0.70 | 0.90 | 0.95 | 0.99 | 0.995 |
| 0.995 | 336 | 488 | 777 | 947 | 1,325 | 1,483 |
| 0.99 | 168 | 244 | 388 | 473 | 662 | 740 |
| 0.95 | 34 | 49 | 77 | 93 | 130 | 146 |
| 0.90 | 17 | 24 | 38 | 46 | 64 | 72 |
| 0.85 | 11 | 16 | 25 | 30 | 42 | 47 |
| 0.80 | 9 | 12 | 18 | 22 | 31 | 34 |
| 0.75 | 7 | 10 | 15 | 18 | 24 | 27 |
| 0.70 | 6 | 8 | 12 | 14 | 20 | 22 |
| 0.60 | 4 | 6 | 9 | 10 | 14 | 16 |
| 0.50 | 3 | 5 | 7 | 8 | 11 | 12 |

*Reproducida de las tablas A-25d de Wilfrid J. Dixon y Frank J. Massey, Jr., *Introduction to Statistical Analysis*, 3a. ed., McGraw-Hill Book Company, Nueva York, 1969. Utilizada con permiso de McGraw-Hill Book Company.

Tabla A.20* Tamaño muestral para límites de tolerancia no paramétricos unilaterales

| $1 - \alpha$ | $1 - \gamma$ | | | | |
|--------------|--------------|------|------|------|-------|
| | 0.50 | 0.70 | 0.95 | 0.99 | 0.995 |
| 0.995 | 139 | 241 | 398 | 919 | 1,379 |
| 0.99 | 69 | 120 | 299 | 459 | 688 |
| 0.95 | 14 | 24 | 59 | 90 | 135 |
| 0.90 | 7 | 12 | 29 | 44 | 66 |
| 0.85 | 5 | 8 | 19 | 29 | 43 |
| 0.80 | 4 | 6 | 14 | 21 | 31 |
| 0.75 | 3 | 5 | 11 | 7 | 25 |
| 0.70 | 2 | 4 | 9 | 13 | 20 |
| 0.60 | 2 | 3 | 6 | 10 | 14 |
| 0.50 | 1 | 2 | 5 | 7 | 10 |

*Reproducida de las tablas A-25e de Wilfrid J. Dixon y Frank J. Massey, Jr., *Introduction to Statistical Analysis*, 3a. ed., McGraw-Hill Book Company, Nueva York, 1969. Utilizada con permiso de McGraw-Hill Book Company.

Tabla A.21* Valores críticos del coeficiente de correlación de rangos de Spearman

| n | $\alpha = 0.05$ | $\alpha = 0.025$ | $\alpha = 0.01$ | $\alpha = 0.005$ |
|-----|-----------------|------------------|-----------------|------------------|
| 5 | 0.900 | - | - | - |
| 6 | 0.829 | 0.886 | 0.943 | - |
| 7 | 0.714 | 0.786 | 0.893 | - |
| 8 | 0.643 | 0.738 | 0.833 | 0.881 |
| 9 | 0.600 | 0.683 | 0.783 | 0.833 |
| 10 | 0.564 | 0.648 | 0.745 | 0.794 |
| 11 | 0.523 | 0.623 | 0.736 | 0.818 |
| 12 | 0.497 | 0.591 | 0.703 | 0.780 |
| 13 | 0.475 | 0.566 | 0.673 | 0.745 |
| 14 | 0.457 | 0.545 | 0.646 | 0.716 |
| 15 | 0.441 | 0.525 | 0.623 | 0.689 |
| 16 | 0.425 | 0.507 | 0.601 | 0.666 |
| 17 | 0.412 | 0.490 | 0.582 | 0.645 |
| 18 | 0.399 | 0.476 | 0.564 | 0.625 |
| 19 | 0.388 | 0.462 | 0.549 | 0.608 |
| 20 | 0.377 | 0.450 | 0.534 | 0.591 |
| 21 | 0.368 | 0.438 | 0.521 | 0.576 |
| 22 | 0.359 | 0.428 | 0.508 | 0.562 |
| 23 | 0.351 | 0.418 | 0.496 | 0.549 |
| 24 | 0.343 | 0.409 | 0.485 | 0.537 |
| 25 | 0.336 | 0.400 | 0.475 | 0.526 |
| 26 | 0.329 | 0.392 | 0.465 | 0.515 |
| 27 | 0.323 | 0.385 | 0.456 | 0.505 |
| 28 | 0.317 | 0.377 | 0.448 | 0.496 |
| 29 | 0.311 | 0.370 | 0.440 | 0.487 |
| 30 | 0.305 | 0.364 | 0.432 | 0.478 |

*Reproducida de E. G. Olds, "Distribution of Sums of Squares of Rank Differences for Small Samples". *Ann. Math. Stat.*, vol. 9, 1938, con permiso del editor.

Tabla A.22 Puntos porcentuales superiores de la distribución de rango studentizado: valores de $q(0.05; k, v)$

| Grados de libertad, v | Número de tratamientos k | | | | | | | | |
|-------------------------|----------------------------|------|------|-------|-------|-------|-------|-------|-------|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 18.0 | 27.0 | 32.8 | 37.2 | 40.5 | 43.1 | 45.4 | 47.3 | 49.1 |
| 2 | 6.09 | 8.33 | 9.80 | 10.89 | 11.73 | 12.43 | 13.03 | 13.54 | 13.99 |
| 3 | 4.50 | 5.91 | 6.83 | 7.51 | 8.04 | 8.47 | 8.85 | 9.18 | 9.46 |
| 4 | 3.93 | 5.04 | 5.76 | 6.29 | 6.71 | 7.06 | 7.35 | 7.60 | 7.83 |
| 5 | 3.64 | 4.60 | 5.22 | 5.67 | 6.03 | 6.33 | 6.58 | 6.80 | 6.99 |
| 6 | 3.46 | 4.34 | 4.90 | 5.31 | 5.63 | 5.89 | 6.12 | 6.32 | 6.49 |
| 7 | 3.34 | 4.16 | 4.68 | 5.06 | 5.35 | 5.59 | 5.80 | 5.99 | 6.15 |
| 8 | 3.26 | 4.04 | 4.53 | 4.89 | 5.17 | 5.40 | 5.60 | 5.77 | 5.92 |
| 9 | 3.20 | 3.95 | 4.42 | 4.76 | 5.02 | 5.24 | 5.43 | 5.60 | 5.74 |
| 10 | 3.15 | 3.88 | 4.33 | 4.66 | 4.91 | 5.12 | 5.30 | 5.46 | 5.60 |
| 11 | 3.11 | 3.82 | 4.26 | 4.58 | 4.82 | 5.03 | 5.20 | 5.35 | 5.49 |
| 12 | 3.08 | 3.77 | 4.20 | 4.51 | 4.75 | 4.95 | 5.12 | 5.27 | 5.40 |
| 13 | 3.06 | 3.73 | 4.15 | 4.46 | 4.69 | 4.88 | 5.05 | 5.19 | 5.32 |
| 14 | 3.03 | 3.70 | 4.11 | 4.41 | 4.64 | 4.83 | 4.99 | 5.13 | 5.25 |
| 15 | 3.01 | 3.67 | 4.08 | 4.37 | 4.59 | 4.78 | 4.94 | 5.08 | 5.20 |
| 16 | 3.00 | 3.65 | 4.05 | 4.34 | 4.56 | 4.74 | 4.90 | 5.03 | 5.15 |
| 17 | 2.98 | 3.62 | 4.02 | 4.31 | 4.52 | 4.70 | 4.86 | 4.99 | 5.11 |
| 18 | 2.97 | 3.61 | 4.00 | 4.28 | 4.49 | 4.67 | 4.83 | 4.96 | 5.07 |
| 19 | 2.96 | 3.59 | 3.98 | 4.26 | 4.47 | 4.64 | 4.79 | 4.92 | 5.04 |
| 20 | 2.95 | 3.58 | 3.96 | 4.24 | 4.45 | 4.62 | 4.77 | 4.90 | 5.01 |
| 24 | 2.92 | 3.53 | 3.90 | 4.17 | 4.37 | 4.54 | 4.68 | 4.81 | 4.92 |
| 30 | 2.89 | 3.48 | 3.84 | 4.11 | 4.30 | 4.46 | 4.60 | 4.72 | 4.83 |
| 40 | 2.86 | 3.44 | 3.79 | 4.04 | 4.23 | 4.39 | 4.52 | 4.63 | 4.74 |
| 60 | 2.83 | 3.40 | 3.74 | 3.98 | 4.16 | 4.31 | 4.44 | 4.55 | 4.65 |
| 120 | 2.80 | 3.36 | 3.69 | 3.92 | 4.10 | 4.24 | 4.36 | 4.47 | 4.56 |
| ∞ | 2.77 | 3.32 | 3.63 | 3.86 | 4.03 | 4.17 | 4.29 | 4.39 | 4.47 |

Table A.23 Factores para la elaboración de gráficas de control

| Observaciones
en la
muestra, <i>n</i> | Gráfica para promedios | | | Gráfica para las desviaciones estándar | | | | | | Gráfica para rangos | | | | | |
|---|---|-------|--------|--|-------|-------|---|-------|-------|-----------------------------------|-------|-------|---|-------|-------|
| | Factores para los
límites de control | | | Factores para la
línea central | | | Factores para los
límites de control | | | Factores para la
línea central | | | Factores para los
límites de control | | |
| | A_2 | A_3 | c_4 | Ue_4 | B_3 | B_4 | B_5 | B_6 | d_5 | Id_5 | d_6 | D_3 | D_4 | D_5 | D_6 |
| 2 | 1.880 | 2.659 | 0.7979 | 1.2533 | 0 | 3.267 | 0 | 2.696 | 1.128 | 0.8865 | 0.853 | 0 | 3.267 | 0 | 3.267 |
| 3 | 1.023 | 1.954 | 0.8862 | 1.1284 | 0 | 2.568 | 0 | 2.276 | 1.693 | 0.5907 | 0.888 | 0 | 2.574 | 0 | 2.574 |
| 4 | 0.729 | 1.628 | 0.9213 | 1.0854 | 0 | 2.266 | 0 | 2.088 | 2.059 | 0.4857 | 0.880 | 0 | 2.282 | 0 | 2.282 |
| 5 | 0.577 | 1.427 | 0.9400 | 1.0638 | 0 | 2.089 | 0 | 1.964 | 2.326 | 0.4299 | 0.864 | 0 | 2.114 | 0 | 2.114 |
| 6 | 0.483 | 1.287 | 0.9515 | 1.0510 | 0.030 | 1.970 | 0.029 | 1.874 | 2.534 | 0.3946 | 0.848 | 0 | 2.084 | 0 | 2.084 |
| 7 | 0.419 | 1.182 | 0.9594 | 1.04230 | 0.118 | 1.882 | 0.113 | 1.806 | 2.704 | 0.3698 | 0.833 | 0.076 | 1.924 | 0.076 | 1.924 |
| 8 | 0.373 | 1.099 | 0.9650 | 1.0363 | 0.185 | 1.815 | 0.179 | 1.751 | 2.847 | 0.3512 | 0.820 | 0.136 | 1.864 | 0.136 | 1.864 |
| 9 | 0.337 | 1.032 | 0.9693 | 1.0317 | 0.239 | 1.761 | 0.232 | 1.707 | 2.970 | 0.3367 | 0.808 | 0.184 | 1.816 | 0.184 | 1.816 |
| 10 | 0.308 | 0.975 | 0.9727 | 1.0281 | 0.284 | 1.716 | 0.276 | 1.669 | 3.078 | 0.3249 | 0.797 | 0.223 | 1.777 | 0.223 | 1.777 |
| 11 | 0.285 | 0.927 | 0.9754 | 1.0252 | 0.321 | 1.679 | 0.313 | 1.637 | 3.173 | 0.3152 | 0.787 | 0.256 | 1.744 | 0.256 | 1.744 |
| 12 | 0.266 | 0.886 | 0.9776 | 1.0229 | 0.358 | 1.646 | 0.346 | 1.610 | 3.258 | 0.3069 | 0.778 | 0.283 | 1.717 | 0.283 | 1.717 |
| 13 | 0.249 | 0.850 | 0.9794 | 1.0210 | 0.382 | 1.618 | 0.374 | 1.585 | 3.336 | 0.2998 | 0.770 | 0.307 | 1.693 | 0.307 | 1.693 |
| 14 | 0.235 | 0.817 | 0.9810 | 1.0194 | 0.406 | 1.594 | 0.399 | 1.563 | 3.407 | 0.2935 | 0.763 | 0.328 | 1.672 | 0.328 | 1.672 |
| 15 | 0.223 | 0.789 | 0.9823 | 1.0180 | 0.428 | 1.572 | 0.421 | 1.544 | 3.472 | 0.2880 | 0.756 | 0.347 | 1.653 | 0.347 | 1.653 |
| 16 | 0.212 | 0.763 | 0.9835 | 1.0168 | 0.448 | 1.552 | 0.440 | 1.526 | 3.532 | 0.2831 | 0.750 | 0.363 | 1.637 | 0.363 | 1.637 |
| 17 | 0.203 | 0.739 | 0.9845 | 1.0157 | 0.466 | 1.534 | 0.458 | 1.513 | 3.588 | 0.2787 | 0.744 | 0.378 | 1.622 | 0.378 | 1.622 |
| 18 | 0.194 | 0.718 | 0.9854 | 1.0148 | 0.482 | 1.518 | 0.475 | 1.496 | 3.640 | 0.2747 | 0.739 | 0.391 | 1.608 | 0.391 | 1.608 |
| 19 | 0.187 | 0.698 | 0.9862 | 1.0140 | 0.497 | 1.503 | 0.490 | 1.483 | 3.689 | 0.2711 | 0.734 | 0.403 | 1.597 | 0.403 | 1.597 |
| 20 | 0.180 | 0.680 | 0.9869 | 1.0133 | 0.510 | 1.490 | 0.504 | 1.470 | 3.735 | 0.2677 | 0.729 | 0.415 | 1.588 | 0.415 | 1.588 |
| 21 | 0.173 | 0.663 | 0.9876 | 1.0126 | 0.523 | 1.477 | 0.516 | 1.459 | 3.778 | 0.2647 | 0.724 | 0.425 | 1.575 | 0.425 | 1.575 |
| 22 | 0.167 | 0.647 | 0.9882 | 1.0119 | 0.534 | 1.466 | 0.528 | 1.448 | 3.810 | 0.2618 | 0.720 | 0.434 | 1.566 | 0.434 | 1.566 |
| 23 | 0.162 | 0.633 | 0.9887 | 1.0114 | 0.545 | 1.455 | 0.539 | 1.438 | 3.838 | 0.2592 | 0.716 | 0.443 | 1.557 | 0.443 | 1.557 |
| 24 | 0.157 | 0.619 | 0.9892 | 1.0109 | 0.555 | 1.445 | 0.549 | 1.429 | 3.895 | 0.2567 | 0.712 | 0.451 | 1.548 | 0.451 | 1.548 |
| 25 | 0.153 | 0.606 | 0.9896 | 1.0105 | 0.565 | 1.435 | 0.559 | 1.420 | 3.931 | 0.2544 | 0.708 | 0.459 | 1.541 | 0.459 | 1.541 |

Tabla A.24 Tabla de la función gamma incompleta: $F(x; \alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} dy$

| $x \backslash \alpha$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 0.632 | 0.264 | 0.080 | 0.019 | 0.004 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.865 | 0.594 | 0.323 | 0.143 | 0.053 | 0.017 | 0.005 | 0.001 | 0.000 | 0.000 |
| 3 | 0.950 | 0.801 | 0.577 | 0.353 | 0.185 | 0.084 | 0.034 | 0.012 | 0.004 | 0.001 |
| 4 | 0.982 | 0.908 | 0.762 | 0.567 | 0.371 | 0.215 | 0.111 | 0.051 | 0.021 | 0.008 |
| 5 | 0.993 | 0.960 | 0.875 | 0.735 | 0.560 | 0.384 | 0.238 | 0.133 | 0.068 | 0.032 |
| 6 | 0.998 | 0.983 | 0.938 | 0.849 | 0.715 | 0.554 | 0.394 | 0.256 | 0.153 | 0.084 |
| 7 | 0.999 | 0.993 | 0.970 | 0.918 | 0.827 | 0.699 | 0.550 | 0.401 | 0.271 | 0.170 |
| 8 | 1.000 | 0.997 | 0.986 | 0.958 | 0.900 | 0.809 | 0.687 | 0.547 | 0.407 | 0.283 |
| 9 | | 0.999 | 0.994 | 0.979 | 0.945 | 0.884 | 0.793 | 0.676 | 0.544 | 0.413 |
| 10 | | 1.000 | 0.997 | 0.990 | 0.971 | 0.933 | 0.870 | 0.780 | 0.667 | 0.542 |
| 11 | | | 0.999 | 0.995 | 0.985 | 0.962 | 0.921 | 0.857 | 0.768 | 0.659 |
| 12 | | | 1.000 | 0.998 | 0.992 | 0.980 | 0.954 | 0.911 | 0.845 | 0.758 |
| 13 | | | | 0.999 | 0.996 | 0.989 | 0.974 | 0.946 | 0.900 | 0.834 |
| 14 | | | | 1.000 | 0.998 | 0.994 | 0.986 | 0.968 | 0.938 | 0.891 |
| 15 | | | | | 0.999 | 0.997 | 0.992 | 0.982 | 0.963 | 0.930 |

Respuestas a los ejercicios con número impar

Capítulo 1

PÁGINA 9

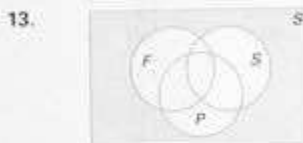
- (a) Media = 124.3, mediana = 120;
(b) 175 es una observación extrema.
- No fumadores. Fumadores
(a) 30.32 43.70
(b) 7.13 16.93
- (d) Los fumadores parecen tomar más tiempo para quedar dormidos. Para los fumadores el tiempo para quedar dormido es más variable.
- Si, valor $P = 0.03125$; probabilidad de obtener $HHHHH$ con una moneda legal.

Capítulo 2

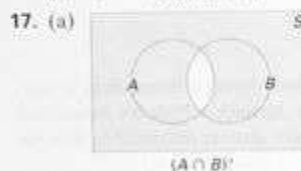
PÁGINA 17

- (a) $S = \{8, 16, 24, 32, 40, 48\}$;
(b) $S = \{-5, 1\}$;
(c) $S = \{T, HT, HHT, HHH\}$;
(d) $S = \{\text{Norte América, Sudamérica, Europa, Asia, África, Australia, Antártida}\}$;
(e) $S = \emptyset$.
- $A = C$
- $S = \{1HH, 1HT, 1TH, 1TT, 2H, 2T, 3HH, 3HT, 3TH, 3TT, 4H, 4T, 5HH, 5HT, 5TH, 5TT, 6H, 6T\}$
- $S_1 = \{MMMM, MMMF, MMFM, MFMM, FMMM, MMFF, MFME, MFFM, FMFM, FFMM, FMFF, MFFF, FMFF, FEME, FFFM, FFFF\}$;
 $S_2 = \{0, 1, 2, 3, 4\}$
- (a) $A = \{1HH, 1HT, 1TH, 1TT, 2H, 2T\}$;
(b) $B = \{1TT, 3TT, 5TT\}$;
(c) $A' = \{3HH, 3HT, 3TH, 3TT, 4H, 4T, 5HH, 5HT, 5TH, 5TT, 6H, 6T\}$;
(d) $A' \cap B = \{3TT, 5TT\}$;
(e) $A \cup B = \{1HH, 1HT, 1TH, 1TT, 2H, 2T, 3TT, 5TT\}$
- (a) $S = \{M_1M_2, M_1F_1, M_1F_2, M_2M_1, M_2F_1, M_2F_2, F_1M_1, F_1M_2, F_1F_2, F_2M_1, F_2M_2, F_2F_1\}$;
(b) $A = \{M_1M_2, M_1F_1, M_1F_2, M_2M_1, M_2F_1, M_2F_2\}$;

- (c) $B = \{M_1F_1, M_1F_2, M_2F_1, M_2F_2, F_1M_1, F_1M_2, F_2M_1, F_2M_2\}$
 (d) $C = \{F_1F_2, F_2F_1\}$
 (e) $A \cap B = \{M_1F_1, M_1F_2, M_2F_1, M_2F_2\}$
 (f) $A \cup C = \{M_1M_2, M_2M_1, M_1F_1, M_1F_2, M_2F_1, M_2F_2, F_1F_1, F_2F_1, F_1F_2, F_2F_2\}$



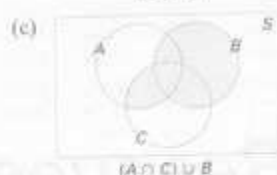
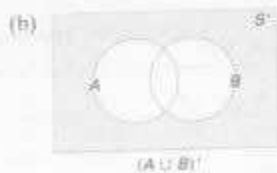
15. (a) [nitrógeno, potasio, uranio, oxígeno];
 (b) [cobre, sodio, zinc, oxígeno];
 (c) [cobre, sodio, nitrógeno, potasio, uranio, zinc];
 (d) [cobre, uranio, zinc];
 (e) \emptyset (f) [oxígeno]


PÁGINA 25

1. 18
 3. 156
 5. 20
 7. 48
 9. 210
 11. (a) 1024; (b) 243
 13. 72
 15. 362,880

PÁGINA 33

1. (a) La suma de las probabilidades excede 1.
 (b) La suma de las probabilidades es menor que 1.
 (c) Una probabilidad negativa.
 (d) Probabilidad de que una carta de corazones y una negra sean cero
3. $S = \{\$10, \$25, \$100\}$; 17/20



19. (a) La familia experimentará problemas mecánicos pero no recibirá una multa por violación al reglamento de tránsito y no llegará a un sitio para acampar que tenga lugares disponibles.
 (b) La familia recibirá una multa de tránsito y llegará a un sitio para acampar que no tiene lugares vacíos pero no tendrá problemas mecánicos.
 (c) La familia sufrirá problemas mecánicos y llegará a un sitio para acampar que no tiene lugares disponibles.
 (d) La familia recibirá una multa de tránsito pero llegará a un sitio para acampar que no tenga lugares disponibles.
 (e) La familia no tendrá problemas mecánicos.

17. 2880
 19. (a) 40,320; (b) 336
 21. 360
 23. 24
 25. 3360
 27. 7920
 29. 56

5. (a) 0.3; (b) 0.2
 7. (a) 5/26; (b) 9/26; (c) 19/26
 9. 10/117
 11. 95/663
 13. (a) 94/54, 145; (b) 143/39,984
 15. (a) 22/25; (b) 3/25; (c) 17/50
 17. (a) 0.32; (b) 0.68; (c) oficina o estudio

PÁGINA 42

1. (a) La probabilidad de que un convicto que promueve el uso de drogas también cometa robo a mano armada.
 (b) La probabilidad de que un convicto que comete robo a mano armada no promueva el consumo de drogas.
 (c) La probabilidad de que un convicto que no promovió el consumo de drogas tampoco cometió robo a mano armada.
3. (a) 14/39; (b) 95/112

PÁGINA 48

1. 0.0960
 3. 0.2097

Capítulo 3
PÁGINA 61

1. Discreto; continuo; continuo; discreto; discreto; continuo.

3. Espacio muestral w

| | |
|-----|----|
| HHH | 3 |
| HHT | 1 |
| HTH | 1 |
| THH | 1 |
| HTT | -1 |
| THT | -1 |
| TTH | -1 |
| TTT | -3 |

5. (a) 1/30; (b) 1/10
 7. (a) 0.68; (b) 0.375
 9. (b) 19/80

11.

| x | 0 | 1 | 2 |
|--------|---------------|---------------|---------------|
| $f(x)$ | $\frac{2}{7}$ | $\frac{4}{7}$ | $\frac{1}{7}$ |

$$13. F(x) = \begin{cases} 0 & \text{para } x < 0 \\ 0.41 & \text{para } 0 \leq x < 1 \\ 0.78 & \text{para } 1 \leq x < 2 \\ 0.94 & \text{para } 2 \leq x < 3 \\ 0.99 & \text{para } 3 \leq x < 4 \\ 1 & \text{para } x \geq 4 \end{cases}$$

5. (a) 5/34; (b) 3/8
 7. (a) 0.018; (b) 0.614; (c) 0.166; (d) 0.479
 9. (a) 0.35; (b) 0.875; (c) 0.55
 11. (a) 9/28; (b) 3/4; (c) 0.91
 13. 0.27
 15. 5/8
 17. (a) 0.0016; (b) 0.9984
 19. (a) 1/5; (b) 4/15; (c) 3/5
 21. (a) 91/323; (b) 91/323

5. 0.1124
 7. (a) 0.045; (b) 0.116; (c) 0.182; (d) 0.517

$$15. F(x) = \begin{cases} 0 & \text{para } x < 0 \\ \frac{2}{7} & \text{para } 0 \leq x < 1 \\ \frac{4}{7} & \text{para } 1 \leq x < 2 \\ 1 & \text{para } x \geq 2 \end{cases}$$

17. (b) 1/4; (c) 0.3

19. $F(x) = (x-1)/2$; 1/4

21. (a) 3/2; (b) $F(x) = x^{3/2}$; 0.3004

$$23. F(w) = \begin{cases} 0 & \text{para } w < -3 \\ \frac{1}{27} & \text{para } -3 \leq w < -1 \\ \frac{2}{27} & \text{para } -1 \leq w < 1 \\ \frac{10}{27} & \text{para } 1 \leq w < 3 \\ 1 & \text{para } w \geq 3 \end{cases}$$

- (a) 20/27; (b) 2/3

25.

| i | 20 | 25 | 30 |
|----------|---------------|---------------|---------------|
| $P(T=i)$ | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

PÁGINA 68

1. (a) Los troncos son 1, 2, 3, 4, 5, 6, 7, 8, 9, con frecuencias 3, 2, 3, 4, 5, 11, 14, 14, 4.
 (b) Los intervalos de clase son 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99 con frecuencias relativas 0.05, 0.03, 0.05, 0.07, 0.08, 0.18, 0.23, 0.23, 0.07.
 (c) Sesgado a la izquierda.
 (d) Las frecuencias relativas acumuladas son 0, 0.05, 0.08, 0.13, 0.20, 0.28, 0.46, 0.69, 0.92, 0.99.
 (f) Primer cuartil = 55.5; séptimo decil = 80
3. (a) Los troncos son 0*, 0-, 1+, 1-, 2+, 2-, 3+ con frecuencias 2, 17, 16, 10, 3, 1, 1.
 (b) Las frecuencias relativas son 0.04, 0.34, 0.32, 0.20, 0.06, 0.02, 0.02.
 (c) Sesgado a la derecha.
 (d) Las frecuencias relativas son 0, 0.04, 0.38, 0.70, 0.90, 0.96, 0.98, 1.00.
 (f) Percentil 75 = 15.9.
5. (a) Los troncos son 0*d*, 0*e*, 1*a*, 1*b*, 1*c*, 1*d*, 1*e*, 2*a*, 2*b*, 2*c*, con frecuencias 1, 1, 1, 2, 4, 13, 8, 5, 3, 2.
 (b) Las frecuencias relativas son 0.025, 0.025, 0.025, 0.050, 0.100, 0.325, 0.200, 0.125, 0.075, 0.050.

PÁGINA 79

1. (a) 1/36; (b) 1/15

3. (a)

| $f(x, y)$ | x | | | |
|-----------|----------------|-----------------|-----------------|----------------|
| | 0 | 1 | 2 | 3 |
| 0 | $\frac{1}{30}$ | $\frac{2}{30}$ | $\frac{2}{30}$ | $\frac{1}{30}$ |
| y 1 | $\frac{2}{30}$ | $\frac{18}{30}$ | $\frac{18}{30}$ | $\frac{2}{30}$ |
| 2 | $\frac{1}{30}$ | $\frac{2}{30}$ | $\frac{2}{30}$ | |

(b) 1/2

5. (a) 1/16;

(b) $g(x) = 12x(1-x)^2, 0 \leq x \leq 1$;

(c) 1/4

7. (a) 3/64; (b) 1/2

9. 0.6534

11. (a) Dependiente; (b) 1/3

13. (a)

| x | 1 | 2 | 3 |
|--------|------|------|------|
| $g(x)$ | 0.10 | 0.35 | 0.55 |

(b)

| y | 1 | 2 | 3 |
|--------|------|------|------|
| $h(y)$ | 0.20 | 0.50 | 0.30 |

(c) 0.2

15. (a)

| $f(x, y)$ | x | | |
|-----------|-----------------|----------------|----------------|
| | 0 | 1 | 2 |
| 0 | $\frac{17}{36}$ | $\frac{5}{36}$ | $\frac{1}{36}$ |
| y 1 | $\frac{5}{36}$ | $\frac{7}{36}$ | |
| 2 | $\frac{1}{36}$ | | |

(b) 11/12

17. (a)

| $f(x, y)$ | x | | | |
|-----------|----------------|----------------|----------------|----------------|
| | 0 | 1 | 2 | 3 |
| 0 | $\frac{1}{55}$ | $\frac{2}{55}$ | $\frac{6}{55}$ | $\frac{1}{55}$ |
| y 1 | $\frac{2}{55}$ | $\frac{3}{55}$ | $\frac{7}{55}$ | |
| 2 | $\frac{6}{55}$ | $\frac{7}{55}$ | | |
| 3 | $\frac{1}{55}$ | | | |

(b) 42/55

19. 3/4

21. Independiente

23. (a) 3; (b) 21/512

25. Dependiente

Capítulo 4

PÁGINA 90

1. 0

3. 25 centavos

5. 0.88

7. \$500

9. \$1.23

11. \$6900

13. $(\ln 4)/\pi$

15. 100 horas

17. 209

19. \$1855

21. \$167

23. (a) 35.2; (b) $\mu_X = 3.20$; $\mu_Y = 3.00$

25. 2

PÁGINA 100

- 55,250,000
- 0.74
- 1/18; en las utilidades reales la variancia es $\frac{1}{18}(5000)^2$.

PÁGINA 111

- 10.33; 6.66
- 80 centavos
- 209
- $\mu = 7/2; \sigma^2 = 15/4$
- 3/14
- 0.03125
- 0.9340

Capítulo 5
PÁGINA 122

- 3/10
- $\mu = 5.5; \sigma^2 = 8.25$
- (a) 0.0480; (b) 0.2375;
(c) $P\{x=5 | p=0.3\} = 0.1789$,
 $P = 0.3$ es ciertamente razonable.
- (a) 0.0474; (b) 0.0171
- (a) 0.7073; (b) 0.4613; (c) 0.1484
- 0.1240
- 0.8369

PÁGINA 131

- (a) 0.3246; (b) 0.44964
- 5/14
- $$h(x; 6, 3, 4) = \frac{\binom{4}{x} \binom{2}{2-x}}{\binom{6}{3}}, x = 1, 2, 3;$$

$$P(2 \leq X \leq 3) = 4/5$$
- 0.9517

PÁGINA 139

- 0.0515
- (a) 0.3840; (b) 0.0067
- 63/64
- (a) 0.0630; (b) 0.9730
- (a) 0.1429; (b) 0.1353
- (a) 0.1638; (b) 0.032

- 1/6
- 118.9
- $\mu_Y = 10; \sigma_Y^2 = 144$
- $\sigma_{XY} = 0.005$
- 0.0062
- 52
- (a) A lo más 4/9; (b) al menos 5/9;
(c) al menos 21/25; (d) 10
- (a) 7; (b) 0; (c) 12.25
- 1
- $P\{-2\sigma < X < 2\sigma\} \approx 0.75$ (valor verdadero = 0.95);
 $P\{-3\sigma < X < 3\sigma\} \approx 0.89$ (valor verdadero = 0.997)

- (a) 0.0778; (b) 0.3370; (c) 0.0870

- $\mu \pm 2\sigma = 3.5 \pm 2.05$

- $f(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} \frac{x_1}{0.35} \frac{x_2}{0.05} \frac{x_3}{0.60}$
- 0.0095
- 0.0077
- 0.8670
- (a) 0.2852; (b) 0.9887; (c) 0.6683

- (a) 0.6815; (b) 0.1153

- 3.25; de 0.52 a 5.98

- 0.9453
- 0.60776
- (a) 4/33; (b) 8/165
- 0.2315
- (a) 0.3991; (b) 0.1315

- (a) 0.3840; (b) 0.1395; (c) 0.0553

- 0.2657
- (a) $\mu = 4; \sigma^2 = 4$; (b) De 0 a 8
- (a) 0.2650; (b) 0.9596
- (a) 0.8243; (b) 14
- 4

Capítulo 6

PÁGINA 158

1. (a) 0.9236; (b) 0.8133;
 (c) 0.2424; (d) 0.0823;
 (e) 0.0250; (f) 0.6435
3. (a) -1.72; (b) 0.54; (c) 1.28
5. (a) 0.1151; (b) 16.1;
 (c) 20.275; (d) 0.5403
7. (a) 0.8980; (b) 0.0287; (c) 0.6080
9. (a) 0.0548; (b) 0.4514;
 (c) 23; (d) 189.95 mililitros

PÁGINA 165

1. (a) 0.8006; (b) 0.7803
3. (a) 0.1574; (b) 0.0108
5. (a) 0.9514; (b) 0.0668
7. (a) 0.1171; (b) 0.2049

PÁGINA 177

1. $2.8e^{-1.8} - 3.4e^{-2.4} = 0.1545$
5. (a) $\mu = 6; \sigma^2 = 18;$
 (b) de 0 a 14.485 millones de litros
7. $\sum_{i=4}^6 \binom{6}{i} (1 - e^{-3/4})^i (e^{-3/4})^{6-i} = 0.3968$
9. (a) $\sqrt{\pi}/2 = 1.2533;$ (b) e^{-2}
11. e^{-4}

Capítulo 7

PÁGINA 195

1. $g(y) = 1/3; y = 1, 3, 5$

$$3. g(y_1, y_2) = \begin{pmatrix} 2 \\ \frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}, 2 - y_1 \end{pmatrix} \\ \times \left(\frac{1}{4}\right)^{(y_1 + y_2)/2} \left(\frac{1}{3}\right)^{(y_1 - y_2)/2} \left(\frac{5}{12}\right)^{2 - y_1}$$

$$y_1 = 0, 1, 2; y_2 = -2, -1, 0, 1, 2;$$

$$y_2 \leq y_1; y_1 + y_2 = 0, 2, 4$$

11. (a) 0.0571; (b) 99.11%;
 (c) 0.3974; (d) 27.952 minutos
 (e) 0.0092
13. 6.24 años
15. (a) 58.77%; (b) \$17.46
17. (a) 0.0401; (b) 0.0244
19. 26
21. (a) 0.6; (b) 0.7; (c) 0.5

9. 0.1357
11. (a) 0.0778; (b) 0.0571; (c) 0.6811
13. (a) 0.9158; (b) 0.0109

13. (a) $\mu = \alpha\beta = 50;$
 (b) $\sigma^2 = \alpha\beta^2 = 500; \sigma = \sqrt{500};$
 (c) 0.8155
15. (a) 0.1889; (b) 0.357
17. Media = e^6 , varianza = $e^{12}(e^4 - 1)$
19. (a) e^{-10} ; (b) $\beta = 0.10$

7. Distribución gamma con $\alpha = 3/2$ y $\beta = m/2b$

9. (a) $g(y) = 32/y^3, y > 4;$ (b) 1/4

11. $h(z) = 2(1 - z), 0 < z < 1$

13. $h(w) = 6 + 6w - 12w^{3/2}, 0 < w < 1$

$$15. g(y) = \begin{cases} \frac{2}{9\sqrt{y}}, & 0 < y < 1 \\ \frac{(\sqrt{y} + 1)}{9\sqrt{y}}, & 1 < y < 4 \end{cases}$$

19. Ambas igual a μ

Capítulo 8

PÁGINA 207

- (a) Respuestas de todas las personas en Richmond que tienen un teléfono
(b) Resultados para un número grande o infinito de lanzamientos de una moneda
(c) Duración de tales tenis cuando se utilizan en el torneo profesional
(d) Todos los intervalos de tiempo para que esta abogada maneje de su casa a su oficina

PÁGINA 223

- 0.3159
- (a) Se reduce de 0.7 to 0.4;
(b) Aumenta de 0.2 to 0.8.
- Si
- (a) $\mu = 5.3$; $\sigma^2 = 0.81$;

PÁGINA 236

- (a) 27.488; (b) 18.475; (c) 36.415
- (a) 13.277; (b) 32.852; (c) 46.928
- (a) 0.05; (b) 0.94
- (a) 0.975; (b) 0.10; (c) 0.875; (d) 0.99
- (a) 2.500; (b) 1.319; (c) 1.714

Capítulo 9

PÁGINA 251

- $0.3097 < \mu < 0.3103$
- (a) $22,496 < \mu < 24,504$; (b) error ≤ 1004
- 11
- 56

PÁGINA 262

- $2.9 < \mu_1 - \mu_2 < 7.1$
- 2.80 $\mu_1 - \mu_2 < 3.40$
- $1.5 < \mu_1 - \mu_2 < 12.5$
- $0.70 < \mu_2 - \mu_1 < 3.30$

PÁGINA 270

- (a) $0.498 < p < 0.642$;
(b) error ≤ 0.072
- $0.194 < p < 0.262$
- (a) $0.739 < p < 0.961$; (b) no
- (a) $0.644 < p < 0.690$;
(b) error ≤ 0.023

- (a) $\bar{x} = 2.4$; (b) $\bar{x} = 2$; (c) $m = 3$
- (a) $\bar{x} = 3.2$ segundos; (b) $\bar{x} = 3.1$ segundos
- (a) 53.75; (b) 75.100
- (a) El rango es 10; (b) $s = 3.307$
- (a) 2.971; (b) 2.971
- $s = 0.585$
- (a) 45.9; (b) 5.1

- (b) $\mu_Y = 5.3$; $\sigma_Y^2 = 0.0225$;
(c) 0.9082
- (a) 0.6898; (b) 7.35
- 0.5596
- 0.9052

- No; $\mu > 20$
- (a) 2.71; (b) 3.51; (c) 2.92;
(d) 0.47; (e) 0.34
- La razón F es 1.44. Las varianzas no son significativamente diferentes.

- $0.978 < \mu < 1.053$
- $47,722 < \mu < 49,278$
- 323,946, 326,154
- 11,426 a 35,574

- $-6522 < \mu_1 - \mu_2 < 2922$
- $(-0.74, 6.29)$
- $(-6.92, 36.70)$

- 2576
- 160
- 16,577
- $-0.0136 < p_F - p_M < 0.0636$
- $0.0011 < p_1 - p_2 < 0.0869$
- $(-0.0817, -0.0019)$; significativamente diferente

PÁGINA 275

- $0.293 < \sigma^2 < 6.736$; afirmación válida
- $1.863 < \sigma < 3.578$
- $6.22 < \sigma < 12.42$

PÁGINA 281

- $p^* = 0.173$
- (a) $f(p|x=1) = 40p(1-p)^3/0.2844$;
(b) $p^* = 0.106$

PÁGINA 286

- $\frac{\sum x_i}{n}$
- (a) $L(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sigma^n} \left[\prod \left(\frac{1}{x_i} \right) \right] e^{1 - \sum \ln(x_i) - \mu^2/(2\sigma^2)}$

- $0.549 < \sigma_1/\sigma_2 < 2.690$
- $0.016 < \sigma_1^2/\sigma_2^2 < 0.454$; no

- $8.077 < \mu < 8.692$
- (a) 0.2509; (b) $68.71 < \mu < 71.69$
(c) 0.0174

$$(b) \hat{\mu} = \frac{\sum_{i=1}^n (\ln x_i)}{n}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \left[\ln x_i - \left(\frac{\sum \ln x_i}{n} \right)^2 \right]}{n-1}$$

- $x \ln p + (1-x) \ln(1-p)$.
Hacer la derivada con respecto a $p = 0$
 $\hat{p} = x = 1.0$.

Capítulo 10

PÁGINA 304

- (a) Concluir que menos de 30% del público es alérgico a algunos productos de queso cuando, de hecho, 30% o más son alérgicos.
(b) Concluir que al menos 30% del público es alérgico a algunos productos de queso cuando, de hecho, menos del 30% son alérgicos.
- (a) La empresa es inocente;
(b) la empresa es culpable.
- (a) $\alpha = 0.1498$; (b) $\beta = 0.0793$
 $\beta = 0.0618$
- (c) La probabilidad de cometer un error tipo I es algo grande.

PÁGINA 325

- $z = -1.64$; no rechazar la hipótesis nula.
- $z = -2.76$; sí, $\mu < 40$ meses.
- $z = 8.97$; sí, $\mu > 20.000$ kilómetros, $P < 0.001$.
- $t = 0.77$; no rechazar H_0 .
- $z = 12.72$; valor $P < 0.0005$
- $t = -1.98$; no rechazar H_0 .
- $z = 2.60$; rechazar H_0 , $\mu_A - \mu_B > 12$ kilogramos.
- $t = 1.50$; no
- $t = 0.70$; no efectiva
- $t = 2.55$; rechazar H_0 ; $\mu_1 - \mu_2 > 4$ kilómetros.

- (a) 0.0418; (b) $\beta = 0.0027$
 $\beta = 0.1251$
 $\beta = 0.6103$
- (a) $\alpha = 0.0032$; (b) $\beta = 0.0062$
- (a) $\alpha = 0.1357$; (b) $\beta = 0.2578$
- $\alpha = 0.0094$; $\beta = 0.0122$
- (a) $\alpha = 0.0718$; (b) $\beta = 0.1151$
- (a) $\alpha = 0.0384$; (b) $\beta = 0.5$; $\beta = 0.2776$

- $t' = 0.22$; no rechazar H_0 .
- $t' = 2.76$; no, $\mu_1 > \mu_2$
- $t = 2.45$; sí.
- $t = -2.53$; afirmación válida.
- $n = 6$
- 79
- 10
- (a) $H_0: M_{caliente} - M_{frio} = 0$;
 $H_1: M_{caliente} - M_{frio} \neq 0$;
(b) pareada t , $t = 0.99$; no rechazar. $P = 0.36$

PÁGINA 335

- $P = 0.4044$ (suponer una prueba de una cola); la afirmación no se rechaza.
- $P = 0.0207$; sí, la moneda no está balanceada.
- $P < 0.0001$; concluir que menos de $1/5$ de las casas se calientan con petróleo.

PÁGINA 339

- $\chi^2 = 18.12$; no rechazar $\sigma^2 = 0.03$.
- $\chi^2 = 63.75$; rechazar H_0 ; $\sigma^2 \neq 4.2$ ppm
- $\chi^2 = 42.37$; la máquina está fuera de control.
- $f = 1.33$; no rechazar H_0 ; $\sigma_1 = \sigma_2$.

PÁGINA 353

- $\chi^2 = 4.47$; sí.
- $\chi^2 = 10.14$; rechazar H_0 , la razón no es 5:2:2:1.
- $\chi^2 = 2.33$; no rechazar H_0 , distribución binomial.
- $\chi^2 = 2.57$; no rechazar H_0 , distribución geométrica.

Capítulo 11
PÁGINA 363

- (a) $a = 64.52916$, $b = 0.56090$;
(b) $\hat{y} = 81.4$
- (a) $6.4136 + 1.8091x$;
(b) $\hat{y} = 9.580$
- (a) $\hat{y} = 5.8254 + 0.5676x$;

PÁGINA 375

- (a) $s^2 = 176.362$;
(b) $t = 2.04$; no rechazar $\beta = 0$
- (a) $s^2 = 0.40$;
(b) $4.324 < \alpha < 8.503$;
(c) $0.446 < \beta < 3.172$
- (a) $s^2 = 6.626$;
(b) $2.684 < \alpha < 8.968$;
(c) $0.498 < \beta < 0.637$

- $z = 1.44$; afirmación válida.
- $z = 2.40$; sí, $P = 0.01$.
- Sí; $P = 0.0359$

- $f = 0.75$; no rechazar H_0 ; $\sigma_1 = \sigma_2$.
- $f = 1.18$; no rechazar H_0 ; $\sigma_A = \sigma_B$, $P = 0.82$.
- $F = 5.54$; $P\{F > 5.54\} = 0.0005$; las varianzas no son iguales.

- $\chi^2 = 5.19$; no rechazar H_0 , distribución normal.
- $\chi^2 = 5.40$; no rechazar H_0 .
- $\chi^2 = 124.59$; sí.
- $\chi^2 = 31.17$; las actitudes no son homogéneas.
- $\chi^2 = 5.92$; no rechazar H_0 .
- $\chi^2 = 1.84$; no rechazar H_0 .

- (c) $\hat{y} = 34.205$
- (b) $\hat{y} = 343.706 + 3.221x$;
(c) $\hat{y} = \$456$
- (a) $\hat{y} = 153.175 - 6.324x$;
(b) $\hat{y} = 123$

- $t = -2.24$; rechazar H_0 , $\beta < 6$;
- (a) $24.444 < \mu_{Y|24.5} < 27.112$;
(b) $21.889 < y_0 < 29.668$
- $7.815 < y_0 < 10.801$
- (a) 17.1812;
(b) no, (95% C.L. sobre las mpg media es [27.95; 29.60]);
(c) las millas por galón probablemente excederán 18

PÁGINA 385

$$1. (a) b = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2};$$

$$(b) \hat{y} = 2.003x$$

$$3. E(B) = \beta + \gamma \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1)x_{2i}}{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2}$$

5. (a) $a = 10.81153$, $b = -0.34370$;
 (b) $f = 0.43$; la regresión es lineal
 7. $f = 1.12$; la regresión es lineal
 9. $f = 1.71$, $P = 0.2517$; la regresión es lineal

PÁGINA 400

$$1. r = 0.240$$

$$5. (a) r = 0.392;$$

- (b) $t = 2.04$; no rechazar $\rho = 0$. El rechazo se concluye en el nivel de significancia de 0.053.

Capítulo 12

PÁGINA 414

$$1. (a) \hat{y} = 27.547 + 0.922x_1 + 0.284x_2;$$

$$(b) \hat{y} = 84$$

$$3. \hat{y} = 0.5800 + 2.7122x_1 + 2.0497x_2$$

$$5. (a) \hat{y} = 56.4633 + 0.1525x - 0.00008x^2;$$

$$(b) \hat{y} = 86.7\%$$

$$7. \hat{y} = 141.6118 - 0.2819x + 0.0003x^2$$

$$9. (a) \hat{y} = -102.7132 + 0.6054x_1 + 8.9236x_2 + 1.4374x_3 + 0.0136x_4;$$

$$(b) \hat{y} = 287.6$$

$$11. \hat{y} = 3.3205 + 0.4210x_1 - 0.2958x_2 + 0.0164x_3 + 0.1247x_4$$

$$13. \hat{y} = -6.5122 + 1.9994x_1 - 3.6751x_2 + 2.5245x_3 + 5.1581x_4 + 14.4012x_5$$

$$15. (a) \hat{y} = 350.9943 - 1.2720x_1 - 0.1539x_2;$$

$$(b) \hat{y} = 140.9$$

PÁGINA 426

$$1. 34.3699$$

$$3. 0.000995$$

$$5. (a) \hat{\sigma}_{\beta_1}^2 = 0.000002; \quad (b) \hat{\sigma}_{\beta_1, \beta_2} = -0.000003$$

$$7. 29.93 < \mu_{Y|10,5} < 31.97$$

$$9. t = 2.86; \text{ rechazar } H_0, \beta_2 > 0$$

$$11. t = 3.55; \text{ rechazar } H_0, \beta_1 > 2, P = 0.01.$$

$$13. (a) t = -1.09, \quad P = 0.3562;$$

$$(b) t = -1.72, \quad P = 0.1841;$$

$$(c) \text{ Sí; ni } x_1 \text{ ni } x_2 \text{ son significativas.}$$

PÁGINA 434

$$1. R^2 = 0.9997$$

$$3. f = 13,409; \text{ la regresión es significativa.}$$

$$5. f = 20.07; \text{ rechazar } H_0, \beta_1 < 0.$$

$$7. f_{2,7} = 10.18, P < 0.01. x_1 \text{ y } x_2 \text{ son significativas en presencia de } x_3 \text{ y } x_4.$$

$$9. \text{ El modelo de dos variables es mejor.}$$

PÁGINA 453

$$1. (b) \hat{y} = 4.690 \text{ segundos;}$$

$$(c) 4.450 > \mu_{Y|180,280} < 4.930$$

$$3. \hat{y} = 2.1833 + 0.9576x_1 + 3.3253x_2$$

$$5. (a) \hat{y} = -587.211 + 428.433x_1;$$

$$(b) \hat{y} = 1180 - 191.691x + 35.20945x^2;$$

$$(c) \text{ modelo cuadrático}$$

$$7. \hat{\sigma}_{\beta_1}^2 = 20.58804; \hat{\sigma}_{\beta_2}^2 = 62.6502$$

$$9. \text{ El modelo de intersección es mejor.}$$

$$11. (a) \hat{y} = 3.1368 + 0.6444x_1 - 0.0104x_2 + 0.5046x_3 - 0.1197x_4 - 2.4618x_5 + 1.5044x_6;$$

- (b) $\hat{y} = 4.6563 + 0.5113x_3 - 0.1242x_4$;
 (d) $\hat{y} = 4.6563 + 0.5113x_1 - 0.1242x_4$;
 (e) las dos observaciones tienen valores R de Student grandes y se deben verificar.

Capítulo 13

PÁGINA 473

3. $f = 0.31$; no hay una diferencia significativa.
 5. $f = 14.52$; sí, significativa.

PÁGINA 485

1. (a) $f = 14.27$; rechazar H_0 ;
 (b) $f = 23.23$; rechazar H_0 ;
 (c) $f = 2.48$; no rechazar H_0 .
 3. (a) $f = 13.50$; las medias de tratamiento difieren.
 (b) $f(1$ versus $2) = 29.35$; significativa.
 $f(3$ versus $4) = 3.59$; no significativa.

| \bar{x}_3 | \bar{x}_1 | \bar{x}_4 | \bar{x}_2 |
|-------------|-------------|-------------|-------------|
| 56.52 | 59.66 | 61.12 | 61.96 |

7. (a) $f = 9.01$; sí, significativa.
 (b) Agotamiento y Hess modificado son significativamente diferentes de los otros tres procedimientos.

PÁGINA 503

3. (a) $f(\text{fertilizantes}) = 6.11$; significativa.
 (b) $f = 17.37$; significativa.
 $f = 0.96$; no significativa.
 5. $f = 5.99$; el porcentaje de aditivos externos no es el mismo para las tres marcas de jamón.
 7. $f(\text{estaciones}) = 26.14$; significativa.
 9. $f(\text{dieta}) = 11.86$; significativa.

PÁGINA 522

1. (a) $f = 14.9$; los operadores difieren significativamente;
 (b) $\hat{\sigma}_a^2 = 28.91$; $s^2 = 8.32$.
 3. (a) $f = 3.33$; no hay diferencia significativa;
 (b) $\hat{\sigma}_a^2 = 1.08$; $\hat{\sigma}_b^2 = 2.25$.
 7. 9

Capítulo 14

PÁGINA 538

1. (a) $f = 8.13$; significativa.
 (b) $f = 5.18$; significativa.

13. (a) $\hat{y} = 125.8655 + 7.7586x_1 + 0.0943x_2 - 0.0092x_1x_2$;
 (b) $\hat{y} = 399.9742 - 0.2307x_2$.

7. $f = 2.25$; no rechazar H_0 .
 9. $f = 8.38$; las actividades específicas promedio difieren.

9. $d_1 = 9.0878$; significativa;
 $d_2 = 6.8498$; significativa;
 $d_3 = 2.3059$; significativa;
 $d_4 = 2.5093$; significativa.
 11. $f = 70.27$, valor $P \leq 0.0001$. Rechazar H_0 .

| x_6 | x_{25} | x_{100} | x_{75} | x_{30} |
|--------|----------|-----------|----------|----------|
| 55.167 | 60.167 | 64.167 | 70.500 | 72.833 |

Las temperaturas son importantes. 75 y 50°C producen baterías con vida activa significativamente más grande.

13. La absorción media para la mezcla 4 es significativamente menor que para las otras mezclas.
 11. $f = 0.58$; no significativa.
 15. $f = 5.03$; las calificaciones resultan afectadas por los diferentes profesores.
 17. $p \leq 0.001$.
 $f = 122.37$; la cantidad de tinte tiene un efecto sobre el color de la tela.

9. (a) $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $\alpha_i \sim N(0, \sigma_a^2)$;
 (b) $\hat{\sigma}_a^2 = 0$ (el componente estimado de la varianza es -0.00027); $\hat{\sigma}^2 = 0.0206$.
 11. (a) $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $\alpha_i \sim N(0, \sigma_a^2)$;
 (b) sí;
 (c) hay un componente de varianza del telar significativo.

- (c) $f = 1.63$; no significativa.
 3. (a) $f = 14.81$; significativa.

- (b) $f = 9.04$; significativa
 (c) $f = 0.61$; no significativa

5. (a) $f = 34.40$; significativa

PÁGINA 546

1. (a) $AB: f = 3.83$; significativa
 $AC: f = 3.79$; significativa
 $BC: f = 1.31$; no significativa
 $ABC: f = 1.63$; no significativa
 (b) $A: f = 0.54$; no significativa
 $B: f = 6.85$; significativa
 $C: f = 2.15$; no significativa
3. (a) Esfuerzo $f = 45.96$, $P < 0.0001$,
 recubrimiento $f = 0.05$, $P = 0.8299$,
 humedad $f = 2.13$, $P = 0.1257$,
 recubrimiento \times humedad, $f = 3.41$, $P = 0.0385$,
 recubrimiento \times esfuerzo, $f = 0.08$, $P = 0.9277$,
 humedad \times esfuerzo, $f = 3.15$, $P = 0.0192$,
 recubrimiento \times humedad \times esfuerzo, $f = 1.93$,
 $P = 0.1138$

PÁGINA 553

1. (a) $f = 1.49$; no hay una interacción significativa
 (b) $f(\text{operadores}) = 12.45$; significativa
 $f(\text{filtros}) = 8.39$; significativa
 (c) $\hat{\sigma}_a^2 = 0.1701$ (filtros)
 $\hat{\sigma}_b^2 = 0.3514$ (operadores)
 $s^2 = 0.1867$
3. (a) $\hat{\sigma}_a^2, \hat{\sigma}_b^2, \hat{\sigma}_{ab}^2$ son significativos;
 (b) $\hat{\sigma}_a^2, \hat{\sigma}_{ab}^2$ son significativos;
5. 0.59
7. (a) Modelo mezclado;

Capítulo 15

PÁGINA 568

1. $SSA = 2.6667$, $SSB = 170.6667$, $SSC = 104.1667$,
 $SS(AB) = 1.5000$, $SS(AC) = 42.6667$, $SS(BC)$
 $= 0.0000$, $SS(ABC) = 1.5000$
3. Los factores A, B , y C tienen efectos negativos en el compuesto fosforoso, y el factor D tiene un efecto positivo. Sin embargo, la interpretación del efecto de los factores individuales debe incluir el uso de las gráficas de interacción.

- (b) $f = 26.95$; significativa
 (c) $f = 20.30$; significativa

(b) La mejor combinación parece ser sin recubrimiento, humedad media y un nivel de esfuerzo de 20.

| 5. Efecto | f | Valor P |
|-------------|-------|---------------|
| Temperatura | 14.22 | ≤ 0.0001 |
| Superficie | 6.70 | 0.0020 |
| HRC | 1.67 | 0.1954 |
| TXS | 5.50 | 0.0006 |
| T * HRC | 2.69 | 0.0369 |
| S * HRC | 5.41 | 0.0007 |
| T * S * HRC | 3.02 | 0.0051 |

7. (a) sí; marca X tipo, marca X temperatura.
 (b) sí;
 (c) marca Y, detergente en polvo; temperatura caliente

- (b) Material: $f = 47.42$, $P < 0.0001$
 Marca: $f = 1.73$, $P = 0.2875$
 Material \times Marca: $f = 16.06$, $P = 0.0004$

(c) sí

9. (a) $y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + E_{ijk}$,
 $A =$ Ajuste de potencia, $B =$ Tipo de cereal,
 $B_j \sim N(0, \sigma_B^2)$, independiente;
 $(AB)_{ij} \sim N(0, \sigma_{AB}^2)$, independiente;
 $E_{ijk} \sim N(0, \sigma^2)$, independiente;
 (b) no; (c) no

5. Efectos significativos:

$A: f = 9.98, BC: f = 19.3$

Efectos no significativos:

$B: f = 0.20, AC: f = 0.20$

$C: f = 6.54, AD: f = 0.57$

$D: f = 0.02, BD: f = 1.83$

$AB: f = 1.83, CD: f = 0.02$

PÁGINA 578

1. $A, B, C, AC, BC,$ y ABC cada uno con 1 grado de libertad se puede probar con el uso de un cuadrado medio del error con 12 grados de libertad.

3.

| Bloque 1 | Bloque 2 | Bloque 3 | Bloque 4 |
|-------------------------------|------------------------------|------------------------------|------------------------------|
| (1)
ab
acd
bcd | c
abc
ad
bd | d
ac
bc
abd | a
b
cd
$abcd$ |

CD también está confundido.

5.

| Bloque 1 | | Bloque 2 | | Bloque 3 | |
|----------|------|----------|------|----------|------|
| 1 | 2 | 1 | 2 | 1 | 2 |
| abc | ab | abc | ab | (1) | a |
| a | ac | a | ac | c | b |
| b | bc | b | bc | ab | ac |
| c | (1) | c | (1) | abc | bc |

Réplica 1 Réplica 2 Réplica 3
 ABC ABC AB
 confundida confundida confundida

7. (a)

| | Máquina | | | |
|--------|---------|-------|---|--------|
| | 1 | 2 | 3 | 4 |
| (1) | c | a | | ac |
| ab | d | b | | ad |
| cd | e | acd | | ae |
| ce | abc | ace | | bc |
| de | abd | ade | | bd |
| $abcd$ | abe | bcd | | be |
| $abce$ | cde | bce | | $acde$ |
| $abde$ | $abcde$ | bde | | $bcde$ |

(b) $AB, CDE, ABCDE$

PÁGINA 593

1. $A = CDE$ $AE = CD$
 $B = ABCDE$ $BC = ABDE$
 $C = ADE$ $BD = ABCE$
 $D = ACE$ $BE = ABCD$
 $E = ACD$ $ABC = BDE$
 $AB = BCDE$ $ABD = BCE$
 $AC = DE$ $ABE = BCD$
 $AD = CE$
3. Bloque principal = $\{(1), ac, bd, abcd, abc, bce, ade, cde, abf, bcf, adf, cdf, cf, acef, bdef, abcdef\}$.

- $A = BCD = ABDEF = CEF$
 $B = ACD = DEF = ABCEF$
 $C = ABD = BCDEF = AEF$
 $D = ABC = BEF = ACDEF$
 $E = ABCDE = BDF = ACF$
 $F = ABCDF = BDE = ACE$

5. Efecto significativo:

$$E: f = 5.39$$

Efectos no significativos:

$$A: f = 0.48, \quad D: f = 1.94$$

$$B: f = 1.35, \quad F: f = 1.09$$

$$C: f = 3.03, \quad G: f = 4.36$$

PÁGINA 606

1. $\hat{y} = 12.7519 + 4.7194x_1 + 0.8656x_2 - 1.4156x_3$
 Las unidades están centradas y escaladas. Falta de ajuste:
 $F = 81.58, \quad P < 0.0001$

3. AFG $CEFG$ $ACDF$
 BEG $BDFG$ $ABEF$
 CDG $BCDE$ $ABCDEF$
 DEF $ADEG$

Capítulo 16

PÁGINA 617

1. $x = 7$; $P = 0.1719$, no rechazar H_0 .
3. $x = 3$; $P = 0.0244$, rechazar H_0 .
5. $x = 4$; $P = 0.3770$, no rechazar H_0 .
7. $x = 4$; $P = 0.1335$, no rechazar H_0 .

PÁGINA 624

1. $u_1 = 1$; la afirmación es válida.
3. $u_2 = 5$; A opera más tiempo.
5. $u = 15$; no, no rechazar H_0 .

PÁGINA 632

1. $v = 7$; $P = 0.910$, muestra aleatoria
3. $v = 6$; $P = 0.044$, no rechazar H_0 .
5. $z = 1.11$; muestra aleatoria
7. 0.70

9. $w = 43$; no rechazar H_0 .

11. $w_+ = 17.5$; no rechazar H_0 .

13. $z = -2.13$; rechazar H_0 , $\mu_1 - \mu_2 < 8$.

7. $h = 10.47$; los tiempos de operación son diferentes.

9. $h = 1.07$; no hay una diferencia significativa.

9. 0.995

11. (a) $r_s = 0.39$; (b) no rechazar H_0 .

13. (a) $r_s = 0.72$; (b) rechazar H_0 ; $\rho > 0$.

15. (a) $r_s = 0.71$; (b) rechazar H_0 ; $\rho > 0$.

Índice

- A**
- Análisis de varianza, 235, 461, 463
 - de un factor, 461
 - en regresión, 377
 - experimentos factoriales fraccionales, 560
- B**
- Bayes, 239
 - ANOVA, 493
 - Bloque completo aleatorizado, 489
 - estimación de, 276
 - regla de, 44
 - Bloques, 488, 489
 - experimentos factoriales 2k, 571
- C**
- Coefficiente de confianza, 242
 - Coefficiente de correlación, 395, 396
 - Coefficiente de correlación múltiple, 427
 - Coefficiente de correlación simple, 395
 - Coefficiente de correlación del rango, 630
 - Coefficiente de determinación, 427
 - Comparaciones múltiples, 479
 - Componentes de varianza, 509
 - Confusión, 571
 - parcial, 577
- Contraste**, 475
 - ortogonal, 477

Correlación, 394

Covarianza, 92, 97, 99

Cuadrados latinos, 499, 500
 - ANOVA para, 501

Curva característica de operación, 516

D

 - Desviación estándar, 93
 - muestra, 207
 - Diagonales HAT, 441, 442
 - Diagrama de tronco y hojas, 63
 - Diagrama de Venn, 16
 - Diferencia entre medias, 253, 313
 - Diferencia entre proporciones, 268, 333
 - Diseño de parámetro robusto, 599
 - Diseño de Plackett-Burman, 377
 - Diseño experimental, 433, 463
 - Distribución
 - beta, 177
 - binomial, 116, 129
 - binomial negativa, 132, 133
 - de Poisson, 135, 136
 - de Weibull, 174
 - exponencial, 166
 - F, 232, 233
 - gamma, 166, 168
 - geométrica, 132
 - hipergeométrica, 125, 127
 - ji cuadrada, 172
 - logarítmica normal, 173
 - multinomial, 116
 - normal, 7, 145
 - t, 228, 229
 - Distribución beta, 177
 - Distribución binomial, 116
 - Distribución condicional, 74, 79
 - Distribución de frecuencia, 65
 - Distribución de Poisson, 135
 - Distribución de probabilidad conjunta, 70, 71
 - Distribución de probabilidad continua, 58
 - Distribución de probabilidad discreta, 53, 114
 - Distribución de Weibull, 174
 - Distribución F, 232, 233
 - Distribución F no central, 516
 - Distribución hipergeométrica, 125, 126
 - Distribución hipergeométrica multivariada, 130
 - Distribución ji cuadrada, 172, 226
 - grados de libertad, 226
 - Distribución muestral, 215, 235
 - Distribución normal, 145
 - aproximación a la binomial, 160
 - curva normal, 145
 - densidad normal, 145
 - Distribución posterior, 276
 - distribuciones, 53
 - graficación, 213, 391
 - Distribución t, 228, 229
 - Distribuciones acumuladas, 55, 67

E

 - Efectos, 56

- Error estándar, 248
de coeficientes, 586
- Errores tipo I y tipo II, 293
- Estadística C_p , 445, 450
- Estadística no paramétrica, 609
- Estimación, 8
de Bayes, 280
de probabilidad máxima, 282
de proporción, 264
de varianza, 271
puntual, 239
- Estimación por intervalos, 242
(véase también Intervalo de confianza)
- Estimado puntual, 239
- Estimador, 239
de Bayes, 276
insesgado, 240
- Estimador insesgado, 240
- Eventos, 13
intersección de, 15
mutuamente excluyentes, 15
unión de, 15
- Experimento de dos factores, 529
ANOVA, 534
interacción, 529
- Experimento de Poisson, 135
- Experimento de tres factores, 540
- Experimentos factoriales, 559
en bloques incompletos, 571
experimentos 2^k , 559
fraccionales, 588
- Experimentos factoriales fraccionales, 588
- F**
- Falta de ajuste, 381
- Función de densidad conjunta, 71
- Función generadora de momento, 189
para combinaciones lineales, 193
- G**
- Gráfica cusum, 660
- Gráfica de control, 636
atributos, 653
defectos, 657
- gráfica \bar{X} , 640
- gráfica-R, 642, 643, 649
para variables, 638
- Gráfica de percentil, 213
gráfica de percentil normal, 213
- Gráfica de probabilidad normal, 213, 391
- Gráfica percentil-percentil normal, 213
- Gráficas de caja y extensión, 209
- Gráficas de diagnóstico, 389
gráficas de residuos, 390
- Gráficas de efectos, 564
- H**
- Hipótesis (véase Pruebas de hipótesis)
- Hipótesis estadística (véase Pruebas de hipótesis)
- I**
- Independencia, 38, 77, 104, 344
prueba de, 344
- Interacciones, 528, 529
- Intervalo bayesiano, 280
- Intervalo de confianza, 242, 257
para diferencia entre medias, 256, 258
para proporción, 265
para respuesta media, 372, 421
- Intervalo de tolerancia, 249
- L**
- Límites de tolerancia, 249
- M**
- Matriz HAT, 441
- Media, 4, 84, 85
distribución muestral de, 241, 282
intervalo de confianza para, 244, 372, 421
prueba de, 306, 310
- Mediana, 4, 202
- Método bayesiano, 275
- Mínimos cuadrados, 362, 418
- Modelo de efectos aleatorios, 506, 549
- Modelo mezclado, 551
- Muestra, 2, 198
espacio, 11
media, 4
mediana, 4, 202
rango, 6, 205, 641
varianza, 6, 205
- Muestra aleatoria, 198
- O**
- Ortogonal, 431
contrastes, 477
diseño, 584
- P**
- Pérdida de ajuste, 381
- Probabilidad condicional, 35
- Probabilidad máxima, 282
- Problema de un factor, 461
ANOVA para, 461
- Proceso de Bernoulli, 116
- Proporción, 264
estimación de, 264
prueba de, 330
- Prueba de Bartlett, 471
distribución de, 471
- Prueba de Cochran, 472
- Prueba de dos colas, 300
- Prueba de Kruskal-Wallis, 623
- Prueba de la suma de rango de Wilcoxon, 619
- Prueba de la suma de rangos, 620
- Prueba de rango múltiple de Duncan, 482
- Prueba de Tukey, 481
- Prueba de una cola, 300
- Prueba del rango con signo de Wilcoxon, 614
- Prueba del rango con signo, 614
- Prueba del signo, 610
- Pruebas de corridas, 625
- Pruebas de hipótesis
bondad de ajuste, 341

- con observaciones pareadas, 316
 - de homogeneidad, 347
 - de independencia, 344
 - prueba de corridas, 625
 - prueba de Kruskal-Wallis, 623
 - prueba de la suma de rango de Wilcoxon, 619
 - prueba de signos, 610
 - prueba del rango con signo de Wilcoxon, 614
 - sobre dos medias, 313, 314
 - sobre dos proporciones, 333
 - sobre dos varianzas, 336
 - sobre una proporción, 331
 - sobre una sola media, 310
 - varias varianzas, 471, 472
- Pruebas no paramétricas, 609
- R**
- Rango, 6, 641
 - Rango mínimo significativo, 482
 - Región crítica, 292
 - Regresión, 8, 358
 - lineal múltiple, 405
 - Regresión lineal múltiple, 405
 - coeficientes de estimación, 406
 - intervalos de confianza, 421
 - intervalos de predicción en, 422
 - métodos paso a paso, 435, 437
 - pruebas de significancia, 427, 428
 - Regresión lineal, 358
 - coeficientes, 359, 362
 - modelo, 361
 - prueba de linealidad, 381
 - Regresión lineal simple, 358
 - Regresión paso a paso, 435
 - Regresión polinomial, 406, 413
 - Residuo, 361, 389, 441
 - gráficas de, 389
 - Residuos estudentizados, 442
 - bondad de ajuste, 340, 341
 - con observaciones pareadas, 316
 - de homogeneidad, 347
 - de independencia, 344
 - prueba de Kruskal-Wallis, 623
 - prueba de la suma de rango de Wilcoxon, 619
 - prueba de signo, 610
 - prueba del rango con signo de Wilcoxon, 614
 - pruebas de corridas, 625
 - sobre dos medias, 313, 314
 - sobre dos proporciones, 333
 - sobre dos varianzas, 336
- T**
- Tabla de contingencia, 344, 347
 - Taguchi, 599-605
 - Teorema del límite central, 217, 224
 - teorema de Chebyshev, 108, 109
 - Transformación, 387
- V**
- Validación cruzada, 445
 - Valor esperado, 85
 - Variable aleatoria, 51
 - continua, 58
 - covarianza entre, 97
 - discreta, 53
 - independiente, 77
 - media de, 84
 - varianza de, 92
 - Variable aleatoria discreta, 51, 53
 - Varianza, 6, 92
 - análisis de, 235, 461, 463
 - de una muestra, 6, 205
 - estimación puntual de, 6, 205



OCT

**PROGRAMAS EDUCATIVOS, S.A. DE C.V.
CALZ. CHABACANO NO. 65,
COL. ASTURIAS, DELG. CUAUHEMOC.
C.P. 06850, MÉXICO, D.F.**

**EMPRESA CERTIFICADA POR EL
INSTITUTO MEXICANO DE NORMALIZACIÓN
Y CERTIFICACIÓN A.C. BAJO LAS NORMAS
ISO-9002:1994/NMX-CC-004:1995
CON EL NO. DE REGISTRO RSC-048
E ISO-14001:1996/NMX-SAA-001:1998 IMNC/
CON EL NO. DE REGISTRO RSAA-003**

2004



PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS

SEXTA EDICIÓN

La sexta edición de *Probabilidad y estadística para ingenieros* destaca la teoría estadística y hace un énfasis apropiado en el análisis de datos, el cálculo estadístico y la mejora de la calidad.

Otra ventaja del texto es la abundancia de ejercicios de la "vida real" que incluyen conjuntos de datos de estudios realizados por ingenieros y científicos.

Además se proporcionan las respuestas para los ejercicios impares.

Los capítulos 1 a 10 se adaptan a un curso de un semestre que incluya el tratamiento de probabilidad e inferencia estadística. Sin embargo, existe mucha flexibilidad para el uso de estos capítulos.

Los capítulos 11 a 17, proporcionan material suficiente para un segundo semestre.

El temario de esta obra es el siguiente:

Introducción a la estadística y al análisis de datos • Probabilidad • Variables aleatorias y distribuciones de probabilidad • Esperanza matemática • Algunas distribuciones de probabilidad discreta • Algunas distribuciones continuas de probabilidad • Funciones de variables aleatorias • Distribuciones fundamentales de muestreo y descripciones de datos • Problemas de estimación de una y dos muestras • Pruebas de hipótesis de una y dos muestras • Regresión lineal simple y correlación • Regresión lineal múltiple • Experimentos de un factor: general • Experimentos factoriales • Experimentos 2^o factoriales y fracciones • Estadística no paramétrica • Control estadístico de calidad.



Visitenos en:
www.pearsonedlatino.com

