



# Tendencia Reciente en la Educación Probabilística y Estadística

## **Editores**

José Dionicio Zacarias Flores  
Hugo Adán Cruz Sosa  
Fernando Velasco Luna  
Bulmaro Juárez Hernández  
Víctor Hugo Vázquez Guevara  
Hortensia Josefina Reyes Cervantes  
Francisco Solano Tajonar Sanabria  
Gladys Denisse Salgado Suárez

# TENDENCIA RECIENTE EN LA EDUCACIÓN PROBABILÍSTICA Y ESTADÍSTICA

José Dionicio Zacarías Flores

Coordinador



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

2019

Primera edición: 2019

ISBN: 978-607-525-641-2

©Benemérita Universidad Autónoma de Puebla

4 Sur 104, Col. Centro Histórico, Puebla, Pue. CP 72000

Teléfono: (222) 229 55 00

[www.buap.mx](http://www.buap.mx)

Dirección General de Publicaciones

2 Norte 1404, Col. Centro Histórico, Puebla, Pue. CP 72000

Teléfonos: (222) 246 85 59 y (222) 229 55 00 Ext. 5768

[publicaciones.buap.mx](http://publicaciones.buap.mx)

Facultad de Ciencias Físico Matemáticas

Av. San Claudio y 18 Sur, Colonia San Manuel, Puebla, Pue. CP 72570

Teléfonos: (222) 229 55 00 Ext. 7552

[www.fcfm.buap.mx](http://www.fcfm.buap.mx)

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA • Rector José Alfonso Esparza Ortiz  
• Secretario General: José Jaime Vázquez López • Vicerrector de Extensión y Difusión de la Cultura:  
José Carlos Bernal Suárez • Director General de Publicaciones: Hugo Vargas Comsille • Director de  
la Facultad de Ciencias Físico Matemáticas: Martha Alicia Palomino Ovando

Impreso y hecho en México

Printed and made in México

## PRÓLOGO

El Cuerpo Académico de Probabilidad y Estadística, perteneciente a la Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla pone a disposición de todos los interesados la edición del libro “*Tendencia Reciente en la Educación Probabilística y Estadística*”. Este libro es el esfuerzo de investigadores pertenecientes a instituciones educativas tanto a nivel nacional como internacional, con la finalidad de aportar una muestra representativa de las líneas de investigación que actualmente se abordan a nivel internacional, acerca de las problemáticas que se dan en el proceso de la enseñanza y aprendizaje de la probabilidad y estadística. El contenido del libro está elaborado por trabajos inéditos aportados por los autores, después de haber sido sometidos a una rigurosa revisión arbitrada, para ser parte de la edición como capítulos de libro.

Atentamente.

Los editores.

Agradecemos a la M. C. Gladys Denisse Salgado Suárez por su participación en el arduo trabajo que implicó la edición de este libro.

## ÍNDICE

<b>1. <u>EVALUACIÓN DEL APRENDIZAJE EN PROBABILIDAD Y ESTADÍSTICA: REFERENTES PARA LA RECONVERSIÓN DE LA PRÁCTICA EDUCATIVA .....</u></b>	<b>1</b>
1.1 INTRODUCCIÓN.....	3
1.2 CONSIDERACIONES DEL MARCO DEL ESTUDIO, ORIENTADAS A LA BÚSQUEDA DE REFERENTES Y REFLEXIONES PARA UNA PRÁCTICA DOCENTE RESIGNIFICADA.....	3
1.3 PROCEDER DEL ESTUDIO .....	4
1.4 RESULTADOS .....	5
1.5 CONCLUSIONES .....	11
<b>2. <u>DROP THE BETA .....</u></b>	<b>15</b>
2.1 INTRODUCCIÓN.....	17
2.2 OBJETIVOS .....	18
2.3 CONOCIMIENTOS PREVIOS .....	19
2.4 DESARROLLO DEL JUEGO .....	21
2.4.1 MATERIALES.....	21
2.4.2 DESCRIPCIÓN DEL RECURSO.....	21
2.4.3 DESARROLLO DEL CONTENIDO.....	23
2.5 SUGERENCIAS DIDÁCTICAS .....	26
2.6 GLOSARIO.....	27
2.7 ANEXOS .....	30
<b>3. <u>FORMACIÓN VIRTUAL: ENSEÑANZA Y APRENDIZAJE DE LA PROBABILIDAD .....</u></b>	<b>35</b>
3.1 INTRODUCCIÓN.....	37
3.2 INDAGACIÓN BIBLIOGRÁFICA.....	38
3.3 OBJETIVO DE INVESTIGACIÓN .....	42
3.4 MÉTODO .....	42
3.5 RESULTADOS .....	45
3.6 DISCUSIÓN.....	48
3.7 CONCLUSIONES.....	49
<b>4. <u>APLICACIONES DE LA ESTADÍSTICA MULTIVARIADA AL ANÁLISIS DE DATOS ECOLÓGICOS.....</u></b>	<b>55</b>
4.1 INTRODUCCIÓN.....	57
4.2 PROBLEMÁTICA .....	57
4.3 OBJETIVO.....	58
4.4 JUSTIFICACIÓN.....	58
4.5 MATERIAL Y MÉTODO.....	58
4.6 RESULTADOS .....	60
4.7 DISCUSIÓN.....	67
4.8 CONCLUSIÓN .....	69
4.9 AGRADECIMIENTOS .....	69
<b>5. <u>APRENDIZAJE DE ESTADÍSTICA Y PROBABILIDAD EN PROCESOS DE INSPECCIÓN. CASO DE ESTUDIO: CONTROL DE CALIDAD DE LA LECHE. ...</u></b>	<b>71</b>
5.1 INTRODUCCIÓN.....	72
5.2 FUNDAMENTOS.....	73
5.3 CONSIDERACIONES A TENER EN CUENTA .....	76

5.4 ESTUDIO DEL PROBLEMA.....	76
5.4.1 CALIDAD COMO PARTE DEL PROCESO .....	77
5.4.2 VARIABILIDAD.....	78
5.4.3 PRUEBAS DE CONTROL DE CALIDAD.....	79
5.4.4 PASOS DEL MÉTODO ESTADÍSTICO .....	80
5.4.5 ESTUDIO DE CASO: CONTROL DE CALIDAD DE LA LECHE .....	80
5.5 CONSIDERACIONES FINALES .....	90
<b>6. ESTUDIO DE CASOS PARA EL APRENDIZAJE DE CONCEPTOS DE ESTADÍSTICA DESCRIPTIVA: UNA EXPERIENCIA CON ESTUDIANTES DE INGENIERÍA .....</b>	<b>93</b>
6.1 INTRODUCCIÓN.....	95
6.2 MARCO TEÓRICO: EL MODELO TPACK.....	95
6.3 SOBRE LA INTERVENCIÓN DIDÁCTICA .....	96
6.4 ANÁLISIS DE LOS ERRORES MÁS OBSERVADOS .....	98
6.5 PERCEPCIONES DE LOS ESTUDIANTES SOBRE LA MODALIDAD DE TRABAJO.....	99
6.6 CONCLUSIONES.....	100
<b>7. EXPLORACIÓN DE LA INTERPRETACIÓN DE UNA TABLA Y UNA GRÁFICA DE COLUMNAS POR ESTUDIANTES DE LICENCIATURA .....</b>	<b>103</b>
7.1 INTRODUCCIÓN.....	106
7.2 ANTECEDENTES .....	107
7.3 REFERENTE TEÓRICO.....	108
7.4 METODOLOGÍA .....	109
7.5 RESULTADOS .....	112
7.6 DISCUSIÓN Y CONCLUSIONES.....	117
<b>8. SUB&amp;DES. UN JUEGO QUE FACILITA EL APRENDIZAJE DEL DISEÑO DE MEDIDAS REPETIDAS.....</b>	<b>123</b>
8.1 INTRODUCCIÓN.....	125
8.2 OBJETIVO GENERAL.....	126
8.3 CONOCIMIENTOS PREVIOS .....	126
8.4 DESARROLLO DEL JUEGO .....	127
<b>9. ANÁLISIS ESTADÍSTICO EN UNA EVALUACIÓN, UNA APLICACIÓN DE LA TÉCNICA CLÚSTER K-MEDIA.....</b>	<b>133</b>
9.1 INTRODUCCIÓN.....	134
9.2 METODOLOGÍA .....	135
9.3 RESULTADOS .....	137
9.3.1 ANÁLISIS DESCRIPTIVO UNIVARIANTE .....	137
9.3.2 ANÁLISIS MULTIVARIANTE.....	138
9.4 CONCLUSIONES.....	140
<b>10. TÉCNICAS DE CONTEO CON SOFTWARE APLICADAS A LA PROBABILIDAD .....</b>	<b>143</b>
10.1 INTRODUCCIÓN.....	144
10.2 ENFOQUE CLÁSICO .....	145
10.3 ENFOQUE DE FRECUENCIA O FRECUENCIAL.....	146
10.4 DESARROLLO AXIOMÁTICO DE PROBABILIDAD.....	148
10.5 PRIMER PRINCIPIO DE CONTEO.....	148

10.6	SEGUNDO PRINCIPIO DE CONTEO .....	150
10.7	PERMUTACIONES .....	151
10.8	COMBINACIONES .....	153
10.9	REGLA DE LA ADICIÓN.....	155
10.10	CONCLUSIONES.....	156
<b>11.</b>	<b><u>ANÁLISIS ESTADÍSTICO DE CONCEPTOS PROBABILÍSTICOS .....</u></b>	<b><u>159</u></b>
11.1	INTRODUCCIÓN .....	161
11.2	METODOLOGÍA .....	164
11.3	RESULTADOS .....	165
11.4	CONCLUSIONES.....	171
<b>12.</b>	<b><u>MODELOS PARAMÉTRICOS EN EL ANÁLISIS DE SUPERVIVENCIA.....</u></b>	<b><u>173</u></b>
12.1	INTRODUCCIÓN .....	175
12.2	¿QUÈ ES EL ANÁLISIS DE SUPERVIVENCIA? .....	176
12.3	CONCEPTOS Y RESULTADOS BÁSICO DE SUPERVIVENCIA .....	179
12.4	MODELOS PARÁMETRICOS .....	182
12.5	CONCLUSIONES.....	186
<b>13.</b>	<b><u>IMPLEMENTACIÓN DEL APRENDIZAJE BASADO EN PROYECTOS EN CURSOS DE INGENIERÍA: COMPARACIÓN DE LA EFICIENCIA ENTRE LA LOMBRICOMPOSTA Y LA COMPOSTA CASERA .....</u></b>	<b><u>189</u></b>
13.1	INTRODUCCIÓN .....	190
13.2	METODOLOGÍA .....	192
13.3	OBJETIVOS DE LA ELABORACIÓN DE LAS COMPOSTAS: .....	194
13.4	DISEÑO PROTOTIPO 1 Y 2: LOMBRICOMPOSTA .....	198
13.5	DISEÑO PROTOTIPO 3 Y 4: COMPOSTA CASERA .....	201
13.6	MEDICIONES DE PROTOTIPOS 1 Y 2.....	203
13.7	COMPARACIÓN DE LAS TÉCNICAS DE MEDIDA POR PROTOTIPO.....	206
13.8	MEDICIONES PROTOTIPOS 3 Y 4.....	207
13.9	RESULTADOS .....	214
14.9.1	MEDICIONES DE PROTOTIPOS 1 Y 2.....	214
14.9.2	MEDICIONES EN COMPOSTA CASERA.....	216
13.9.3	ANÁLISIS DE LAS OBSERVACIONES PROTOTIPOS 1 Y 2 .....	217
13.9.5	ANÁLISIS DE LAS OBSERVACIONES PROTOTIPOS 3 Y 4 .....	218
13.10	CONCLUSIONES.....	223
<b>14.</b>	<b><u>EL CÁLCULO DE ÁREAS A TRAVÉS DE LA PROBABILIDAD .....</u></b>	<b><u>231</u></b>
14.1	INTRODUCCIÓN .....	232
14.2	SOBRE LA PROPUESTA .....	234
14.2.1	PÚBLICO META.....	234
14.2.2	MARCO TEÓRICO .....	235
14.3	UNA PROPUESTA PARA INTRODUCIR EL CÁLCULO DE ÁREAS UTILIZANDO PROBABILIDAD .....	237
14.3.1	PRIMERA EXPLICACIÓN Y PRIMER EJEMPLO .....	237
14.3.2	EXPERIMENTACIÓN: .....	239
14.3.3	ACTIVIDAD DE CIERRE.....	241
14.3.4	ACTIVIDADES A REALIZAR .....	242
14.4	CONCLUSIONES.....	243



<b>15. “APRENDER HACIENDO”: APLICACIÓN DE TECNOLOGÍA EN LA ENSEÑANZA DE LOS CUATRO COMPONENTES DE SERIES DE TIEMPO .....</b>	<b>247</b>
15.1 INTRODUCCIÓN .....	249
15.2 MARCO CONTEXTUAL .....	250
15.3 ANTECEDENTES .....	252
15.4 DESARROLLO DE LA SECUENCIA .....	254
15.5 CONCLUSIONES.....	257



---

**Evaluación del Aprendizaje en Probabilidad y Estadística: Referentes para la reconversión de la práctica educativa**

---

Hugo Moreno Reyes  
hmoreno@ciidet.edu.mx



**Resumen**

Este trabajo de investigación, que se llevó a cabo desde una óptica más amplia para tener una visión articulada del proceso enseñanza-aprendizaje de la probabilidad y estadística, tiene como propósito presentar a manera de estructura referencial los aspectos contemporáneos vinculados con las tecnologías digitales, las dinámicas comunicativas y socioculturales que enmarcan la práctica evaluativa en la asignatura de probabilidad y estadística que se imparte a nivel licenciatura en las carreras de ingeniería de los Institutos Tecnológicos. También se tocan aspectos clave para clarificar que la evaluación del aprendizaje es un concepto mucho más amplio y profundo que el de calificación y acreditación, útiles para la reconversión de las prácticas evaluativas de los

profesores. Con respecto a la metodología, se recurrió al método de la transdisciplinariedad descriptiva-explicativa a partir del análisis crítico de documentos, la observación y análisis de contextos y circunstancias y, actividades dialógicas con los sujetos actores; privilegiando el carácter interpretativo que busca ante todo una lectura que permita la comprensión y exégesis del sentido y las prácticas humanas en la circunstancia estudiada, más que su mera medición o presentación en términos cuantitativos. Se concluye haciendo énfasis en la importancia que tiene la evaluación como estructura de apoyo al aprendizaje, ya que, si no se evalúa, no se mejora; pero también si se evalúa de forma incorrecta se entorpece el proceso de mejora del hecho educativo.

**Palabras clave:** Evaluación del aprendizaje, Probabilidad y estadística, Cultura Digital, Educación.

### **Abstract**

This research work, which was carried out from a broader perspective to have an articulated vision of the teaching-learning process of probability and statistics, has the purpose of presenting as a referential structure the contemporary aspects linked to digital technologies, communication dynamics and sociocultural that frame the evaluative practice in the subject of probability and statistics that is taught at the undergraduate level in the engineering careers of the Technological Institutes. Key aspects are also addressed to clarify that the evaluation of learning is a much broader and deeper concept than the one of qualification and accreditation, useful for the reconversion of the evaluative practices of teachers. With respect to the methodology, the descriptive-explanatory transdisciplinarity method was used from the critical analysis of documents, the observation and analysis of contexts and circumstances, and dialogical activities with the actor subjects; privileging the interpretative character that seeks above all a reading that allows the understanding and explanation of the human sense and practices in the circumstance studied, more than its mere measurement or presentation in quantitative terms. It is concluded by emphasizing the importance of evaluation as a support structure for learning, because if it is not evaluated, it is not improved; but also if it is evaluated incorrectly the process of improvement of the educational fact is hindered.

**Keywords:** Learning assessment, Probability and statistics, Digital Culture, Education.

## 1.1 Introducción

En el proceso educativo, enseñar e investigar son dos prácticas educativas que tienen saberes-prácticos, complejos, con objetivos particulares: la primera consiste en promover de manera crítica e innovadora conocimientos, habilidades, actitudes y valores; la segunda consiste más bien en generar conocimientos nuevos, en articular antecedentes para la toma de decisiones, en descubrir alternativas para solucionar problemas. Juntas pueden concebirse como un binomio imprescindible orientado hacia el logro de la idoneidad del proceso educativo en el aula. En este sentido, cabe mencionar que el trabajo se desprende de una investigación más amplia acerca de los procesos de enseñanza y aprendizaje contemporáneos en la matemática desde una perspectiva investigativa en educación matemática. Pretende dar una visión a manera de estructura referencial que proporcione elementos al profesor de probabilidad y estadística para resignificar su práctica evaluativa orientada hacia una idoneidad del proceso de estudio que tenga mayor impacto en el aprendizaje de los estudiantes.

## 1.2 Consideraciones del marco del estudio, orientadas a la búsqueda de referentes y reflexiones para una práctica docente resignificada

Algunas de las preguntas centrales que hoy en día se formulan los profesores: ¿qué se entiende por evaluación del aprendizaje?, ¿por qué y para qué evaluar?, ¿qué evaluar? De modo general, la tendencia actual es la de concebir a la evaluación desde una perspectiva comprensiva en cuanto a su objeto, funciones, metodología y técnicas, participantes, condiciones, resultados, efectos y determinantes. La mayoría de las definiciones actuales coinciden en reconocer, como procesos básicos de la evaluación, la recogida de información y la emisión de un juicio valorativo. En este sentido, para los propósitos del estudio se ha considerado que la evaluación, y el acto evaluativo como unidad, suponen operaciones o subprocesos que van desde el establecimiento de los objetivos

o propósitos, la delimitación y caracterización del objeto de evaluación, la definición y aplicación de los instrumentos para la recogida de información, el procesamiento y análisis de dicha información, su interpretación y expresión en un juicio evaluativo, la retroinformación y toma de decisiones derivadas de él, su aplicación y la valoración de resultados.

### 1.3 Proceder del estudio

El proceder metodológico del estudio se llevó a cabo desde una óptica más amplia para tener una visión articulada del proceso enseñanza-aprendizaje de la probabilidad y estadística. En el proyecto macro se recurrió al método de la transdisciplinariedad descriptiva-explicativa a partir del análisis crítico de documentos, la observación y análisis de contextos y circunstancias y, actividades dialógicas con los sujetos actores (Ver figura 1).



Figura 1. Lectura para la comprensión e interpretación del objeto de estudio (Elaboración propia)

Es importante destacar, que a diferencia de los estudios que se respaldan en métodos cuantitativos, en lo que a investigación se refiere, los estudios que asumen un método con carácter interpretativo buscan ante todo una lectura que permita la comprensión e interpretación del sentido y las prácticas humanas en la circunstancia estudiada, más que su mera medición o presentación en términos cuantitativos.

#### **1.4 Resultados**

Con respecto a los resultados encontrados, se organizaron en las siguientes categorías: Docencia contemporánea; Tecnologías digitales, nuevas dinámicas comunicativas y socioculturales; y Evaluación del aprendizaje: referentes y reflexiones para una práctica resignificada.

En lo que respecta a Docencia contemporánea, el quehacer docente podemos entenderlo como un proceso con intención precisa y carácter formal de interacción que pone en juego una serie de métodos y técnicas particulares orientadas al logro de aprendizajes, y que, finalmente no puede ser estudiado al margen de factores institucionales, escolares, ideológicos, sociales y psicológicos (Chadwick y Rivera, 1997). Es en este sentido, que la función docente requiere de organizar el proceso educativo en el aula con el fin de establecer un proceso pedagógico que garantice su eficacia. De allí que una de las principales funciones del profesor es la evaluación, una tarea ineludible y permanente, compleja y dinámica que exige la adopción de una actitud reflexiva y crítica con respecto a la realidad educativa para transformarla creativamente. En esta tesitura, el profesor hace de su práctica educativa objeto de estudio dando inicio al desarrollo de un proceso investigativo sobre este fenómeno educativo de su labor docente (Amigues y Zerbato, 1999).

Diversos aspectos, como el gran avance en las ciencias, tecnologías y la diversidad de paradigmas y enfoques educativos, plantean al profesor un reto que exige conocer estos referentes que le permitan situarlo con mayor fundamento en esta realidad compleja y dinámica.

Desde hace unas décadas se han observado cambios vertiginosos en cuanto a tecnología se refiere, las Tecnologías de la Información y Comunicación (TIC) no se han quedado atrás; más aún, han despertado amplias expectativas sobre la transformación que sufrirán las prácticas sociales en un futuro cercano. En este sentido, el contexto educativo no es ajeno a estas prácticas que los estudiantes han apropiado y traído al aula desde afuera, desde la sociedad global en la que viven y en donde se originan nuevas y cambiantes prácticas socio-culturales resultado de la tecnologización e informatización del mundo actual.

Con relación a la categoría de Tecnologías digitales, nuevas dinámicas comunicativas y socioculturales, en los años recientes se han observado acelerados cambios tecnológicos en todos los campos, incluyendo el de las TIC. En este campo, como lo señala Breton (1992) los cambios cuantitativos (más digitalización, más informatización, más medios) vendrán a provocar cambios de orden cualitativo en lo social asociados al cambio tecnológico. En este sentido, en la sociedad contemporánea, los medios tecnológicos hacen posible que la información pueda transportarse e integrarse a una red global de intercomunicación, que como lo menciona Lévy (1997) crea un espacio para la inteligencia colectiva bajo el signo del encuentro y la emancipación de obstáculos en la producción de conocimientos, concepción retomada de la conectividad global propuesta por McLuhan y Powers (1989). De la misma manera, indica que la inteligencia distribuida puede ponerse en contacto en el espacio virtual, y es a partir del principio de colaboración solidaria, que se da origen a un proceso sinérgico cuyo resultado es el saber colectivo.

En esta tesitura, se han establecido nuevos parámetros espacio-temporales, se modifican y virtualizan procesos sociales, se originan nuevas formas de producir y compartir saberes, y por ende, se transforman además, las dinámicas en el contexto escolar. La nueva forma de sociedad que se está constituyendo está caracterizada por diferentes formas de trabajo y de entretenimiento, diferentes clases de relaciones interpersonales y maneras diversas de comunicarse unos con otros. Por ello la educación debe saber adelantarse y formar a cada individuo de acuerdo a las destrezas, actitudes y contenidos necesarios para responder a este reto, preparándose así para afrontar el

futuro, como lo señala Castellano (2011). Debe prepararse a cada persona para que sepa responder a los nuevos desafíos que la sociedad plantea, en donde lo único constante es el cambio.

Bauman (2008) señala que en la modernidad líquida existen nuevas relaciones entre la sociedad y la educación jamás vistas en la historia humana, ya que las propias condiciones de la modernidad líquida han generado tiempos difíciles para la educación que aún continúa anclada a sus prácticas de la modernidad sólida. En el mundo de la modernidad líquida, la solidez de las cosas, como ocurre con la solidez de los vínculos humanos se diluye, lo duradero ya no lo es, ahora es escurridizo, volátil. En el pasado, la educación adquiría muchas formas y pudo ajustarse a las cambiantes circunstancias, fijándose nuevos objetivos y diseñando nuevas estrategias. El cambio actual no es como los cambios del pasado, ahora se vive un mundo sobresaturado de información, de tecnologías que conlleva a nuevas formas de identidad de los sujetos que serán educados.

Así, el acercamiento que tienen los estudiantes a las tecnologías se ve mediado y significado por formas de socialización tecnológica en distintos entornos, como el del hogar, el educativo, así como por la experiencia previa con otras herramientas cotidianas. En este sentido resulta relevante el estudio de la apropiación social de las tecnologías y en consecuencia el análisis de las dinámicas de aplicación escolar, orientado hacia el diseño de los procesos de aprendizaje. Por otra parte, conocer los procesos por los que los jóvenes estudiantes transitan en su ámbito social, permite al docente tomarlos en cuenta como un elemento que empodere el proceso de aprendizaje y en consecuencia propicie la apropiación de conocimientos con mayor significado, pero sobre todo aprovechando los saberes que la era de la información posibilita.

En ese mismo orden de ideas, la tecnología está tan inmersa en la vida cotidiana, que cada vez resulta más difícil desarrollar una vida social y productiva activa sin un mínimo manejo de los recursos tecnológicos digitales. En este sentido, más allá de la incorporación de las tecnologías digitales en las aulas como recurso didáctico, es de suma importancia incorporar la enseñanza de las habilidades necesarias para desenvolverse en forma autónoma y crítica en esta nueva realidad caracterizada por la inserción de los medios digitales y el acceso masivo a la información. Es de



esta manera, que no se trata de analizar si es o no conveniente incorporar la tecnología en las aulas, sino como señala Cukierman y Virgili (2010) que es tan inevitable hoy utilizar tecnologías de diversos tipos en la educación como siempre lo fue el libro, el lápiz y el pizarrón, al igual que la computadora, la videograbadora o el último dispositivo de comunicación móvil. La pregunta no es si se ha de usar tecnología, ni siquiera qué tecnologías usar, sino por qué y cómo usarlas.

Con respecto a la Evaluación del aprendizaje vista desde sus referentes y reflexiones para una práctica resignificada, existen discrepancias en la extensión del concepto, y por tanto del proceso evaluativo en su totalidad, en lo que se refiere a la inclusión o no del juicio valorativo y de la toma de decisiones derivadas de la información y valoración que se realizan, así como de la ejecución de esas decisiones y sus resultados. Sin duda, la obtención sistemática de información o evidencias sobre el objeto de evaluación y la elaboración y formulación de un juicio fundamentado sobre dicho objeto, constituyen procesos básicos, centrales de la evaluación, pero no agotan el acto evaluativo y por lo tanto no son suficientes para el análisis del mismo.

De esta manera, la evaluación, y el acto evaluativo como unidad, suponen operaciones o subprocesos que van desde el establecimiento de los objetivos o propósitos, la delimitación y caracterización del objeto de evaluación, la definición (selección, elaboración) y aplicación de los instrumentos para la recogida de información, el procesamiento y análisis de dicha información, su interpretación y expresión en un juicio evaluativo, la retroinformación y toma de decisiones derivadas de él, su aplicación y valoración de resultados (Díaz y Hernández, 1998). Para reiniciar en un ciclo ascendente, progresivo, que permita, en su dinámica, imprimir el auténtico significado de esta actividad, contribuir al aprendizaje y a la mejora de la propuesta didáctica que plantea el profesor.

El objetivo de la evaluación del aprendizaje, como actividad genérica, es valorar el aprendizaje en su proceso y resultados. Las finalidades o fines marcan los propósitos que signan esa evaluación. Las funciones se refieren al papel que desempeña para la sociedad, para la institución, para el proceso de enseñanza-aprendizaje, para los individuos implicados en éste (López e Hinojosa,

2000). La determinación de los resultados del aprendizaje y la calidad de éstos. Si se ha alcanzado o no el aprendizaje esperado (u otros no previstos) y qué características o atributos posee, de acuerdo con los criterios asumidos a tal fin.

En tal caso la principal función es la de comprobación de resultados. La constatación de los resultados o productos es una función legítima de la evaluación, aunque no suficiente, también aporta información para acciones de ajuste y mejoras del proceso, a más largo plazo, al contrastar los resultados con las necesidades que le dieron origen, por lo que no se excluye su vínculo con la retroalimentación y regulación de la actividad (Florez, 1999). Se supone que la evaluación de los resultados tenga también una proyección futura y no solo retroactiva. Esto es, sirve de base para hacer predicciones sobre el posterior desempeño académico y profesional de estudiantes. En este sentido se habla de una función de predicción, tan cuestionada como asumida (Mateo, 2000). Además de proporcionar información que permita la orientación y regulación del proceso de enseñanza-aprendizaje. Estas funciones son muy valoradas actualmente y constituyen un pilar para fundamentar la concepción de la evaluación como parte del proceso enseñanza-aprendizaje. En este sentido la evaluación es un elemento necesario para realizar el proceso.

Por otra parte, la delimitación del objeto que se evalúa es un asunto central. De ella se deriva, en gran medida, las decisiones sobre cómo se realiza la evaluación: los instrumentos, procedimientos, momentos, indicadores, criterios, que se utilicen en el proceso evaluativo.

Si el aprendizaje tiene como función principal la potenciación de las capacidades cognitivas del individuo y que involucra procesos complejos de procesamiento y asimilación de información recogida fundamentalmente a través de los sentidos (Caracheo, 2000), la comprobación de estos resultados implica un proceso que permita valorar, juzgar o estimar si se ha logrado el propósito o no y en qué grado, este proceso se conoce como evaluación.

Para Casanova (2007), la evaluación aplicada a la enseñanza y el aprendizaje consiste en un proceso sistemático y riguroso de obtención de datos, incorporado al proceso educativo desde su comienzo, de manera que sea posible disponer de información continua y significativa para conocer la situación, formar juicios de valor con respecto a ella y tomar las decisiones adecuadas para proseguir la actividad educativa mejorándola progresivamente.

De acuerdo a los resultados encontrados en las categorías anteriores, puede decirse que, la práctica introspectiva y crítica de los profesores hacia su propia práctica individualizada y colegiada, estimula un nivel de análisis y reflexión profundos que conduzcan a que cobren total conciencia acerca de cuáles son los sustentos teóricos de sus prácticas evaluativas y, en un momento posterior, puedan establecer los cambios que requieren hacer en éstas, para que realicen una práctica evaluativa congruente con los fundamentos del Modelo educativo vigente en la institución en la que realizan su labor docente. Estos cambios que se requieren en sus nuevas y resignificadas prácticas evaluativas deben integrar elementos tanto conceptuales como de procedimiento y actitudinales para su correcta aplicación, ya que un cambio en su práctica evaluativa implica la necesidad de cambios coordinados en todas las partes del proceso educativo en el aula, como podemos observar en la figura 2 en donde el dominio de las tecnologías digitales, los conocimientos disciplinares, los conocimientos pedagógicos y un conocimiento amplio acerca de las dinámicas comunicativas y socioculturales, conforman aspectos de suma importancia como competencia profesional para el docente.

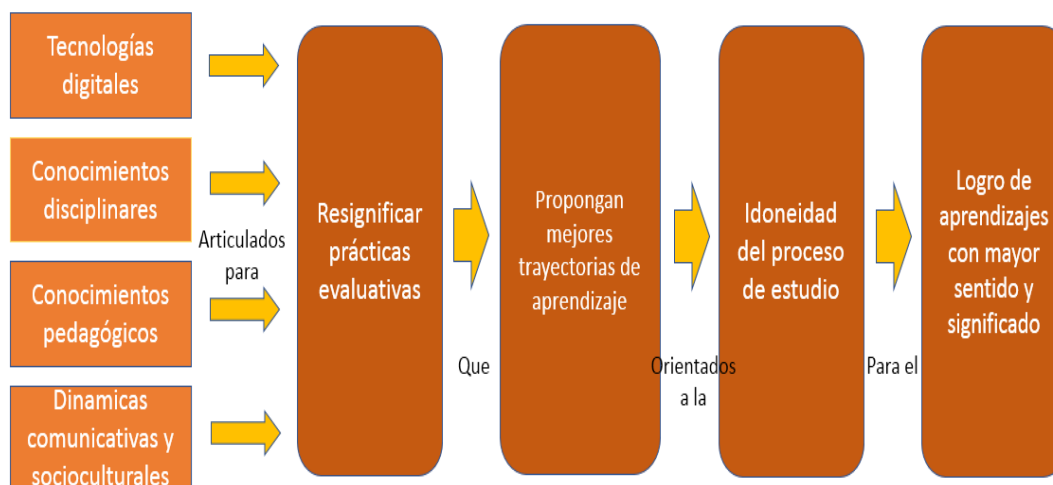


Figura 2. Referentes contemporáneos para las prácticas evaluativas (Elaboración propia).

En esta tesitura, el dominio disciplinar, pedagógico y tecnológico; articulado y orientado hacia el planteamiento de trayectorias de aprendizaje de los alumnos, es un aspecto de suma importancia como competencia profesional para el docente. De manera tal, que plantea y posibilita, como señala Díaz Barriga (2003), una transformación en el proceso de aprendizaje del estudiante, la innovación en los métodos de enseñanza, los materiales de apoyo educativo, el ambiente de aprendizaje y por supuesto la evaluación.

## 1.5 Conclusiones

A manera de cierre, es importante clarificar que la evaluación del aprendizaje es un concepto mucho más amplio y profundo que el de examen, medición o calificación, de los que se apoya en última instancia para la interpretación de los resultados. No debe confundirse a la evaluación con una simple calificación, por ejemplo de un examen. El propósito fundamental de la evaluación es servir como estructura de apoyo al aprendizaje, como también verificar si el aprendizaje ha sido logrado o no, con independencia de los usos formales o institucionales a partir de la información recolectada por el proceso de evaluación, la cual debería permitir valorar no solo el aprendizaje en

sí mismo, sino valorar la eficacia de las actividades de aprendizaje, la dificultad mostrada para el aprendizaje de los contenidos, el tiempo en el que este se ha logrado, la idoneidad del ambiente y materiales de apoyo al proceso educativo así como los estilos de aprendizaje y las estrategias adoptadas por los propios estudiantes.

Por otra parte, dentro de la educación formal en un sistema escolarizado la evaluación tiene doble propósito, además de verificar el aprendizaje de los estudiantes, proporciona los insumos necesarios para la asignación de notas o calificaciones los cuales a su vez permitirán tomar decisiones acerca de la acreditación de la asignatura correspondiente.

Indudablemente, aunque esta última no constituye el fin primordial del proceso de evaluación, al parecer, aún queda un largo camino por recorrer antes de cambiar los paradigmas ligados al proceso de evaluación, calificación y acreditación.

Finalmente se hace necesario un proceso reflexivo que oriente el quehacer del docente en relación al proceso de evaluación, estableciendo apropiadamente los momentos, instrumentos y uso de la información recolectada, en el que invariablemente la naturaleza de los contenidos, el contexto y las características de los estudiantes son factores esenciales que el profesor debe considerar.

## Referencias

- Amigues, R. y Zerbato, P. M. (1999). *Las prácticas escolares de aprendizaje y evaluación*. México: FCE.
- Bauman, Z. (2008). *Los retos de la educación en la modernidad líquida*. Barcelona: GEDISA.
- Breton, A. (1992). *Oeuvres complètes*. Vol. II. Paris: Gallimard.
- Caracheo, F. (2000). *Los principios del aprendizaje*. Documento inédito. CIIDET, México.
- Casanova, M. A. (2007). *Manual de evaluación educativa*. 9ª Edición. Madrid: La Muralla.
- Castellano, H. (2011). *Enseñando con las TIC*. Buenos Aires: Cengage.
- Chadwick, C. B. y Rivera, N. (1997). *Evaluación formativa para el docente*. México: Paidós Educador.
- Cukierman, U. y Virgili, J. (2010). *La tecnología educativa al servicio de la educación tecnológica: experiencias e investigaciones en la UTN*. Buenos Aires: edUTecNe.
- Díaz, F. (2003). Cognición situada y estrategias para el aprendizaje significativo. *Revista Electrónica de Investigación Educativa*, 5 (2). Consultado en: <http://redie.ens.uabc.mx/vol5no2/contenido-arceo.html>
- Díaz, F. y Hernández, R. G. (1998). *Estrategias docentes para un aprendizaje significativo. Una interpretación constructivista*. 2ª. Edición. México: Mc Graw Hill.
- Flórez, O. R. (1999). *Evaluación pedagógica y cognición*. Bogotá: Mc Graw-Hill Interamericana.
- Lévy, P. (1997). *L'intelligence collective: pour une anthropologie du cyberspace*. Paris: Editions La Découverte.
- López, B. e Hinojosa E. (2000). *Evaluación del aprendizaje. Alternativas y nuevos desafíos*. México: Trillas-ILCE-ITESM.
- McLuhan, M. Y Powers, B. (1989). *The Global Village: Transformations in World Life and Media in the 21st Century (Communication and Society)*. Nueva York: Oxford University Press.

Mateo, J. (2000). *La evaluación educativa, su práctica y otras metáforas*. Cuadernos de Educación.  
Barcelona: Universitat de Barcelona - Editorial Horsori.



---

## Drop the Beta

---

María Paula Mesén Méndez, Paola Elena Zamora Charpentier  
mariapaulamesnmendez@gmail.com, pao.zch@gmail.com.



### Resumen

Drop the Beta es una actividad interactiva en la cual se busca que el usuario comprenda, con ayuda del computador y de un tablero de juego, los conceptos de: interacción entre variables categóricas y cuantitativas, procesos de eliminación de variables en modelos de regresión usando el Criterio de Información de Akaike (AIC), la relación entre la verosimilitud y el AIC obtenido al realizar la selección de variables y la relación entre la cantidad de variables de un modelo y su AIC.

Este juego didáctico utiliza tres modelos de regresión, a los cuales se le deben adicionar o eliminar variables e interacciones conforme el jugador avanza en el tablero. Estos tres modelos buscan predecir: (1) el gasto en colones que una familia de clase media costarricense dedicaría por



mes a viajar en Uber; (2) la cantidad de horas semanales que una familia de clase media costarricense dedica a ver televisión; y, (3) el gasto mensual en colones que una familia de clase media costarricense dedicaría por mes en la compra de alimentos. Cada uno de estos modelos está representado por un peón, y pueden jugar hasta tres personas simultáneamente.

El jugador inicialmente conforma su modelo con seis variables cuantitativas y una cualitativa, las cuales debe seleccionar al azar de una pila de tarjetas que contienen el nombre y el número de la variable que debe asignarse a su modelo. Una vez que el jugador tenga su modelo puede comenzar a jugar. En cada turno el jugador debe decidir qué variable eliminar de su modelo usando el Criterio de Información de Akaike, el cual se puede observar en la computadora al correr una simulación creada especialmente con ese propósito. El jugador debe eliminar la variable con el mayor AIC, pero debe respetar la jerarquía del proceso de eliminación de variables. Si se toma una mala decisión, el jugador debe responder correctamente una pregunta para evitar que se le agregue a su modelo una variable o interacción adicional. Una vez que un jugador llegue a la última casilla del tablero los jugadores deben revelar el AIC de su modelo y el ganador es aquel cuyo AIC sea menor.

**Palabras clave:** Criterio de Información de Akaike (AIC), Regresión, Interacción, Simulación, Juego didáctico.

### **Abstract**

Drop The Beta is an interactive activity in which the main objective is for the user to understand, with the aid of a computer and a game board, the concepts of: interaction between qualitative and quantitative variables, the regression model variable elimination process using Akaike Information Criterion (AIC), the relationship between likelihood and the AIC obtained on the process of variable selection and the relationship between the amount of variables present on a regression model and its AIC.

This educational game uses three different regression models, to which variables and interactions must be added or eliminated as the player goes forward on the game board. These three models predict: (1) the amount of money in colones a middle-class Costa Rican family spends monthly on Uber; (2) the number of hours per week a middle-class Costa Rican family watches television.; and (3) the amount of money in colones a middle-class Costa Rican family spends monthly on food. Each model has a distinctive game piece and the game can be played simultaneously by three people.

The player initially forms his models by choosing six quantitative and one qualitative variable, all of them must be selected randomly from a card pack. Each of the variable cards has a distinctive name and number. Once the player has formed his model he can start playing. Each turn the player must choose which variable to eliminate from his model using Akaike Information Criteria, which can be seen on the computer using a simulation created especially for this. The player must choose to eliminate the variable with the highest AIC, but he must respect the hierarchy of the elimination process. If a wrong decision is made, the player must answer a question correctly to prevent another variable or interaction be added to his model. Once a player has reached the last box of the board, the players must reveal their model's AIC. The winning player is the one whose model has the lowest AIC.

**Key Words:** Akaike Information Criterion (AIC), Regression, Interaction, Simulation, Educational game.

## 2.1 Introducción

El método de regresión lineal es una técnica ampliamente utilizada, la cual permite evaluar con minuciosidad la existencia de relaciones entre distintas variables, facilitando la conexión de diversas áreas de estudio. Este método tiene infinidad de aplicaciones prácticas en áreas como la producción industrial, el comercio, la economía, la biología, entre otras.

Al ser un método tan minucioso, requiere de gran esfuerzo por parte de la persona que lo está aplicando, ya que requiere de fuerte conocimiento teórico para poder ser aplicado de la manera adecuada. Es por esta razón que se ha buscado innovar en la enseñanza del método de regresión, para con ello lograr facilitar el proceso de aprendizaje y que este sea más amigable con el usuario.

En la actualidad, se aprovechan mucho los dispositivos electrónicos para facilitar el aprendizaje, sobre todo a través de simulaciones y aplicaciones, sin importar el área de estudio. Según Salas y Ardanza (1995), “la simulación consiste en situar a un educando en un contexto que imite algún aspecto de la realidad y en establecer en ese ambiente situaciones, problémicas o reproductivas, similares a las que él deberá enfrentar”.

Adicionalmente, se pueden utilizar recursos físicos para reforzar el conocimiento aprendido a través de las simulaciones y aplicaciones, de modo que la persona logre aplicar sus conocimientos de forma más sencilla.

## **2.2 Objetivos**

### *Objetivo general*

Reforzar los conocimientos acerca del proceso de selección de variables mediante el criterio de información de Akaike (AIC)

### *Objetivos específicos*

Identificar, mediante la dinámica del juego, las variables y las interacciones que deben ser eliminadas de un modelo, utilizando el método de AIC.

Verificar los conocimientos del jugador sobre la importancia de realizar un adecuado proceso de selección de variables.

A través de la actividad didáctica que se plantea en este trabajo se busca abarcar los siguientes conceptos:

- I. Variable categórica o predictores cualitativos.
- II. Interacción.
- III. Criterio de información de Akaike (AIC).
- IV. Proceso de selección de variables por pasos.

### 2.3 Conocimientos previos

La actividad está planteada para estudiantes de las áreas de estadística, economía, ingeniería, negocios y matemática que tengan o requieran conocimiento sobre los siguientes temas:

- I. Concepto de predictores cualitativos en un modelo.
- II. Concepto de interacción entre una variable categórica y una numérica.
- III. Concepto de eliminación hacia atrás en procesos de selección de variables.
- IV. Relación entre la función de máxima verosimilitud y la suma de cuadrados residual.
- V. Concepto de devianza.
- VI. Concepto de criterios de información.

La presente sección aborda diversos conceptos necesarios en un análisis de regresión, los cuales serán descritos a continuación:

- I. Predictores cualitativos: los predictores cualitativos son aquellas variables que se refieren a características, cualidades o estado de un objeto que no pueden ser medidas con números, por ejemplo: sexo, estado civil, entre otras. El objetivo de utilizar variables categóricas en un análisis de regresión es poder cuantificar los datos categóricos mediante la asignación de valores numéricos a las categorías (IBM, s.f.).
- II. Interacción: se dice que hay interacción entre una variable categórica y una numérica cuando la pendiente de una categoría es muy diferente a la pendiente de otra categoría, lo

que implica que si no hay interacción es porque las rectas de ambas categorías son paralelas y ambas variables decrecen o aumentan en la misma cantidad.

- III. Criterio de Información de Akaike (AIC): es un estimador muestral de la esperanza del logaritmo natural de la verosimilitud, cuyo objetivo principal es la penalización por la tenencia de un exceso de parámetros ajustados (Caballero, 2011).

La fórmula para calcular el AIC de un modelo es la siguiente:  $-2 \cdot \ln(\text{verosimilitud}) + 2 \cdot p$ . Como se puede notar, el modelo utiliza el logaritmo del estimador de máxima verosimilitud con signo negativo, al cual se le adiciona un “castigo” que depende de la cantidad de parámetros con la que cuente el modelo. Por lo tanto, si el excedente de parámetros no es muy grande, al restarle 2 veces el logaritmo natural de la verosimilitud se obtiene un menor AIC.

- IV. Proceso de selección de variables paso a paso: existen dos formas de llevar esto a cabo, ya sea eliminando variables de un modelo “grande” o adicionando variables a un modelo que solo cuenta con el intercepto.

La selección hacia atrás es el proceso en el cual se inicia con un modelo completo con todos los predictores, y además se establece un valor alfa ( $p$ -para remover) el cual se utilizará como referencia para decidir si eliminar o no, alguna variable. Se eliminará el predictor que tenga el valor más alto de “ $p$ ” en comparación al alfa. Este proceso se repite hasta que ya no haya más predictores que eliminar.

La selección hacia adelante es el proceso en el cual se agregan variables secuencialmente en el modelo que cumplan con tener un valor “ $p$ ” menor en comparación con el alfa. Este proceso se repite hasta que ya no se puedan agregar más predictores que cumplan con dicho criterio.

Cabe destacar que estos métodos de selección pueden dar como resultado modelos distintos.

## 2.4 Desarrollo del Juego

### 2.4.1 Materiales

Los materiales necesarios para llevar a cabo la dinámica planteada en este trabajo son:

- I. Juego de simulación con el software R Studio.
- II. Tablero del juego.
- III. 1 dado.
- IV. 3 figuras (peones) referentes a las variables respuesta (carro, hamburguesa, televisor).
- V. 12 tarjetas de recompensa (6 “Elimine una variable” y 6 “Elimine una interacción”).
- VI. Tarjetas con los nombres y números correspondientes de cada variable que puede usar el jugador para la construcción del modelo.
- VII. 12 preguntas referentes al tema de selección de variables

### 2.4.2 Descripción del recurso

#### *a. Juego de simulación*

Se creó una base de datos utilizando funciones de R para generar aleatoriamente datos con distribuciones binomial, normal o Poisson dependiendo de cada variable. Algunas de estas variables se generaron basándose en otras, para forzar una cierta correlación entre las mismas. Las variables respuesta se generaron determinando arbitrariamente cuánto iba a modificar la respuesta cada una de las variables predictoras en una escala del 0 al 5.

La simulación requiere que el jugador ingrese el número de identificación de las variables, el número de interacciones obtenido al lanzar el dado y el nombre de las variables para realizar las actualizaciones correspondientes para eliminar variables, utilizando el AIC. Sin embargo, el usuario no debe realizar ningún tipo de programación. Referirse al manual del juego para más detalles (anexo 1).

El jugador debe aplicar su conocimiento para poder responder adecuadamente las preguntas planteadas y para eliminar de forma correcta las variables o interacciones que correspondan según el criterio de información de Akaike.

Para la implementación de la simulación se requieren 4 archivos:

- I. Generación de base de datos
- II. Funciones
- III. Creación del modelo
- IV. Drop the Beta

Para iniciar la dinámica, el jugador debe abrir los cuatro archivos anteriormente mencionados, correr el *script* completo de “Generación de base de datos” seguido de “Funciones”. Una vez hecho esto, el jugador puede cerrar esas pestañas o las puede mantener abiertas, de cualquier forma, ya no las necesitaría. Seguidamente, abre el archivo “Creación del modelo”, donde debe seguir las instrucciones para crear apropiadamente el modelo de regresión que escogió al inicio. Finalmente, abre el archivo “Drop the Beta” para revisar el AIC de su modelo y eliminar o agregar una variable al mismo.

El juego finaliza una vez que algún jugador llegue a la casilla de final, el ganador es el jugador que obtuvo el modelo con el AIC menor.

#### *b. Tablero*

Superficie cuadrangular compuesta por distintas casillas, una casilla de inicio y una de final, además de estas, hay algunas en blanco, cinco con recompensas y cinco con castigos por las cuales el jugador deberá pasar. Al inicio de la dinámica, el jugador debe seleccionar uno de los dos caminos a seguir en el tablero (a partir de la tercera casilla). En el anexo 2 se encuentra el tablero diseñado para esta dinámica.

*c. Preguntas referentes al tema de selección de variables*

Preguntas seleccionadas al azar por los jugadores, las cuales deberán contestar como método de verificación para así medir qué tanto conocen sobre el tema abordado en el juego. Las preguntas planteadas buscan que el jugador utilice los conocimientos que tiene sobre el proceso de eliminación de variables en un modelo de regresión para responderlas adecuadamente; sin embargo, existen algunas preguntas “de trampa”. En la sección “Preguntas Generadoras” se encuentran las preguntas planteadas para esta dinámica.

### **2.4.3 Desarrollo del contenido**

*a. Situación problema*

Se proponen 3 modelos, cada uno de los cuales debe componerse por 6 variables cuantitativas y 1 categórica. Se tiene que definir la cantidad de interacciones que cada modelo tendrá a través del lanzamiento de un dado. La pregunta generadora inicial es: ¿Qué sucede con el AIC al agregar o eliminar variables e interacciones en los modelos?

*b. Preguntas generadoras*

Según indica Castillero (2016), el conocimiento es el resultado del proceso de adquisición de información interna o externa. En palabras de Gagné (1970), “el aprendizaje es la interrelación entre persona y ambiente, siendo un cambio de tipo comportamental, conductual e incluso de disposición o actitud respecto a una parte o la totalidad de la realidad”.

Ambos autores resaltan la importancia de la motivación en el proceso de aprendizaje, ya que esta exige utilizar el aprendizaje almacenado a través de estímulos o situaciones que generen conductas y expectativas en la persona. Es justo a través de esto que se busca llevar a cabo una dinámica que motive al estudiante y propicie el aprendizaje de conceptos estadísticos mencionados anteriormente.



Si bien la dinámica está diseñada para que máximo 3 personas interactúen entre sí conforme se lleva a cabo la actividad, se espera que el docente o moderador actúe como apoyo en el proceso de aprendizaje. El juego incluye las siguientes preguntas:

- I. ¿Falso o Verdadero? Si el AIC es menor a “p”, entonces se puede decir que el modelo no es sesgado. Respuesta: Falso. El sesgo no está relacionado con que el AIC sea menor al valor p.
- II. ¿Falso o Verdadero? Antes de eliminar una variable con interacción usando el AIC se debe eliminar primero las variables individuales de la interacción. Respuesta: Falso. Se debe eliminar primero la interacción.
- III. ¿Falso o Verdadero?  $\alpha$  puede estar entre 0.15 y 0.20 para un modelo de predicción. Respuesta: Verdadero.
- IV. ¿Falso o Verdadero? El mejor modelo siempre es el que tiene menos variables. Respuesta: Falso. Depende de su capacidad predictiva.
- V. Conteste lo siguiente: ¿Qué influye en que el AIC del modelo aumente o disminuya? Respuesta: La cantidad de variables, la calidad predictiva de las variables, la devianza del modelo.
- VI. ¿Falso o Verdadero? Se intenta encontrar un modelo con el mayor AIC. Respuesta: Falso. Se busca el de menor AIC ya que con este se maximiza la verosimilitud del modelo.
- VII. ¿Falso o Verdadero? Si el AIC es alto, la verosimilitud es alta. Respuesta: Falso. El AIC y la verosimilitud tienen una relación inversa.
- VIII. ¿Falso o Verdadero? La maximización de la función de verosimilitud es equivalente a maximizar su logaritmo. Respuesta: Verdadero.
- IX. ¿Falso o Verdadero? El Criterio de Información de Akaike se usa solamente en selección de variables hacia adelante. Respuesta: Falso. Se usa para selección de variables hacia adelante y hacia atrás.

- X. ¿Falso o Verdadero? Se prefiere el Criterio de Información de Akaike por ser más conservador que el BIC. Respuesta: Falso. Al contrario, el BIC es más conservador.
- XI. ¿Falso o Verdadero? Se prefiere el uso del AIC sobre el BIC porque el primero no requiere del cumplimiento del supuesto de linealidad. Respuesta: Falso. Ninguno requiere del cumplimiento de dicho supuesto.
- XII. ¿Falso o Verdadero? El principio de Parsimonia indica que si se tienen dos modelos se prefiere al más pequeño. Respuesta: Falso, se prefiere el más pequeño si el poder predictivo de este es igual o mayor que el del grande.

*c. Desarrollo de las situaciones de aprendizaje*

Según González (s.f.), las situaciones de aprendizaje involucran la creación de espacios físicos y virtuales con condiciones que estimulen las actividades por realizar y el pensamiento de los alumnos. Para efectos del desarrollo del juego didáctico se abordará en tres pasos centrales: propuesta del problema (1), trabajo estudiantil independiente (2), clausura (3).

1. Propuesta del problema (10 minutos)

El docente o moderador hará presentación del tema y para apoyo de esto se hará uso de los elementos descritos en la sección 6 (materiales). El moderador hará una breve explicación de la situación inicial de cada jugador, presentará la simulación por utilizar y hará un breve recorrido por el diseño del tablero de juego.

El moderador podrá propiciar un espacio en el que él mismo puede hacer uso rápido de la simulación (introducción de variables) para que al iniciar la dinámica el jugador esté más familiarizado con el uso de la misma.

2. Trabajo estudiantil independiente (40 minutos)

Una vez que el moderador haya hecho una introducción al tema, se dispondrá a dividir a los estudiantes en grupos de máximo 3 integrantes para llevar a cabo la dinámica propuesta. Se explicarán las reglas del juego de forma general y se debe evacuar todas las dudas que haya antes

de iniciar la dinámica. Se hará entrega de una hoja en blanco a cada estudiante, la cual debe ser usada como apoyo y método de control de los cambios que se realicen en el modelo a lo largo del juego.

De aquí en adelante se propone dejar que los estudiantes sean los que se encarguen de llevar a cabo la dinámica, y que el docente funja como apoyo a lo largo del proceso. Los estudiantes serán los responsables de agregar o quitar variables e interacciones a sus modelos bajo su propio criterio.

Se debe fomentar un espacio de discusión abierto en cada grupo, en el cual los participantes estén al tanto de los cambios que se realicen en todos los modelos que se estén trabajando; no obstante, cada participante debe responder individualmente y de forma correcta las preguntas que se le planteen.

### 3. Clausura (15 minutos)

Una vez concluido el juego se realiza una revisión de los modelos con el que cada jugador finaliza, para proceder con la verificación de los valores del AIC correspondiente a cada modelo y así conocer el ganador de la dinámica.

Además, se dispone de un espacio para realizar una retroalimentación sobre las respuestas de las preguntas formuladas durante el juego, donde el docente sea el que explique el porqué de las respuestas. Con esto se busca reforzar aquellos conceptos, ideas o conocimientos sobre el tema evaluado (selección de variables) que no hayan quedado claros durante la dinámica, tales como: criterios de información, verosimilitud, interacción entre variables, entre otros.

## **2.5 Sugerencias didácticas**

Se recomienda al docente para el trabajo de la unidad didáctica en el aula lo siguiente:

- I. Permitir que los estudiantes se formen en grupos de máximo 3 personas, donde cada uno seleccione el modelo que más le interese.

- II. Asignar una computadora a cada grupo para poder utilizar el recurso digital (simulación) como apoyo del juego físico.
- III. Entregar hojas blancas a cada uno de los participantes de modo que cuenten con un soporte físico para llevar a cabo la actividad y puedan llevar un mejor control de los cambios que realicen a sus modelos.
- IV. Propiciar la discusión grupal a lo largo de la actividad, sin embargo, las respuestas a las preguntas planteadas en el juego deben ser respondidas de manera individual.
- V. El docente debe finalizar la actividad al declarar ganador al participante cuyo modelo tenga menor AIC. Además, puede generar un espacio de comentarios y de aclaración de posibles dudas al final de la dinámica.

## 2.6 Glosario

Predictores cualitativos: los predictores cualitativos son aquellas variables que se refieren a características, cualidades o estado de un objeto que no pueden ser medidas con números, por ejemplo: sexo, estado civil, entre otras. El objetivo de utilizar variables categóricas en un análisis de regresión es poder cuantificar los datos categóricos mediante la asignación de valores numéricos a las categorías (IBM, s.f.).

Interacción entre una variable categórica y una numérica: se dice que hay interacción entre una variable categórica y una numérica cuando la pendiente de una categoría es muy diferente a la pendiente de otra categoría, lo que implica que si no hay interacción es porque las rectas de ambas categorías son paralelas y ambas variables decrecen o aumentan en la misma cantidad.

Eliminación hacia atrás en procesos de selección de variables: proceso utilizado en la selección de variables en el cual se inicia con un modelo completo con todos los predictores, y además se establece un valor alfa (p-para remover) el cual se utilizará como referencia para decidir si eliminar o no, alguna variable. Se eliminará el predictor que tenga el valor más alto de “p” en comparación al alfa. Este proceso se repite hasta que ya no haya más predictores que eliminar. Además, es importante recordar que en el proceso de eliminación de variables no se debe eliminar variables

sin antes eliminar las interacciones en las que están participan, esto para propiciar la obtención del modelo más “óptimo”.

Relación entre la función de máxima verosimilitud y la suma de cuadrados residual: a través de la búsqueda de la función de máxima verosimilitud se obtiene la estimación con menor suma de cuadrados residual, esto evidencia que la relación existente entre ellas es inversa. Al obtener la máxima verosimilitud, se puede realizar la transformación de aplicarle logaritmo natural a dicho resultado, de modo que el logaritmo también corresponde a un máximo (Vela, s.f.), manteniendo la relación mencionada anteriormente.

El impacto de esto en el uso del Criterio de Información de Akaike (AIC) es que se utiliza dicha transformación multiplicada por un coeficiente, de modo que se obtiene una expresión llamada devianza, para conocer el AIC del modelo. El AIC y el logaritmo natural de la verosimilitud tienen una relación inversa, de modo que a mayor verosimilitud, menor es el AIC y menor es la suma de cuadrados residual del modelo.

Devianza: se conoce como devianza la expresión  $-2 \cdot \ln(\text{verosimilitud})$ . Esta posee la misma dirección que la suma de cuadrados residual, por lo que en un modelo se busca minimizar la devianza.

Criterios de información: el objetivo de los criterios de información es calcular una medida que indique que tan próximos están los modelos alternativos al verdadero modelo, para así seleccionar al más “óptimo”.

Según Solera (2000), “los criterios de selección de modelos miden el “ajuste” de un modelo dado, maximizando el valor de la función de máxima verosimilitud con el uso de diferentes funciones de costos para tomar en cuenta el hecho de que diferentes números de parámetros desconocidos pudieron haber sido estimados para diferentes modelos bajo consideración.”

## Referencias

Caballero D, F.F. (2011) *Selección de modelos mediante criterios de información en el análisis factorial. Aspectos teóricos y computacionales*. Recuperado desde: <https://hera.ugr.es/tesisugr/19964808.pdf>

Castillero M, O (2016) *La teoría del aprendizaje de Robert Gagné*. Recuperado desde: <https://psicologiaymente.net/desarrollo/teoria-aprendizaje-robert-gagne>

Gagné, R (1970) *The conditions of learning*. 4a.ed. Holt, Rinehart and Winston.

González, C (s.f.) *Diseño de situaciones de aprendizaje*. Recuperado desde: <https://sites.google.com/site/situacionesdeaprendizaje/home> IBM (s.f.) *Regresión Categórica*. Recuperado desde: [https://www.ibm.com/support/knowledgecenter/es/SSLVMB\\_22.0.0/com.ibm.spss.statistics.help/spss/categories/idh\\_catr.htm](https://www.ibm.com/support/knowledgecenter/es/SSLVMB_22.0.0/com.ibm.spss.statistics.help/spss/categories/idh_catr.htm)

Kutner, M, Nachtsheim, C, Neter, J, Li, William (2005). *Applied Linear Statistical Models*. 5a. ed. WCB, McGraw -Hill.

Salas P, R.S y Ardanza Z, P (1995) *La simulación como método de enseñanza y aprendizaje*. Recuperado desde: [http://www.bvs.sld.cu/revistas/ems/vol9\\_1\\_95/ems03195.htm](http://www.bvs.sld.cu/revistas/ems/vol9_1_95/ems03195.htm)

Solera R, A (2000) *Criterios para la selección de modelos estadísticos*. Recuperado desde: <http://www.bccr.fi.cr/ndie/NT-07-2000.PDF>

Vela P, F (s.f.) *Máxima Verosimilitud*. Recuperado desde: <https://ceconometria.files.wordpress.com/2011/04/maxima-verosimilitud.pdf>

## 2.7 Anexos

### Anexo 1. Manual de instrucciones del juego Drop the Beta

#### Drop the Beta

##### Introducción:

El juego consiste en seleccionar la mejor aproximación posible de tres distintos modelos de regresión lineal para las variables respuesta: cuánto dinero gasta en comida la familia mensualmente, cuánto dinero gasta la familia en Uber mensualmente y cuántas horas de televisión se ven en total entre todos los miembros de la familia semanalmente. Esto para una muestra de familias de clase media costarricenses.

##### Instrucciones para jugar:

1. Primeramente se abren los archivos: *Generacion de la base de datos.R*, *Funciones.R*, *Creacion del modelo.Rmd* y *Drop the beta.Rmd*. Se corre todo el *script* en los primeros dos y estos no se necesitarán más.
2. Cada jugador deberá lanzar el dado una vez. Se busca que todos obtengan un número diferente así que, en caso de que dos o más jugadores obtengan el mismo número, volverán a lanzar el dado todos los jugadores hasta que todos obtengan un número diferente.
3. Los jugadores tomarán turnos para crear sus modelos. Los turnos se asignan de acuerdo a el orden, de mayor a menor, de los números obtenidos de lanzar los dados. Es decir, el jugador con el valor mayor será el primero en crear su modelo y el jugador con el valor menor será el último. La creación del modelo se hace de la siguiente forma:
  - a. Se abre la pestaña en RStudio de Creación del modelo.
  - b. Se selecciona una de las fichas que representan el modelo por estimar: hamburguesa para el modelo que estima cuánto dinero gasta en comida la familia mensualmente, carro para el modelo que estima cuánto dinero gasta la familia en Uber

mensualmente y televisor para el modelo que estima cuántas horas de televisión se ven en total entre todos los miembros de la familia semanalmente.

- c. Se escogen seis tarjetas de variables predictoras no categóricas. En esta tarjeta viene el nombre de la variable dentro de la base de datos, el número asignado a esta variable y una descripción de lo que significa esta variable. Para las seis variables, se procede a anotar el número asignado a cada variable dentro del vector correspondiente a cada modelo (VariablesComida, VariablesUBER o VariablesTV); cada número separado por una coma del otro, de la siguiente forma: VariablesUBER=c(2,16,5,3,14,10).
  - d. Se barajan de nuevo las tarjetas para que el siguiente jugador haga la selección correspondiente.
  - e. Se selecciona una tarjeta de variables categóricas y se apunta el número asignado después del igual para el valor de variable categórica correspondiente a cada modelo (CategoricaComida, CategoricaUBER o CategoricaTV).
  - f. Se lanza el dado y el número que indique es el número de interacciones que va a tener el modelo. Ese número se anota después del igual para los valores de interacciones para cada modelo (InteraccionesComida, InteraccionesUBER o InteraccionesTV).
4. Cuando todos los jugadores hayan creado su modelo, se corren todos los comandos dentro de la pestaña. Esta pestaña no se necesitará más.
  5. Se abre la pestaña Drop the beta. Ahora comenzará el juego.
  6. Todos ponen su ficha en la casilla de inicio.
  7. Cada jugador tomará turnos para lanzar el dado y avanzar las casillas que indique el mismo, respetando el mismo orden secuencial en el que crearon los modelos. Cada vez que el jugador llega a una nueva casilla, podrá ver el AIC de su modelo y tomar la decisión de

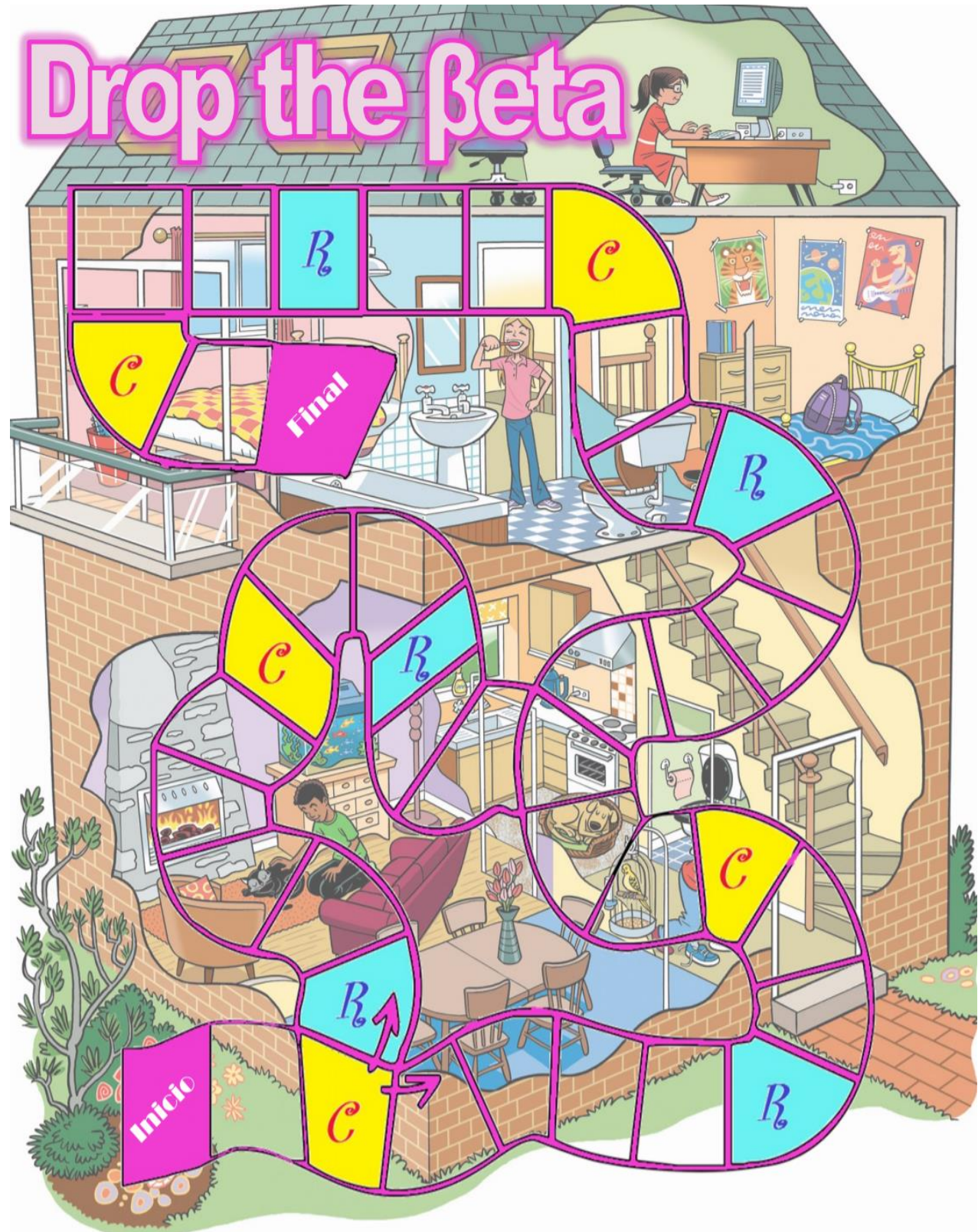


eliminar o no una variable de su modelo, si así lo desea; excepto si cae en casillas especiales; en ese caso deberá respetar lo que represente esta casilla. Existen sólo cuatro tipo de casillas especiales:

- a. Inicio: indica dónde deben empezar los jugadores dentro del tablero.
  - b. Final: cuando un jugador llega acá, termina el juego después de que el jugador haya decidido eliminar una última variable o no.
  - c. Castigo: el jugador deberá seleccionar una tarjeta de castigo y contestar la pregunta que viene en ella. Los otros jugadores servirán de jueces para verificar si la pregunta fue contestada correctamente o no. El jugador puede refutar la decisión negativa de los jueces, si tiene evidencia (materia dada por el profesor con anterioridad o preguntar al mismo) de que sí contestó correctamente. Si contesta erróneamente, deberá de agregar una variable al su modelo; si contesta correctamente no deberá agregar una variable pero tampoco podrá eliminar ninguna.
  - d. Recompensa: el jugador puede seleccionar una tarjeta de recompensa y eliminar una variable o una interacción, según indique esta.
8. Si un jugador elimina una variable que está en interacción sin haber eliminado primero la interacción y otro jugador lo nota, este deberá agregar nuevamente esta interacción y otra más dentro de su modelo.
  9. Para agregar o quitar variables o interacciones se usará la función `update` de RStudio, se deberá poner el nombre de la variable o interacción después del `+` o el `-` según corresponda dentro de la función para cada modelo (los objetos de interacción se nombraron con el formato nombre de la variable no categórica, un punto, nombre del valor de referencia de la variable categórica ej: `Miembros.Hombre`).

10. El juego termina cuando algún jugador haya llegado a la casilla final y decidida si elimina una última variable o interacción de su modelo. En este punto se revisarán los AIC para cada estimación de modelo y gana el jugador con el AIC menor.

Anexo 2. Tablero





# 3

---

## **Formación Virtual: Enseñanza y Aprendizaje de la Probabilidad**

---

Christian Camilo López Mora, Nury Alexandra Bulla Buitrago, Pedro Gómez Guzmán, Diana Marcela

Rodríguez Garzón

clopezcamilom@gmail.com, alexandrabulla128@gmail.com, argeifontes@gmail.com,

dianamaroga@hotmail.com



### **Resumen**

La reciente inclusión de la probabilidad en los currículos de primaria en diversos países ha puesto de manifiesto la debilidad en la formación de los profesores en este campo. La preocupación se centra en la falta de conocimiento de los profesores al enseñar temas de probabilidad en primaria. En este trabajo, presentamos un esquema de formación virtual para profesores de primaria enfocado en el conocimiento, el aprendizaje y la enseñanza de la probabilidad. Esta propuesta forma parte de un curso de formación virtual en matemáticas que considera el aprendizaje y la enseñanza de los cinco pensamientos matemáticos y tiene por objetivo proporcionar conceptos, procedimientos,

técnicas y herramientas útiles para mejorar las prácticas pedagógicas de los profesores. Detallamos los referentes teóricos que dieron origen a esta idea y describimos los temas matemáticos enseñados, y las herramientas y recursos utilizados. Analizamos las producciones de los participantes. Con base en ese análisis, establecemos las dificultades de algunos de los participantes. Estas dificultades se refieren su concepción parcial de la noción de evento (consideran que un evento muy probable es un evento seguro) y a una aproximación muy general al análisis cognitivo de la actividad (los objetivos que propusieron no incluyen conceptos puntuales de probabilidad, y los errores y ayudas propuestos no son específicos a esos conceptos).

**Palabras clave:** Formación de profesores; Enseñanza de la probabilidad, Educación primaria

### **Abstrac**

Probability has been recently included in the primary mathematics syllabi of several countries. This situation has raised concern about primary teachers education on this topic. This concern alludes to teachers' lack of pedagogical knowledge for teaching the subject. In this paper, we present a virtual primary teacher education program which has as one of its foci the knowledge for learning and teaching probability. This proposal is part of a larger virtual teacher education program that considers the learning and teaching of the five mathematical strands and seeks to offer concepts, procedures, techniques and tools that can improve teachers' classroom practices. We present the theoretical framework in which the program is based, and describe the mathematical topics treated, and the tools and resources used. We analyzed the participants' work on probability and identified the main difficulties they put forward in their responses. These difficulties refer to their partial conception of the notion of event (they consider a highly probable event as a sure event) and to a very general approach to the cognitive analysis of the activity (the learning goals they proposed did not include the specific concepts of probability, and the mistakes and help suggestions were not specific to those concepts).

**Keywords:** Teachers' training; Teaching of probabilistic; Primary educations

### 3.1 Introducción

La calidad de la educación en matemáticas depende de las oportunidades que los estudiantes tienen en el aula para desarrollar una formación matemática acorde con los fines de la educación (Ball, Lubienski y Mewborn, 2001, p. 435). La manera casi autodidacta con la que una gran cantidad de los maestros de primaria aprendieron matemáticas (usualmente como parte de su formación en el colegio y con muy poca profundización en su formación de pregrado o como normalistas superiores) explica que su conocimiento de esta área se centre principalmente en la memorización de procesos de tipo simbólico para la resolución de ejercicios rutinarios. Como consecuencia, ellos tienen pocas herramientas conceptuales y metodológicas para ofrecer a sus estudiantes oportunidades en las que puedan fortalecer sus competencias matemáticas. Pero, para enseñar matemáticas, el profesor de educación primaria debe tener un conocimiento profundo de los temas objeto de esa enseñanza, con el nivel de reflexión y la amplitud que le permitan comprender y prever el proceso de aprendizaje de sus estudiantes y diseñar e implementar actividades que promuevan ese proceso.

En los últimos años, la enseñanza de la probabilidad se ha incorporado a los currículos escolares y universitarios, dada la importancia, en las diversas situaciones en contextos de la vida diaria, de los conceptos relacionados con Probabilidad. Por esta razón, se hace necesario tener profesores que comprendan y enseñen este tema. Por consiguiente, la incorporación de la enseñanza de la probabilidad en educación primaria es un gran reto en la actualidad y existe un alto interés por diseñar propuestas de formación de profesores que destaquen la relevancia de este tema (Vásquez y Alsina, 2014).

### 3.2 Indagación bibliográfica

En Colombia, la Ley General de Educación 115 de 1994 (Ley 115, 1994) establece la autonomía curricular de los centros educativos que se responsabilizan de la formulación y registro de un Proyecto Educativo Institucional (PEI). El PEI debe incluir un plan de estudios en el que se debe concretar el diseño curricular de las áreas. El MEN asume la responsabilidad de formular y difundir lineamientos curriculares para guiar el proceso de formulación del PEI. Estos lineamientos se expresan en unas expectativas de aprendizaje en términos de competencias (Ministerios de Educación Nacional, 2006). El MEN espera que sean los centros escolares y particularmente los profesores quienes diseñen el currículo con base en los lineamientos anteriores, en un contexto en el que, en la mayoría de las instituciones públicas no se usan libros de texto. Esto implica que la responsabilidad del diseño y desarrollo del currículo lo tienen los colegios y los profesores. Por lo tanto, es necesario proporcionar a los profesores las herramientas necesarias para realizar apropiadamente estos procesos.

Desde el marco legal colombiano, el documento de los Estándares Básicos de Competencias en matemáticas gestionado por el Ministerio de Educación Nacional (MEN, 2006), contempla cinco pensamientos matemáticos (numérico, espacial, métrico, variacional y aleatorio) para la enseñanza y aprendizaje de las matemáticas en educación primaria, secundaria y media para alumnos de todo el país. En nuestro estudio, nos interesamos en el pensamiento aleatorio para la educación primaria. Ese pensamiento se relaciona con la toma de decisiones en situaciones que involucren azar o incertidumbre, con conceptos y procedimientos de la probabilidad y la estadística inferencial directamente y con la estadística descriptiva y la combinatoria indirectamente. Aunque, en el pensamiento aleatorio, se pueden identificar conceptos y procedimientos de la probabilidad y la estadística, nuestro interés se centra en lo relacionado directamente con la probabilidad.

Como lo menciona Gómez (2014), el alumno debe recibir una formación en el pensamiento aleatorio desde temprana edad. Para ello, se hace necesario la preparación de profesores en este nivel educativo, con el objetivo de enseñar los conceptos y procedimientos relacionados con probabilidad. En algunas investigaciones que se enfocan en el desarrollo profesional y conocimiento del profesor de matemáticas en probabilidad, se ha identificado que: “algunos profesores mantienen inconscientemente una variedad de dificultades y errores en probabilidad que podrían transmitir a sus estudiantes” (Gómez, 2014, p.7). En términos del contenido matemático, los conceptos probabilísticos tratados en la educación primaria no utilizan una terminología formal o definiciones rigurosas matemáticas. Sin embargo, el desarrollo de nociones y cuestiones intuitivas frente a los conceptos y procedimientos de la probabilidad se hacen necesarios y útiles para la continuación de este estudio en secundaria y universidad. Al mismo tiempo, debe desarrollarse con los alumnos la habilidad para solucionar problemas sencillos que involucren el azar.

En el estudio realizado en Vásquez y Alsina (2014), los profesores de primaria manifiestan que en su formación no han contado con cursos o estudios relacionados con la enseñanza de la probabilidad y menos aún con capacitaciones matemáticas de la misma probabilidad. Este tipo de capacitaciones son pertinentes y útiles, ya que, la educación primaria es la base para la formación de los alumnos en estudios de secundaria y universitarios. Estos espacios de formación de profesores deben estar orientados hacia el conocimiento matemático y didáctico, los profesores deben comprender que deben enseñar y lo más importante cómo enseñarlo.

Se reconoce que la probabilidad forma parte del currículo de matemáticas para primaria. Sin embargo, la formación disciplinar para la enseñanza de la probabilidad aún debe ser fuertemente trabajada. Se sugiere que la formación de profesores requiere de entornos y contextos en que se trabajen sobre problemas significativos relacionados con su desarrollo profesional, así como de la reflexión sobre dichas actividades (Batanero y Díaz, 2012). En términos del contenido matemático, la probabilidad es útil para entender conceptos en otras ciencias, el razonamiento sobre la toma de decisiones, la relación con fenómenos biológicos, económicos, meteorológicos, políticos y



actividades sociales, razones por las cuales es necesario incluir el estudio de la probabilidad desde la educación primaria, continuando en la educación secundaria y en cursos universitarios. En la actualidad, existe un interés para que, en educación primaria, se estudien y manejen datos con la enseñanza de la probabilidad. Se espera que se desarrollen experimentos, simulaciones y predicciones; se recojan y analicen datos de esos experimentos; y se establezcan conclusiones a partir de los datos. En resumen, se busca que el alumno desarrolle un pensamiento estocástico y no solo conocimientos estocásticos. Este pensamiento estocástico es primordial para que los estudiantes comprendan su importancia en la vida diaria. No obstante, en primera instancia, se debe contar con profesores capacitados para cumplir con este objetivo.

Aunque el desarrollo de procesos de formación de profesores para la enseñanza en probabilidad es un aspecto que se ha ido fortaleciendo, el reto es aún más grande en relación con la formación de profesores de primaria.

Segovia y Romero (2011) consideran que la educación en primaria entorno a la probabilidad debe enfocarse en los principios y Estándares para la Educación Matemática (NCTM, 2000). Por ejemplo, para un ciclo específico de primaria, el estudiante debería desarrollar las siguientes capacidades.

- ✓ Comprender y utilizar la terminología apropiada para describir sucesos complementarios y mutuamente excluyentes.
- ✓ Utilizar la proporcionalidad y una comprensión básica de la probabilidad para formular y comprobar conjeturas sobre los resultados de experimentos y simulaciones.
- ✓ Calcular probabilidades de sucesos compuestos sencillos utilizando métodos como listas organizadas, diagramas de árbol. (p. 431)

En cuanto a los propósitos que tenemos presentes consideramos lo propuesto por Batanero, Godino y Roa (2004) que consideran componentes primordiales para la formación de profesores, que deben: (a) promover una reflexión sobre el significado de los conceptos que enseñan, esto

implica, conocimiento de aspectos culturales, filosóficos, históricos, epistemológicos y relación con otras disciplinas, (b) desarrollar una habilidad que les permita a los profesores ajustar el contenido según las características de los alumnos que tengan en el aula, (c) fortalecer su capacidad crítica sobre información en libros de texto o uso de materiales y recursos en el aula, (d) fomentar el interés por el desarrollo de tareas específicas que permitan abordar errores y dificultades en los estudiantes, (e) mejorar sus propuestas de actividades en pro de la mejora del aprendizaje de sus alumnos, y (f) establecer instrumentos de evaluación que permitan evidenciar aprendizajes de sus estudiantes.

Utilizamos el modelo del análisis didáctico como conceptualización de los procesos de formación del profesor de matemáticas (Gómez y González, 2013). Este modelo nos permite constituir los conocimientos teóricos, técnicos y prácticos (González y Gómez, 2014) que un profesor debería tener idealmente a la hora de planificar una clase didáctica sobre un tema concreto de las matemáticas escolares.

El modelo del análisis didáctico es una conceptualización de las actividades que el profesor realiza para planificar, llevar a la práctica y evaluar unidades didácticas (Gómez, 2007). El análisis didáctico se configura alrededor de cuatro análisis que conforman un ciclo: de contenido, cognitivo, de instrucción y de actuación. Cada uno de estos análisis pone en juego unas nociones, los organizadores del currículo (Rico, 1997), a partir de las cuales es posible identificar y organizar los múltiples significados de un tema matemático concreto —análisis de contenido—, seleccionar los significados relevantes para la instrucción y prever la actuación de los estudiantes al abordar tareas —análisis cognitivo—, describir, analizar y seleccionar las tareas que pueden contribuir al logro de los objetivos de aprendizaje —análisis de instrucción— y evaluar la planificación con el propósito de producir información que sea relevante para ciclos posteriores —análisis de actuación—.

En el análisis de contenido se abordan los conceptos pedagógicos de estructura conceptual, sistemas de representación y fenomenología, con el propósito de identificar y organizar los diversos

significados de un tema de las matemáticas escolares. Los aspectos cognitivos de la planificación se estudian por medio de tres conceptos pedagógicos: (a) expectativas de aprendizaje — competencias, estándares, objetivos, capacidades—, (b) limitaciones de aprendizaje —errores y dificultades— y (c) hipótesis de aprendizaje —previsión de la actuación de los estudiantes ante una tarea—. El análisis de instrucción se fundamenta en dos procedimientos: (a) descripción, análisis, selección y mejora de tareas y (b) descripción, análisis, selección y mejora de secuencias de tareas. Finalmente, en el análisis de actuación se busca finalizar el diseño de una unidad didáctica y utilizar la información que surge de la puesta en práctica de las actividades de enseñanza y aprendizaje para producir información que permita determinar la comprensión de los estudiantes en ese momento, los contenidos a tratar en el aula y los objetivos de aprendizaje que se deben buscar en un nuevo ciclo de planificación.

### **3.3 Objetivo de investigación**

El objetivo de investigación consiste en indagar sobre las dificultades que los profesores participantes en los módulos manifestaron en relación con su conocimiento didáctico acerca del tema de probabilidad.

### **3.4 Método**

Este trabajo surge y forma parte de una propuesta de formación en matemáticas para profesores de primaria gestionada por “una empresa docente”, de la Facultad de Educación de la Universidad de los Andes<sup>1</sup>. El programa tiene modalidad virtual y está compuesto por tres módulos que abordan las tres cuestiones clave de la formación del profesor de primaria en matemáticas: (a) el contenido de las matemáticas escolares, (b) el aprendizaje y la enseñanza de las matemáticas escolares y (c) el diseño curricular (plan de aula y plan de área). Este programa tiene como objetivo principal

---

<sup>1</sup><https://goo.gl/eg4FSO>

permitir que los profesores fortalezcan los temas de las matemáticas escolares de primaria, reconozcan la especificidad del aprendizaje y enseñanza de esos temas, y desarrollen su capacidad para diseñar e implementar planes de aula y para analizar y mejorar el plan de área de su institución.

En este apartado, nos centramos en la descripción de los temas relacionados con probabilidad. En el curso, el estudio de la probabilidad forma parte de los dos primeros módulos. Estos módulos se centran en el estudio de contenidos de la probabilidad y la enseñanza y aprendizaje de la probabilidad. En cuanto al contenido en el primer módulo, el profesor tiene la oportunidad de capacitarse y fortalecer sus conocimientos sobre probabilidad, estudiar conceptos y procedimientos, sistemas de representación, los fenómenos y situaciones que les dan sentido y por supuesto, la historia de la probabilidad. En el segundo módulo, el profesor profundiza sobre el conocimiento didáctico de la probabilidad en la educación primaria. Para ello, él profundiza sobre las competencias, conocimientos, capacidades y actitudes relacionados con el aprendizaje de la probabilidad que le permitan analizar y organizar el contenido, y diseñar e implementar actividades que promuevan el aprendizaje de sus estudiantes.

En la figura 1, presentamos el diseño curricular de los módulos de probabilidad.

## 6. Probabilidad

Este módulo está a cargo de Antonio Moreno de la Universidad de Granada (España) y Camilo López de la Universidad de los Andes.

### Videos

A continuación, se encuentran los videos de este módulo.

1. Inicio módulo
2. Currículo y experimentos aleatorios y deterministas
3. Probabilidad
4. Cálculo de la probabilidad
5. Dependencia e independencia de sucesos
6. Presentación de la actividad
7. Juego de dados. Parte 1
8. Juego de dados. Parte 2

### Documentos y actividades

A continuación, se encuentran los documentos y actividades de este módulo.

1. Actividad
2. Juego de los dados
3. Hoja registro de dudas
4. Foro
5. Reacciones a las dudas
6. Reflexiones de los formadores
7. Actividad de progreso
8. Bibliografía

Figura 1. Esquema módulos de Probabilidad

El diseño curricular de los módulos aborda los siguientes contenidos: currículo y experimentos aleatorios y deterministas, probabilidad, cálculo de probabilidad, y dependencia e independencia de sucesos. Los profesores participantes tienen acceso a estos contenidos por medio de unos videos y de un libro de texto. Durante el desarrollo de cada módulo, el participante inicialmente debía revisar los diferentes videos con el fin de dar respuesta al texto de la actividad. Si durante la solución a la actividad, surgían dudas y dificultades, podría interactuar en un foro virtual, en el que podía proponer temas de discusión entre sus compañeros y el formador o participar de las discusiones generadas por sus compañeros. También, podía registrar sus dudas en un documento que fue revisado por el formador de la sesión quien publicó su reacción ante estas dificultades y dudas en el portal del módulo. Al terminar el módulo, cada participante cargaba su trabajo (solución a la actividad) en la plataforma. Para finalizar el tema, el formador producía un texto en el que destaca las dificultades y fortalezas que ha detectado en el trabajo individual y grupal de los participantes. La actividad propuesta para desarrollar los módulos consistió en dos problemas matemáticos. En este trabajo, nos centramos en el primer problema que se describía con la siguiente situación.

**En una ciudad latinoamericana, se encuentran 15 norteamericanos y 10 colombianos en el edificio A. En frente de este edificio, se encuentran 12 norteamericanos y 15 colombianos en el edificio B. La alarma contra incendios del edificio A se activa de manera automática. En ese momento, una persona sale del edificio A y entra al edificio B para avisar del posible incendio. Las personas del edificio B deciden evacuar. ¿Cuál es la probabilidad que la primera persona que salga del edificio B sea algún colombiano?**

En estas actividades, se esperaba que los participantes resolvieran el problema propuesto y produjeran la información correspondiente a los análisis de contenido, cognitivo y de instrucción. Los trabajos presentados por los participantes fueron la fuente de información para este estudio.

Nuestro propósito consistió en identificar las principales dificultades manifestadas por los participantes. Para ello, analizamos sus respuestas y las clasificamos de acuerdo con su validez y los tipos de errores y dificultades manifestadas.

### **3.5 Resultados**

Con respecto al contenido, los participantes identificaron que el contenido matemático requerido para solucionar el problema era el espacio muestral y la probabilidad. En los trabajos entregados por los participantes, constatamos el esfuerzo de los participantes por desglosar minuciosamente el proceso realizado. Esto les permitió adquirir información relevante sobre las capacidades matemáticas que los estudiantes podrían utilizar para resolver la tarea con éxito. Sin embargo, en su mayoría, los participantes plantearon una única estrategia de solución (ver figura 2), en la que no contemplaron la información adicional que modifica el espacio muestral y condiciona la probabilidad solicitada. No respondieron acertadamente, ya que se solicitaba la probabilidad de qué la primera persona que saliera del edificio B fuera colombiano y la respuesta fue “La mayor probabilidad es que salga: Un colombiano”. Esto evidencia un error relacionado con el lenguaje probabilístico, al considerar que un evento muy probable es un evento seguro.

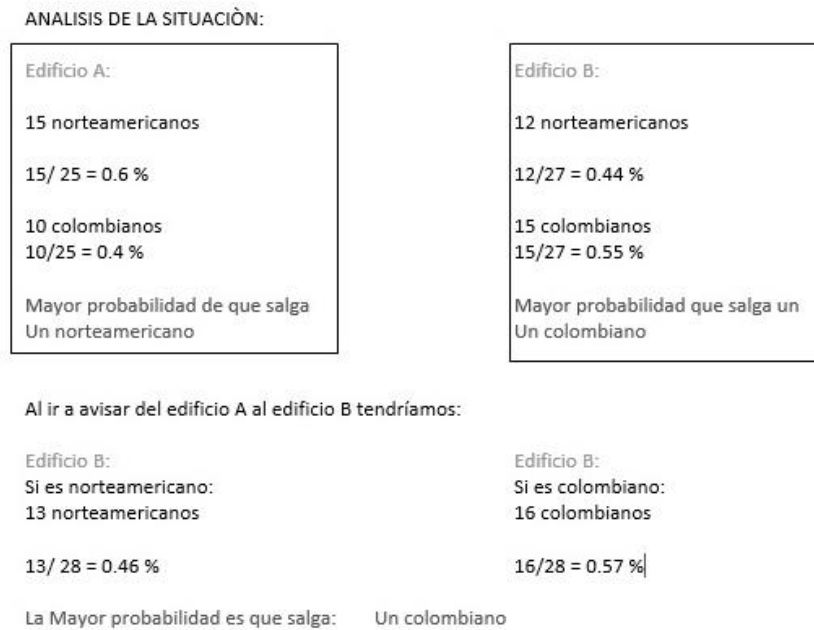


Figura 2. Solución propuesta por un participante del curso.

En las soluciones exitosas, observamos que los participantes utilizaron una tabla de contingencia o un diagrama de árbol (ver figura 3). Esto confirma la relevancia del uso de uno de estos sistemas de representación para identificar, comprender y utilizar correctamente la probabilidad marginal, la probabilidad conjunta, la probabilidad condicional, el principio aditivo, principio multiplicativo y, por tanto, tener un mejor entendimiento de conceptos más estructurados como el teorema de Bayes.

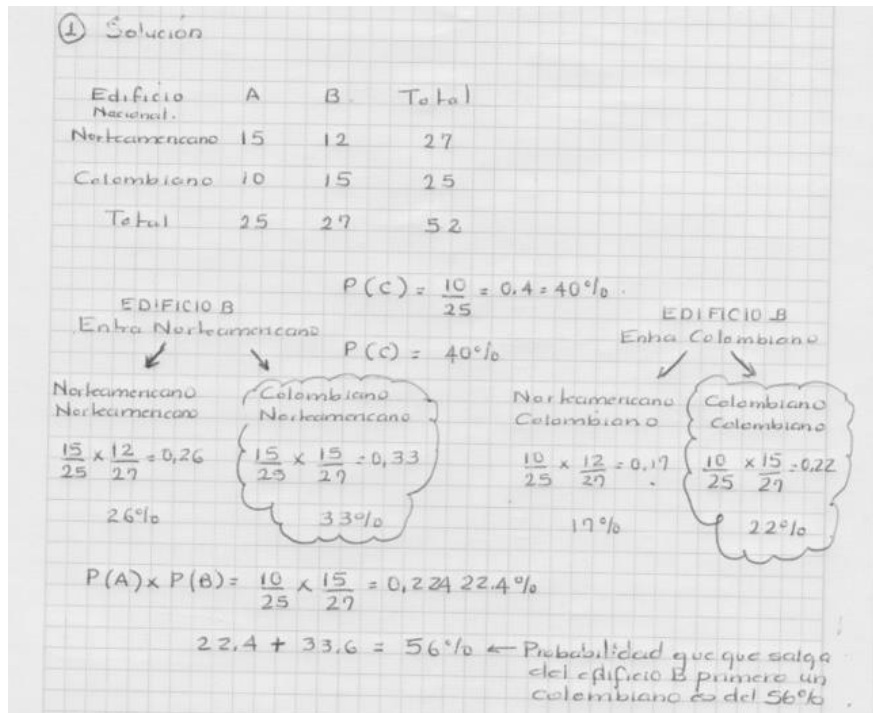


Figura 3. Solución propuesta por un participante del curso.

Con respecto al análisis del problema en la dimensión cognitiva, resaltamos la dificultad de los participantes por proponer diversas estrategias de solución. Ellos plantearon objetivos muy generales que no permiten evidenciar la identificación de conceptos puntuales de probabilidad, lo que genera a su vez, listados de errores y ayudas redactados en términos del aprendizaje en general. Como presentamos en la figura 4, se evidencia un objetivo enfocado a los conceptos básicos de probabilidad. Sin embargo, no se especifican los conceptos que se requieren abordar por medio del problema matemático propuesto. La redacción del objetivo podría utilizarse en cualquier otra situación relacionada con la enseñanza de la probabilidad y no expresa la expectativa del problema propuesto. Consideramos que los profesores pudieron enfocar este objetivo al cálculo de probabilidad simple a partir del uso de diagrama de árbol. También, se evidencia que los errores que plantean los profesores no se encuentran enfocados a las habilidades que no desarrollan los estudiantes o a conocimientos incompletos. De igual manera, los profesores consideran que los



pensamientos matemáticos generan representaciones. Por ejemplo, hacen referencia a la existencia de un lenguaje probabilístico. Por último, los profesores ponen de evidencia su falta de conocimiento para generar ayudas que contribuyan a un mejor aprendizaje de la probabilidad en sus estudiantes. Como se muestra en la figura la idea de ayuda se enfoca en proporcionar un listado de material manipulativo que no tiene un fin relacionado con el objetivo y los errores en los que el estudiante podría incurrir.

2. **OBETIVO DE APRENDIZAJE**  
Aplicar conceptos básicos de probabilidad en diferentes situaciones de problemas planteados.
3. **LISTADO DE ERRORES**
  - No hacer buen uso del lenguaje probabilístico
  - Pensar que las situaciones de incertidumbre no es llegar a la probabilidad de ocurrencia, si no predecir con éxito un resultado
  - Generalizar a partir de pocos ensayos
  - Tener concepciones erróneas sobre las secuencias aleatorias.
  - Privilegiar algunos resultados
  - El sesgo de independencia: No acepta la independencia de resultados en un experimento aleatorio
  - El sesgo de equiprobabilidad: Pensar que todos los sucesos de un experimento aleatorio tiene la misma probabilidad
4. **ERROR FRECUENTE**  
**No hacer buen uso del lenguaje probabilístico**  
Se deben utilizar diferentes tipos de materiales manipulables: dados, urnas con peloticas, naipes etc. De tal forma que se permita a los estudiantes la realización de actividades de tipo explorativo.

Figura 4. Solución propuesta por un participante del curso.

### 3.6 Discusión

En este trabajo, hemos puesto de manifiesto la importancia de la formación de profesores de primaria en probabilidad y estadística y hemos presentado la primera versión de un programa de formación virtual de profesores que aborda esta problemática. El análisis de las producciones de los participantes puso en evidencia que algunos de ellos manifiestan dificultades importantes en relación con su conocimiento didáctico del tema. En particular, encontramos que al menos la mitad de los profesores participantes propusieron una única estrategia de solución a la actividad que evidencia su concepción parcial de la noción de evento: consideran que un evento muy probable

es un evento seguro. Además, ellos tuvieron una aproximación muy general al análisis cognitivo de la actividad: los objetivos que propusieron no incluyen conceptos puntuales de probabilidad, y los errores y ayudas propuestos no son específicos a esos conceptos.

### **3.7 Conclusiones**

Nuestro interés de nuestro trabajo era permitir a nuestros profesores una formación virtual de calidad. El esquema del curso nos permitió identificar que los profesores manifiestan que sus conocimientos en probabilidad son muy escasos en relación con otros contenidos de las matemáticas. Este curso es una oportunidad de aprendizaje para aquellos profesores que se interesen por comprender los conceptos, procedimientos y sistemas de representación que puedan permitirles lograr un mejor aprendizaje. Aunque logramos establecer cuestiones metodológicas que permitieran a los estudiantes de nuestro curso comprender conceptos y procedimientos relacionados con probabilidad, somos conscientes que existe una responsabilidad e interés de proponer y proporcionarles a los profesores diversas actividades que les sea útiles para ahondar en estos conocimientos.

Las dificultades manifestadas por los profesores nos permitieron identificar algunas debilidades del programa de formación con el propósito de mejorarlo. En ese sentido, nos encontramos trabajando en un nuevo diseño alrededor de la noción de deconstrucción de problemas con base en un acompañamiento sincrónico más intenso. La deconstrucción de un problema matemático (Gómez y Henao, 2016) consiste en el proceso en virtud del cual se establecen (a) los conceptos y procedimientos, los sistemas de representación y los fenómenos que están implicados (análisis de contenido); (b) las expectativas de aprendizaje que se espera lograr, los errores y dificultades que puede promover y las demandas cognitivas que conlleva (análisis cognitivo); y (c) las metas, la formulación, las ayudas, la agrupación y la interacción de la tarea de aprendizaje que se puede implementar en el aula (análisis de instrucción).

Con el nuevo diseño, buscamos que los participantes conozcan y se hagan conscientes de la complejidad de los temas de las matemáticas escolares de primaria y reconozcan la especificidad del aprendizaje y enseñanza de esos temas, con el propósito de mejorar sus prácticas pedagógicas para proporcionar oportunidades de aprendizaje a sus estudiantes. El curso se compone de dos módulos enfocados al conocimiento del contenido matemático y al aprendizaje y enseñanza de las matemáticas escolares. Las actividades hacen mayor énfasis en el análisis de un problema matemático que los participantes deben deconstruir. Por ejemplo, para el análisis de contenido, se solicita a los participantes que produzcan la siguiente información en relación con el problema matemático propuesto.

a. Identificar y justificar el tema matemático inmerso en el problema.

b. Analizar de contenido del tema matemático inmerso en el problema.

i. Realizar un esquema que permita evidenciar los conceptos y procedimientos relacionados con el tema matemático y la relación entre ellos.

ii. Realizar un esquema que permita evidenciar las diferentes representaciones que tienen sentido para el tema y la relación entre ellas.

iii. Realizar un listado de los contextos que serían válidos para el tema.

c. Identificar la ubicación del problema dentro del análisis de contenido del tema.

i. ¿Cuáles son los conceptos y procedimientos que se utilizan para dar solución al problema?

ii. ¿Por medio de cuáles representaciones se puede solucionar el problema?

iii. ¿Cuál es el o los contextos del problema?

Este nuevo diseño se está implementando en la actualidad con un grupo de 90 profesores de primaria. Hemos diseñado un esquema de evaluación de este diseño que esperamos nos permita constatar el efecto del curso en las prácticas curriculares de los profesores.

## Referencias

- Ball, D. L., Lubienski, S. T. y Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. En V. Richardson (Ed.), *Handbook of research on teaching* (4 ed., pp. 433-456). Washington, DC: American Educational Research Association.
- Batanero, C., Godino, J. D., y Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education*, 12. Disponible en: [www.amstat.org/publications/jse/](http://www.amstat.org/publications/jse/).
- Batanero, C., y Díaz, C. (2012). Training school teachers to teach probability: Reflections and challenges. *Chilean Journal of Statistics*, 3(1), 3-13. Disponible en: [chjs.socde.cl](http://chjs.socde.cl) Congreso de Colombia. (8 de febrero de 1994) Ley General de Educación. [Ley 115 de 1994]. DO: 41.214
- Gómez, P. (2007). *Desarrollo del conocimiento didáctico en un plan de formación inicial de profesores de matemáticas de secundaria*. Granada, España: Departamento de Didáctica de la Matemática de la Universidad de Granada.
- Gómez, P., & González, M. J. (2013). Diseño de planes de formación de profesores de matemáticas basados en el análisis didáctico. En L. Rico, J. L. Lupiañez & M. Molina (Eds.), *Análisis didáctico en Educación Matemática. Formación de profesores, innovación curricular y metodología de investigación* (pp. 121-139). Granada: Comares.
- Gómez, E. (2014). *Evaluación y desarrollo del conocimiento matemático para la enseñanza de la probabilidad en futuros profesores de educación primaria* (Doctoral dissertation, Universidad de Granada)
- Gómez, P., & Henao, G. (2016). *Deconstrucción de una tarea*. Documento no publicado. Bogotá: Universidad de los Andes. Disponible en <http://funes.uniandes.edu.co/8491/>

Ministerio de Educación Nacional. (2006). Estándares básicos de competencias en lenguaje, matemáticas, ciencias y ciudadanas. Bogotá: Autor.

Rico, L. (1997). Los organizadores del currículo de matemáticas. En L. Rico (Ed.), *La Educación Matemática en la enseñanza secundaria* (pp. 39-59). Barcelona, España: ICE-Horsori.

Segovia, I., & Romero, L. R. (Eds.). (2011). *Matemáticas para maestros de Educación Primaria*. Pirámide.

Vásquez, C., & Alsina, Á. (2014). Enseñanza de la probabilidad en Educación Primaria. Un desafío para la formación inicial y continua del profesorado. *Números. Revista de Didáctica de las Matemáticas*, 85, 5-23.





---

## Aplicaciones de la Estadística Multivariada al Análisis de Datos Ecológicos

---

Diana Teresa Ortiz Vargas, Armando Cervantes Sandoval, Patricia Rivera García.

dianat1995@gmail.com, arpacer@unam.mx, patyriv@unam.mx.



### Resumen

Se plantea desarrollar, mediante el uso de las tecnologías de la Información y Comunicación (TICs), un entorno virtual de enseñanza aprendizaje que sirva de guía para el autoaprendizaje del usuario. Para esto se generó un edublog donde se abordan la aplicación de las técnicas multivariadas de ordenación por medio del software estadístico R en complemento con los paquetes ade4 y vegan, este blog contiene información sobre los aspectos básicos de cada técnica, ejemplos resueltos que muestran el manejo de datos, las instrucciones en R, resultados con su interpretación, vínculos a otros sitios con la misma temática y material de apoyo. Esta herramienta debe aportar opciones que permitan responder a las necesidades de análisis de los estudiantes e investigadores, acordes a su pregunta ecológica de investigación. La creación de este entorno virtual considera de



manera preponderante el diseño instruccional, para la definición de los temas abordados, la información que contendrá y el nivel de profundización. Si se realiza adecuadamente todo lo anterior, se entiende que el usuario tendrá disponible las herramientas y conocimientos básicos para que genere su propio aprendizaje de acuerdo con sus necesidades y a su propio ritmo. El blog actualmente es solo de guía introductoria, ya que es difícil que el usuario comprenda a la perfección los fundamentos teóricos y la complejidad de las técnicas abordadas, pero está mostrando su utilidad y aceptación por los usuarios a los que va dirigido, tanto a nivel de estudiantes como de investigadores.

**Palabras clave:** R, edublog, ordenación, análisis multivariado.

### **Abstract**

We propose to develop, using Information and Communication Technologies (ICTs), a virtual teaching-learning environment that serves as a guide for the self-learning of the user. For this, an edublog was generated where the application of the multivariate management techniques is addressed by means of statistical software R using the ade4 and vegan packages. This blog contains information on the basic aspects of each technique, solved examples that show the management of data, the instructions in R, results with their interpretation, links and support material. This tool should provide options to answer the ecological research question. The creation of this virtual environment considers predominantly the instructional design, for the definition of the topics addressed, the information it will contain and the level of deepening. If all the above is properly done, the user will have available the basic tools and knowledge to generate their own learning according to their needs and at their own pace. The blog currently serves only as a guide, since it is difficult for the user to understand perfectly the theoretical foundations and the complexity of the techniques revised, but it is showing its usefulness and acceptance by the users to whom it is addressed, both at the students as researchers.

**Keywords:** R, edublog, ordination, multivariate analysis.

## 4.1 Introducción

Los métodos estadísticos se han utilizado durante más de un siglo para analizar los resultados de los estudios del medio ambiente, la vida silvestre y las especies, pero los datos ecológicos de hoy han alcanzado niveles sin precedentes de detalle, complejidad y tamaño. (Sanet, 2016). Es por eso que surge la ecología estadística, la cual se ocupa del desarrollo de nuevas metodologías para analizar datos ecológicos (King, 2014).

Por otro lado, cada vez es mayor la difusión y el aprovechamiento de las Tecnologías de la Información y Comunicación (TICs) en los centros educativos. Estas han irrumpido en la vida cotidiana y sobre todo en el proceso de enseñanza-aprendizaje, arrasando con la metodología tradicional y apostando por un nuevo futuro educativo que promueve el uso de la tecnología como máximo exponente. (Arnaiz & Azorín, 2012).

De acuerdo con esto, mediante el empleo de las TICs se está desarrollando un entorno virtual de aprendizaje que tiene como objetivo servir de herramienta para el proceso de enseñanza-aprendizaje en el análisis de datos ecológicos, desde un enfoque multivariado y mediante el uso del software estadístico R en complemento con los paquetes `ade4` y `vegan` para realizar las complejas tareas de cálculo numérico.

## 4.2 Problemática

El análisis de datos ecológicos presenta una naturaleza multivariable, con amplia gama de posibilidades de análisis que el software actual pone a su disposición, por lo que el problema suario es saber cuál es la mejor técnica estadística y definir los criterios para seleccionar la más adecuada en el manejo e interpretación de un conjunto de datos a procesar. Debido a esto se propone la creación mediante el uso de las Tecnologías de la información y comunicación (TICs) la creación de un entorno virtual que sirva como herramienta para la enseñanza y aprendizaje de los criterios de uso para un adecuado análisis estadístico en ecología, así como el manejo de herramientas software para el análisis estadístico.

### **4.3 Objetivo**

Desarrollar un entorno virtual, que centrado en el manejo de ejemplos guíe y apoye al usuario en la aplicación de las técnicas multivariadas de ordenación ecológica para el análisis de sus datos mediante el uso del software de análisis estadístico R.

### **4.4 Justificación**

La mayoría de material de consulta en materia estadística se encuentra en un enfoque matemático, lo cual dificulta la comprensión del usuario debido a que no está familiarizado con el lenguaje utilizado o aun conociendo la técnica estadística a aplicar no entiendo los fundamentos y los criterios de interpretación. Esto hace evidente la necesidad de proponer diferentes recursos que apoyen el proceso de enseñanza-aprendizaje en materia de estadística ecológica.

Es por eso por lo que mediante el uso de las TICs se busca desarrollar un entorno virtual, el cual aborde las técnicas de análisis estadístico tomando como base el paquete de R: ade4, se inicia con un blog que pretende guiar al usuario en la comprensión de estas técnicas estadísticas con el fin ampliar su conocimiento en el tema, y con esto mejorar la selección y uso de las principales técnicas estadísticas con enfoque ecológico para un adecuado análisis de los datos reales.

### **4.5 Material y Método**

En una primera fase se realizó la búsqueda de información referente al tema en diversas fuentes: artículos, libros, sitios web y tesis. Una vez recopilada la información se procedió a su selección, síntesis y depuración.

Se procedió a la elaboración del diseño instruccional y se seleccionó el servidor BLOGGER como herramienta para la creación del entorno virtual (EV), una vez definida la dirección electrónica para acceder al blog, se definieron los temas que se abordan en este espacio:

- I. Inicio
- II. Ecología estadística
- III. Enfoque multivariado para el análisis de datos
- IV. Ordenación
- V. R en ecología
- VI. ade4 y vegan
- VII. Análisis de Componentes Principales
- VIII. Análisis de Correspondencias
- IX. Análisis de Correspondencias Canónicas

Con estos temas se comenzó con la elaboración de los archivos de información en diferentes formatos: .doc para información; .pdf para la difusión, además de la selección de imágenes y material en línea que pudiera servir de complemento para cada tema.

Mediante el uso de Google drive se procedió a cargar los archivos tales como: guía de instalación de R y los paquetes base, los archivos en formato xls o .doc de datos de cada ejemplo.

Para la elaboración de los ejemplos se realizó una búsqueda en base datos, sitios web, libros, tesis para poder recolectar los datos necesarios y así generar los ejemplos requeridos para cada técnica: Análisis de Componentes Principales (ACP), Análisis de Correspondencias (CA), y Análisis de Correspondencias Canónicas (ACC). Estos se trabajaron con los paquetes ade4TkGUI y vegan, realizando ejemplos para cada técnica a abordar.

Para completar cada tema se seleccionó, con base en la búsqueda inicial, material de apoyo para que el usuario pueda consultarlo, dichos materiales se colocaron en la parte final de cada sección, escribiendo el tema o descripción de este seguido del enlace donde se encuentra.

## 4.6 Resultados

Se desarrolló un edublog; la dirección para consultar el blog es <https://ecoestadefesz.blogspot.mx/> se desarrolló con *BLOGGER* y lleva por título “**Ecología estadística: Aplicaciones de la estadística multivariada al análisis de datos ecológicos**” (Figura 1), dicho sitio se encuentra disponible en cualquier momento que el usuario lo requiera.



Figura 1. Página de inicio y presentación del Blog.

Contiene ocho apartados cada uno con aspectos teóricos, ejemplos de cada técnica multivariada de ordenación abordada, una pequeña guía sobre cómo descargar R y los paquetes base de trabajo (ade4 y vegan), así como los archivos de datos que se utilizaron para generar los ejemplos, todo esto con el fin de que el usuario pueda practicar rehaciendo los ejercicios, por último, en cada sección se incluyen enlaces que direccionan al usuario a material de apoyo del tema.

En la parte superior se encuentra la barra de menú, en cada casilla se muestra un tema, comenzando desde la presentación del sitio (Figura 2). Cada tema tiene como objetivo ser la base al tema siguiente, se busca primero dar una visión general del uso de la estadística multivariada en análisis de datos ecológicos y con esa información previa partir hacia el concepto de ordenación.



Figura 2. Barra de menú del entorno virtual.

Cada apartado cuenta con una breve introducción al tema, en el caso de los apartados de las técnicas multivariadas (ACP, CA, ACC) se encuentra una parte teórica, así como ejemplos, cada ejemplo está desarrollado en `vegan` y `ade4TkGUI`. De igual manera cada uno de los apartados cuentan con enlaces a material de apoyo (artículos, libros, tesis, sita web).

También se incluye un archivo `.pdf`, el cual contiene paso a paso desde la instalación de `R`, y de los paquetes en específico `ade4TkGUI`, `vegan`; los comandos necesarios para “correr” las funciones de cada técnica, consejos para cargar datos y modificar gráficas, al igual que los archivos de datos analizados para que el usuario practique y pueda re-hacer estos ejemplos.

El sitio cuenta con 10 apartados, dos de estos tratan de la presentación del blog y de los créditos del sitio respectivamente, por lo que quedan ocho apartados a desarrollar a continuación se presenta una breve descripción de cada uno:

**I.- Ecología estadística:** En esta sección se presenta una introducción acerca de lo qué es la Ecología estadística su objetivo, desarrollo y enlaces a otros sitios afines al tema, así como material de estudio. (Figura 3).



Figura 3. Vista del contenido del apartado Ecología estadística.

**II.- Enfoque multivariado en el análisis de datos ecológicos:** Se aborda la necesidad de contar con un enfoque multivariado en el análisis de datos ecológicos, se presentan ventajas de este sobre las técnicas univariadas (Figura 3), además se hace una breve introducción a los dos principales enfoques de las técnicas multivariadas: clasificación y ordenación. Con el fin de que el usuario pueda tener las bases para entender los siguientes temas.

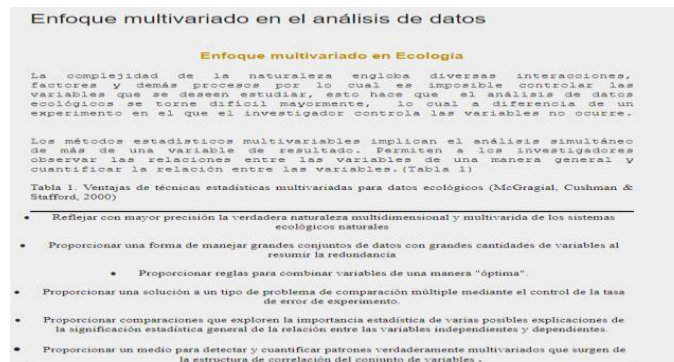


Figura 3. Vista del contenido del apartado Enfoque multivariado en el análisis de datos ecológicos.

**III.- Ordenación:** Dicha sección aborda el concepto, uso y las diferentes técnicas de la ordenación ecológica.



Figura 4. Parte del contenido del apartado de Ordenación.

**IV.- R en ecología:** Se habla acerca del software estadístico de uso libre R, los paquetes que contiene con enfoque ecológico, al igual que se incluye vínculos donde el usuario puede descargar una guía de descarga e instalación de R y los paquetes base de trabajo (ade4 y vegan).

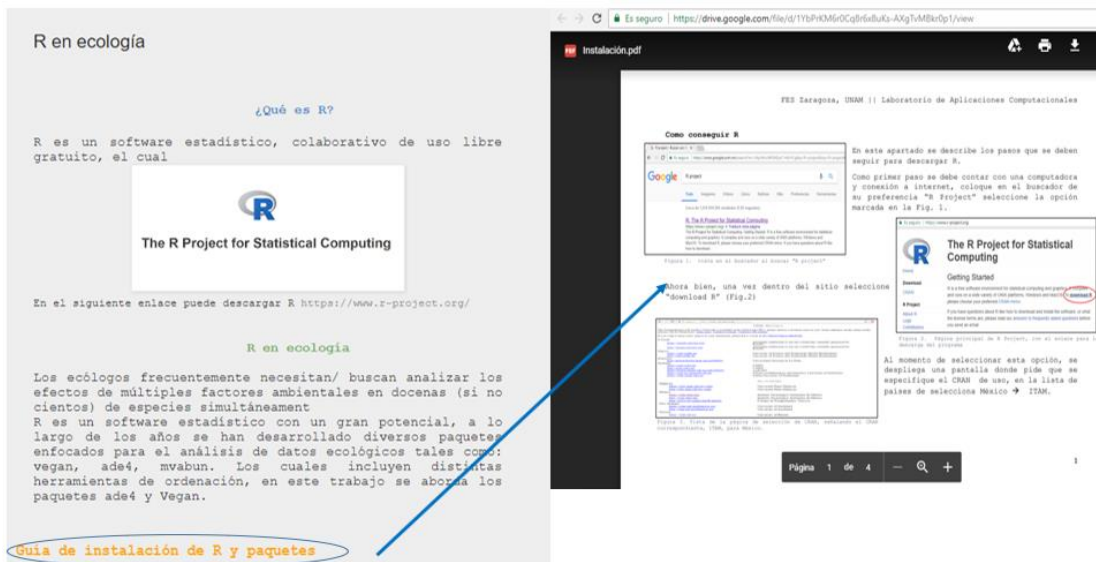


Figura 6. Contenido del apartado R en ecología, se muestra la guía de instalación disponible.



**V.- ade4 y vegan:** Sección donde se menciona en que consiste cada paquete a trabajar, usos, así como material de apoyo, referente al tema.



Figura 7. Contenido del apartado ade4 y vegan.

**VI.- Análisis de Componentes Principales:** En esta sección se da la definición de lo que es el análisis de componentes principales, los criterios de uso, contiene cuatro ejemplos del uso de esta técnica, los cuales fueron elaborados en los paquetes antes mencionados, al igual que los vínculos para la descarga de hoja de datos de cada ejemplo con el fin de que el usuario pueda recrear los ejemplos y así iniciarse en el manejo de este software estadístico. (Figura 8).

Se describe un ejemplo dentro del blog, los otros tres ejemplos se encuentran en formato .doc., se dirige a ellos mediante enlaces respectivamente señalados (Figura 9). Al final de esta sección como en las anteriores se encuentra una sección de material de apoyo, en la cual se muestra diversos enlaces que dirigen al usuario al lugar de consulta de éste, con el objetivo de servir de complemento para el tema.



Figura 8. Apartado del Análisis de Componentes Principales, sección de ejemplos y archivos de datos que contiene.

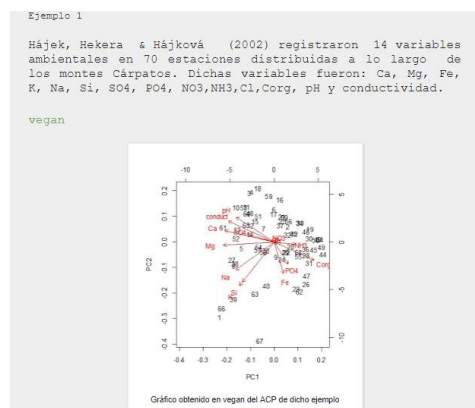


Figura 9. Ejemplo del apartado de ACP desarrollado en vegan.

**VII. Análisis de Correspondencias:** se presenta la definición de lo que es el análisis de correspondencias, los criterios de uso, contiene cuatro ejemplos del uso de esta técnica elaborados en los paquetes base de trabajo, al igual que los hipervínculos para la descarga de los datos de cada ejemplo con el fin de que el usuario pueda recrear los ejemplos y así iniciarse en el manejo de este software estadístico. (Figura 10.)

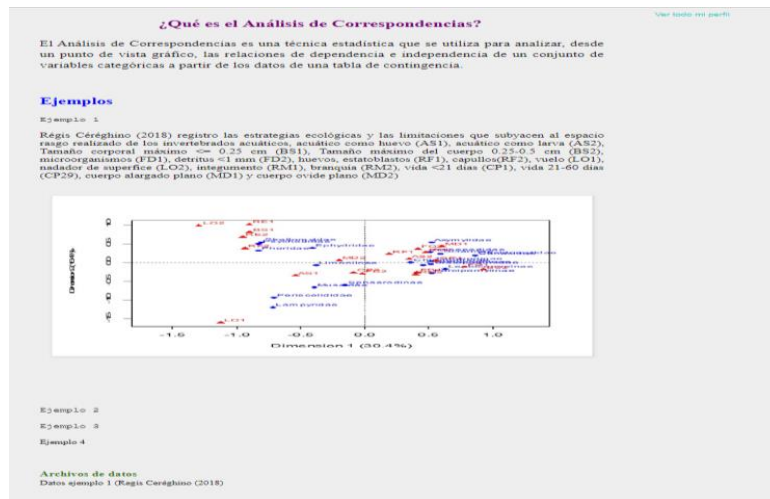


Figura 10. Vista del contenido del apartado Analisis de Correspondencias.

**VIII. Análisis de Correspondencias Canónicas:** Contiene la definición de lo que es el análisis de correspondencias canónicas, los criterios de uso, contiene tres ejemplos del uso de esta técnica, trabajados en los paquetes ade4 y vegan, al igual que los hipervínculos para la descarga de hoja de datos de cada ejemplo con el fin de que el usuario pueda rehacer los ejemplos y así introducirse en el manejo de este software estadístico. (Figura 11).

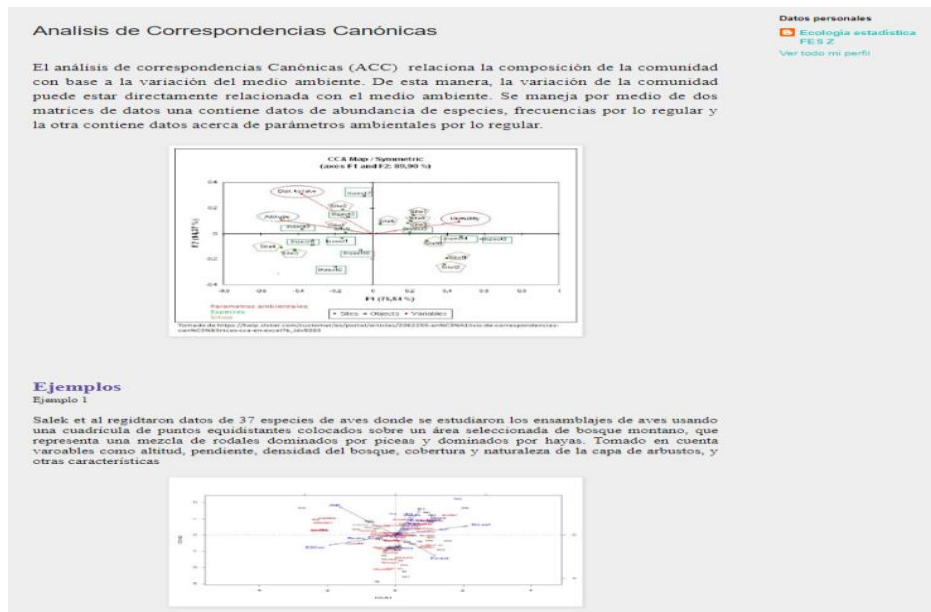


Figura 11. Parte del contenido y ejemplos del Análisis de Correspondencias Canónicas.

## 4.7 Discusión

Se debe señalar que, dentro de la creación de un entorno virtual, es de suma importancia el diseño instruccional, considerando la definición de los temas abordados, la información que contendrán y el nivel de profundización. Si se realiza adecuadamente todo lo anterior, se entiende que el usuario tendrá disponible las herramientas y conocimientos básicos para que genere su propio aprendizaje de acuerdo con sus necesidades y a su propio ritmo.

Por otro lado, en la enseñanza, la tecnología permite orientar los procesos de innovación hacia los diferentes entornos que tienden a promover la construcción de espacios de aprendizaje más dinámicos e interactivos como es el caso del desarrollo de blog como herramienta de enseñanza-aprendizaje. Autores como Rodríguez y Barbosa (2010) están a favor del uso de las TICs debido a que proporcionan nuevas posibilidades de innovación y mejora de los métodos tradicionales de enseñanza y aprendizaje, pero señalan que se debe ser cauteloso en la implementación de éstas para

que su aplicación garantice en algún modo la mejora del proceso, de manera que no se constituyan únicamente en una herramienta de apoyo que refuerce el modelo de aprendizaje tradicional.

En los últimos años han surgido diversas herramientas para el entendimiento en materia estadística; varios autores han informado acerca de las mejoras que la tecnología puede aportar en materia estadística, por ejemplo, Everson & Garfield (2008) informaron sobre el uso de la tecnología en el fomento del aprendizaje colaborativo entre los estudiantes y en la promoción de un entorno en el que los estudiantes aprenden unos de otros diferentes temas estadísticos y de probabilidad.

Por el momento, el blog se centra en enseñar al usuario como trabajar en estos programas estadísticos, mostrar pasó a pasó los comandos que se utilizan para generar los ejemplos aunado a su interpretación.

Uno de los principales problemas al generar los ejemplos para el blog es la parte de encontrar datos para elaborar los diferentes ejemplos pues no se tiene una base de registro de datos de proyectos, la mayoría de los datos presentados en artículos se encuentran incompletos, y dentro de la Facultad, los alumnos tienen datos, pero la mayoría de estos no saben el porqué de estos, el por qué los obtuvieron o ¿para qué? Motivo por el cual también es difícil analizarlos pues no saben “Que quieren saber o que quieren preguntarles a estos”.

El blog, ya está en uso y accesible a todos los internautas, pero aún no ha sido sometido a una evaluación por parte de la comunidad estudiantil o usuarios en general, aspecto sobre el que se está trabajando.

Sólo se presentan algunas técnicas multivariadas para el análisis de datos ecológicos, como una muestra, pues existen diversas técnicas de este tipo para analizar este tipo de datos. Ya que la ordenación en ecología es un concepto que engloba diversas técnicas multivariadas,

Este espacio por el momento solo sirve de guía, para que el usuario tenga una primera aproximación a los fundamentos teóricos y a la complejidad de las técnicas abordadas, pero está mostrando su utilidad y aceptación por los usuarios a los que va dirigido, tanto a nivel de estudiantes como de investigadores.

#### **4.8 Conclusión**

Este blog es una propuesta para emplear las Tecnologías de la Información y Comunicación, contribuyendo a la creación de una opción más para la enseñanza de las herramientas de manejo estadístico para el análisis de datos ecológicos desde un enfoque multivariado.

Los temas estadísticos multivariados son complejos, tanto en su comprensión como en la aplicación a datos reales, por lo que esta es una opción para que sus usuarios revisen a su ritmo y necesidades, pero sobre todo en castellano.

Este sitio proporciona los elementos necesarios para aplicar técnicas complicadas con una herramienta software de actualidad como lo es R, centrándose en la enseñanza del funcionamiento del programa mediante rehacer secuencias claras de manejo de datos a través de ejemplos. Teniendo como meta final contar con elementos para la toma de decisiones que conduzca a la correcta selección y aplicación de las técnicas multivariadas de ordenación.

#### **4.9 Agradecimientos**

Al proyecto PAPIME-UNAM: PE208816, por el financiamiento para la realización de este proyecto y la difusión de resultados.

## Referencias

Arnaiz Sánchez, P., & Azorín Abellán, C. M. (2012). El edublog como herramienta de aprendizaje para todos en el entorno virtual. *Didáctica, Innovación y Multimedia*, (24), 0001-12.

Everson, M. G., & Garfield, J. (2008). An innovative approach to teaching online statistics courses. *Technology innovations in statistics education*, 2(1).

King, R. (2014). Statistical ecology. *Annual Review of Statistics and its Application*, 1, 401-426.

Rodríguez, K., & Barboza, L. (2010). Las TIC como apoyo al proceso de enseñanza-aprendizaje en Bibliotecología. *Costa Rica*.

Sanet, H. (2016). World economic forum. *What is statistical ecology - and what can we learn from it?* Recuperado de: <https://www.weforum.org/agenda/2016/05/what-is-statistical-ecology-and-what-can-we-learn-from-it>



---

## **Aprendizaje de Estadística y Probabilidad en Procesos de Inspección. Caso de Estudio: Control de Calidad de la Leche.**

---

María Alejandra Osorio Angarita, Augusto Bimberto Suárez Parra, Carmen Constanza Uribe Sandoval  
maaosorioan@unal.edu.co, augustosuarez@uniboyaca.edu.co, ccuribe@uniboyaca.edu.co



### **Resumen**

La importancia de la enseñanza de la Probabilidad y de la Estadística a nivel de educación superior, radica en su aplicabilidad a la solución de problemas contextualizados. Es por esto que en este capítulo se revisa la propuesta metodológica PPDAC que lleva al estudiante, en una secuencia de pasos, a la solución de problemas, a partir de situaciones que él mismo identifique. Un caso de control de calidad de la leche que llega a una empresa de lácteos, se toma como objeto de análisis para que él entienda la importancia del uso de la estadística y de la probabilidad y para que el docente encuentre una propuesta metodológica que ayude al estudiante a aprender mediante la solución de un problema.



**Palabras Clave:** Aprendizaje, Educación, Enseñanza, Estadística, Pensamiento aleatorio y Probabilidad.

### **Abstract**

The importance of the probability of statistics in the level of higher education, lies in its application for the solution of contextualized problems. That is the reason that in this chapter is revised methodological proposals as PPDAC that carried out a study in a sequence of steps for the solution of problems, from situations that he identifies. A case of quality control of the milk that reaches a dairy company, is taken in order to students analyze the importance of use of statistics and probability and the teacher take an alternative to teach from the resolution of a problem.

**Key words:** Learning, Education, Teaching, Statistics, Random Thinking, Probability.

## **5.1 Introducción**

La importancia del aprendizaje de la Estadística y Probabilidad en contexto y por competencias es una de las necesidades más sentidas en la formación de profesionales de diversas áreas. De acuerdo a lo afirmado por Zapata (2014), el pensamiento estadístico requiere el desarrollo de habilidades especializadas para leer, interpretar, evaluar críticamente y apreciar información estadística del contexto en que se está inmerso.

Este planteamiento exige que todos los docentes de Estadística estén identificados totalmente con las políticas y estrategias del proceso enseñanza y aprendizaje, comprendan lo que significa enseñar y aprender Estadística y Probabilidad y orienten sus esfuerzos al cumplimiento de las competencias establecidas previamente, para formar a los estudiantes en aprendizajes duraderos, es decir, formar en pensamiento estadístico. (Osorio y Suárez, 2018).



Figura 1. ¿Cómo formar en Pensamiento Estadístico?

Por tal razón en este capítulo se presenta un estudio de caso que permite analizar algunos procesos que se realizan en una industria de lácteos, su control de calidad y la aplicación de la Estadística y Probabilidad, de una manera integrada, con el fin de que el estudiante consolide sus conocimientos y aprenda de forma significativa; lo cual se debe manifestar con su disposición para relacionar el material nuevo con su estructura cognoscitiva (Ausubel y otros, 2009).

## 5.2 Fundamentos

La enseñanza de los conceptos de Estadística y Probabilidad no se debe centrar en el desarrollo de la teoría, sino que complementariamente dichas bases teóricas se deben llevar a la práctica, así el estudiante encuentra más sentido a lo que aprende y se prepara para los retos profesionales. En este caso se debe buscar que el estudiante se involucre en los procesos de medición y obtención de datos para que aprecie el error y en general, evalúe la calidad de los datos.

En concordancia con la anterior, para el aprendizaje de la Probabilidad, se tiene el modelo PPDAC propuesto por Mackay y Oldford (1995) en alusión al método que siguen los Estadísticos para solucionar problemas en su desempeño profesional. PPDAC como estrategia de enseñanza de la Estadística corresponde a una sigla que reúne los conceptos de problema, plan, datos, y análisis; esta estrategia la puede emplear como guía u orientación el docente para ejecutar acciones en bien del aprendizaje a largo plazo de la Estadística (Zapata, 2011).

Específicamente en el ámbito de los programas de Ingeniería, el profesional requiere heurísticas y herramientas que le permitan resolver y decidir en ambientes de riesgo e incertidumbre, lo que se constituye como una de las actividades más importantes del quehacer de este profesional, como lo afirma Koen (1985). Dichas herramientas y métodos estadísticos las puede encontrar en el contexto de la Estadística y la Probabilidad, dentro de las que se pueden mencionar: métodos gráficos y tabulares, el muestreo, procesos inferenciales, regresión, análisis de varianza y el diseño de experimental (Osorio et al, 2013). Ver figura 2.



Figura 2. Algunas herramientas para la toma de decisiones.

En trabajo previo realizado por los autores (2011), se confirmó que la problemática alrededor del aprendizaje de la probabilidad y su aplicación se presenta en forma general y que aparece un reto para los docentes, en cuanto al planteamiento de propuestas innovadoras y la utilización de las existentes que faciliten este proceso de aprendizaje. Por esta razón, más adelante se presenta una alternativa de trabajo para lograr que este aprendizaje sea para la vida y para su quehacer profesional.

La toma de decisiones en las empresas frente a diversas situaciones problémicas es una actividad relativamente frecuente y se constituye en un activo importante para ellas; en épocas pasadas, los encargados de tomar las decisiones tenían como base la experiencia e intuición, pero esta situación ha cambiado, ya sea por el incremento en el uso de herramientas tecnológicas o por la evolución natural del conocimiento. Se entienden dichas situaciones problémicas como las que son susceptibles de estudio y análisis, las cuales pueden afectar la calidad de un producto o servicio, dentro de ellas se pueden mencionar: incumplimiento en los pedidos, retraso en los procesos, insatisfacción de los clientes, incremento en los desperdicios y en general incrementos en los costos que generan pérdidas económicas (Gutiérrez y de la Vara, 2004). Un integrante del personal, ya sea técnico, ingeniero o funcionario, tendrá la responsabilidad de estudiar las diferentes opciones de solución y analizar si se encuentra frente a un problema donde está presente la incertidumbre, el azar y el riesgo, lo cual conlleva a tener un problema de tipo estadístico.

Antes de obtener información sobre un problema o situación, se debe tener muy claro y delimitado el objetivo que se persigue, el tiempo y los recursos disponibles (Gutiérrez, 2010). Posteriormente se debe establecer qué tipo de información se necesita, cómo obtenerla y cómo analizarla. Esto anterior con el fin de llegar a la fase de conclusiones y toma de decisiones.

### 5.3 Consideraciones a tener en cuenta

A partir de un problema cotidiano en el ámbito de la producción se da inicio a una serie de interrogantes y de ahí surge la necesidad de buscar el método adecuado para encontrar una solución. Es posible que el problema en términos generales posea las siguientes características:

- ✓ se encuentre contextualizado, de manera que sea real y palpable, aplicado a una situación del ámbito local o regional.
- ✓ genere datos, en cuya solución se debe acudir a ellos como materia prima para obtener información y poder tomar decisiones.
- ✓ presente el componente de aleatoriedad, es decir que al realizar un experimento bajo las mismas condiciones se obtengan resultados distintos.
- ✓ sea susceptible de hacer planes de muestreo y hacer inferencias a partir de estadísticos.

Es decir que, por la naturaleza de los datos, su procedencia y la interpretación que de ellos se haga para encontrar una solución, es necesario que problemas como el mencionado, sean abordados siguiendo el método que emplearía un profesional en Estadística.

### 5.4 Estudio del problema

Teniendo en cuenta lo anterior, se presentará un caso de estudio que se puede abordar con las 7 herramientas de control de calidad, en las cuales se utilizan los métodos estadísticos apropiados, estas herramientas se presentan en la figura 3.



Figura 3. Herramientas de control de calidad

#### 5.4.1 Calidad como parte del proceso

La Calidad, de acuerdo a las normas ISO 9000, corresponde al “*grado en el que un conjunto de características inherentes a un objeto (producto, servicio, proceso, persona, organización, sistema o recurso) cumple con los requisitos*”. Otra definición que puede ser apropiada en este contexto, corresponde a lo expuesto por Deming (1989), quien considera que Calidad es la satisfacción del cliente.

Es decir que, la calidad de un producto o servicio es satisfactoria cuando responde a las necesidades del consumidor, es lo que esperaba o superior a lo que esperaba recibir (Colomer, 1997); por esto, las empresas cada día se preocupan, en mayor medida, por tener altos estándares de calidad que garanticen el posicionamiento de sus productos en un ambiente de competencia.

De ahí surge la necesidad de implementar un conjunto de técnicas y procedimientos que le sirven a la dirección para orientar, supervisar y controlar todas las etapas que garantizan la obtención de un producto con la calidad deseada; a esto se le denomina Control de Calidad (Colomer, 1997).

El Dr Deming, llamado padre de la calidad dio una guía de cómo administrar una organización para asegurar el éxito por medio de la calidad y además aportó los 14 principios para transformar la gestión (Gutiérrez, 2005).

Las empresas son conscientes de que la calidad, además del precio, es un factor determinante para la competitividad en el mercado. Es así como el actuar sobre las variables influyentes en la variabilidad de la calidad de los productos o servicios, son el camino para garantizar una mejora en indicadores de eficiencia y eficacia, es decir en la productividad de la organización. En este momento cobra importancia el uso de técnicas estadísticas para comprender y controlar la variabilidad en los procesos, cuya fuente está asociada con la maquinaria, la mano de obra, las mediciones, el medio ambiente, los métodos y los materiales que intervienen en su producción (Gutiérrez & De la Vara 2004).

#### **5.4.2 Variabilidad**

La variabilidad es una característica inherente a todo proceso de producción, que si bien es imposible eliminarla, es preciso que se mantenga controlada. Una variabilidad en niveles mínimos, asegura la productividad y la satisfacción de los clientes. Si se tiene como meta en toda empresa el control de la variabilidad, es imperativo reducir la incertidumbre mediante el uso de métodos estadísticos sobre aquellos elementos que la generan. Ver figura 4.



Figura 4. Elementos de los procesos que aportan a la variabilidad

Las empresas productoras de alimentos no son la excepción a los requerimientos de calidad, ya que los productos alimenticios deben pasar por pruebas rigurosas que den confianza al consumidor de que han sido elaborados o procesados de la manera adecuada.

### **5.4.3 Pruebas de Control de Calidad**

Resulta importante desarrollar pruebas de control de calidad para proteger al consumidor y obtener de manera constante, productos de superior calidad. En muchos países existen reglamentos y regulaciones que exigen que el producto pase por pruebas de control de calidad con el fin de proteger a los consumidores. El grado y la cantidad de pruebas que deban realizarse varían de un país a otro, por lo tanto, debe solicitarse a las autoridades locales información sobre la materia.

Hoy en día son cada vez más los países que exigen que los productos alimentarios a comercializar – incluidos los productos tradicionales- pasen por pruebas de control de calidad. En el caso de las empresas productoras de lácteos, se realizan diversas pruebas para asegurar la calidad del producto, desde las fases iniciales y la calidad de la materia prima, que es la leche.



Las pruebas y controles (sensoriales, microbiológicos, físicos y químicos) son necesarios en las tres etapas del proceso:

- La prueba de la leche cruda, para establecer su frescura, pureza y condiciones higiénicas.
- Los controles en el procesamiento, para asegurar que se cumpla con ciertas etapas claves que contribuyen con la calidad del producto final.
- El control del producto final, para garantizar que este cumpla con las normas de calidad establecidas.

#### 5.4.4 Pasos del Método Estadístico

Teniendo en cuenta que se debe realizar un estudio apropiado se puede acudir al Método Estadístico el cual puede contener básicamente las etapas que se presentan en la figura 5.

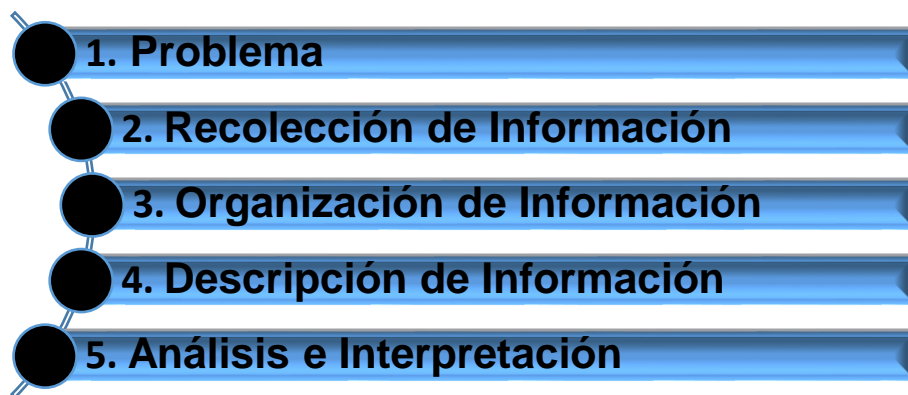


Figura 5. Pasos del Método Estadístico

#### 5.4.5 Estudio de Caso: Control de calidad de la leche

Para ilustrar el proceso de Aplicación de la Estadística en los procesos de inspección, se presenta el Estudio de Caso: “Control de Calidad de la leche cruda en una empresa productora de lácteos”.

Al encargado de realizar el proceso de estudio de acá en adelante se le ha denominado “investigador”.

A continuación, se describe a grandes rasgos la aplicación del Método Estadístico para este caso:

#### **5.4.5.1 Establecer el Problema de investigación**

“Determinar si la leche que suministran los proveedores de leche cruda cumple con los estándares de calidad exigidos”.

**Objetivo:** Realizar control de la variabilidad en la leche que se recepciona, de forma que se pueda ofrecer un producto de alta calidad.

Para dar desarrollo al proceso se requiere responder los siguientes interrogantes a través de las etapas:

1. ¿Qué pruebas se van a realizar?
2. ¿Qué tipo de muestreo se va a emplear?
3. ¿Qué características de la leche cruda se van a medir?
4. ¿Qué valores están establecidos como normales? ¿Cuáles son los límites de aceptación?
5. ¿Qué instrumentos se utilizarán para cada una de las mediciones?
6. ¿Cuál es el plan de muestreo que se llevará a cabo?

**Meta:** Tomar decisiones con respecto a las características de la leche

#### **5.4.5.2 Planear la Recolección de la Información**

Se debe responder a los siguientes interrogantes:

¿Qué técnicas de recolección se pueden emplear?

¿Qué se debe tener en cuenta al utilizar la técnica, para garantizar la confiabilidad y validez de los datos?

### Pruebas que se pueden emplear

El *investigador* del proceso debe decidir cuáles pruebas aplicar para determinar la calidad de la leche, las cuales se constituyen en *experimentos aleatorios* dentro del Análisis Químico de la leche. En la tabla 1. se presenta una lista de algunas de estas pruebas y su proceso de aplicación.

Tabla 1. Algunos Experimentos aleatorios

Prueba a Aplicar	Proceso a Realizar
Tomar el valor del PH	Se toma por medio de un potenciómetro.
Tomar el valor de la Acidez	Se toma por neutralización
Establecer el punto de congelación	Se utiliza para determinar el aguado ilegal. Se utiliza un crioscopio.
Tomar la densidad	Se utiliza un picnómetro y un lacteodensímetro
Establecer el porcentaje de Cenizas	Están representadas por la parte mineral residual
Establecer los sólidos totales	Se miden los residuos obtenidos después de calentar la leche a cierta temperatura.
Establecer la presencia de impurezas macroscópicas	Se determina por filtración
Establecer la grasa	Se utiliza un butirómetro
Identificar si hay presencia de almidones y harinas	se utiliza una solución de reactivo de Lugol.

A diferencia de la leche cruda, para la leche pasteurizada que es el producto final, además del análisis químico anterior, se pueden generar otros experimentos:

- Realizar prueba de la fosfatasa
- Cuantificación de la Lactosa
- Establecer la presencia de Aditivos.

El *investigador* podrá realizar los siguientes experimentos aleatorios dentro del Análisis Microbiológico, tanto a muestras de leche cruda como pasteurizada:

- Análisis Cuantitativo: Recuento de bacterias mesófilas, psicrófilas, termofílicas, coliformes.
- Análisis Cualitativo: Se analiza la presencia de bacterias tales como: Bacilo Tuberculoso, Brucella sp, Streptococcus, etc.

Al tener la leche pasteurizada, se tomarán muestras para detectar si el proceso de pasteurización cumple con los estándares de calidad que se deben ofrecer al consumidor, los cuales aparecerán en la etiqueta del producto. Es necesario establecer si el proceso destruyó en forma eficiente los gérmenes potencialmente patógenos para la salud humana. Este sería un caso de estudio para un módulo posterior de la aplicación de la estadística en el control de calidad de un producto.

Al abordar el problema de variabilidad de las características de la materia prima de un producto, se estarán utilizando algunos conceptos estadísticos con sus correspondientes componentes. Ver figura 6.

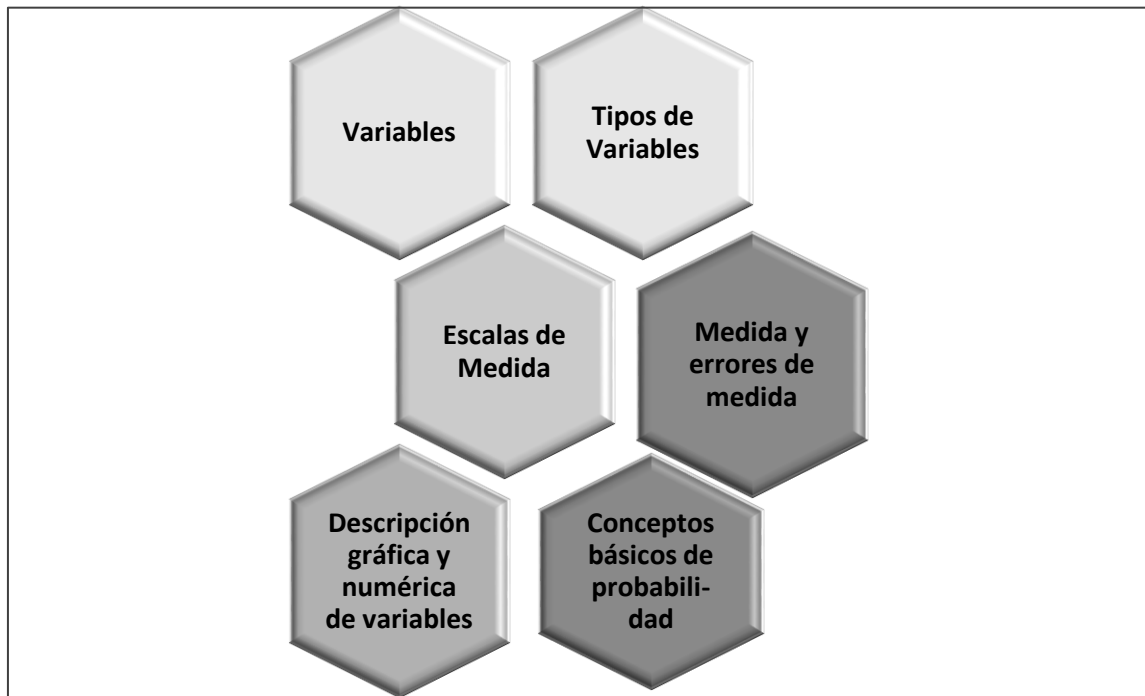


Figura 6. Conceptos de Estadística y Probabilidad a estudiar

Dentro de las múltiples opciones de pruebas que se tienen para determinar la calidad de la leche cruda, en la figura 7 se presentan algunas de ellas, cuáles se utilizan depende de los estándares de calidad estipulados en la empresa.

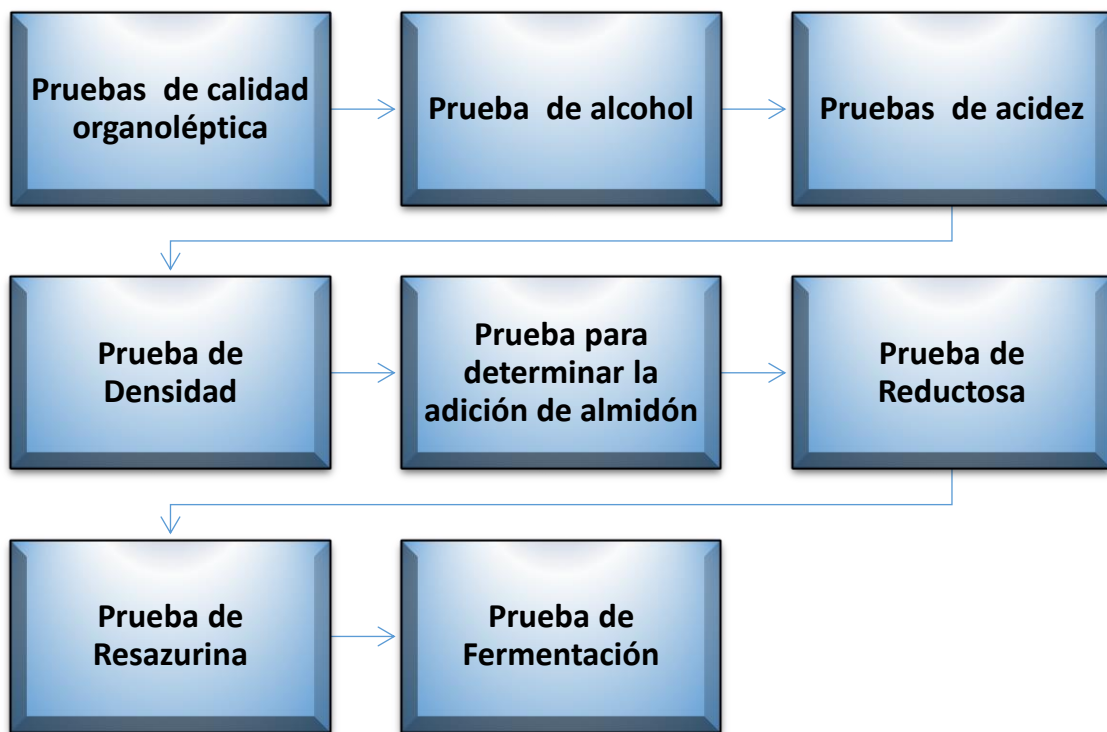


Figura 7. Algunas pruebas para el control de calidad de la leche

Para cada una de las pruebas presentadas se debe inicialmente establecer el propósito de su aplicación en la etapa del proceso. Posteriormente se debe realizar un análisis que se compone de tres pasos. Ver figura 8.

En la primera fase llamada de Conceptualización se debe entender la característica a estudiar, a la cual llamaremos *variable aleatoria*, se estudian las unidades de medida utilizadas y se establece la forma de obtener los valores finales de dicha característica o variable aleatoria.

Para poder realizar la segunda fase, denominada de Operatividad, adecuadamente, es necesario tener claridad en:

- Tipos de variables
- Concepto de medición
- Posibles errores de medición
- Escalas de medida



Figura 8. Etapas a desarrollar al realizar cada experimento aleatorio relacionado con el control de calidad de la leche

Si hay varios métodos para medir la variable aleatoria, se selecciona el que se considere más conveniente, el instrumento de medida apropiado según sea su precisión y exactitud y finalmente se desglosa el procedimiento a seguir.

En la última etapa, denominada de Análisis e Interpretación de resultados, se comparan las mediciones con los valores aceptados como normales y se concluye el resultado de la prueba, es decir, si el resultado es aceptado o rechazado.

**Aplicación:**  $x$ : densidad de la leche. (Corresponde a la variable aleatoria). En física y química, la **densidad** (del latín densitas, -ātis) es una magnitud escalar referida a la cantidad de masa en un

determinado volumen de una sustancia. Es una propiedad que permite medir la ligereza o pesadez de una sustancia.

Unidades de medida: gr/ml o gr/cc (las más usuales).

Para establecer la forma para obtener sus valores y los aspectos a tener en cuenta para realizar el muestreo, determinaremos el instrumento a utilizar:

- **Lactodensímetro**

Los lactodensímetros son instrumentos de vidrio utilizados para la medición de la densidad de la leche y así poder determinar si ha sido mezclada con agua, o si ha sido parcialmente descremada.

Algunos lactodensímetros son de tipo Quevenne cuyo vástago con escala graduada comprende valores entre 15 y 40 que corresponden a las milésimas de densidad por encima de la unidad. (Por ejemplo, el número 32 del lactodensímetro indica la densidad  $1032 \text{ Kg/M}^3$ ) (DILABO,2018). Es un aparato de no más de 40 centímetros de longitud que cabe en una probeta o recipiente delgado en el que se introduce la leche. Con este medidor de densidad, se conoce, en menos de un minuto, la cantidad de grasa, sólidos no grasos y agua presentes en el alimento. Ver figura 9.



Figura 9 Lactodensímetro.



Al tener los resultados tomados con el lactodensímetro se comparan con los valores establecidos como normales.

Los posibles errores pueden estar asociados al instrumento de medida, a la nata o grasa de la leche acumulada, a la toma de la medida o a otros factores. Es por esto que dichos errores se catalogan como sistémicos y operativos.

### 5.4.5.3 ¿Cómo organizar la Información?

El investigador tendrá que analizar: ¿cómo se organiza la Información? Qué herramientas utilizar de acuerdo a lo que se quiere indagar y a la forma en que quieren presentar los resultados. Para ello puede guiarse por la tabla 2 que se presenta a continuación, pero es muy conveniente ampliar la información.

Tabla 2. Guía para descripción de casos pertinentes para el uso de Algunas Herramientas estadísticas

<b>Herramienta Estadística</b>	<b>Conviene usarla cuando...</b>
Diagrama de Pareto	Se quieren detectar los problemas que tienen más relevancia en una situación. Permite establecer en donde se deben concentrar los mayores esfuerzos.
Histograma	Se quiere establecer de manera visual y ordenada datos numéricos estadísticos...
Hojas de Verificación	Se necesitan recolectar datos de cumplimiento de diferentes factores
Gráficos de control	Se requiere mostrar los resultados de una medición como una serie cronológica. Se colocan unos valores límite.
Muestreo estratificado	Cuando una población presenta subgrupos o estratos y se quiere tener representación proporcional de dichos subgrupos.
Diagramas de Dispersión	Se necesita contrastar dos variables de tipo numérico, los datos se muestran como un conjunto de puntos.
Gráficos de causa-efecto	Se quieren identificar las posibles causas de un problema específico.

#### 5.4.4 Descripción de la información

Es una etapa en la cual los datos pasan de ser solo el resultado de la medición particular de una variable, a representar información resumida en tablas, gráficas y estadísticos, que son de interés para el tomador de decisiones. La descripción puede ser de manera univariada, bivariada o multivariada, dependiendo de la naturaleza del problema que se esté abordando.

#### 5.4.5 Análisis e Interpretación de la Información

Las preguntas que se deben responder para poder analizar la información que se ha obtenido y procesado, se pueden resumir en los siguientes interrogantes:

¿Cuáles herramientas estadísticas son las más apropiadas para hacer los análisis de los datos?

¿El resultado de las diferentes pruebas utilizadas para el análisis de la leche fue satisfactorio?



Figura 10. Análisis de los resultados

Si la calidad de la leche está por debajo de los estándares especificados por la empresa, ¿qué proceso se lleva a cabo?

Si la leche que llevan los proveedores no cumple con las exigencias se tomará la decisión de descartarla, porque estaría afectando la calidad del producto que se emplea como materia prima en los procesos. ¿El proveedor puede rechazarse por incumplimiento en la calidad de la leche? Estas decisiones dependen de los estándares de calidad establecidos por la empresa.

## **5.5 Consideraciones Finales**

En este capítulo se presenta de manera general la forma de analizar un problema específico en una industria, se utilizan las etapas del método estadístico para abordarlo, debido a que cumple con las condiciones para ser estudiado desde este punto de vista. No pretende ser una guía al pie de la letra sino por el contrario el docente y estudiante podrán proponer novedades y formas de aplicar estas ideas.

## Referencias

- Ausubel D.P., Novak J.D. y Hanesian H. (1983). Psicología Educativa, un punto de vista cognoscitivo. Editorial Trillas, 2ª Edición, México.
- Colomer M. (1997). Estadística en el control de calidad. Ediciones de la Universidad de Lleida. Cataluña. España.
- Deming, W. (1989). Calidad, productividad y competitividad. Madrid, Ediciones Díaz de Santos.
- DILABO S.A. (2018). Suministros para laboratorio. Consultado el 10 de abril de 2018 de, [http://www.dilabo.com/producto\\_28517\\_NombreProd.html](http://www.dilabo.com/producto_28517_NombreProd.html)
- Gutiérrez, H. y De la Vara R. (2004). Control Estadístico de la Calidad y Seis Sigma. México: McGraw-Hill.
- Gutiérrez, H. (2010). Calidad Total y Productividad. 3ª. Edición. México: McGrawHill.
- International Organization for Standardization. ISO (2018). Calidad Integral. Consultado el 22 de abril de 2018 de, [https://www.calidadintegral.com/iso\\_9000.php](https://www.calidadintegral.com/iso_9000.php)
- Kohen, B. (1985). Definition of the engineering method . Washington D.C., EEUU: American Society for Engineering Education
- Mackay, R.J y Olford, R.W. (2010). Scientific method, statistical method and the speed of light. Statistical Science, 15 (3), 254–278.
- Osorio, M. y Suárez, A. (2011). Aplicaciones de las TIC en Educación. Probabilidad: de la teoría a la práctica. Ponencia presentada en el II Seminario Internacional Gestión Tecnológica en Ingeniería, Tecnologías de la información y Comunicaciones (TIC), Industria y Medio Ambiente. Tunja

Osorio, M., Suárez, A. y Uribe, C. (Septiembre-diciembre 2011). Revisión de aspectos asociados a la problemática del aprendizaje de la Probabilidad. Revista Virtual Universidad Católica del Norte, 34, 360-384. Disponible en: <http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/342/65>

Osorio, M., Suárez, A. y Uribe, C. (Febrero-mayo 2013). Revisión de alternativas propuestas para mejorar el aprendizaje de la Probabilidad. Revista Virtual Universidad Católica del Norte, (38), 127-142.

Osorio, M. Suárez, A. (2018). Aprendiendo Probabilidad y Estadística en Contexto, Preprint.

Zapata, L. (Mayo-agosto 2011). ¿Cómo contribuir a la alfabetización estadística?

Revista Virtual Universidad Católica del Norte, 33, 234-247.

Disponible en: <http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/4Zapata>, L. (2014) Alcance de las tareas propuestas por los profesores de estadística. Uni-pluri/versidad, Vol. 14, N.º 1, Medellín.



---

**Estudio de Casos para el Aprendizaje de conceptos de Estadística Descriptiva:  
Una experiencia con estudiantes de Ingeniería**

---

Maria Cristina Kanobel  
mckanobel@gmail.com



**Resumen**

En la actualidad, la Estadística es un saber necesario para interpretar, comunicar o producir información a partir de datos. Este trabajo describe una experiencia desarrollada durante los años 2016 y 2017 en cursos de Probabilidad y Estadística de la Universidad Tecnológica Nacional-Facultad Regional Avellaneda (Argentina) para abordar la enseñanza de conceptos centrales de Estadística descriptiva a partir del uso del uso de un software para el tratamiento de datos y del estudio de casos como estrategia para la evaluación de los aprendizajes de los estudiantes. Esta técnica busca introducir a los estudiantes en un problema de solución abierta y de carácter complejo con datos reales a partir de situaciones que puedan presentarse en la vida profesional, para motivar

el análisis, la reflexión y el planteo de soluciones posibles. Pretende también plantear formas alternativas de enseñanza y aprendizaje, mediante el trabajo colaborativo en grupos de pares. Para evaluar los aprendizajes, se solicitó a los estudiantes la entrega de un informe, surgido del procesamiento e interpretación de información cuantitativa proporcionada por la cátedra. Para analizar la motivación de los estudiantes, se utilizó el cuestionario MLSQ como herramienta para relevar las percepciones de los estudiantes sobre la metodología utilizada, se realizó una encuesta cuyos resultados se describen en el presente trabajo general.

**Palabras Clave:** Estadística descriptiva, TIC, estudio de casos, evaluación

### **Abstract**

Nowadays, Statistics is a necessary knowledge to interpret, communicate or produce information from data. This work describes an experience developed during 2016 and 2017 in all of the courses of Probability and Statistics of the Universidad Tecnológica Nacional – Avellaneda (Argentina) to teach central concepts of descriptive statistics using Excel as software for the treatment of data and the study of cases as a strategy to evaluate learning. This technique allows to introduce students to an open solution problem of a complex nature with real data from situations that may appear in professional life and motivate the analysis, the reflection and the discussion of possible solutions. It also aims to propose alternative ways of teaching and learning, through collaborative work in peer groups. To evaluate students were asked to submit a report, arising from the processing and interpretation of quantitative information provided by the professorship. To analyze the motivation of the students, the tool was used to survey the perceptions of the students about the methodology used, a survey was conducted whose results are described in the present general work.

## 6.1 Introducción

Desde hace algunos años, la cátedra de Probabilidad y Estadística de la Universidad Tecnológica Nacional Regional Avellaneda (UTN FRA) viene desplegando diversas estrategias para que los estudiantes se acerquen al mundo de la Estadística, no solo para proveerlos de una estructura de conocimientos para aplicar en contenidos de años posteriores sino, especialmente, para posibilitar que desarrollen algunas competencias asociadas a su futura vida profesional. Esta visión, requiere nuevos abordajes de enseñanza de los contenidos de la asignatura como así también, formas alternativas de evaluación.

Particularmente, este trabajo describe la implementación de un diseño instruccional mediado por el uso de tecnologías para la enseñanza de conceptos de Estadística descriptiva para introducir a los alumnos universitarios en el análisis exploratorio de datos estadísticos.

La experiencia piloto fue llevada a cabo en todos los cursos de la cátedra durante el año 2016 y luego de posteriores análisis, evaluación y ajustes, se implementó nuevamente en el ciclo lectivo 2017. Para el abordaje de la enseñanza se usó el software Excel que, aunque no es de uso gratuito, es un recurso conocido y bastante accesible para los estudiantes. Para la evaluación de saberes, se utilizó la técnica de estudio de casos acompañada por el uso del software para el tratamiento de datos.

## 6.2 Marco teórico: el modelo TPACK

Este modelo teórico cuya sigla significa Conocimiento Técnico Pedagógico del Contenido, fue desarrollado por Punya Mishra y Matthew J. Koehler (2009) basados en la idea de Lee Shulman sobre la integración de conocimientos pedagógicos y curriculares que deberían poseer los docentes, teniendo en cuenta que la didáctica debe contextualizarse en la asignatura que se enseña y, en consecuencia, debe estar impregnada y condicionada por ella. Mishra y Koehler amplían la idea original de Shulman e integran las TIC a la dupla planteada. Definen así el modelo TPACK como



un marco conceptual para integrar las llamadas Nuevas Tecnologías en el proceso de enseñanza según se describe en la figura 1:

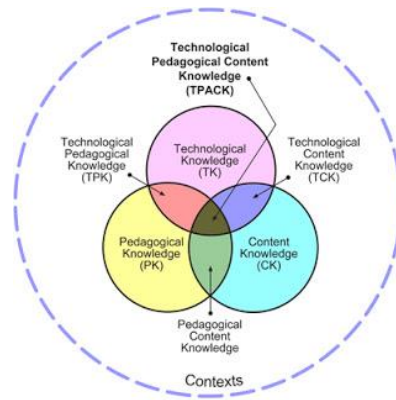


Figura 1. Esquema del modelo TPACK

Según los autores del modelo TPACK, los conocimientos pedagógicos, disciplinares y tecnológicos del docente interactúan entre sí cuando se construye un diseño instruccional que tenga en cuenta el conocimiento tecnológico relativo a los recursos que se emplearán, el conocimiento disciplinar se refiere a los contenidos que se deben enseñar para que los estudiantes aprendan, mientras que el conocimiento pedagógico implica de qué forma abordar dichos contenidos a través de diferentes medios. El profesor debe articular dichos conocimientos de manera que esta interacción suponga una mejora en la calidad de la enseñanza.

### 6.3 Sobre la intervención didáctica

Aunque los tópicos sobre Estadística descriptiva no están contemplados dentro de los contenidos mínimos asignados para la asignatura, consideramos que es necesario incluirlos ya que la mayoría de los estudiantes no los tiene incorporados como conocimientos previos a pesar de estar incluidos en las curriculas de la escuela media, que es el nivel anterior de enseñanza. Por otro lado, incorporar nuevos contenidos al programa de la asignatura fue un desafío ya que este accionar no podía provocar la necesidad de excluir a otros. El criterio utilizado fue la jerarquización de los

contenidos. En tal sentido, a partir del año 2015 decidimos plantear la inclusión, como primera unidad temática, de los contenidos de Estadística Descriptiva con el tratamiento mediado por un software, sin que esto provocase la eliminación de algún bloque temático. La experiencia desarrollada nos indica que la incorporación de una introducción al análisis exploratorio de datos fue un anclaje para abordar el estudio de la Teoría de la Probabilidad y del aspecto Inferencial de la Estadística. Hacer distintos análisis de frecuencias relativas, favoreció la comprensión de la definición frecuencial de probabilidad y contribuyó a la mejor comprensión del concepto de probabilidad por parte de los estudiantes. Por otra parte, se pudo observar que la enseñanza de las medidas de posición y de variabilidad favoreció la posterior comprensión de los conceptos de estimación de parámetros y de test de hipótesis, y la diferenciación entre parámetros poblacionales y variables muestrales, una dificultad frecuente observada en los estudiantes.

Para llevar a cabo la enseñanza, se implementó un diseño instruccional para desarrollar en un período de tres clases de tres horas cátedra en todos los cursos de la asignatura realizadas en el laboratorio de computación: esto es, unos diez cursos con un total de 255 estudiantes. Para el diseño de la intervención pedagógica, se elaboró un material didáctico para usar en la clase a modo de aula taller. Para ello eran necesarios *conocimientos pedagógicos, disciplinares y tecnológicos* según el marco TPACK. Partiendo de las preguntas: “¿qué enseñar?” y “¿para qué?” construimos un diseño didáctico de la unidad temática con inclusión de EXCEL para propiciar el procesamiento y análisis de datos. En el primer encuentro, los docentes a cargo de los cursos hicieron una introducción para dar una visión general del concepto de Estadística relacionándolo con la Teoría de Probabilidades, en el sentido de Hans van Buren (2010) que afirma que debe enseñarse con un enfoque integrado. Luego, para indagar los conocimientos previos, se les propuso a los estudiantes la resolución de una actividad. Luego del espacio de discusión en grupos, se realizó la puesta en común donde se analizaron distintas respuestas para acordar conclusiones que, según pudimos observar, fueron realizadas sin dificultades. Una vez finalizada dicha tarea, los profesores introdujeron los contenidos con presentaciones PowerPoint y los estudiantes trabajaron en sus computadoras a partir de tutoriales elaborados por el equipo docente. Para abordar los contenidos

de esta unidad elegimos Excel, por ser un programa que todos los estudiantes conocen. Una vez desarrollados los contenidos de la unidad temática y para evaluar los contenidos, se les propuso a los estudiantes una actividad para resolver en grupos. Los estudiantes contaron con un lapso de cuatro semanas para realizar una actividad integradora: se les dio un caso para analizar y una matriz de datos para procesar la información. Durante ese período, contaron con distintos canales para realizar consultas: en forma presencial o en forma virtual a través de diversas herramientas como mensajería y foros del aula virtual, correo electrónico y grupos de Facebook.

La corrección del trabajo práctico fue llevada a cabo por los auxiliares docentes bajo la guía del profesor. Los estudiantes recibieron las devoluciones para hacer ajustes en los casos necesarios.

Luego de las devoluciones a cada grupo, se fijó una fecha para la defensa oral de los trabajos para que cada equipo pudiera hacer su presentación y defender el informe.

Para analizar y evaluar la actividad se tuvo en cuenta una rúbrica.

#### **6.4 Análisis de los errores más observados**

El caso proporcionado a los estudiantes presentaba algunas preguntas divididas en varias secciones:

En la primera parte, las preguntas referían a la diferencia entre población y muestra y a la identificación y clasificación de las variables puestas en juego.

En la segunda parte, se hacía necesario recurrir a las tablas de frecuencia y gráficos estadísticos para resumir y visualizar la información.

Para la resolución de la tercera parte, los estudiantes necesitaban poner en juego algunos conceptos como homogeneidad y heterogeneidad, como así también el de medidas de centralización y de dispersión.

En la sección cuatro se profundizó el análisis a partir de la comparación de muestras,

En la parte cinco se solicitó realizar una conclusión a partir de los resultados obtenidos en ítems anteriores y en concordancia al planteo inicial del caso.

Los docentes encargados de la corrección hicieron una primera devolución con una orientación a los estudiantes sobre los ítems a mejorar o modificar evitando marcar los errores cometidos para que cada grupo pudiera proseguir con el análisis promoviendo una instancia de metacognición de saberes entre los integrantes de cada grupo.

Algunos grupos realizaron una nueva entrega tanto en los casos de trabajos incompletos como en aquellos casos en los que se habían señalado algunas observaciones para mejorar.

Una vez aprobados los informes, cada grupo dispuso de una instancia de exposición de sus trabajos. En esa instancia, el docente a cargo de cada curso hizo devoluciones generales y particulares sobre las producciones.

Uno de los errores más frecuentes hallados en la resolución del trabajo práctico es la representación gráfica de los datos: un 50% de los grupos cometieron este error. Es importante tener en cuenta estas dificultades ya que la interpretación correcta y la lectura crítica de datos es una componente de la alfabetización estadística y una necesidad en nuestra sociedad tecnológica por lo que nos preocupa que los estudiantes puedan desarrollarla.

## **6.5 Percepciones de los estudiantes sobre la modalidad de trabajo**

Para relevar información sobre las percepciones de los estudiantes sobre la modalidad de trabajo adaptamos algunos de los ítems del cuestionario MSLQ y desarrollamos un instrumento que fue respondido por los estudiantes a través de un cuestionario en line, realizado con Google Forms. De las respuestas obtenidas obtuvimos la siguiente información:

- el 68,9% de los estudiantes respondió que no tuvo dificultades para trabajar los contenidos de Estadística Descriptiva con software
- un 79,7% expresó que los materiales didácticos facilitados para el desarrollo de las actividades en el laboratorio de computación le resultaron útiles como apoyo al aprendizaje
- solamente un 7,45% tuvo dificultades para comprender el material de estudio.

También observamos que:

- un 69,23% afirma no haber tenido dificultades para resolver las tareas propuestas
- solamente un 7,69% considera que las tareas propuestas carecen de interés.

También resulta importante para nuestra cátedra indagar si la modalidad de cursada y las estrategias implementadas en la asignatura, pueden producir avances sobre el conocimiento de los estudiantes sobre su propio desempeño:

- Un 93,10% de los estudiantes es consciente de la utilidad de los contenidos de la asignatura en su futuro rol profesional
- más del 65% considera que los contenidos son importantes.

Al indagar si la forma de abordaje utilizada como estrategia para el aprendizaje fue utilizada en otras asignaturas, un 89.7% respondió en forma negativa.

## **6.6 Conclusiones**

Los resultados obtenidos de los dos años de implementación de la actividad nos permiten ensayar algunas afirmaciones:

- i. observamos que, en su gran mayoría, los estudiantes involucrados en este estudio no están habituados a abordar problemas con respuestas abiertas que requieran un análisis que se desprenda del estudio de los resultados obtenidos a partir del cálculo,
- ii. Es importante destacar también, la predisposición y motivación de los estudiantes que redundó en la participación y en el trabajo activo en cada uno de los encuentros, en coincidencia con las afirmaciones de Batanero en cuanto a que los proyectos estadísticos aumentan la motivación de los estudiantes.
- iii. En la instancia oral, observamos con grata sorpresa que los grupos, además de haber utilizado los conocimientos aprendidos en la presentación oral, investigaron y recuperaron información adicional, que enriquecía el informe del caso analizado.

En el abordaje del estudio de la Estadística, la incorporación de las tecnologías, aunque vigentes desde hace mucho tiempo, nuevas para el desarrollo de la asignatura, se introduce la posibilidad de evolucionar didácticamente desde la algoritmia hacia el análisis de los datos. La posibilidad de procesar más datos permite trabajar en contextos más cercanos a la realidad, efectuar comparaciones, formular hipótesis sobre la población y tomar decisiones acerca de las formas de presentación de los datos, posibilita un acercamiento de los estudiantes a los métodos científicos usados en los grupos de investigación y no a los recortes didácticos de los problemas de los libros de texto. Pensamos que, trabajar en este sentido en esta asignatura posibilita promover modos de pensamiento relacionados con la futura práctica profesional del ingeniero. Por los resultados obtenidos, tanto en el desempeño académico de los estudiantes como en las percepciones relevadas, podemos afirmar que nuestra propuesta puede resultar un aporte importante para promover aprendizajes significativos de los estudiantes.

## Referencias

Almiron, M. G. Lopes, B.; Oliveira, A. L., Medeiros, A. C., Frery, A. C. (2010). On the numerical accuracy of spreadsheets. *Journal of Statistical Software*, Vol.34, No. 4, pp. 1-29

Ausubel, D. y Novak, J. (2009) *Psicología educativa, un punto de vista cognitivo*. México: Trillas

Batanero, C. (2000) ¿Hacia dónde va la educación estadística?, *Blaix*, Vol.15, No.2, p.13.

Barrio del Castillo, I. (2012). *El estudio de casos, Métodos de Investigación educativa*. recuperado el 1 de junio de 2015 de [https://www.uam.es/personal\\_pdi/stmaria/jmurillo/InvestigacionEE/Presentaciones/Est\\_Casos\\_doc.pdf](https://www.uam.es/personal_pdi/stmaria/jmurillo/InvestigacionEE/Presentaciones/Est_Casos_doc.pdf)

Chrobak, R. (2006) *La metacognición y las herramientas didácticas*. Neuquén: UNCOMA

Mishra, M., Koelher, M. (2009) *TPACK*, Recuperado el 1 de marzo de 2016 de <http://www.tpack.org/>

Novak, J.D. y Gowin, D.B. (1988) *Aprendiendo a Aprender*. Martínez Roca: Barcelona.

Duval, R. (2006) Un tema crucial en la educación matemática: La habilidad para cambiar el registro de representación. *La Gaceta de la Real Sociedad Matemática Española*, Vol. 9, No.1, pp. 143-168

Koehler, M. J.; Mishra, P.(2009) *What is technological pedagogical content knowledge*. *Contemporary issues in technology and teacher education*, 9,1,60-70

Panko, R. R.(1998) *What we know about spreadsheet errors*. *Journal of Organizational and End User Computing (JOEUC)*,10, 2,15-21



---

## **Exploración de la interpretación de una tabla y una gráfica de columnas por estudiantes de licenciatura**

César López Calvario, Jaime Israel García García, Elizabeth Hernández Arredondo  
nass1\_2012@hotmail.com, jaime.garcia@ulagos.cl, elizabeth.hernandez@ulagos.cl



### **Resumen**

En esta investigación exploramos la interpretación de una tabla y una gráfica de columnas por estudiantes de pregrado, considerando que esta habilidad es parte de la cultura estadística que un ciudadano debe poseer para tomar decisiones y comprender su entorno social. Por ello un elemento presente en la exploración de este trabajo es la lectura de estas representaciones de forma abierta, la cual es la más cercana al modo de divulgación de la información política, social, cultural, deportiva.



Los resultados del estudio exploratorio provienen de un grupo de 36 estudiantes de primer semestre de la Licenciatura en Matemáticas de la Universidad Autónoma de Guerrero, quienes realizaron la tarea de leer e interpretar la misma información representada en dos tipos diferentes de representación, mediante una tabla y una gráfica de columnas comparativa. Esto con el objetivo de analizar los niveles de lectura que exhiben cuando interpretan los datos estadísticos, e identificar si el tipo de representación promueve niveles superiores propuestos por Curcio (1989) y Friel, Curcio y Bright (2001). Asimismo, se consideró la jerarquía de Aoyama (2007) cuando el estudiante realiza una valoración crítica de la información. Nuestro estudio siguió una metodología de tipo cualitativa – descriptiva.

En general, la mayoría de las interpretaciones de los estudiantes se clasifican en el nivel 2, leer dentro de los datos, al enfocarse en la comparación de los datos de los valores de la variable; no obstante, sólo la cuarta parte de los participantes alcanzan los niveles 3 y 4, leer más allá de los datos y leer detrás de los datos, respectivamente, al dar una predicción sobre tendencia de los datos o al integrar la información con el contexto para extraer conclusiones. Asimismo, observamos que el tipo de representación, tabular o gráfica, no influye las interpretaciones de los estudiantes para alcanzar niveles superiores.

Finalmente observamos que en la lectura e interpretación de información en estas representaciones, los estudiantes hacen uso de saberes matemáticos elementales (como la comparación de dos datos numéricos) y del contexto, de manera individual o conjunta; siendo el conocimiento sobre el contexto de la información el que determina su nivel crítico.

**Palabras clave:** interpretación, niveles de lectura, tabla, gráfica de columnas.

### **Abstract**

In this research, we explore the interpretation of a table and a column graph by undergraduate students, considering that this ability is part of the statistic culture that citizens must have to make decisions and understand their social environment. Therefore, reading these representations in an

open way constitutes a key element in the exploration of this work; it is the closest reading to the divulgation of political, social, cultural, and sports information, among others.

The results of this exploratory study come from a group of 36 students in the first semester of a degree in Mathematics at Universidad Autonoma de Guerrero. The students read and interpreted the same information represented in two different kinds of representation: a table and a comparative column graph. The objective was to analyze the levels of reading exhibited by the students when they interpret the statistical data. In addition, we sought to identify whether the type of representation promotes higher levels as proposed by Curcio (1989) and Friel, Curcio and Bright (2001). We also considered the hierarchy proposed by Aoyama (2007) when the students made a critical assessment of the information. This study was carried out following a qualitative-descriptive methodology.

In general, most of the students' representations were classified in level 2, reading within the data, since they focused on comparing the data of the values of the variable. However, only a fourth of the participants reached levels 3 and 4, reading beyond the data and reading behind the data, respectively. These students predicted a data trend or integrated the information and the context to obtain conclusions. Additionally, we observed that the type of representation, table or graph, did not affect the students' interpretations to reach higher levels.

Finally, we observed that, during the reading and interpretation of the information in these representations, the students used contextual and elementary mathematical knowledge, as the comparison of numerical data, individually or in-group. In this sense, the knowledge of the context of the information determined the students' critical level.

**Keywords:** interpretation, reading levels, table, column graph.

## 7.1 Introducción

En la actualidad, es común que diversos medios de comunicación recurran al uso de tablas o gráficas para presentar algún tipo de información, esto implica que el ciudadano interprete, de manera adecuada y crítica, datos estadísticos. Al respecto, Eudave (2009, p. 6) menciona que “la capacidad para leer y entender datos estadísticos es una necesidad social y educativa relativamente nueva”. Díaz-Levicoy, Batanero, Arteaga y López-Martín (2015) hacen hincapié que el ciudadano debe interpretar y evaluar de manera crítica la información estadística que usualmente se presenta en gráficas estadísticas para tomar decisiones y comprenda el mundo que le rodea. Por su parte, Gal (2002) menciona que la competencia gráfica es parte de la cultura estadística, necesaria en la sociedad actual.

Hasta el momento, los estudios sobre interpretación de datos estadísticos han tenido como objetivo analizar los niveles de lectura de gráficas que se muestran en las tareas de los libros de textos (e.g. Díaz-Levicoy, Batanero, Arteaga y Gea, 2015; Díaz-Levicoy, Arteaga y Batanero, 2015; Sánchez y Arteaga, 2013), o bien, aquellos que presentan los estudiantes cuando responden preguntas ligadas a la información representada (e.g. Arteaga, Vigo y Batanero, 2017; Vigo, 2016; Fernandes y Morais, 2011; Eudave, 2009; Pagan, Leite, Magina y Cazorla, 2008; Monteiro y Ainley, 2007; Buendía, 2004; Lisboa, 2002). Sin embargo, hemos encontrado pocas investigaciones (e.g. Gea, Arteaga y Cañadas, 2017; Arteaga, 2011) que aborden la interpretación de tablas y gráficas sin que se guíe al alumno en su lectura, siendo este un aspecto importante; esto debido a que en su cotidianidad, el estudiante se enfrenta a la lectura de información divulgada en libros, revistas, periódicos, televisión o internet. Con base en lo anterior se percibe que hace falta explorar más sobre esta problemática, más aún, dentro de nuestro del contexto educativo universitario; esto nos conlleva a establecer las siguientes preguntas de investigación: ¿Qué niveles de lectura exhiben los estudiantes de primer semestre de licenciatura cuando realizan la interpretación de una tabla y una gráfica de columnas? ¿Qué tipo de representación, bajo un mismo contenido contextual, promueve niveles superiores de lectura en los estudiantes de primer semestre de licenciatura?

## 7.2 Antecedentes

Enseguida se presentan algunas investigaciones sobre la forma en la que los estudiantes, con características similares a los estudiados aquí, interpretan tablas y gráficas estadísticas; las cuales nos sirven para situar nuestro trabajo.

Monteiro y Ainley (2007) estudian la competencia de lectura de gráficos tomados de la prensa diaria con futuros profesores, encontrando que muchos no tenían conocimientos matemáticos suficientes para llevar a cabo dicha lectura; incluso algunos no leían correctamente el gráfico y otros lo leían pero no interpretaban correctamente su significado en el contexto de la noticia, por lo que no llegan al nivel de lectura crítica de los datos. En Eudave (2009) se reporta un estudio con 28 estudiantes, de distintas edades, que cursaban la primaria o secundaria en alguna de las modalidades de educación para adultos que se imparte en México; emanando que únicamente cinco personas poseen los tres niveles de comprensión contemplados por Curcio, ya que pudieron realizar una lectura completa y adecuada de la tabla de frecuencias y de la gráfica de líneas. Por su parte, Arteaga (2011), en su tesis doctoral, reporta que la mayoría de los futuros profesores de su muestra alcanzan los primeros tres niveles de lectura de Curcio, sin embargo, son pocos los que llegan al nivel superior 4 (leer detrás de los datos). Asimismo, Gea, Arteaga y Cañadas (2017) evalúan la interpretación de tres gráficos estadísticos por futuros profesores de matemáticas, encontrando un alto índice de éstos de manera correcta; sin embargo, la mayor parte de las interpretaciones se encuentran en un nivel inicial (leer los datos) o intermedio (leer entre los datos) de comprensión gráfica, y sólo en el diagrama de caja una proporción algo mayor de participantes alcanza al nivel superior de lectura más allá de los datos.

Con relación a los trabajos de investigación mencionados aquí, se observa que la mayoría de los estudiantes leen literalmente la gráfica, hacen comparaciones, pero no todos, llegan a una lectura crítica.

### 7.3 Referente Teórico

Una manera de describir la interpretación de datos estadísticos, es analizando y organizando las respuestas de los estudiantes a la tarea de leer tablas y gráficas a partir de niveles de comprensión. Curcio (1989) y Friel et al. (2001) distinguen cuatro niveles de lectura de gráficas que permiten describir las diferencias en sus habilidades para interpretarlas:

- *Nivel 1. Leer los datos.* Implica la lectura literal de la información representada en la gráfica, sin realizar su interpretación ni cálculos adicionales. Por ejemplo, leer la frecuencia asociada a un valor de la variable, describir el contenido de la gráfica sin interpretarla, indicar palabras o expresiones presentes como el título o la fuente, entre otros.
- *Nivel 2. Leer dentro de los datos.* Es la interpretación e integración de la información presente en la gráfica, pero que no está representada explícitamente; esto implica la comparación de datos o la realización de cálculos matemáticos con ellos. Un ejemplo de ello sería identificar la moda de la variable al comparar todas las frecuencias para encontrar la mayor.
- *Nivel 3. Leer más allá de los datos.* Consiste en realizar predicciones o inferencias a partir de los datos sobre información que no se presenta directamente en la gráfica. Por ejemplo, predecir tendencias o valores considerando la información que se puede observar.
- *Nivel 4. Leer detrás de los datos.* Corresponde a la valoración crítica del uso de la gráfica, la recogida y organización de los datos, su validez y fiabilidad; así como a la integración de la información con el contexto para extraer conclusiones. Un ejemplo sería cuestionarse sobre la calidad de los datos y la forma de recolección.

Si bien, estos niveles fueron establecidos para lectura de gráficos, también pueden aplicarse para tablas (Batanero, 2001), tal como lo presenta Eudave (2009) en su estudio. En nuestra investigación surge la necesidad de considerar la valoración crítica de la información, por ello nos apoyamos en la jerarquía propuesta por Aoyama (2007) con el fin de identificar los elementos que recuperan los

estudiantes cuando presentan el nivel superior “leer detrás de los datos”, el cual puede categorizarse en:

- *Nivel racional / literal.* Los estudiantes pueden leer valores, realizar comparaciones, detectar tendencias particulares. Explican significados contextuales literalmente en términos de las características que muestra la gráfica, pero no pueden sugerir interpretaciones alternativas; sólo usan significados presentados. Generalmente son incapaces de cuestionar la confiabilidad de la información.
- *Nivel crítico.* Los estudiantes pueden leer la gráfica, comprender el contexto, evaluar la fiabilidad del significado contextual descrito y cuestionar la información presentada.
- *Nivel hipotético.* Los estudiantes pueden leer, aceptar y evaluar la información representada en la gráfica, formando sus propias hipótesis explicativas.

Hemos agregado el nivel 0, nombrándolo como *lectura idiosincrática*, aunque Curcio (1989) y Friel et al. (2001) no lo consideran debido a que se enfocan en la evidencia de la comprensión gráfica. Los estudiantes de este nivel no pueden leer valores o tendencias, o bien, no pudieron conectar algunas características extraídas de los gráficos con el contexto. Por lo general, sus lecturas se basaron en su experiencia individual limitada o en perspectivas puramente personales (Aoyama, 2007).

## 7.4 Metodología

La metodología de nuestra investigación es de tipo descriptiva cualitativa, ya que se analizan las respuestas de los estudiantes a la tarea de lectura de una tabla estadística y una gráfica de columnas comparativa, utilizando los niveles de Curcio (1989) y Friel et al. (2001) y la jerarquía propuesta por Aoyama (2007).

## Participantes

Participaron 36 estudiantes de primer semestre de la Licenciatura en Matemáticas de la Facultad de Matemáticas de la Universidad Autónoma de Guerrero, seleccionados de manera intencionada, cuyas edades oscilaban entre 18 y 31 años, quienes no recibieron información alguna del propósito del estudio; y el profesor titular, quien colaboró con la aplicación de la tarea.

### **Instrumento**

Se elaboró una tarea de estudio que consistía en dos actividades. La primera, se refiere a la lectura de una tabla que muestra el porcentaje de hogares con televisión y con televisión de paga en México durante el periodo de 2013 a 2016; y la segunda, a la lectura de los mismos datos estadísticos pero representados en una gráfica de columnas comparativa. En ambas actividades, se les solicitó a los estudiantes leer la información representada, comparar y observar las tendencias, generar conclusiones y realizar críticas, a través de la redacción de varios enunciados donde interpretarían los datos estadísticos.

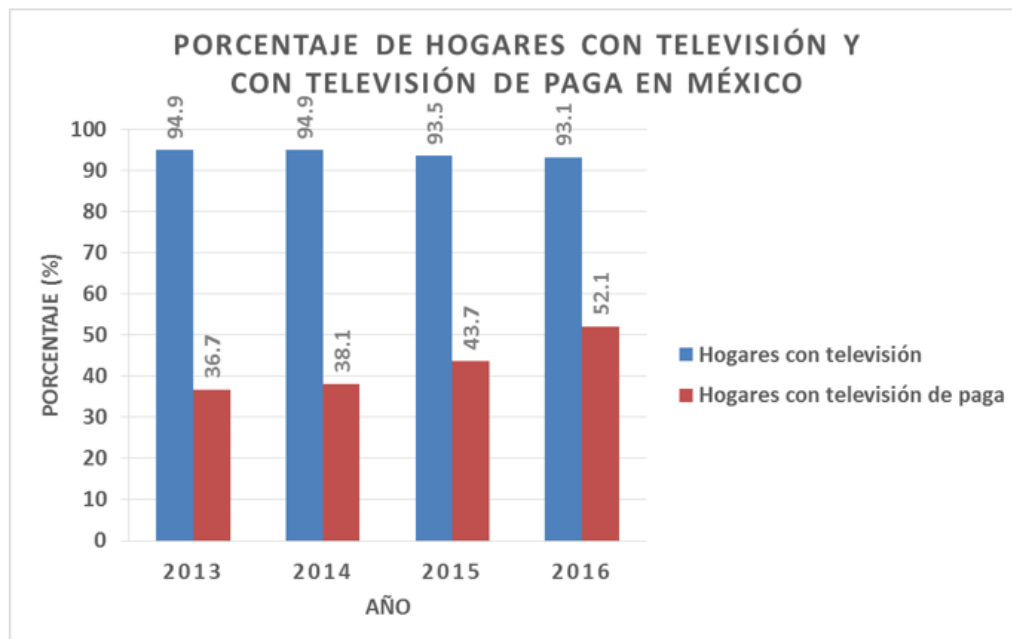
Consideramos no abordar la lectura de la tabla y la gráfica a partir de preguntas debido a que, de manera cotidiana, el estudiante se enfrenta a la necesidad de interpretar información (en tablas o gráficas) de cualquier índole (política, social, cultural, deportiva, entre otros) presentada por los diversos medios de comunicación. En la Figura 1, se muestra la tabla y la gráfica de columnas correspondientes a la tarea de estudio.

**Actividad 1. Lectura de la tabla**

Porcentaje de hogares con televisión y con televisión de paga en México		
Año	Con televisión	Con televisión de paga
2013	94.9 %	36.7 %
2014	94.9 %	38.1 %
2015	93.5 %	43.7 %
2016	93.1 %	52.1 %

Fuente: INEGI. Encuesta Nacional sobre Disponibilidad y Uso de TIC en Hogares, ENDUTIH.

**Actividad 2. Lectura de una gráfica de columnas comparativa**



Fuente: INEGI. Encuesta Nacional sobre Disponibilidad y Uso de TIC en Hogares, ENDUTIH.

Figura 1. Actividades correspondientes a la tarea de estudio



## **Procedimientos**

La tarea de estudio se aplicó en dos sesiones, aproximadamente de 20 minutos cada una. En la primera, los estudiantes realizaron la actividad de lectura de la tabla, y en la segunda, la lectura de la gráfica de columnas comparativa. Consideramos dos sesiones de trabajo para evitar la influencia de los datos de la tabla en la interpretación de la gráfica, o viceversa, en una misma sesión; esto con el objetivo de identificar qué tipo de representación, bajo un mismo contenido contextual, promueve niveles superiores de lectura en los estudiantes.

## **Análisis de datos**

Para el análisis de los datos se llevaron a cabo los siguientes pasos: mediante una primera revisión, se realizó un análisis comparativo para observar incidentes presentes en las respuestas de los estudiantes que hicieran referencia a una misma idea; después, se crearon categorías cuyo concepto corresponde a la idea en común del incidente observado; posteriormente, se llevó a cabo una segunda revisión para clasificar cada una de las respuestas de los estudiantes de acuerdo a los niveles de lectura descritos por Curcio (1989) y Friel et al. (2001) y la jerarquía propuesta por Aoyama (2007); y finalmente, se establecieron relaciones entre las categorías emergentes de las respuestas de los jóvenes y los niveles de lectura.

## **7.5 Resultados**

A continuación, en la Tabla 1 se presentan algunos ejemplos del tipo de respuestas dadas por los estudiantes a la tarea de leer e interpretar información estadística representada en una tabla y una gráfica de columnas comparativa, de acuerdo a los niveles de lectura y las categorías emergentes, seguida de una breve justificación sobre esta clasificación.

Tabla 1. Ejemplificación del tipo de respuestas de los estudiantes

Tipo de representación con respuesta del estudiante a su lectura	Nivel de lectura	Característica	Descripción
<p>Tabla:</p> <p><i>Con las encuestas realizadas se puede manifestar que la economía de los ciudadanos cada año va disminuyendo por falta de empleo en nuestro país.</i></p> <p>Gráfica de columnas:</p>	Nivel 0. Lectura idiosincrática	Perspectiva personal	La lectura está basada en la experiencia o perspectiva del estudiante, no lee valores o tendencias extraídas de la tabla, ni conecta la información con el contexto.
<p><i>Gracias al INEGI nos da una gráfica para saber el porcentaje de TV en hogares y TV de paga.</i></p> <p>Tabla:</p> <p><i>En esta encuesta se puede observar que ha <u>disminuido</u> el porcentaje de hogares con televisión con el paso de los años pero el porcentaje de hogares con televisión de paga ha <u>aumentado</u>.</i></p> <p>Gráfica de columnas:</p>	Nivel 1. Leer los datos	<u>Lectura</u> <u>literal:</u> del título y fuente	La lectura está basada en la descripción del contenido de la gráfica sin interpretarla.
<p><i>Entre los años 2013 y 2016 existe una <u>diferencia de menos 1.8%</u> en el aspecto de hogares con televisión y entre los años 2013 a 2016, se observa que existe un <u>aumento del 15.4%</u> en hogares con TV. de paga</i></p> <p>Gráfica de columnas:</p>	Nivel 2. Leer dentro de los datos	<u>Comparación:</u>  por columnas	La lectura muestra la comparación de datos por columnas, tomando como base el avance de los años, identificando el aumento o disminución en los valores de las variables.
<p><i>En esta gráfica se muestra el porcentaje de gente que tiene televisión y televisión de paga en los años 2013, 2014, 2015 y 2016. <u>En el año 2013 había más personas con televisión, mientras que en el 2016 aumento el porcentaje de hogares con televisión de paga.</u></i></p>	Nivel 2. Leer dentro de los datos	<u>Comparación:</u> de ambas maneras (horizontal y vertical)	La lectura muestra la comparación de datos de manera horizontal tomando como base el avance de los años, y de manera vertical al observar la variable con mayor porcentaje en el año 2013.

*También se muestra que entre en los años 2013 - 2016 como va aumentando y disminuyendo los porcentajes.*

Gráfica de columnas:

*En esta grafica se observa que el porcentaje de hogares con televisión en 2013 y 2014 fue el mismo con un 94.4%. Pero en el caso del porcentaje con televisión de paga aumento en ese año de 2013 a 2014 de un 36.7% a un 38.1%. Podemos decir que en el 2017 disminuirá el número de hogares con televisión pero aumentara el número de hogares con televisión de paga.*

Gráfica de columnas:

*Se puede observar que con el paso de los años ha ido aumentando el porcentaje de hogares que cuentan con televisión de paga, al mismo tiempo ha disminuido un mínimo porcentaje de hogares con televisión. De acuerdo con estos datos puedo concluir que hoy en día, en su gran mayoría la televisión de paga se ha vuelto una manera de entretenimiento necesaria en los hogares. Por ello es muy probable que siga aumentando con el tiempo.*

Tabla:

*Las casas con televisión han ido disminuyendo durante los últimos 4 años, pero las que tienen televisión de paga han aumentado. Algo un poco extraño pero entendible por el apagón analógico que se dio en el último año, por eso la televisión cada año va disminuyendo y la televisión de paga aumenta debido a la poca*

Nivel 3. Leer más allá de los datos

Predicción:  
de una tendencia

La lectura presenta la predicción de tendencias sobre las variables para el año 2017 considerando la información representada en la gráfica.

Nivel 4. Leer detrás de los datos.

Integración con el contexto:  
racional/literal

La lectura presenta comparaciones, predicciones de tendencias, y la integración de la información con el contexto al dar una conclusión basada en significados contextuales en términos de los datos que muestra la gráfica.

Nivel 4. Leer detrás de los datos.

Integración con el contexto:  
hipotético

La lectura muestra comparaciones y la integración de la información con el contexto; esto al evaluarse los datos estadísticos representados en la gráfica cuando se emiten hipótesis propias explicativas.

*y mucha demanda que tiene cada una.*

En la Tabla 2, se presentan las frecuencias de las interpretaciones por nivel y categoría emergente de los datos en cada uno de los niveles de lectura.

Tabla 2. Clasificación de acuerdo a nivel de lectura y característica

Nivel	Lectura de la gráfica de columnas		Total	Nivel	Lectura de la tabla		Total
	Categoría	Frecuencia			Categoría	Frecuencia	
Lectura idiosincrática	Perspectiva personal	3	3	Lectura idiosincrática	Perspectiva personal	4	4
Leer los datos	Lectura literal del título y fuente	1	1	Leer los datos	Lectura literal del título y fuente	0	0
	Comparación horizontal	17			Comparación por filas	1	
Leer dentro de los datos	Comparación vertical	3	23	Leer dentro de los datos	Comparación por columnas	17	23
	Comparación por ambas maneras	3			Comparación por ambas maneras	5	
Leer más allá de los datos	Predicción de una tendencia	2	2	Leer más allá de los datos	Predicción de una tendencia	1	2
	Predicción de un valor	0			Predicción de un valor	1	
	Integración con el contexto, racional/literal	1			Integración con el contexto, racional/literal	1	
Leer detrás de los datos	Integración con el contexto, hipotético	6	7	Leer detrás de los datos	Integración con el contexto, hipotético	6	7
	Total		36		Total		36

Como podemos observar, la mayoría de los estudiantes (23 de 36) muestran el nivel 2, leer dentro de los datos, en la tarea de interpretación de la gráfica de columnas y de la tabla; esto al efectuar la comparación de los datos estadísticos: 17 de 36 la efectúan de manera horizontal o por columnas identificando el aumento o disminución en los porcentajes de los valores de la variable tomando como base el avance de los años; 3 de 36 y 1 de 36, respectivamente, identifican el valor

de la variable con mayor porcentaje al compararlos por año; mientras que 3 de 36 y 5 de 36, respectivamente, realizan de la comparación de ambas maneras. Por otro lado, 7 de 36 se clasificaron en el nivel 4, leer detrás de los datos, en la lectura de la gráfica de columnas y de la tabla, al integrar la información con el contexto extrayendo conclusiones: 6 de 36 leen, aceptan y evalúan la información formando sus propias hipótesis explicativas, respectivamente; y sólo uno, en cada caso, realiza comparaciones, detecta tendencias y explica significados contextuales sin sugerir hipótesis sobre lo observado. En el nivel 3, leer más allá de los datos, se presentan 2 de 36 respuestas a la lectura de cada representación, gráfica y tabla: 2 de 36 y 1 de 36 estudiantes, respectivamente, proporcionan una predicción sobre la tendencia de los datos; y sólo uno da una predicción de un valor, “*en algún año la televisión de paga ocupará el 90% o más en México*”. Únicamente la respuesta de un estudiante se clasifica en el nivel 1, leer los datos, al mencionar en su lectura solo el título y la fuente de la gráfica, sin interpretarla. Además, se identificaron lecturas basadas en la experiencia individual limitada o en perspectivas personales del estudiante (3 de 36 y 4 de 36, en la lectura de la gráfica y tabla, respectivamente), clasificándolas en el nivel 0, lectura idiosincrática.

Para organizar y analizar las respuestas a las dos actividades de manera conjunta, se realiza una tabla de contingencia (Tabla 3) que divide las interpretaciones en los niveles de lectura expuestos en nuestro marco de referencia.

Tabla 3. Niveles de lectura alcanzados en cada actividad de manera conjunta

Lectura	Tabla					Total	
	Nivel 0	Nivel 1	Nivel 2	Nivel 3	Nivel 4		
Gráfica de columnas	Nivel 0	<b>1</b>	0	1	0	<u>1</u>	3
	Nivel 1	0	<b>0</b>	1	0	0	1
	Nivel 2	1	0	<b>19</b>	<u>1</u>	<u>2</u>	23
	Nivel 3	0	0	<u>1</u>	<b>1</b>	0	2
	Nivel 4	<u>2</u>	0	<u>1</u>	0	<b>4</b>	7
Total	4	0	23	2	7	36	

Analizando los niveles de comprensión alcanzados por los estudiantes de manera conjunta, se observa que la mayor parte de ellos reflejan lecturas en el nivel 2, leer dentro de los datos (19/36); pocos reflejan en ambas lecturas el nivel 4, leer detrás de los datos (4/36); sólo un estudiante refleja su lectura en el nivel 3, leer más allá de los datos; y otro una lectura idiosincrática. En general, la mayoría de los estudiantes fueron consistentes en su interpretación de la información representada, independientemente del tipo de representación. No obstante, identificamos que las interpretaciones de 8 estudiantes se ven influenciadas por el tipo de representación promoviendo niveles superiores, 4 por la lectura de la tabla y 4 por la gráfica.

## 7.6 Discusión y Conclusiones

En este estudio se puede apreciar una exploración de cómo los estudiantes de licenciatura interpretan los mismos datos estadísticos representados en una tabla y en una gráfica de columnas, identificando el nivel de lectura y los rasgos característicos que exhiben en sus respuestas; así como si el tipo de representación promueve niveles superiores.

Se observa que la mayoría de nuestros participantes, estudiantes de primer semestre de licenciatura en matemáticas, exhiben en sus respuestas el nivel 2 al efectuar una comparación de los datos estadísticos de manera horizontal, en la gráfica, y por filas, en la tabla, observando la variación de los datos de los valores de la variable tomando como base el tiempo. Pocos jóvenes alcanzan los niveles superiores propuestos por Curcio y cols. al integrar la información con el contexto extrayendo hipótesis explicativas (nivel 4), o al proporcionar una predicción sobre una tendencia observada en el comportamiento de los datos (nivel 3). Sólo un participante hace únicamente la lectura literal de algún componente de la gráfica. Sin embargo, es preocupante identificar respuestas idiosincráticas en estudiantes de pregrado, alrededor del 10% de los estudiantes. Enfatizamos este aspecto porque consideramos que la interpretación de datos estadísticos realizada por los estudiantes de pregrado, como los involucrados en nuestro estudio, deberían contemplar valoraciones críticas sobre la información, independientemente de su contexto.

Además, a partir del análisis, podemos mencionar que el tipo de representación empleado para representar información, tabular o gráfica, no afecta en la interpretación de los estudiantes participantes para alcanzar niveles superiores de lectura. Si bien no podemos generalizar nuestros resultados, nos da un panorama como futura línea de investigación, donde se indague la interpretación de tablas y otros tipos de gráficos, en diversos contextos y sin que se guíe al estudiante; por ejemplo, considerar la gráfica circular o de líneas, siendo estas unas de las más utilizadas en los periódicos de circulación nacional en México, junto con la gráfica de columnas aquí analizada (Ramírez, 2012).

## Referencias

- Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, 2 (3), 298-318.
- Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores*. Tesis de Doctorado no publicada. Universidad de Granada, España.
- Arteaga, P., Vigo, J. M. y Batanero, C. (2017). Niveles de lectura de gráficos estadísticos en estudiantes de formación profesional. En J. M. Muñoz, A. Arnal, P. Beltrán, M.L. Callejo y J. Carrillo (Eds.), *Investigación en Educación Matemática XXI* (pp. 229-238). Zaragoza: SEIEM.
- Batanero, C. (2001). *Didáctica de la Estadística*. Grupo de Investigación en Educación Estadística, Departamento de Didáctica de la Matemática. Granada: Universidad de Granada.
- Buendía, G. (2004). *Una epistemología del aspecto periódico de las funciones en un marco de prácticas sociales*. Tesis de Doctorado no publicada, Centro de Investigación y de Estudios Avanzados del IPN. México.
- Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: NCTM.
- Díaz-Levicoy, D., Arteaga, P. y Batanero, C. (2015). Gráficos estadísticos y niveles de lectura propuestos en textos chilenos de Educación Primaria. En C. Fernández, M. Molina y N. Planas (eds.), *Investigación en Educación Matemática XIX* (pp. 229-238). Alicante: SEIEM.
- Díaz-Levicoy, D., Batanero, C., Arteaga, P. y Gea, M. (2015). Análisis de gráficos estadísticos en libros de texto de educación primaria española. *UNION, Revista Iberoamericana de Educación Matemática*, 44, 90-112.



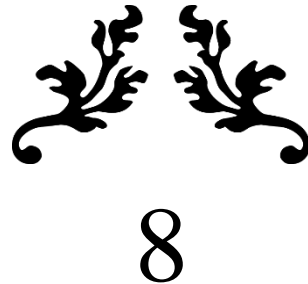
- Díaz-Levicoy, D., Batanero, C., Arteaga, P. y López-Martín, M. (2015). Análisis de los gráficos estadísticos presentados en libros de texto de educación primaria chilena. *Educação Matemática Pesquisa*, 17 (4), 715-739.
- Eudave, D. (2009). Niveles de comprensión de información y gráficas estadísticas en estudiantes de centros de educación básica para jóvenes y adultos de México. *Educación Matemática*, 21 (2), 5-37.
- Fernandes, J. A. y Morais, P. C. (2011). Leitura e interpretação de gráficos estatísticos por alunos do 9º ano de escolaridade. *Educação Matemática Pesquisa*, 13 (1), 95-115.
- Friel, S., Curcio, F. y Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education*, 32 (2), 124-158.
- Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, 70 (1), 1-25.
- Gea, M., Arteaga, P. y Cañadas, G. (2017). Interpretación de gráficos estadísticos por futuros profesores de Educación Secundaria. *Avances de Investigación en Educación Matemática*, 12, 19-37.
- Lisboa, G. (2002). *Interpretando e construindo gráficos de barras*. Tesis de Doctorado no publicada, Universidade Federal de Pernambuco. Brasil.
- Monteiro, C. y Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. *International Electronic Journal of Mathematics Education*, 2 (3), 188-207.
- Pagan, A., Leite, A. P., Magina, S. y Cazorla, I. (2008). A leitura e interpretação de gráficos e tabelas no Ensino Fundamental e Médio. *2º Simpósio Internacional de Pesquisa em Educação Matemática (SIPEMAT)*. Recife, Brasil.

Ramírez, T. (2012). *Tipos de gráficas que usan los periódicos para comunicar información*. Tesis de maestría no publicada. Universidad Autónoma de Guerrero. México.

Sánchez, T. y Arteaga, P. (2013). Los gráficos estadísticos en las directrices curriculares para la Educación Primaria en España y Colombia. En J. M. Contreras, G. R. Cañadas, M. M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 397-404). Granada: Departamento de Didáctica de la Matemática de la Universidad de Granada.

Vigo, J. (2016). *Comprensión de gráficos estadísticos por alumnos de formación profesional*. Tesis de maestría no publicada, Universidad de Granada. España.





---

**SUB&DES. Un juego que facilita el aprendizaje del diseño de medidas repetidas.**

---

Edwin Alonso Abarca Araya, Freisel Enrique Calvo Calderon, Alejandro Alberto Salmerón Arce.  
abaray94@hotmail.com, beche94@hotmail.com, alesalme123@gmail.com



**Resumen**

Uno de los mayores retos que se presentan al enfrentarse a un conjunto de datos es determinar el tipo de análisis idóneo que mejor se adapte a estos, mayormente en el caso de las personas que no son especializadas en Estadística. El diseño de medidas repetidas tiene un método de recolección que lo hace particular y por ende un análisis adecuado genera resultados óptimos del conjunto de datos, el problema es cuando no se aplican las herramientas correctas y se utiliza un análisis que no es el mejor para el tipo de datos recolectados. Se realizó un juego didáctico basado en materia referente al diseño de medidas repetidas con el fin de enseñar y a la vez evaluar en las personas interesadas este tipo de diseño. Para llevarlo a cabo se necesita una computadora y que el jugador

tenga conocimientos previos en regresión, pues se trata de un video juego, en donde éste tiene que utilizar un personaje con el cual va a atravesar cada una de las secciones que lo componen, en estas etapas el jugador se adentra en un ambiente estadístico enfrentándose primeramente a un repaso de los modelos de regresión lo cual es básico para aplicar el diseño de medidas repetidas, posteriormente a la diferencia entre efectos fijos y efectos aleatorios para finalmente arribar en el diseño de medidas repetidas, enfatizando en conceptos como sujeto, estudios longitudinales, pendiente y tendencia. El jugador se enfrenta a situaciones tales como tener que saber si en algunos ejemplos se utilizan variables con efectos fijos o aleatorios o bien determinar si el diseño más adecuado para desarrollar un análisis es el de medidas repetidas o no esto dependiendo de la forma en cómo se recolectaron los datos. También otro de los propósitos del juego es saber la diferencia entre un modelo de regresión simple y uno de medidas repetidas en términos de supuestos y la parte aleatoria.

**Palabras clave:** juego didáctico, modelos de regresión, medidas repetidas, efectos fijos y aleatorios.

### **Abstract**

One of the biggest challenges that arise when facing a data set is to determine the type of analysis that best suits them, especially in the case of people who are not specialized in Statistics. The repeated measures design has a particular collecting method that creates optimum results in a data set, however the problem is when the incorrect tools are used and the analysis it's not the best for the collected data. An educational computer game based on repeated measures designs contents was developed with the goal of teaching and evaluating Statistics students. The video game requires a computer and a player with prior knowledge of regression models. Through the video game platform the player must use his/her character to go through different stages, each of these being a statistical environment. In the first one, the player must face the regression model, which demonstrates the basics for the application of repeated measures designs. The following stage examines the difference between fixed and random effects. Then, at the final stage he/she arrives

to the repeated measures designs, where the game emphasizes various concepts such as longitudinal studies, slopes, and trends. The player comes across a variety of situations such as having to know if, in certain problems, he/she should use variables with fixed or random effects or even determining if the most appropriate design for developing an analysis is repeated measures or not, depending on the way data were collected. Another purpose of the game is to teach the differences between a simple regression model and one of repeated measures, in terms of assumptions and the random parameters.

**Keywords:** educational game, regression models, repeated measures, fixed and random effects.

## 8.1 Introducción

En muchos estudios experimentales, la identificación del tipo de diseño que se debe utilizar puede llegar a ser confuso para los investigadores, ya que algunos pueden poseer conocimientos muy básicos en estadística y la documentación existente de las técnicas en la experimentación, en sus campos de interés, no suele profundizar en la buena adaptación de estas al tipo de dato con el que se cuenta. (Serrano, 2003).

Para realizar un análisis adecuado de los datos experimentales obtenidos, se debe tener una metodología de recolección de datos sólida, además de un completo conocimiento de las distintas formas estadísticas de cómo se pueden analizar estos datos, siempre enfocado al objetivo del estudio sobre la relación de causa y efecto que se quiere medir. (Ávila, 2006)

El presente juego didáctico se plantea con la intención de ayudar a investigadores o estudiantes, con conocimientos en regresión, en la identificación del diseño de medidas repetidas como un método viable para el análisis de sus datos, siempre y cuando estos cumplan con un conjunto de características necesarias para poder aplicarlo; así como en el reconocimiento de los conceptos asociados a este tipo diseño. No abarca propiamente la forma de análisis que proveen estos diseños, sino la estructura en la recolección de datos requerida para poder aplicarlos y las características que tienen estos modelos.

En los diseños de medidas repetidas se ve el efecto intrasujeto, con lo que se consigue reducir la varianza del error, ya que se elimina la variabilidad que pertenece al sujeto, actuando como control propio. Los efectos de orden pueden confundir los efectos en los tratamientos. Se pueden ver efectos de período, los cuales ocurren cuando el sujeto responde al período o posición que ocupa el tratamiento en la secuencia, independientemente del tratamiento aplicado. Así como los efectos residuales, que es la persistencia de un tratamiento después del tiempo de aplicación. Este tipo de diseños se basan en modelos de regresión que deben cumplir con ciertos supuestos como por ejemplo, que los errores provengan de una distribución normal con media cero y varianza ( $S^2$ ), que exista homogeneidad en la varianzas, que no haya valores atípicos y que no se de la multicolinealidad, es decir que no exista correlación entre las variables predictoras. (Marín, s.f).

## 8.2 Objetivo general

Identificar el diseño de medidas repetidas en investigaciones o estudios experimentales.

La meta de este juego interactivo, es que la persona que acceda a este, sea capaz de:

- Identificar un diseño de medidas repetidas, como un método de análisis, una vez que se tienen los datos, o bien, saber la forma de plantear la recolección de observaciones para poder aplicar este tipo de diseño.
- Poner en práctica conceptos referentes a los modelos de regresión lineal.
- Reconocer los supuestos de un modelo de regresión lineal.
- Ser capaz de distinguir entre variables con efectos fijos y efectos aleatorios.

## 8.3 Conocimientos previos

La persona que interactuara con el juego deberá tener conocimientos en los análisis de regresión lineal, que abarcan desde las características de este tipo de modelos hasta los supuestos que deben cumplir para su validez.

En los análisis de medidas repetidas se toman mediciones de los individuos en el tiempo, para observar los posibles cambios en la respuesta tomando en cuenta la variabilidad del individuo. Esto es, según Arnau & Bono (2008):

Metodológicamente hablando, los elementos que son observados o medidos en diversas ocasiones se denominan unidades, individuos o sujetos. Los intervalos de tiempo en que se observa o registra la respuesta de las unidades de observación se denominan puntos de tiempo u ocasiones y pueden variar desde unos cuantos minutos a muchos años. A su vez, el conjunto de estas respuestas forma el perfil de respuesta (curva o tendencia) de cada unidad. Términos tales como diseño o estudio longitudinal suelen ser sinónimos de diseño de medidas repetidas. (pág. 33)

Como en los diseños de medidas repetidas se observan tendencias de los individuos, esto se puede ver como un modelo de regresión, con lo que los investigadores pueden confundirse al realizar un análisis de medidas repetidas.

Con este juego se pretende, que el jugador aprenda a reconocer el diseño de medidas repetidas en los estudios experimentales para el correcto análisis de los datos, o sea, que cuando la persona llegue a realizar un experimento, este pueda saber cómo llevar a cabo la metodología de recolección de datos, o bien si ya los tienen, saber identificar estos diseños con el fin de realizar un correcto análisis.

## **8.4 Desarrollo del Juego**

### **Materiales**

Para poder acceder al juego se requiere de una computadora con sistema operativo Windows.



### **Descripción del recurso**

Para el desarrollo de este trabajo se realizó un videojuego de rol, ya que este tipo de recursos llaman la atención de la población meta del juego. La escogencia del nombre y ambientación del videojuego, como pantallas, gama de colores, personajes y diálogos, se realizó con el fin de generar interés en la población meta de nuestro trabajo.

En el videojuego desarrollado, la persona interactúa con un personaje y está ambientado en un aula donde el profesor se encuentra impartiendo clases. El jugador debe ir respondiendo a las preguntas que se le brinden en los diferentes escenarios, las cuales tendrán relación con los diseños de medidas repetidas y los modelos de regresión lineal.

### **Desarrollo del contenido**

Al inicio se mencionan los controles necesarios para que la persona sea capaz de jugar. El juego transcurre en un aula donde la persona toma el lugar de uno de los estudiantes, tendrá que contestar las preguntas que se le realicen, siendo estas de respuesta múltiple, dándole la posibilidad a la persona de elegir entre las diversas opciones; se mencionara cuando la persona conteste de forma correcta o incorrecta, y luego de las preguntas el profesor dará la respuesta correcta correspondiente a cada caso (fig. 1).

También se encuentra un mini juego (fig. 2) donde se pondrá en práctica si se entendieron o no los conceptos de la variable con efectos fijos o efectos aleatorios, para el cual se muestran ejemplos de experimentos donde se le preguntará por una variable en específico. La persona tendrá que interactuar con el entorno para completar el mini juego, también se da la opción de salir del mini juego en caso de que la persona ya sepa identificar entre esas variables.

En un momento del juego, el personaje se queda dormido y se transporta a una instalación donde están llevando a cabo un experimento, pero con un análisis completamente inadecuado (fig. 3). Al final de este apartado, el personaje tiene la oportunidad de corregir al investigador en el análisis

que se está realizando reuniendo los conceptos que se van abarcando a lo largo del juego. De esta manera se busca que la persona sea capaz de ir comprendiendo el diseño de medidas repetidas.

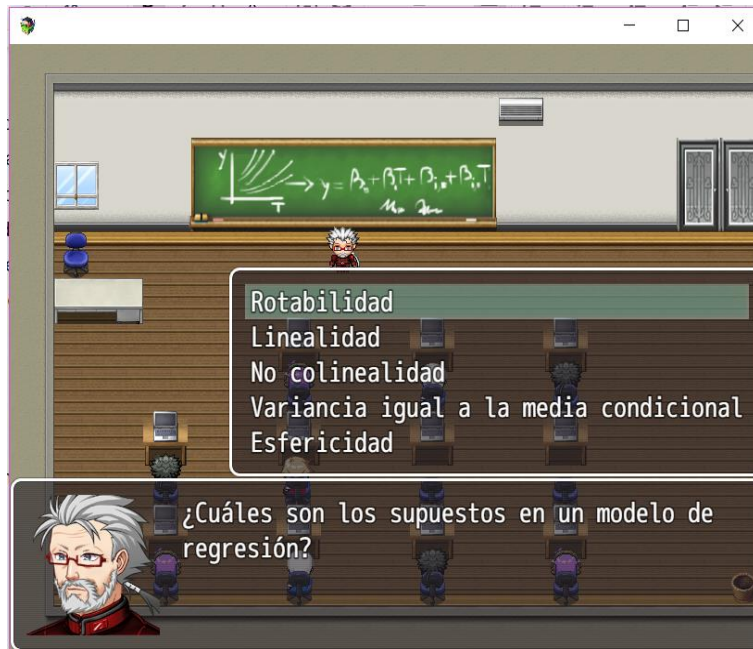


Figura 1. Ambientación del videojuego, desarrollo de preguntas por el profesor.



Figura 2. Ambientación del videojuego, desarrollo del mini juego sobre efectos fijos y aleatorios.



Figura 3. Ambientación del videojuego, desarrollo de experimento de medidas repetidas.

Al despertar, como refuerzo se le presenta un estudio donde se puede aplicar el análisis de medidas repetidas (fig. 4), donde tiene que reconocer las variables del estudio e identificar qué tipo de variables son, según el conocimiento que adquiere en el juego, siendo esto un repaso de algunos conceptos que se ven durante el juego.

Al final del juego (fig. 5) el profesor hace un breve resumen de las características de los diseños de medidas repetidas, además de los conceptos que se van adquiriendo durante el juego, para reforzar los conocimientos adquiridos de la persona y enlazar los puntos que pueden haber quedado dispersos para el jugador.

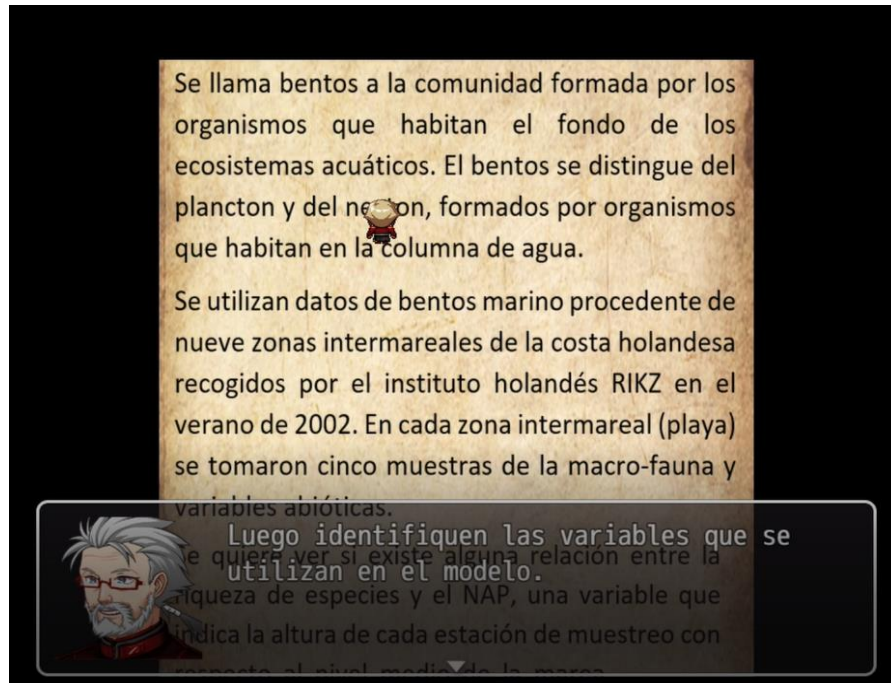


Figura 4. Ambientación del videojuego, estudio donde se aplica el análisis de medidas repetidas.

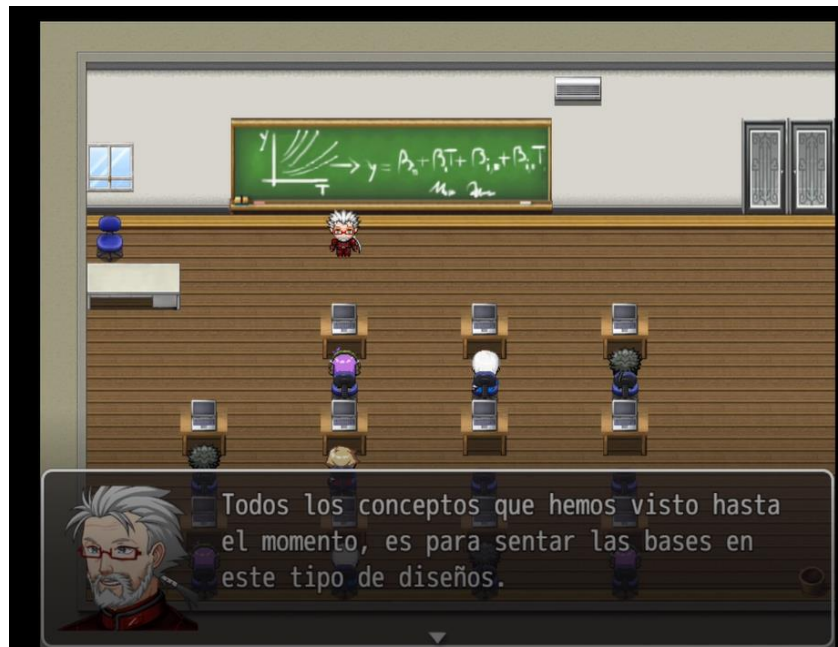


Figura 5. Ambientación del videojuego, resumen de los conceptos del diseño de medidas repetidas presentes en el juego.

## Referencias

- Ávila, H. (2006). *Introducción a la metodología de la investigación*. Obtenido de <http://www.eumed.net/libros/2006c/203/>
- Arnau, J. & Bono, R. (2008). *Estudios longitudinales de medidas repetidas: Modelos de diseño y análisis*. *Escritos de Psicología (Internet)*, 2(1), 32-41. Recuperado en 21 de mayo de 2018, de [http://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1989-38092008000300005&lng=es&tlng=es](http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1989-38092008000300005&lng=es&tlng=es).
- Marín, J. (s.f). [*Análisis de varianza con medidas repetidas*]. Curso de expertos de U.C.M. Obtenido de <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/GuiaSPSS/16anovar.pdf>
- Serrano, R. (2003). *Introducción al análisis de datos experimentales: tratamiento de datos en bioensayos*.
- RPG Maker (Versión MV) {software}. (2015). Mundial: Kadokawa Games. Obtenido de <http://www.rpgmakerweb.com/products/programs/rpg-maker-mv>



---

## **Análisis estadístico en una evaluación, una aplicación de la técnica clúster k-media.**

Juana Elisa Escalante Vega, Francisco Sergio Salem Vega, Emmanuel Morales García.

[jescalante@uv.mx](mailto:jescalante@uv.mx), [fsergios@gmail.com](mailto:fsergios@gmail.com), [emmanuelgarciam94@gmail.com](mailto:emmanuelgarciam94@gmail.com)



### **Resumen**

El trabajo de investigación tiene el propósito de realizar un análisis estadístico de los resultados obtenidos en el examen de ingreso a diversos programas del área económica administrativa ofrecida por la Universidad Veracruzana. Estos exámenes son aplicados por el Centro Nacional de Evaluación para la Educación Superior (Ceneval) especializado en el diseño y la aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias, así como el análisis y la difusión de los resultados que arrojan estas pruebas. Dichas pruebas se centran en 4 áreas. Estas son: Comprensión lectora, pensamiento analítico, pensamiento matemático, estructura de la lengua. Se analizó la base de datos de los alumnos de nuevo ingreso 2017. En primer lugar se realiza un

análisis univariado donde se observó que los alumnos presentan dificultades en todas las áreas evaluadas. Se utilizó la técnica clúster k-medias para formar grupos divididos por eficiencia. Esta técnica conformo tres grupos: Baja eficiencia, regular y alta eficiencia. Además permitió caracterizar a los alumnos de nuevo ingreso de cada programa educativo en términos de sus puntuaciones en cada área.

**Palabras clave:** Análisis estadístico, clúster k-medias, habilidades del pensamiento, Ceneval.

### **Abstract**

The purpose of this work is the perform a statistical analysis of the results obtained in the application test to various programs of the administrative economic area offered by the Universidad Veracruzana. These exams are applied by the Centro Nacional de Evaluación para la Educación Superior (Ceneval) specialized in the design and application of tools for assessing knowledge, skills and competencies, as well as the analysis and dissemination of the result of these test. These test focus on 4 areas. These are: reading, analytical thinking, mathematical thinking and language structure. The database of the new 2017 student was analyzed. First, a univariate analysis was carried out, showing that the student presented difficulties in all the areas evaluated. The k- means cluster technique was used to form groups divided by efficiency. This technique formed three groups: low efficiency, regular efficiency and high efficiency. It also characterized the new students of each educational program in terms of their scores in each area.

**Keywords:** Statistical analysis, K –means cluster, thinking skills.

## **9.1 Introducción**

En el área profesional las habilidades que desarrolla el ser humano son significativas, dado que con estas se logra un desempeño eficiente y de calidad. La universidad pública que se analiza en este trabajo se caracteriza por realizar un proceso de selección anualmente para la admisión a los diferentes programas educativos que oferta, los estudiantes que solicitan ingreso a esta universidad,

en primera instancia tienen que seleccionar el programa educativo al que desean incorporarse, continúan con un proceso administrativo requerido por la institución, posteriormente son evaluados por el Centro Nacional de Evaluación para la Educación Superior (Ceneval), cuya actividad principal es el diseño y la aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias, así como el análisis y la difusión de los resultados que arrojan las pruebas. Ceneval aplica el examen de admisión a todos los aspirantes con el fin de evaluar las habilidades que se consideran sustanciales para el desarrollo académico. Entre las habilidades que se evalúan se encuentran, comprensión lectora, pensamiento analítico, pensamiento matemático y estructura de la lengua. Esta investigación tiene como finalidad examinar los resultados de esta evaluación y detectar áreas de mejora para los estudiantes aceptados, de ser necesario, proponer el desarrollo de un programa de apoyo para los estudiantes que presenten eficiencia baja en las habilidades, así como generar estrategias que permitan reforzar los conocimientos durante su transitar por el programa educativo seleccionado. Para el estudio de los datos, se efectúa un análisis descriptivo general de las variables de estudio, posteriormente para reconocer patrones de comportamiento de los individuos se desarrolla el método estadístico multivariante análisis clúster (K-medias).

## 9.2 Metodología

Para poder analizar los datos y proponer la técnica estadística, se necesita conocer el tipo de datos que se estudiarán utilizados en esta investigación, por lo que a continuación se realiza una descripción general de estos.

La matriz de datos que se ocupó para poder realizar este estudio fue obtenida del registro de ingreso a las licenciaturas, Ingeniería del software, Redes y servicios de cómputo, Tecnologías computacionales y Estadística, todas pertenecientes a la Facultad de Estadística e informática de la Universidad Veracruzana, dicho ingreso fue del año 2017. A continuación se muestra la tabla 1 para dar más características de los datos.



Tabla 1. *Descripción de las variables.*

Clasificación de la variable	Tipo de la variable	Escala de medición
Cuantitativa	Numérica	Razón

*Nota: En la tabla 1 se muestran las características principales de la matriz de datos que se ocupó para llevar a cabo el análisis estadístico. Teniendo en cuenta la clasificación de las variables, el tipo (numérica) y para finalizar su escala de medición.*

Esta sección se concentra en el análisis estadístico, en donde se hablará de la técnica multivariante Clúster (K-medias), fue elegida por que tenemos una matriz de datos multivariante y se necesita saber el patrón de comportamiento de los alumnos que presentaron el examen para ingresar a estas licenciaturas, así mismo identificar las habilidades que estos presentan. Clúster: Es una técnica multivariante que tiene como finalidad crear grupos homogéneos de las variables, según la similitud que exista entre ellos (Peña, 2002).

Clúster K-Medias: Es un método que agrupa los casos que están contenidos en las variables, basándose en las distancias existentes entre ellos. Está destinado a situaciones en las cuales todas las variables son del tipo cuantitativo, y la distancia cuadrática Euclídea (Giménez, 2010).

El procedimiento del análisis clúster de K-medias empieza con la construcción unos centros de conglomerados iniciales. Podemos asignar estos nosotros mismos o tener un procedimiento de selección de k observaciones bien situadas para los centros de conglomerados.

Después de la obtención de los centros de los conglomerados, el procedimiento:

- Asigna casos a los conglomerados basándose en la distancia de los centros de los conglomerados.

- Actualizar las posiciones de los centros de los conglomerados basándose en los valores medios de los casos en cada conglomerado.

Esto se repite hasta que los conglomerados tengan la misma similitud entre sus variables.

## 9.3 Resultados

### 9.3.1 Análisis descriptivo univariante

En la siguiente figura logramos la descripción estadística para cada variable, la cual podemos observar en los histogramas de frecuencia.

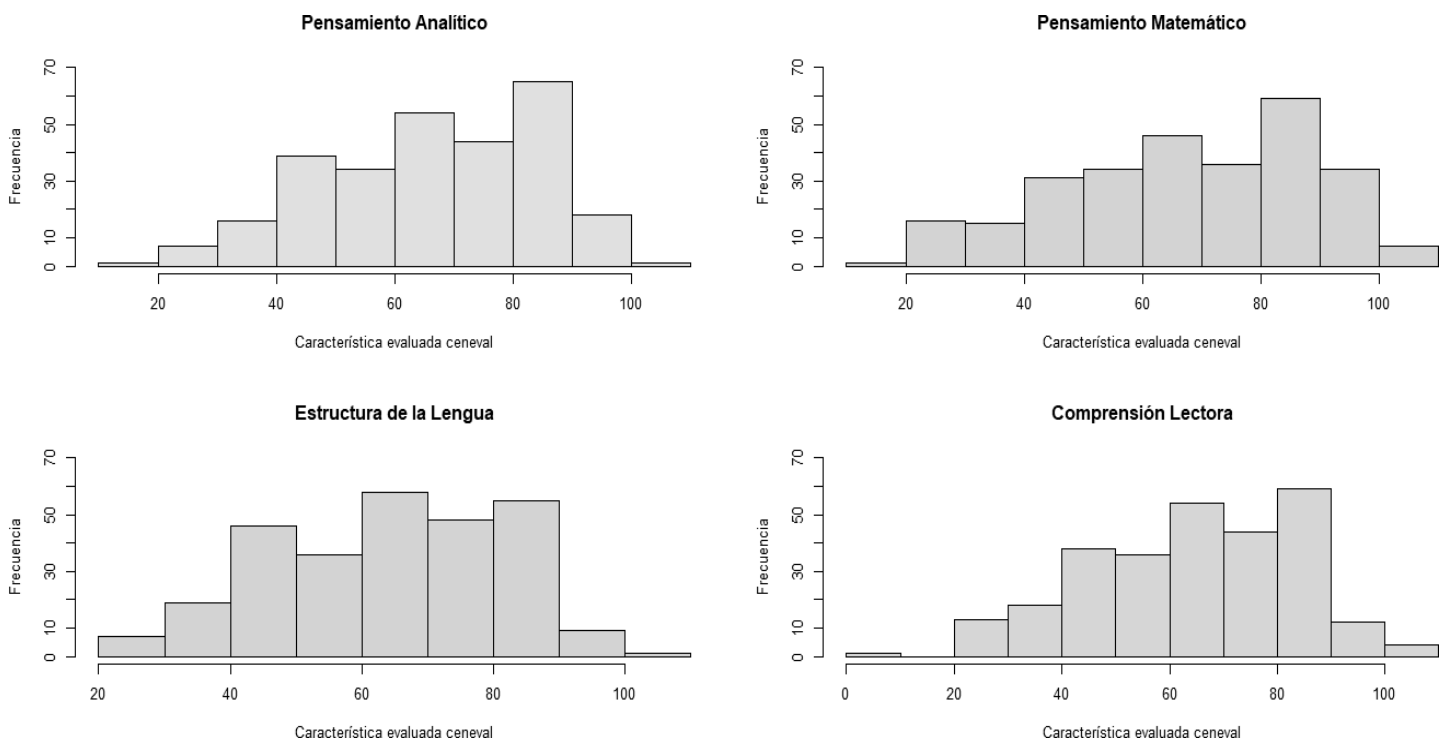


Figura 1. Histogramas de frecuencias de las variables en estudio.  
Elaboración propia usando: R Project 3.4.2

Podemos apreciar (véase figura 1) que en pensamiento analítico, comprensión lectora y estructura de la lengua solo el 25% de los estudiantes poseen la habilidad. Mientras que pensamiento matemático, aunque tiene un porcentaje considerable de estudiantes que la demuestra de manera general el 75% exhibe deficiencia para esta, el 25% de los estudiantes es considerado en un promedio de 7-9 mientras que el resto es menor o igual a 6.

### 9.3.2 Análisis multivariante

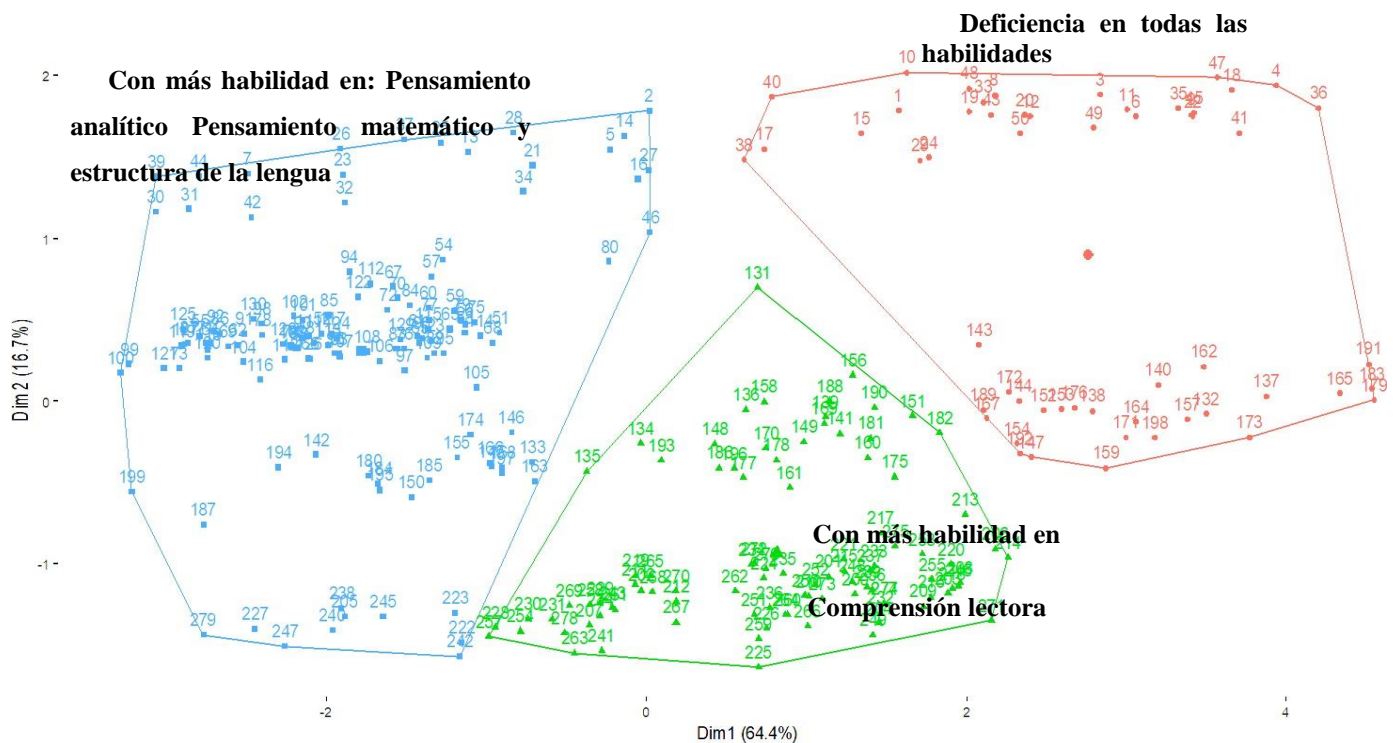


Figura 2. Grupos Homogéneos: Clúster k-medias.  
Elaboración propia usando: R Project 3.4.2

Como se puede observar en el clúster (Figura 2) existen tres agrupaciones, a continuación, se describirán cada una de ellas.

**Grupo 1:** Este grupo está caracterizado por todos los alumnos que claramente tienen una deficiencia en las habilidades evaluadas en el examen de ingreso a la Universidad Veracruzana. Las licenciaturas que están dentro de este grupo son, Estadística con 58% de los alumnos que

realizaron el examen para este programa educativo, mientras que la licenciatura en redes y servicio de computo cuenta con 36.6% de alumnos con deficiencia.

**Grupo 2:** Tiene como característica principal a todos aquellos alumnos que son regulares en las habilidades, aunque de acuerdo con los promedios de cada sección se puede observar que les hace falta más conocimiento y desarrollo de dichas habilidades, cabe de mencionar que este grupo tiene un alto rendimiento en la comprensión lectora. En este bloque las licenciaturas que tienen mayor concentración son, tecnologías computacionales con un 87.5% de los alumnos que eligieron este programa educativo, nuevamente redes y servicio de computo tiene un 36.2% de alumnos regulares.

**Grupo 3:** Este bloque concentra a todos los alumnos que obtuvieron un promedio total de 70 puntos a 99 puntos general del examen, es decir que son alumnos que tienen mayor habilidad en pensamiento analítico, pensamiento matemático y estructura de la lengua. En esta sección se tienen un porcentaje más bajo en cuanto a la habilidad comprensión lectora. En esta agrupación se concentran alumnos de todas las licenciaturas la única diferencia que, con porcentajes diferentes, Licenciatura en ingeniería del software 100% de los alumnos que han elegido este programa se concentran en dicho grupo de alto conocimiento en las diferentes habilidades, estadística con un 42%, redes y servicio de computo con 26.08% y por último solo con 12.5% de alumno dentro de este grupo se encuentra la licenciatura en tecnologías computacionales.

Como se puede apreciar en los resultados el patrón de comportamiento es revelador, siendo las licenciaturas en redes y servicio de cómputo y estadística los programas educativos que fueron seleccionados por alumnos que no cumplen con las diferentes habilidades. Mientras que los alumnos de ingeniería del software son alumnos que poseen un nivel más alto en el desarrollo de estas habilidades, y observamos que los estudiantes que fueron aceptados en el programa de Tecnologías computacionales están en un nivel regular.

Para fines de la técnica estadística multivariante empleada se tomará en cuenta la explicación total de la varianza ambas dimensiones. Con un 81.1% se explica toda la información es decir que no existe mucha perdida de la información original y solo se contemplan dos dimensiones.

#### **9.4 Conclusiones**

Para el estudio de los datos, se desarrolló un análisis descriptivo general de las variables de estudio, donde obtenemos los siguientes resultados que muestran que solo el 25% de los estudiantes poseen la habilidad de pensamiento analítico, comprensión lectora y estructura de la lengua y que el 75% de los estudiantes que muestran la habilidad de pensamiento matemático exhiben deficiencia en ella. Para reconocer patrones de comportamientos de los individuos se empleó el método estadístico multivariante análisis clúster (K-medias), lo que nos permitió caracterizar tres grupos;

El grupo uno, este grupo está caracterizado por todos los estudiantes que claramente tienen una deficiencia en las habilidades, y está conformado por el 58% de los estudiantes aceptados en la licenciatura de Estadística y el 36.6% estudiantes aceptados en la licenciatura en Redes y Servicio de Computo.

El grupo dos tiene como característica principal a todos aquellos estudiantes que son regulares en las habilidades pero con áreas de mejora potencial, conformado por el 87.5% de estudiantes que ingresan a la licenciatura de Tecnologías Computacionales y un 36.2% de los estudiantes que cursaran la licenciatura en Redes y Servicio de Computo.

El grupo tres, concentra a todos los estudiantes que tienen mayor habilidad en pensamiento analítico, pensamiento matemático y estructura de la lengua, la licenciatura en Ingeniería de Software con un 100%, Estadística con un 42%, Redes y Servicio de Computo con 26.08% y con 12.5% estudiantes de la licenciatura en Tecnologías Computacionales.

Estos resultados nos permitirán proponer el desarrollo de estrategias que permitan reforzar los conocimientos de los estudiantes durante su transitar por el programa educativo correspondiente.

Para complementar esto sería conveniente llevar a cabo investigaciones que nos permitan conocer los estilos de aprendizaje de estos estudiantes así como construir un programa de seguimiento.

## Referencias

Giménez, Y. (23 de Marzo de 2010). Tesis de Licenciatura. *Clasificación no supervisada: El método de k-medias*. Buenos Aires, Argentina: Universidad de Buenos Aires.

Lê, S., Josse, J. y Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. **25(1)**. pp. 1-18.

Peña, D. (2002). *Análisis de datos multivariantes*. España: Mcgraw-Hill / Interamericana de España, s.a. ISBN9788448136109.

R Development Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0URL <http://www.R-project.org>.



# 10

---

## **Técnicas de conteo con software aplicadas a la Probabilidad**

---

Mary Carmen Gómez Conde, Margarita Amaro Aranda, Eduardo Moreno Barbosa

mcgzc@yahoo.com.mx, mgmaro1@gmail.com, eduardo.morenoba@correo.buap.mx



### **Resumen**

La Probabilidad y Estadística son dos áreas de gran aplicación en diversos campos de la ciencia (matemáticas, matemáticas aplicadas, física y física aplicada), la tecnología (procesos de control de diseño) e inclusive en el área industrial y de procesos (Líneas de producción). Sin embargo, durante la educación de científicos o ingenieros, resulta, desde tedioso hasta complicado aplicar las teorías para poder resolver problemas prácticos en cada una de las áreas mencionadas. Uno de los principales, es que el estudiante requiere del uso de técnicas de conteo para poder en lo sucesivo definir una probabilidad. En este trabajo se discuten problemas clásicos que le permitirán al alumno



entender de mejor manera esta área y se propone el uso de herramientas matemáticas como son WolframAlpha y GeoGebra ambos, software de acceso libre y el uso de Excel, esto, con la finalidad de apoyar en el entendimiento de técnicas de conteo.

**Palabras clave:** educación, probabilidad, conteo, problemas, software.

### **Abstract**

Probability and Statistics are two areas of great application in various fields of science (mathematics, applied mathematics, physics and applied physics), technology (design control processes) and even in the industrial and process area (Production lines). However, during the education of scientists or engineers, it is tedious to complicated to apply the theories to solve practical problems in each of the areas mentioned. One of the main ones is that the student requires the use of counting techniques to be able to define a probability in the future. In this paper we discuss classical problems that will allow the student to better understand this area and propose the use of mathematical tools such as WolframAlpha and GeoGebra, both free access software and the use of Excel, in order to support in the understanding of counting techniques.

**Keywords:** education, probability, counting, problems, software.

## **10.1 Introducción**

Contar consiste en enumerar eventos difíciles de cuantificar, la enumeración de los eventos que se pueden presentar, en ocasiones es difícil y laboriosa por la cantidad de puntos a contar o enumerar, propiciando que se puedan cometer errores al emprender esa tarea.

En la experiencia propia, el primer problema que se ha encontrado en estos años dedicados a la docencia es que a los alumnos les cuesta trabajo pasar del lenguaje común al lenguaje matemático, usando la teoría de conjuntos y las técnicas de conteo para poder ocupar las reglas básicas de probabilidad, enfatizándose más este problema en estudiantes de ingeniería o incluso de carreras del área económico administrativas.

Antes de empezar a plantear algunos problemas, recordemos los siguientes conceptos básicos de probabilidad, encontrados en los libros de texto clásicos (Seymour, 1998) (Jay, 2008) (Hines & Montgomery, 2005).

**Experimento:** es el procedimiento mediante el cual se somete a verificación una o varias hipótesis relacionadas con un determinado fenómeno.

**Experimento Aleatorio:** es aquel que es regido por el azar y se puede repetir indefinidamente, siempre en las mismas condiciones. También antes de realizar el experimento, se conocen todos los resultados posibles, pero no se tiene certeza de cuál será el resultado de éste.

**Espacio muestral o Muestra ( $\Omega$ ):** Se basa en el conjunto de los posibles resultados de un experimento aleatorio.

**Evento (Medible):** En teoría de la probabilidad, a cada evento aleatorio se le puede asignar una medida de probabilidad, y el conjunto de todos los sucesos aleatorios constituye una  $\sigma$ -álgebra de conjuntos.

## 10.2 Enfoque Clásico

Al tener un experimento equiprobable. La probabilidad de que ocurra el evento  $A$  se define de la siguiente manera:

$$P(A) = \frac{n_A}{n} = \frac{\text{Número de elementos de un evento}}{\text{Número de elementos del espacio muestral}}$$

En donde  $n_A$  es el número de casos que cumple determinada condición y  $n$  el número total de casos posibles en el experimento. Notemos que en el enfoque clásico (cuando es aplicable) se determinan los valores de probabilidad antes de observar los resultados experimentales, por esta razón se le denomina *enfoque a priori*. En la tabla 1 se muestra un ejemplo de este enfoque y el

calculo de probabilidad del uso de un determinado procesador en un conjunto de computadoras personales.

Tabla 1. Uso de un determinado procesador en un lote de computadoras personales.

Procesador	Con CD-ROM	Sin CD-ROM	Total
Pentium IV	12	18	<b>30</b>
Celeron	33	37	<b>70</b>
TOTAL	<b>45</b>	<b>55</b>	100

Al realizar la razón definida en la ecuación anterior, obtenemos los siguientes valores numéricos, debido a que se está manejado proporciones, la multiplicación de este valor por cien nos da el porcentaje de total de elementos con determinada característica, enunciándose de la siguiente manera.

$$\frac{30}{100} = 0.3 \Rightarrow \frac{n_A}{n} \qquad \frac{37}{55} = 0.67 \Rightarrow \frac{n_A}{n}$$

Del total de ordenadores, 30% utilizan un procesador Pentium IV

De los ordenadores que no tienen CD-ROM un 67% tendrán procesador Celeron.

### 10.3 Enfoque de Frecuencia o Frecuencial

Se determina la probabilidad con base en la proporción de veces que ocurre un resultado favorable en un determinado número de observaciones o experimentos. No hay implícita ninguna suposición previa de igualdad de probabilidades. En este caso es necesaria la observación y recopilación de datos para determinar los valores de probabilidad.

La probabilidad del enfoque Frecuencial se asocia directamente al concepto de frecuencia relativa, y se define como:

$$P(A) = \frac{n_f}{n} = \frac{\text{Número de resultados favorables}}{\text{Número de veces que se repite el experimento}}$$

El siguiente ejemplo muestra el cálculo de la probabilidad mediante este enfoque. Se tiene una caja con 9 transistores semiconductores, 7 tipo NPN (4 de ellos de un modelo diferente) y dos tipo PNP. El experimento consiste en sacar uno a uno de estos, una vez verificado el tipo de transistor se calcula la probabilidad para un transistor en específico.

$$PNP \rightarrow \frac{2}{9} = 0.22 \Rightarrow 22\%$$

$$NPN \rightarrow \frac{7}{9} = 0.77 \Rightarrow 77\%$$

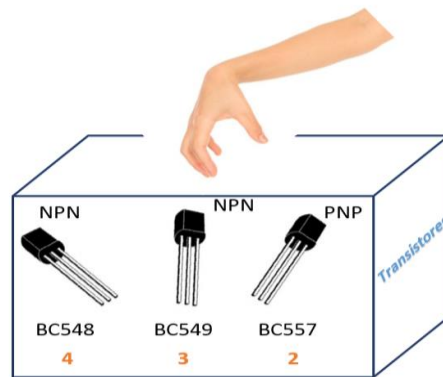


Figura 1. Elección de Transistores

Ambos enfoques expuestos anteriormente, solo contemplan el caso clásico de probabilidad y es cuando se tiene un número finito de elementos en el espacio muestral.

### 10.4 Desarrollo Axiomático de Probabilidad.

La medida de probabilidad ( $P$ ), está construida con argumentos de Teoría de Medida. Retomando las definiciones previas de experimento aleatorio, espacio muestra y suceso o evento, se establece lo siguiente:

$P: \Omega \rightarrow [0,1]$ , es una función que asigna a cada uno de los sucesos un valor numérico como medida de probabilidad y satisface los siguientes axiomas;

$$1) \quad 0 \leq P(A) \leq 1$$

$$2) \quad P(\Omega) = 1$$

$$3) \quad \text{Si } A \text{ y } B \text{ son eventos disjuntos, es decir, } A \cap B = \emptyset \text{ entonces} \quad P(A \cup B) = P(A) + P(B)$$

$$4) \quad \text{Sea } A_1, A_2, \dots, A_n, \dots \text{ una sucesión de eventos disjuntos dos a dos, es decir, } A_i \cap A_j = \emptyset, \text{ si } i \neq j, \text{ entonces } P[\cup_{k=1}^{\infty} A_k] = \sum_{k=1}^{\infty} P(A_k)$$

*NOTA:* Denótese por  $A, B, C$  eventos diferentes, entonces por la teoría de conjuntos el conectivo “**o**” representa unión de eventos ( $A \cup B$ ), y el conectivo “**y**” representa intersección de eventos ( $A \cap B$ ).

### 10.5 Primer principio de conteo.

Si un evento puede realizarse de  $n_1$  maneras diferentes y si continuando el procedimiento, un segundo evento se puede realizar de  $n_2$  maneras diferentes, y si, después de efectuados, un tercer evento puede realizarse de  $n_3$  maneras diferentes, y así sucesivamente hasta el  $k$ -ésimo evento, entonces el número de maneras en que los eventos pueden realizarse en el orden indicado es el producto

$$n_1 \cdot n_2 \cdot n_3 \cdots n_k$$

Una forma de que el estudiante entienda el uso de este principio es en problemas que se razonan como una secuencia de pasos, esto es, formas diferentes de efectuar el paso 1, seguido de formas diferentes de realizar el paso 2 y así sucesivamente o problemas modelados como el llenado de casillas. Las aplicaciones a estos problemas pueden ser: generación de placas, códigos de acceso, talonarios de boletos, etc.

*Ejemplo 1:* Si una persona tiene 5 camisas, 3 pantalones y dos pares de zapatos, ¿De cuántas formas distintas puede vestirse?

Solución:

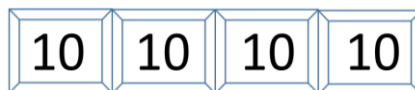
Por el primer principio de conteo, se tiene  $(5)(3)(2) = 30$ , esto es existen treinta formas distintas de vestirse.

*Ejemplo 2:* ¿Cuántos códigos de 4 dígitos pueden generarse, si se permite repetición?

Solución:

Visto este problema como llenado de casillas se tiene

$(10)(10)(10)(10) = 10^4 = 10000$  códigos **con repetición**.



Mientras que los códigos **sin repetición** serán:

$(10)(9)(8)(7) = 5040$  códigos



## 10.6 Segundo Principio de Conteo

Si en una actividad, ésta puede realizarse de  $k$  formas disjuntas, en donde la primera forma puede hacerse de  $n_1$  subformas posibles, denotada por  $n_1$ , la forma 2 puede hacerse de  $n_2$  subformas posibles, y así de manera sucesiva hasta la forma  $k$ -ésima de  $n_k$  subformas posibles, entonces el número de formas distintas estará determinada por:

$$n_1 + n_2 + \dots + n_k$$

En este caso para que el estudiante diferencie el primer principio de conteo respecto al Segundo principio, es que el primer principio se refiere a secuencia de pasos para cubrir el proceso y el Segundo se refiere a efectuar un proceso que tiene formas disjuntas de realizarse.

Ejemplo: Un restaurante tiene en su menú de postres, 4 clases de panqués, 2 clases de galletas y 3 clases de helado. Encuentre el número de formas en las que una persona puede seleccionar: a) uno de los postres, b) uno de cada clase de postre.

*Solución 1:*

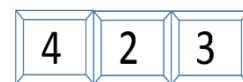
a) Por el Segundo Principio de Conteo, puede elegir un panqué o una galleta o un helado (formas disjuntas).

$$4+2+3 = 9 \text{ formas distintas de elegir un postre}$$

*Solución 2:*

b) Por el Primer Principio de Conteo (Figura 2), Puede elegir un panqué y una galleta y un helado (se puede ver como llenado de casillas).

$$(4)(2)(3) = 24 \text{ formas distintas de elegir un postre de cada tipo.}$$



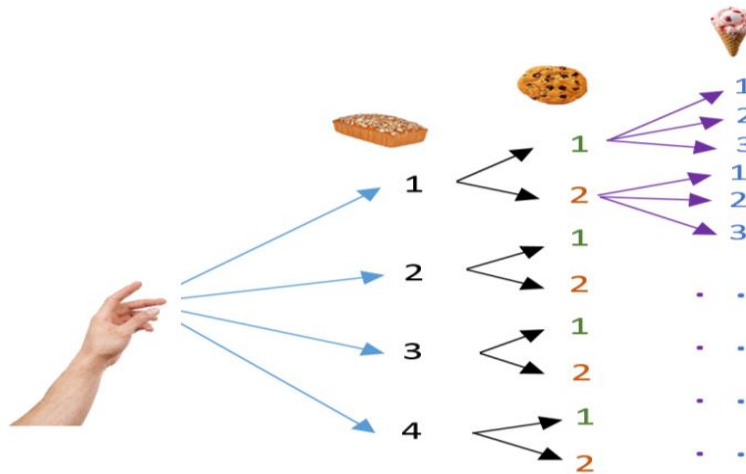


Figura 2. Diagrama de árbol para la demostración del principio de multiplicación.

Una problemática muy común es que el estudiante confunda el concepto de permutación (orden importa) con el de combinación (el orden no importa). En la teoría de Probabilidad y Estadística se ocupa mucho el concepto de muestra la cual claramente representa una combinación.

### 10.7 Permutaciones.

El número de permutaciones de un conjunto de  $n$  objetos distintos es

$$n! = n(n-1)(n-2)(n-3) \cdots (2)(1)$$

El número de permutaciones  $r$  de un conjunto de  $n$  objetos distintos es

$$nPr = n(n-1)(n-2)(n-3) \cdots (n-r+1)$$

**Ejemplo:** Al tener 10 símbolos diferentes de cuántas maneras diferentes se pueden ordenar 5 de ellos.

$$10P5 = (10)(9)(8)(7)(6) = 30240$$



Este cálculo de permutación respectivamente en WolframAlpha, GeoGebra y Excel (Gámiz Casarrubias & Gámiz Casarrubias, 2012) queda como se muestra en la Figura 3:

The figure displays three different interfaces used for calculating the permutation of 10 objects taken 5 at a time. At the top is the WolframAlpha interface, where the input 'permutation(10,5)' is entered into the search bar. Below the search bar, the formula  $\frac{10!}{(10-5)!}$  is shown, and the result is 30240. The middle part of the image shows the GeoGebra interface, where the command 'a = nPr(10,5)' is entered, resulting in '→ 30240'. The bottom part shows the Microsoft Excel interface, where the formula '=PERMUT(10,5)' is entered into cell B8, resulting in the value 30240.

Figura 3. Cálculo de permutaciones.

### 10.8 Combinaciones.

Una combinación de  $n$  objetos tomados  $r$  a la vez, es un subconjunto de  $r$  objetos tomados de entre los  $n$  (el orden no importa).

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

Ejemplo: Una empresa emplea 18 trabajadores en el turno del día, 17 en el turno de la tarde y 11 en el turno de media noche. Al seleccionar a 5 de estos trabajadores, ¿Cuál es la probabilidad que por lo menos dos turnos diferentes estén representados en los trabajadores seleccionados?

Solución:

Por los menos dos turnos en la muestra significa incluir exactamente dos o tres turnos, pero obviamente esto incluye muchos casos entonces lo ideal será plantearlo por complemento:

$$P(\text{al menos dos turnos}) = 1 - P(\text{mismo turno}) = 1 - \left[ \frac{\binom{18}{5} + \binom{17}{5} + \binom{11}{5}}{\binom{46}{5}} \right] = 0.999$$

Este problema en WólframAlpha ,Excel y geogebra respectivamente se calcula como se ve en la Figura 4:

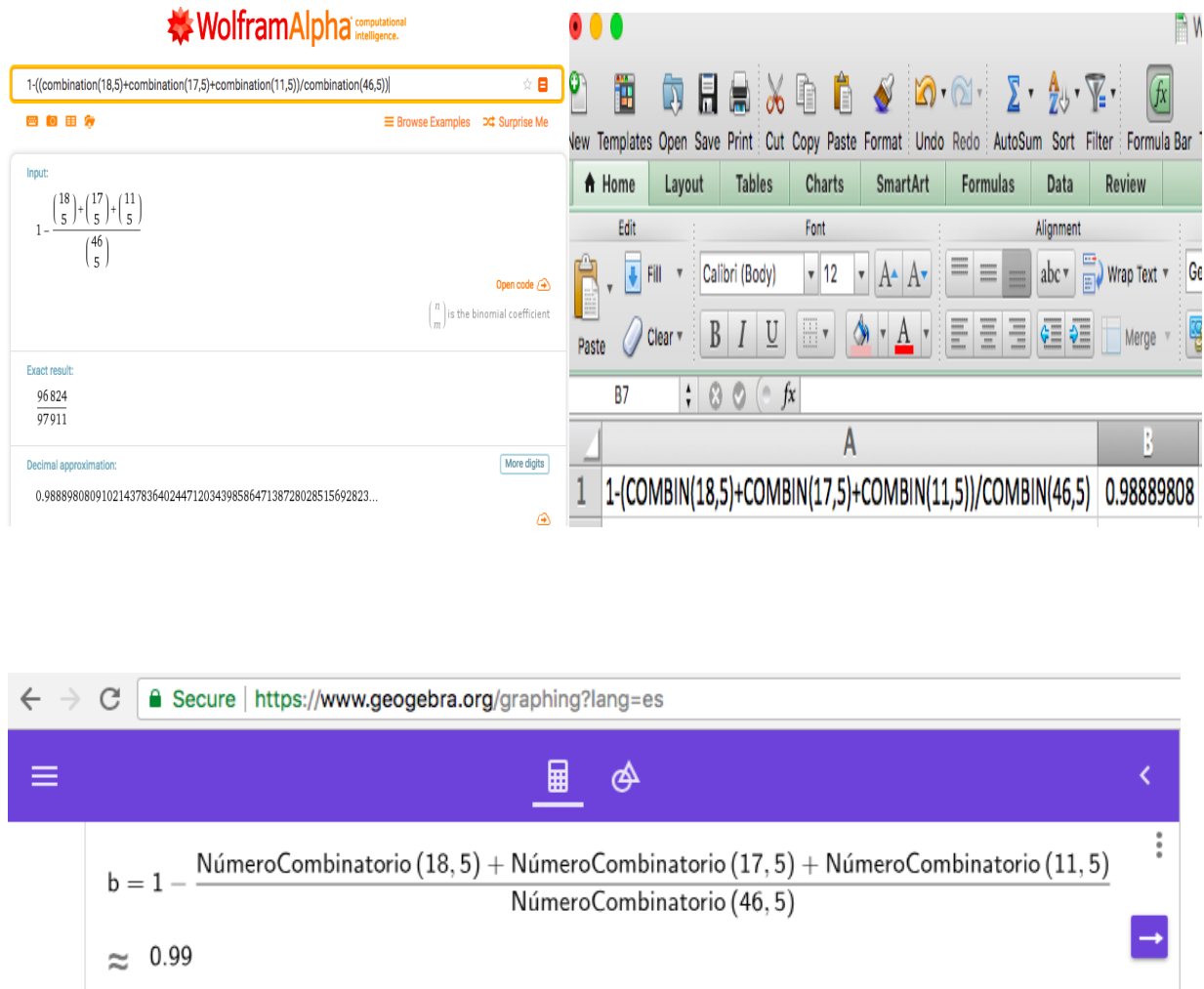


Figura 4. Cálculo de Combinaciones

La diferencia entre permutación y combinación, es que en la permutación el orden de los elementos importa, mientras que en la combinación el orden de éstos no importa.

### 10.9 Regla de la adición.

Si  $A$ ,  $B$  y  $C$  son eventos en un espacio muestral  $\Omega$

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

Ejemplo:

En una población de 2000 habitantes se tiene que 1000 tienen seguro médico, 500 seguro de vida y 250 seguro de auto, 270 tienen seguro médico y de vida, 150 seguro de vida y de auto, 100 seguro médico y de auto y 50 tienen los tres tipos de seguros, calcule la probabilidad de tener cualquiera de los tres seguros (médico, de vida o de auto)

Solución: En este tipo de ejemplo lo primero que hay que indicarles a los alumnos es que **cada una** de las características indicadas **representan un evento**, de forma que retomando la teoría de conjuntos:  $A, B, C$  son **eventos diferentes**, conectivo “o” representa **unión** de eventos ( $A \cup B$ ) y conectivo “y” representa **intersección** de eventos ( $A \cap B$ ).

Desglose de operaciones:

$$A = \text{tener seguro médico}, P(A) = \frac{1000}{2000},$$

$$B = \text{tener seguro de vida}, P(B) = \frac{500}{2000}$$

$$C = \text{tener seguro de auto}, P(C) = \frac{250}{2000},$$

$$A \cap B = \text{tener seguro médico y de vida}, P(A \cap B) = \frac{270}{2000}$$

$$B \cap C = \text{tener seguro de vida y de auto}, P(B \cap C) = \frac{150}{2000}$$

$$A \cap C = \text{tener seguro médico y de auto, } P(A \cap C) = \frac{100}{2000}$$

$$A \cap B \cap C = \text{tener los tres tipos de seguro, } P(A \cap B \cap C) = \frac{50}{2000}$$

Sustituyendo en la regla de la adición para tres eventos

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C) \\ &= \frac{1000 + 500 + 250 - 270 - 150 - 100 + 50}{2000} = \frac{1280}{2000} = \mathbf{0.64} \end{aligned}$$

$\Rightarrow$  **que el 64% de la población cuenta con algún seguro**

Mediante un diagrama de Venn también se puede visualizar cada una de las intersecciones y los valores adquiridos.

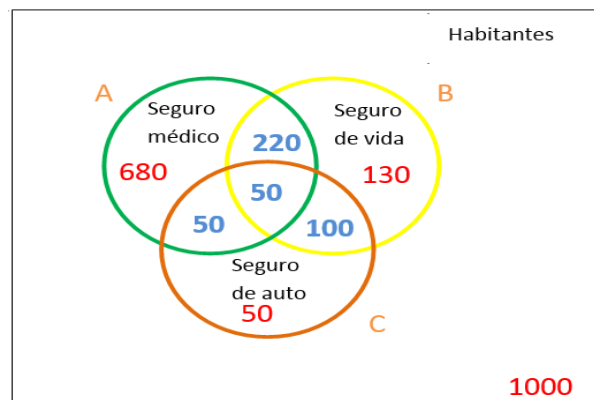


Figura 5. Diagrama de Venn

### 10.10 Conclusiones.

Al impartir cualquier curso de Probabilidad y Estadística en distintos programas educativos, ha sido de gran ayuda para los alumnos el plantearles en palabras coloquiales cómo distinguir las

distintas estrategias del conteo, facilitando el pasar del lenguaje común a conjuntos para poder contar los casos posibles y los casos favorables, enseguida introducirse al cálculo de la probabilidad para su adecuada solución; agregando el cálculo de combinaciones y permutaciones en WolframAlpha, Excel y GeoGebra, favoreciendo el entendimiento de los conceptos probabilísticos. Del mismo modo han aumentado los porcentajes de aprobación de exámenes departamentales en los grupos impartidos en los últimos años, este tipo de examen es un punto importante para las acreditaciones, como las que aplica el CACEI (Consejo de Acreditación de la Enseñanza de la Ingeniería, A.C.), fomentando mejor y mayor preparación a los egresados, con bases fuertes en Probabilidad.

### **Referencias.**

Gámiz, B. E., & Gámiz Casarrubias, O. T. (2012). *Probabilidad y Estadística, con prácticas en Excel* (Tercera Ed). México: JIT PRESS.

Hines, W. W., & Montgomery, D. (2005). *Probabilidad y Estadística para Ingeniería* (Tercera Ed). México: CECSA.

Jay, L. D. (2008). *Probabilidad y Estadística para Ingeniería y Ciencias* (Séptima Ed). Cengage Learning.

Seymour, L. (1998). *PROBABILIDAD*. México: Mc. Graw-Hill.



# 11

---

## **Análisis Estadístico de Conceptos Probabilísticos**

---

Fernando Velasco Luna, Francisco Solano Tajonar Sanabria,

Hugo Adán Cruz Suárez, José Dionicio Zacarías Flores

fvelasco@cfm.buap.mx, ftajonar@cfm.buap.mx, hcs@cfm.buap.mx,

jzacarias@cfm.buap.mx



### **Resumen.**

La Estadística es de gran utilidad y una herramienta fundamental para la realización de procesos de investigación en ciencias fácticas que utilizan la investigación cuantitativa, Las bases teóricas de la Estadística recaen en la teoría de la Probabilidad, por lo que la formación probabilística es importante para la formación de estudiantes universitarios, más aún en aquellos que cursan alguna licenciatura del área de matemáticas. A pesar de su importancia, no todos los estudiantes asimilan los conceptos que se abordan en los cursos universitarios, como sería lo ideal. En este trabajo se lleva a cabo un análisis estadístico en base a graficas de sectores y tablas de contingencia con el



objetivo de conocer que tanto los alumnos comprenden algunos conceptos básicos de probabilidad, tanto de manera individual así como conjunta, tales como espacio muestral, sigma álgebra, variable aleatoria, entre otros. Algunos de los resultados obtenidos muestran que respecto al concepto de sigma Álgebra un alto porcentaje de los alumnos respondió que lo entendieron solo un 50%, respecto al concepto de variable aleatoria solo un 15% menciona haberlo entendido un 100%. Respecto a comprender en forma conjunta el concepto de sigma álgebra y variable aleatoria solo aproximadamente un 7% contestó que los comprende al 100%. Los conceptos de función de densidad y sigma álgebra en forma conjunta ningún alumno los comprendió al 100% ambos, lo mismo ocurrió con los conceptos de sigma álgebra y Función de Distribución Acumulativa.

**Palabras Clave:** Aprendizaje, Probabilidad, Variable Aleatoria.

### **Abstract**

Statistics is very useful and a fundamental tool for the realization of research processes in factual sciences that use quantitative research. The theoretical bases of statistics fall on the theory of probability, so that probabilistic training is important for the training of university students. , even more so in those who have a degree in the area of mathematics. Despite its importance, not all students assimilate the concepts taught in university courses, as would be ideal. In this paper a statistical analysis is carried out based on graphs of sectors and contingency tables with the aim of knowing that both students understand some basic concepts of probability, both individually and jointly, such as sample space, sigma algebra, random variable, among others. Some of the results obtained show that with respect to the concept of sigma Algebra a high percentage of the students answered that they understood it only 50%, regarding the concept of random variable only 15% mentioned having understood it 100%. Regarding the joint understanding of the concept of sigma algebra and random variable, only approximately 7% answered that it comprises 100%. The concepts of density function and sigma algebra jointly no student understood 100% both, the same happened with the concepts of sigma algebra and Cumulative Distribution Function.

## 11.1 Introducción

La Estadística es de gran utilidad en todas las áreas de la investigación fáctica. Además la formación tanto probabilística como estadística, es importante para la formación de estudiantes universitarios para poder dar una interpretación adecuada, eficiente y suficiente de datos. Se tiene que recordar que la estadística es una herramienta fundamental para la realización de procesos de investigación en ciencias fácticas que utilizan la investigación cuantitativa, ya que la competencia estadística proporciona recursos para analizar datos críticamente. Actualmente se considera esencial la competencia estadística para ciudadanos capaces de orientarse en un entorno de fuertes interdependencias sociales, políticas y económicas, donde con frecuencia las decisiones con impacto social se toman sobre la base de estudios estadísticos. Las bases teóricas de la Estadística recaen en la teoría de la Probabilidad, es la razón por la cual se le ha dado una creciente importancia a la enseñanza de la Probabilidad en la educación universitaria. Pero la investigación didáctica está mostrando las dificultades que tienen los estudiantes para aprender de forma significativa los conceptos relacionados con el azar y la probabilidad (Borovcnik et al., 1991). Como mencionan Batanero y Díaz, (2007), un conocimiento genuino de probabilidad solo se alcanza con el estudio de alguna probabilidad formal, aunque este estudio debe ser gradual y apoyado en la experiencia estocástica de los estudiantes.

A pesar de la importancia de la probabilidad en la formación estadística, no todos los estudiantes asimilan o aplican los conceptos de manera adecuada, por esto se han estudiado los obstáculos que se presentan en su aprendizaje, entre los cuales podemos mencionar: la dificultad para el manejo de conceptos básicos por bases insuficientes en la formación matemática (Cochran, 2005). Se ha venido reportando como parte de la problemática actual de la educación, el hecho de que los esfuerzos que se realizan están principalmente enfocados a las metodologías para enseñar más que para entender y actuar en el cuánto y de qué forma se aprende. Se dice que el problema está en que el enfoque se orienta más a la enseñanza que al aprendizaje (Behar, 2001). Respecto a la educación en probabilidad y estadística, la enseñanza de la Probabilidad en muchas universidades generalmente se limita a seguir los temas que se encuentran en varios textos de Estadística, los

cuales son impuestos por el profesor desde el inicio del semestre. Sin embargo, esta estrategia tradicional para enseñar Probabilidad, a veces no es la mejor y los estudiantes se pueden ver frustrados en su aprendizaje a la hora de entender la aplicabilidad de la Probabilidad en diferentes situaciones reales (Cochran, 2005). Generalmente, el proceso de la enseñanza de la Probabilidad comienza con la presentación de conceptos. (Ver: Osorio, 2013). Investigaciones señalan que los estudiantes tienen dificultades para lograr un aprendizaje con comprensión de los conceptos y procedimientos formales relacionados con el azar (Batanero *et al.*, 1997; Borovcnik *et al.*, 1991).

Barragués y Guisasola (2009), describen el diseño, la implementación y la evaluación de una secuencia de enseñanza destinada a introducir los conceptos y procedimientos probabilísticos elementales en la enseñanza técnica universitaria. La propuesta se basa en los resultados de las investigaciones sobre las dificultades de enseñanza-aprendizaje, la perspectiva social constructivista del aprendizaje de las matemáticas y el concepto de indicadores de aprendizaje. Proporcionamos pruebas de que esta secuencia de enseñanza, junto con su metodología de aplicación en el aula, puede lograr que los estudiantes adquieran una mayor capacidad de razonamiento probabilístico.

Barragués y Guisasola (2012), llevan a cabo un estudio en condiciones reales de docencia y la evaluación, de una secuencia de enseñanza y su correspondiente modelo de aplicación destinados a introducir las conceptualizaciones subjetiva y axiomática de la probabilidad en la enseñanza técnica universitaria. La propuesta se basa en los resultados de las investigaciones sobre las dificultades de aprendizaje de los conceptos relacionados con la probabilidad y en la perspectiva social constructivista del aprendizaje de las matemáticas. Los resultados obtenidos muestran una mejora significativa en diversos aspectos del razonamiento probabilístico, respecto a los estudiantes que reciben una enseñanza convencional. Además, hemos observado en los estudiantes hábitos metodológicos avanzados para la resolución de problemas y cierta perspectiva de las problemáticas que encierran cada una de las concepciones de la probabilidad.

Caracterizar qué es aprender la teoría de la probabilidad consiste en definir las capacidades (competencias) que los estudiantes deben adquirir para su aplicación en cierto contexto. Barragués y Guisasaola (2009) definen una colección de 8 indicadores de aprendizaje para valorar si se ha logrado un aprendizaje con comprensión de los conceptos probabilísticos.

1. Saber por qué son necesarios los modelos probabilísticos. Es decir, comprender las carencias de los modelos deterministas a la hora de estudiar fenómenos reales. En concreto, debe comprenderse que el modo en que se desarrolla un fenómeno real depende de una cantidad inconcreta de factores y contingencias, cuya interacción es demasiado compleja para que pueda ser determinada por unas pocas variables y por relaciones de causa/efecto.
2. Entender la estrategia general que utiliza la teoría probabilística para la resolución de problemas. Es decir, si bien el enfoque probabilístico formal no niega necesariamente la existencia de mecanismos causales subyacentes, se adopta una posición según la cual se ignoran tales posibles mecanismos, centrándose en las regularidades que ocurren en una serie de ensayos realizados en\* las mismas condiciones.
3. Conocer las estrategias particulares de resolución de problema y el alcance de las soluciones probabilísticas. Es decir:
  - 3.1. Entender las condiciones en las que existe una regularidad en la frecuencia relativa de un suceso.
  - 3.2. Realizar predicciones probabilísticas acerca de una población.
  - 3.3. Reconocer los conceptos probabilísticos elementales y proponer soluciones probabilísticas. En particular, deben reconocerse estos conceptos en:
    - 3.3.1. Situaciones alejadas del problema tipo.
    - 3.3.2. Problemas mal o insuficientemente definidos.

3.3.3. Situaciones en las que la intuición o el sentido común parezcan ser suficientes para proporcionar una solución.

4. Saber utilizar los aspectos procedimentales fundamentales relacionados con el cálculo de la probabilidad:
5. Emplear la simulación de modelos probabilísticos mediante computadora como medio de investigación acerca del significado de los conceptos y sus propiedades, así como de búsqueda de soluciones a problemas.
6. Comprender el alcance y las limitaciones del marco teórico probabilístico para la resolución de problemas en el ámbito del azar.
7. Ser capaz de reflexionar acerca del proceso de construcción de un marco teórico científico, captar una visión de la matemática que, al igual que la mayoría de los asuntos humanos, arranca de situaciones problemáticas y se desarrolla impulsada por la controversia y el debate.
8. Valorar positivamente la teoría de la probabilidad como marco útil para resolver problemas.

Osorio *et. al.* 2013, mencionan que diversas herramientas y estrategias que se han propuesto para apoyar los procesos de enseñanza y aprendizaje de la Probabilidad, así como su aplicación. Tauber *et.al.* (2005) diseñan una secuencia de aprendizaje basada en el uso de ordenadores.

En este trabajo se lleva a cabo un análisis descriptivo en base a graficas de sectores con el objetivo de conocer que tanto los alumnos comprenden algunos conceptos básicos de probabilidad tales como espacio muestral, variable aleatoria, valor esperado, varianza, entre otros.

## 11.2 Metodología

**Aspectos Generales.** Los datos analizados se obtuvieron a partir de un cuestionario que se aplicó a alumnos que cursaron la materia de Probabilidad I en la FCFM, durante el periodo primavera 2018. Son preguntas relacionadas al porcentaje de comprensión de conceptos de Probabilidad tales como Sigma Álgebra, Espacio Muestral, Variable Aleatoria, entre otras.

*Análisis Estadístico.* Se obtuvieron gráficas de sectores con el objetivo de conocer el comportamiento de las variables bajo estudio. Además se obtuvieron tablas de contingencia para ver el comportamiento conjunto de las variables bajo estudio.

### 11.3 Resultados

En primer lugar se presentan graficas de sectores para conocer la distribución de las variables bajo estudio.

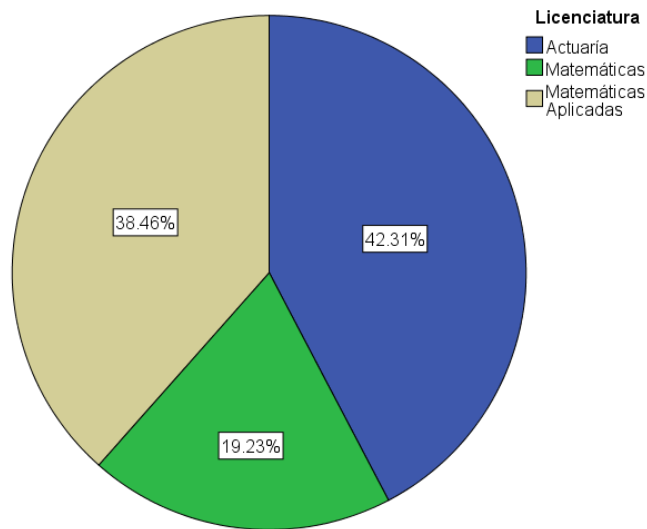


Figura 1. Distribución de alumnos por licenciatura.

De la Figura 1 se tiene que de los alumnos encuestados aproximadamente un 42% cursan la Licenciatura en Actuaría, mientras aproximadamente un 38% la Licenciatura en Matemáticas Aplicadas.

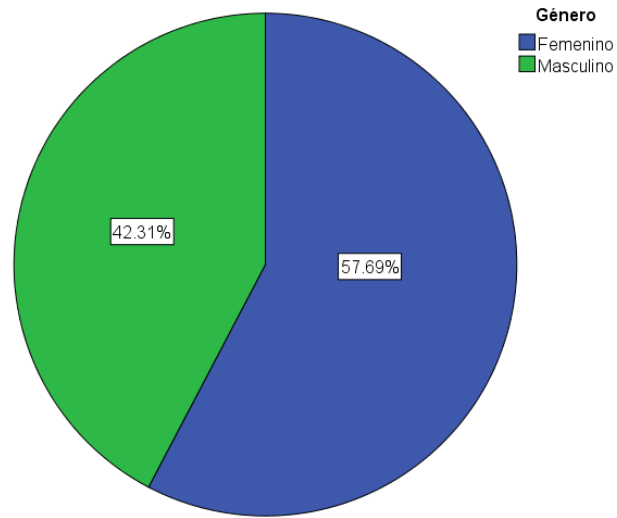


Figura 2. Distribución de alumnos por género.

De la Figura 2 se tiene que de los alumnos encuestados aproximadamente un 58% son del género Femenino.

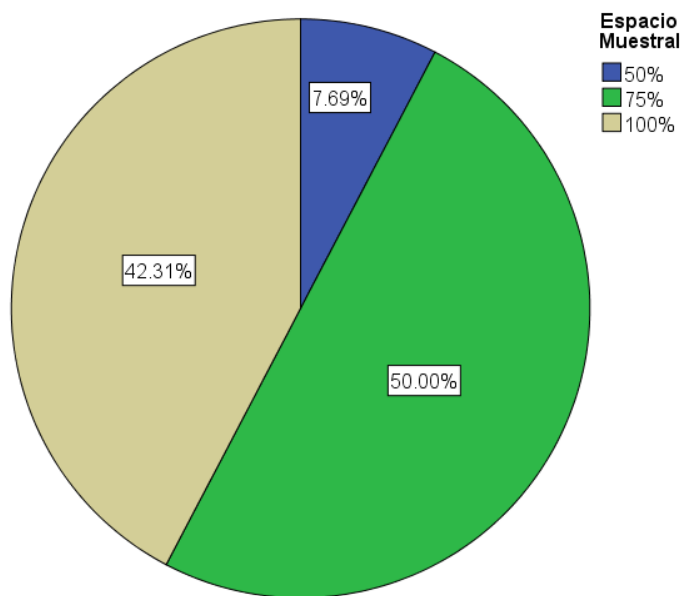


Figura 3. Espacio Muestral.

De la Figura 3 se tiene que de los encuestados el 50% respondieron que comprenden un 75% el concepto de espacio muestral, mientras que alrededor del 43% lo entienden al 100%.

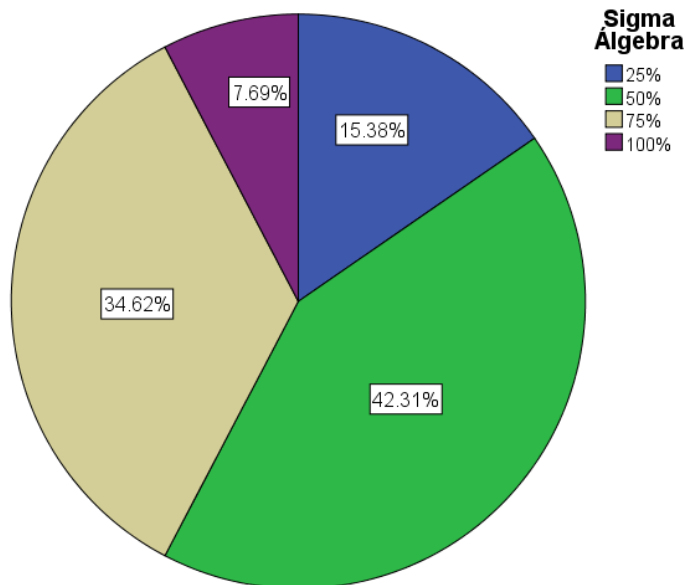


Figura 4. Sigma Álgebra.

Respecto al concepto de Sigma Álgebra en la Figura 4 se observa que alrededor de un 43% de los alumnos encuestados mencionaron comprender el concepto un 50%, mientras que sólo alrededor del 8% lo comprenden un 100%. Alrededor de un 35% lo comprende un 75%.

En la Figura 5 se observa que alrededor del 65% de los alumnos encuestados mencionaron comprender el concepto de variable aleatoria 75%, mientras aproximadamente un 15% respondió comprenderlo un 100%.



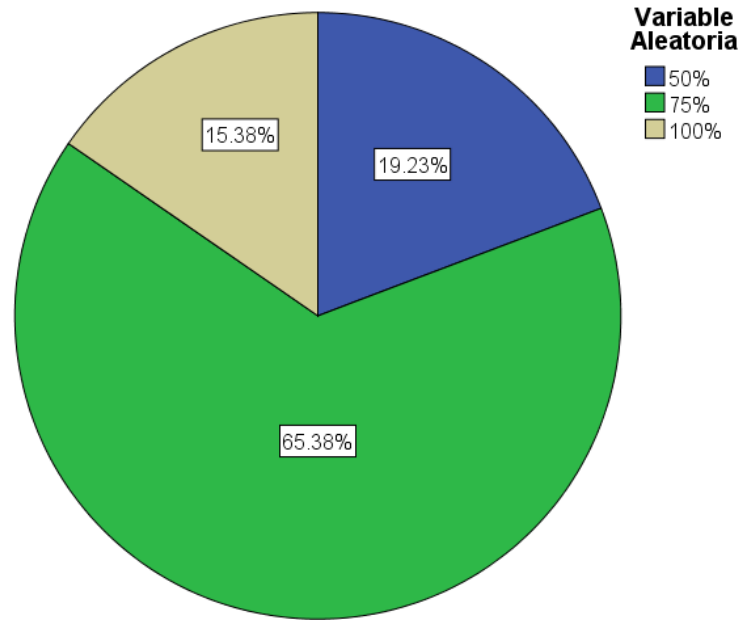


Figura 5. Variable Aleatoria.

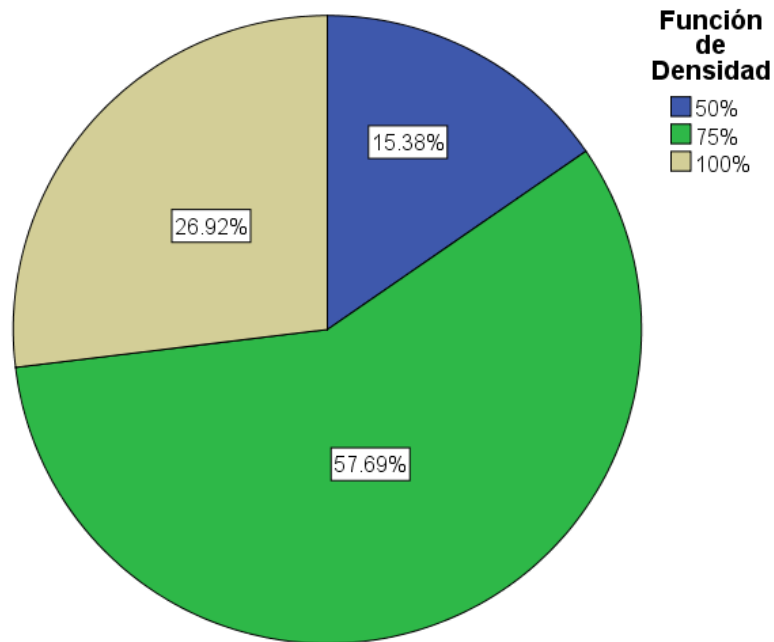


Figura 6. Función de Densidad.

En la Figura 6 se observa que alrededor del 58% de los alumnos encuestados mencionaron comprender el concepto de función de densidad un 75%, mientras aproximadamente un 27% respondió comprenderlo un 100%.

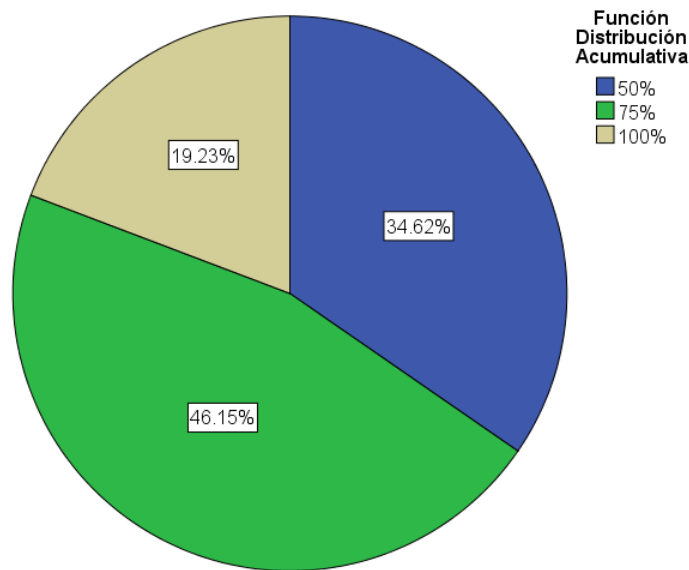


Figura 7. Función de Distribución Acumulativa.

Respecto al concepto de FDA, en la Figura 7 se observa que alrededor de 47% de los alumnos encuestados mencionaron comprender el concepto un 75%, mientras aproximadamente un 35% respondió comprenderlo un 50%.

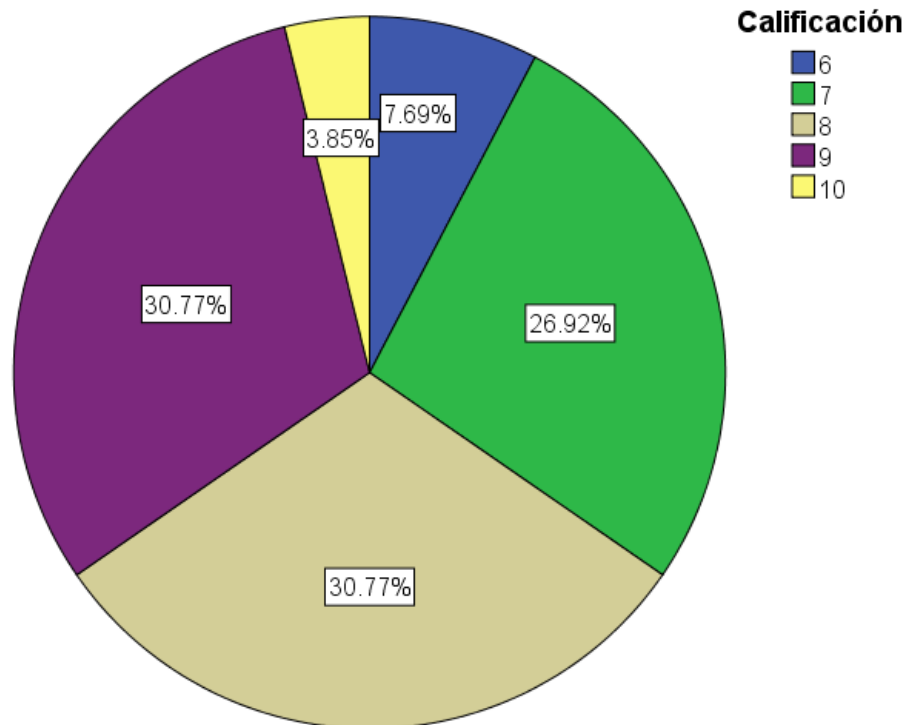


Figura 8. Calificación del curso.

Respecto a la calificación obtenida en el curso, en la Figura 8 se observa que alrededor de 31% de los alumnos encuestados obtuvieron una calificación de 9, igual porcentaje una calificación de 8, mientras aproximadamente solo un 4% una calificación de 10.

A continuación se presentan los resultados obtenidos de las tablas de contingencia para ver el comportamiento conjunto de las variables bajo estudio.

De los alumnos encuestados aproximadamente un 31% contestó que comprendieron tanto un 75% el concepto de espacio muestral y un 75% el concepto de variable aleatoria. Destaca el hecho de que solo el 11% de los encuestados comprendieron ambos conceptos al 100%. De los alumnos encuestados aproximadamente un 39% contestó que comprendieron tanto un 75% el concepto de espacio muestral y un 75% el concepto de función de densidad, mientras que solo un 30% contestó que comprenden a la vez al 100% los conceptos espacio muestral y FDA.

Respecto a comprender en forma conjunta el concepto de sigma álgebra y variable aleatoria solo aproximadamente un 7% contestó que los comprende al 100%. Los conceptos de función de densidad y sigma álgebra en forma conjunta ningún alumno los comprendió al 100% ambos, lo mismo ocurrió con los conceptos de sigma álgebra y FDA.

Respecto a comprender en forma conjunta el concepto de variable aleatoria y función de densidad solo aproximadamente un 8% contestó que los comprende al 100%, mientras que aproximadamente un 35% contestó que los comprenden ambos al 75%. Respecto a los conceptos de variable aleatoria y FDA comprendidos en forma conjunta solo aproximadamente un 4% contestó comprenderlos al 100%.

#### **11.4 Conclusiones**

Cuando el alumno se enfrenta por primera vez a conceptos del área de probabilidad se presenta un grado de dificultad para comprender tales conceptos. En algunos cursos universitarios se presenta la probabilidad de maneja muy sencilla ya que no es tan importante para los fines del curso que el alumno comprenda, como sería lo ideal, tales conceptos, ya que los fines del curso son de aplicación de la estadística. No es el caso en las escuelas de matemáticas a nivel universitario. De los resultados obtenidos de este trabajo se tiene que es alto el porcentaje de estudiantes que no comprendieron ciertos conceptos, por ejemplo, un bajo porcentaje de alumnos comprendió el concepto de sigma álgebra. En la literatura, estos conceptos no se encuentran de manera común y además no se profundiza lo suficiente en ellos, por esta razón, los profesores y facilitadores deben tomar ese trabajo en sus manos, tratando de explicar de forma más detallada estos contenidos que resultan importantes para la construcción del conocimiento posterior del estudiante.

## Referencias

- Behar, R. (2001). *Aportaciones para la Mejora del Proceso de Enseñanza-Aprendizaje de la Estadística*. Tesis Doctoral. Universidad Politécnica de Cataluña, Barcelona, España.
- Barragués F.J.I. y Guisasola A.J. (2009). Una propuesta para la enseñanza de la probabilidad en la universidad basada en la investigación didáctica. *Educación Matemática*, 21(3), 127-162.
- Barragués F.J.I. y Guisasola A.J. (2012). Una propuesta para la introducción de las concepciones subjetiva y axiomática de la probabilidad en la universidad. *International Journal for Studies in Mathematics Education*, 5(1), 131-157.
- Batanero, C y Díaz C. (2007) Probabilidad, grado de creencia y proceso de aprendizaje. Ponencia Invitada en las *XIII Jornadas Nacionales de Enseñanza y Aprendizaje de las Matemáticas*. Granada, Julio, 2007. Federación Española de Profesores de Enseñanza de las Matemáticas.
- Borovcnik, M., Bentz, H. J. & Kapadia, R. (1991). A Probabilistic Perspective. *En R. Kapadia & M. Borovcnik (Comps.), Chance encounters: probability in education (27-70)*. Dordrecht (The Netherlands): Kluwer Academic Publishers.
- Cochran, J. (2005). Can you really learn basic probability by playing a sports board game? *The American Statistician*, 59(3), 266–272.
- Osorio A.M.A.; Suárez P.A.; Uribe S.C.C. (2013) Revisión de alternativas propuestas para mejorar el aprendizaje de la Probabilidad *Revista Virtual Universidad Católica del Norte*, 38, 127-142.
- Tauber, L., Batanero, C., y Sánchez, V. (2005). Diseño, implementación y análisis de enseñanza de la distribución normal en un curso universitario, *EMA*, 9(2), 82-105.



# 12

---

## Modelos Paramétricos en el Análisis de Supervivencia

---

Francisco Solano Tajonar Sanabria, Guadalupe Yoanna Arenas Martínez, Fernando Velasco Luna

Hugo Adán Cruz Suárez, José Dionisio Zacarías Flores,

ftajonar@fcfm.buap.mx, guadalupe\_yam@hotmail.com, fvelasco@fcfm.buap.mx, hcs@fcfm.buap.mx,

jzacarias@fcfm.buap.mx



### Resumen.

El objetivo principal del presente trabajo es destacar la importancia de la Estadística en los distintos campos de la investigación y de sus aplicaciones, dígase biología, física, química, ingenierías, matemáticas, en las ciencias sociales, medicas, etc. El trabajo hace notar que la estadística inferencial se puede aplicar en la línea de investigación del análisis de supervivencia presentando algunos modelos de distribución que se estudian en los cursos de probabilidad.

El Análisis de Supervivencia es una línea de investigación de la Estadística Inferencial que surge de una teoría paralela llamada Teoría de Confiabilidad la cual tiene como propósito el estudio del tiempo de vida o el tiempo de falla de un objeto o un conjunto de objetos sometidos a un cierto tipo de tensión. De esta forma se puede decir que el análisis de supervivencia se encarga de estudiar el tiempo de vida o de falla de una persona (objeto) o un grupo de personas.

La función de riesgo es un componente importante en el análisis de los eventos-tiempo ya que describe el riesgo inminente de falla en un momento dado.

**Palabras Claves:** Función de supervivencia, Función de distribución, Función de riesgo.

### **Abstract**

The main objective of this work is to highlight the importance of statistics in the different fields of research and its applications, say biology, physics, chemistry, engineering, mathematics, social sciences, medicine, etc. The paper notes that inferential statistics can be applied in the survival analysis research line by presenting some distribution models that are studied in probability courses.

Survival analysis is a line of investigation of inferential statistics that arises from a parallel theory called Theory of Reliability which has the purpose of studying the time of life or the time of failure of an object or a set of objects subjected to a certain type of tension. In this way it can be said that the survival analysis is responsible for studying the time of life or failure of a person (object) or a group of people.

The risk function is an important component in the analysis of time-events since it describes the imminent risk of failure at a given moment.

**Keywords:** Survival function, Distribution function, Risk function.

## 12.1 Introducción

Generalmente, al término estadística se le relaciona con el cálculo de promedios, porcentajes, etc., y con la presentación de los datos en forma tabular y en forma gráfica. Aunque las técnicas para resumir y presentar los datos son importantes, comprenden solo una parte de la estadística, llamada Estadística Descriptiva; la otra parte es la Inferencia Estadística, la cual se encarga de analizar datos recolectados, y posteriormente inducir algún conocimiento (característica) sobre una población, donde los datos analizados son solo un subconjunto o muestra; esto es, el proceso va de lo particular a lo general.

La estadística es una herramienta ampliamente utilizada en la investigación científica. También se emplea en instituciones gubernamentales y educativas, en los negocios y en la industria. El empleo juicioso de las técnicas estadísticas permite obtener conclusiones útiles a partir de datos numéricos.

En los últimos años el análisis de supervivencia ha desarrollado tópicos importantes en muchas áreas, especialmente en biología, medicina, ingeniería, teoría de confiabilidad, epidemiología, etc., con aplicaciones que van desde estudiar el desarrollo de alguna enfermedad, tiempos de vida en individuos con enfermedades terminales, falla en maquinaria, la durabilidad de electrodomésticos, hasta estudiar la resistencia de materiales al calor, etc., son solo algunas de las aplicaciones en donde el papel del análisis de supervivencia es importante.

Los fenómenos mencionados anteriormente no pueden predecirse o modelarse con algún modelo determinista, para más detalles ver (Rohatgi, and Ehsanes, 2001; Kalbfleisch, and Prentice, 1980). Por ejemplo, supóngase que estamos interesados en determinar el tiempo de vida útil de un fusible puesto a funcionar en algún circuito, es claro que, este fallará de forma imprevista, en el sentido de que en un momento está funcionando bien y en el momento siguiente falla. Por el contrario, una barra de acero bajo una carga pesada se irá debilitando gradualmente en el transcurso de un largo periodo de tiempo. Se observa entonces que, la falla dependerá del fenómeno en estudio.



De forma que, según los comentarios anteriores, el uso de un modelo probabilístico parece ser un planteamiento realista para modelar cada uno de los fenómenos anteriores, ( Garthwahwaite, Jolliffe and Jones, 2002; Rohatgi, and Ehsanes, 2001;Hogg and Craig, 1995).

Se debe señalar que existen varios modelos de probabilidad que se han propuesto para describir datos relacionados con tiempos de vida. Sin embargo, existen situaciones en donde los datos de tiempos de vida, no pueden ser ajustados por los modelos paramétricos clásicos. Entonces el problema básico dentro del análisis de supervivencia es dar y representar las distribuciones de tiempos de vida mediante modelos estadísticos y hacer inferencia sobre los parámetros de estos modelos.

## **12.2 ¿Qué es el análisis de supervivencia?**

El Análisis de Supervivencia es una línea de la estadística inferencial que estudia y modela tiempos de falla o de vida de un individuo o un grupo de individuos sometido(s) a un tipo de “tensión” que puede ser un tratamiento, cirugía, etc., de esta forma se puede decir que el análisis de supervivencia se encarga de estudiar eventos-tiempo.

En la actualidad el uso de modelos paramétricos para representar fenómenos relacionados con tiempos de vida, tiempos de supervivencia o tiempos de falla, se han desarrollado ampliamente con aplicaciones que van desde estudiar tiempos de vida en individuos con enfermedades terminales, la comparación de la efectividad de  $k$  tratamientos buscando prolongar el tiempo restante de vida de estos individuos, así como en artículos de uso cotidiano bajo condiciones normales o de vida acelerada, garantía del funcionamiento de electrodomésticos durante  $k$  unidades de tiempo, entrega de suplementos electrónicos confiables, dar confiabilidad o determinarla en el transporte público, describir el tiempo de vida de una amplia clase de productos como función de la temperatura, tales como aislantes eléctricos y dieléctricos, acumuladores, lubricantes y grasas, plásticos y filamentos de lámparas incandescentes, son solo algunos de los casos en donde el papel del Análisis de supervivencia o Teoría de confiabilidad juega un papel importante.

Recientemente el análisis de supervivencia se ha utilizado con buenos resultados en situaciones de enfermedades crónico – degenerativas tales como cáncer de prostate, diabetes, trasplante de organos, estudiando principalmente los cambios que puedan tener los pacientes, al llevar a cabo un tratamiento o una cirugía.

Los modelos que se han propuesto en la literatura para describir datos de tiempos de vida, entre ellos los más usados se encuentran los modelos Exponencial, Gama, Weibull y Lognormal, sin embargo, no todos los conjuntos de datos se ajustan a dichos modelos, por lo que es importante proponer o estudiar modelos más generales para explicar el comportamiento de este tipo de datos.

Un modelo que han seguido varios autores, para datos de tiempos de vida, es el que esta dado por la distribución Gama generalizada, ver (Cox and Oakes, 1984; Hogg and Craig, 1995; Klein and Moeschberger, 2005; Mood, Graybill and Boes, 1974), el cual es un modelo paramétrico general que incluye entre otros modelos a los ya mencionados anteriormente, desafortunadamente dicha distribución tiene limitantes significativas como lo es la estimación de sus parámetros por máxima verosimilitud (Cox and Oakes, 1984; Lawless, 2002).

En el análisis estadístico, muchas veces los datos se refieren a observaciones de tiempos de vida o tiempos de falla, sin embargo, desde el enfoque del análisis de supervivencia, el análisis de estos datos es diferente a la de los datos comunes, ya que para llevar a cabo el estudio de tiempos de vida o de falla a través de un análisis estadístico dentro de este se deben considerar conceptos como función de supervivencia, función de riesgo y otros conceptos básicos los cuales se mencionan a continuación.

Antes de presentar los conceptos propios del análisis de supervivencia, es bueno considerar que para realizar un buen análisis de tiempos de vida se necesita lo siguiente:

1. Tener bien definido nuestro evento de interés.
2. Definir de forma apropiada el origen o inicio del estudio, es decir, indicar el comienzo de nuestro análisis.

3. Definir la escala de tiempo.

De lo anterior tenemos que el análisis de supervivencia es el conjunto de técnicas que nos permiten analizar, estudiar y modelar la variable (en este caso representa a un evento-tiempo) de interés. Esta variable, en general depende de otras variables que son explicables y que se denominan covariables del fenómeno, ver (Harter, 1967; Lawless, 2002).

Sea  $T$  una variable aleatoria que denota el tiempo hasta que ocurre algún evento específico. Este evento puede ser la muerte, la aparición de un tumor, el desarrollo de alguna enfermedad, la recurrencia de alguna enfermedad, fallo de equipo eléctrico, dejar de fumar, etc.

De forma más específica, en este trabajo,  $T$  será una variable aleatoria continua, no negativa que representará a una población homogénea. Debe señalarse que  $T$  al ser una variable aleatoria continua debe tener una función de densidad de probabilidad, digamos  $f(t)$ .

En seguida se presentan tres funciones que en el análisis de supervivencia generalmente caracterizan a  $T$  :

**Función de Supervivencia:** es la probabilidad de que un individuo sobreviva más del tiempo  $t$

**Función de Riesgo:** es la probabilidad de que un individuo de  $t$  años experimente el evento en el siguiente instante de tiempo.

**Función de densidad de probabilidad:** es el modelo probabilístico que indica el comportamiento límite de la probabilidad en cada evento  $\{t \leq T < t + \Delta t\}$  cuando  $\Delta t \rightarrow 0$ .

Si se conoce alguna de estas funciones, se puede ver fácilmente que las otras dos pueden ser determinadas a partir de la función conocida. Estas funciones nos permitirán analizar diferentes aspectos de  $T$  , por ejemplo, que forma tiene su distribución de probabilidad, su función de riesgo, etc.

### 12.3 Conceptos y Resultados Básico de Supervivencia

Iniciamos la sección presentando el concepto de función de supervivencia el cual es fundamental para definir y estudiar fenómenos aleatorios de tipo evento-tiempo, (Cox and Oakes, 1984, Kalbfleisch, and Prentice, 1980).

**Definición 3.1.** La función de supervivencia de la variable aleatoria (v.a)  $T$  es la probabilidad de que un individuo viva al menos hasta tiempo  $t$ , esta función se denota por  $S(t)$  y está definida por:

$$S(T) = P(T \geq t) . \quad (1)$$

En el caso de estudiar los tiempos de falla en artículos o objetos, a  $S(t)$  se le llama función de confiabilidad.

**Teorema 3.1.** Sea  $T$  una variable aleatoria continua, con función de densidad  $f(t)$  entonces

$$f(t) = -\frac{dS(t)}{dt} \quad (2)$$

**Demostración:** Como  $T$  es una variable aleatoria continua, se tiene que:  $\forall t \in \mathbf{R}$  ,  $F(t-) = F(t)$ , de donde,

$$S(t) = 1 - F(t) \Rightarrow S'(t) = -F'(t).$$

Entonces

$$f(t) = -\frac{dS(t)}{dt} ,$$

que es lo que se afirmó en (2). ◻

También, dado que,

$$\int_t^{t+\Delta t} f(u)du \approx f(t)\Delta t ,$$

entonces,  $f(t)\Delta t$  puede entenderse como la probabilidad aproximada de que el evento ocurra en el intervalo de tiempo  $[t, t + \Delta t]$ .

Otra función fundamental en el análisis de supervivencia es la función de riesgo. Esta función también es conocida como tasa de mortalidad en demografía, la función de intensidad en procesos estocásticos, (Cox and Oakes, 1984; Klein and Moeschberger, 2005). La función de riesgo es predominante en el análisis de datos de tiempo de vida ya que generalmente nos puede ayudar a determinar que forma tiene la función de distribución de los datos.

**Definición 3.2.** La función de riesgo se define como sigue

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (3)$$

**Observación 3.1:**  $P[t \leq T < t + \Delta t | T \geq t]$  esto se puede interpretar como la probabilidad de que un individuo de  $t$  años experimente el evento en las próximas  $\Delta t$  unidades de tiempo.

Entonces la función de riesgo se puede caracterizar como la razón de cambio o tasa de falla instantánea de un individuo al tiempo  $t$ , dado que el individuo ha sobrevivido hasta el momento  $t$ . Esto es,  $h(t)\Delta t$  es aproximadamente igual a la probabilidad de que un individuo experimente el evento o la falla en el intervalo  $[t, t + \Delta t]$  dado que ha sobrevivido hasta el tiempo  $t$ .

**Corolario 3.1.**

$$h(t) = \frac{f(t)}{S(t)}. \quad (4)$$

**Demostración:** Su prueba se sigue directamente de (3) y del concepto de probabilidad condicional, ver Cox and Oakes, 1984.

Notese que:

$$S(t) = \exp\left(-\int_0^t h(u)du\right).$$

La función de riesgo es particularmente útil para determinar apropiadamente la distribución de falla. Utilizando información cuantitativa acerca del mecanismo de fracaso, para describir la forma en la cual la probabilidad de experimentar el evento cambia con el tiempo. Esta tiene varias formas generales para porcentajes de riesgo distintos. La única condición para la función de riesgo es la no negatividad, es decir,  $h(t) \geq 0$ .

**Teorema 3.2.** Sea  $T$  el tiempo para que ocurra una falla,  $T$  es una variable aleatoria continua con densidad  $f$  y f.d.a.  $F$  con  $F(0) = 0$ , entonces  $f$  puede expresarse en términos de la función de riesgo de la siguiente manera:

$$f(t) = h(t) \exp\left(-\int_0^t h(r)dr\right). \quad (6)$$

**Demostración:** Su demostración la puede consultar en (Cox and Oakes, 1984; Klein and Moeschberger, 2005; Lawless, 2002).

El resultado anterior nos dice que la función de riesgo proporciona mayor información acerca del mecanismo subyacente de falla que la función de supervivencia. Por esta razón, consideramos a la función de riesgo como el elemento dominante para el análisis de datos de supervivencia.

De forma más precisa, la función de riesgo determina de forma única a la función de densidad de probabilidad  $f(t)$ .

## 12.4 Modelos Paramétricos

En esta sección se presentan los modelos de distribución de probabilidad más utilizados en supervivencia para ajustar datos de tiempos de vida o de falla. En los cursos de probabilidad se estudian algunos modelos de distribución de probabilidad analizando algunas de sus características, tales como el valor esperado, sus momentos, su varianza y se señala que los parámetros que tiene cada modelo se consideran como constantes, sin embargo, en el análisis de supervivencia los parámetros que tiene cada modelo no se consideran constantes, por lo cual se requiere llevar a cabo la estimación de cada uno de ellos. Iniciamos con el modelo más simple y el cual se ha utilizado en el ajuste de una gran variedad de tiempos de vida.

### 1) MODELO EXPONENCIAL

La distribución exponencial desempeña un papel muy importante en la descripción de fenómenos aleatorios, especialmente en el análisis de supervivencia. Inicialmente se usó en trabajos referentes al análisis de confiabilidad (componentes electrónicas) y con algunas limitaciones se extendió a estudios médicos, (Cox and Oakes, 1984; Garthwaite, et. al., 2002).

Esta distribución tiene un solo parámetro ajustable (la media). Sin embargo, los métodos sobre los que se basa esta distribución son muy sensibles e incluso limitados.

**Definición 3.** Se dice que una variable aleatoria continua  $T$ , la cual toma valores positivos, tiene una distribución exponencial con parámetro  $\alpha$  si su función de densidad de probabilidad está dada por:

$$f(t) = \begin{cases} \frac{1}{\alpha} e^{-\left(\frac{t}{\alpha}\right)}, & \text{si } t > 0, \\ 0, & \text{cualquier otro caso.} \end{cases} \quad (7)$$

Lo anterior se denota por  $T \sim Exp(\alpha)$ .

La distribución exponencial se le suele caracterizar como el tiempo necesario de espera hasta que ocurre el primer éxito o fracaso.

En la literatura actual, algunas veces a la función de densidad de probabilidad (f.d.p) de  $T$ , también la definen como

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{si } t > 0, \\ 0, & \text{cualquier otro caso.} \end{cases}$$

Generalmente se elige en este estudio de tiempos de vida la parametrización dada por (7) debido a que, su valor esperado es  $E(T) = \alpha$ , ver (Cox and Oakes, 1984; Garthwahwaite, et. al., 2002; Lawless, 2002).

Enseguida se presentan algunas propiedades de esta distribución, las cuales las puede consultar en (Kalbfleisch, and Prentice, 1980; Klein and Moeschberger, 2005).

## PROPIEDADES

### 1. Función de Distribución Acumulativa

La f.d.a. de la variable exponencial está dada por:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)}. \quad (8)$$

### 2. Función de Supervivencia

$$S(t) = e^{-\left(\frac{t}{\alpha}\right)}. \quad (9)$$

### 3. Función de Riesgo

A continuación, veremos que esta distribución tiene la propiedad de tener tasa de riesgo

$$h(t) = \frac{1}{\alpha}. \quad (10)$$



El hecho de que el riesgo sea constante, nos indica que después de usar un artículo, la probabilidad de que falle es la misma que cuando era nuevo, en otras palabras, el objeto no envejece. El modelo de falla exponencial es utilizado en el estudio del tiempo de vida de un foco eléctrico, de una televisión, etc., esto dentro de la teoría de confiabilidad, mientras que en el análisis de supervivencia la distribución exponencial es adecuado para ajustar datos de tiempos de vida de grupos de personas homogénea, para tratamientos medicos, etc.

El siguiente resultado nos ayudará a comprender mejor esta situación.

**TEOREMA 4.1:** Si  $T$  es el tiempo transcurrido para que ocurra una falla y  $T$  es una variable aleatoria no negativa, entonces  $T$  tiene una distribución exponencial, si y sólo si, tiene una tasa de riesgo constante.

Para su demostración puede ver (Kalbfleisch, and Prentice, 1980; Klein and Moeschberger, 2005; Lawless, 2002).

## 2) MODELO WEIBULL

Este modelo, es muy usado en el análisis de supervivencia, sobre todo en la resistencia de material eléctrico, aplicaciones industriales y en biomedicina, (Harter, 1967; Rohatgi, and Ehsanes, 2001). Su función de riesgo no es constante, esto permite al modelo ajustar a diferentes conjuntos de tiempos de vida o de falla. La función de riesgo del modelo Weibull esta dada a través de una función polinomial que depende de su exponente, lo que permite que esta función de riesgo sea decreciente, constante o creciente.

El lector puede observar que el modelo exponencial es un caso especial del modelo Weibull.

**Definición 4.** Una variable aleatoria  $T$  tiene una distribución Weibull, si existen parámetros  $\lambda > 0$  y  $\beta > 0$ , tales que la variable  $Z = \left(\frac{T}{\lambda}\right)^\beta$  tiene una distribución exponencial con media 1.

La función de densidad de  $T$  esta dada por

$$f(t) = \begin{cases} \frac{\beta}{\lambda} \left(\frac{t}{\lambda}\right)^{\beta-1} \left(\exp\left(\frac{t}{\lambda}\right)^\beta\right), & \text{si } t > 0, \\ 0, & \text{cualquier otro caso.} \end{cases} \quad (11)$$

Las siguientes propiedades de la distribución Weibull se pueden verificar directamente de las definiciones correspondientes.

## PROPIEDADES

### 1. Función de Distribución Acumulativa

La f.d.a. de la variable Weibull está dada por:

$$F(t) = P(T \leq t) = 1 - \exp\left(-\left(\frac{t}{\lambda}\right)^\beta\right). \quad (12)$$

### 2. Función de Supervivencia

$$S(t) = \exp\left(-\left(\frac{t}{\lambda}\right)^\beta\right). \quad (13)$$

### 3. Función de Riesgo

$$h(t) = \frac{\beta}{\lambda} \left(\frac{t}{\lambda}\right)^{\beta-1}. \quad (14)$$

Además del modelo Exponencial y del modelo Weibull, otros modelos que son usualmente utilizados para modelar tiempos de vida o de falla se encuentra el modelo Gama y Gama generalizado, con la complicación de que este modelo posee tres parámetros y la estimación de sus parámetros mediante el método de máxima verosimilitud no resulta tan simple (Cox and Oakes, 1984; Garthwahwaite, et. al., 2002; Roussas, 1973).

## **12.5 Conclusiones**

En este trabajo se presentó el modelo Exponencial que nos permite modelar tiempos de vida o de falla que no tienen desgaste, también se analizó el modelo Weibull que resulta ser una buena distribución ya que se ajusta a una amplia clase de datos que están relacionados con el análisis de supervivencia.

Poe ejemplo, en el análisis de pacientes con mieloma múltiple, puede ser modelado de manera adecuada por la distribución Weibull, considerando datos sin censura.

El modelo Gama Generalizado, aunque ciertamente es una familia de distribuciones más amplia. No obstante, una limitante significativa se presenta en la estimación de sus parámetros, que no es tan sencilla.

## Referencias

Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*, Chapman and Hall.

Garthwaite, P.H., Jolliffe, I, and Jones B. (2002). *Statistical Inference*. Second Edition: Oxford Science Publications.

Harter, H. L. (1967). Maximum Likelihood Estimation of the Parameters of a Four Parameter Generalized Gamma Population for Complete and Censored Samples. *Technometrics*, 9, 159-165.

Hogg, R. V. and Craig, A. T. (1995). *Introduction to Mathematical Statistics*. Fifth Edition: Prentice Hall.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*: New York John Wiley.

Klein, J.P. and Moeschberger, M.L, (2005), *Survival Analysis: Techniques for Censored and Truncated Data*: New York Springer-Verlag.

Lawless J. F., (2002). *Statistical Models and Methods for Lifetime Data*. Second Edition: Wiley Series.

Mood, A.M., Graybill, F. A. and Boes D. C. (1974). *Introduction to the Theory of Statistics*. Third Edition: McGraw-Hill, Inc.

Rohatgi, K. V. and Ehsanes, A.K., (2001). *An Introduction to Probability and Statistics*. Lawless J. F. *Statistical Models and Methods for Lifetime Data*. Second Edition: Wiley Series.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*: Addison-Wesley Publishing Company.





# 13

---

## **Implementación del aprendizaje basado en proyectos en cursos de ingeniería: Comparación de la eficiencia entre la lombricomposta y la composta casera**

---

Guillermina Sánchez López, José Dionicio Zacarías Flores, Gladys Denisse Salgado Suárez, Hugo Adán Cruz Suárez, Víctor Hugo Vázquez Guevara.

guille.sanlopez@gmail.com, jzacarias@cfm.buap.mx, gladys008@hotmail.com, hcs@cfm.buap.mx, vvazquez@cfm.buap.mx



### **Resumen**

Este trabajo es el resultado de haber aplicado el aprendizaje basado en proyectos, en la asignatura “Diseño de proyectos de bioenergía”, teniendo como finalidad el lograr el desarrollo de la competencia que el programa de estudios marca.

Tras los resultados de la primera implementación, se observa la dificultad que al término del proyecto el estudiante tiene para concluir y presentar soluciones innovadoras como resultado de su

experimentación. Por lo que se propone como mejora para la siguiente aplicación que las variables que deben tomarse en cuenta son: la planeación de la asignatura, el seguimiento al proyecto, el aprendizaje alcanzado por los estudiantes, la retroalimentación oportuna y la evaluación.

**Palabras Clave:** ABP, Estadística, Ingeniería.

### **Abstract**

This work is the result of having applied project-based learning in the subject “Design of bioenergy projects”, with the purpose of achieving the development of the competence that the study program marks.

After the results of the first implementation, we observe the difficulty that at the end of the project the student has to conclude and present innovative solutions as a result of their experimentation. For what is proposed as an improvement for the following application that the variables that must be taken into account are: the planning of the subject, the monitoring of the project, the learning achieved by the students, the timely feedback and the evaluation.

**Keywords:** ABP, Statistics, Engineering.

## **13.1 Introducción**

El lograr que el aprendizaje sea significativo para el estudiante es de suma importancia en el desarrollo de todo curso, para esto, existen diferentes modelos educativos donde se establecen estrategias enseñanza-aprendizaje, para alcanzar este fin.

Si bien el aprendizaje basado en proyectos permite que el estudiante desarrolle sus habilidades destrezas y capacidades, aplicando los conocimientos apropiados; es necesario realizar una metodología correcta que le permita obtener observaciones precisas y sobre todo que pueda tomar decisiones correctas y precisas.

Estas decisiones pueden ayudar a que una empresa pueda salir adelante, o que la producción de algún bien o servicio sea redituable, pero para determinar lo anterior, en todo proceso es necesario analizar todas las variables que se encuentra presentes y que causan modificación en los resultados.

Todo lo anterior resulta ser importante en la formación del estudiante ya que se le está preparando para formar parte del terreno laboral o bien continuar con una preparación académica más especializada.

En este contexto, este documento da una propuesta de estrategia aprendizaje-enseñanza con la finalidad de mejorar la práctica educativa desde una perspectiva de emprendimiento que *promueva el aprendizaje basado en proyectos*

### **Hipótesis**

Las preguntas de investigación por tanto para esta propuesta son:

¿De qué manera las diferentes materias que forman la curricula escolar, pueden “aportar” a lograr este objetivo? ¿Cómo implementar desde las asignaturas la elaboración de proyectos? ¿Cómo darles seguimiento? ¿Cómo desarrollar el emprendimiento a lo largo de la carrera?

Para resolver las preguntas anteriores los autores del presente trabajo se plantean los siguientes objetivos.

### **Objetivo General**

Objetivo general: Implementar el Aprendizaje basado en Proyectos (ABP), como una estrategia de trabajo que desarrolle el emprendimiento en el estudiante de Ingeniería.

### **Objetivos específicos**

Objetivo específico: Emplear el diseño de experimentos (DOE), como una herramienta estadística que permita establecer el orden, la repetibilidad y replicabilidad en el desarrollo de un proyecto.



## 13.2 Metodología

El composteo es una actividad que ha ayudado al hombre en la agricultura desde el comienzo del sedentarismo. Hoy en día, con los conocimientos científicos que se tienen sobre la química se han creado fertilizantes artificiales que son de gran ayuda para el crecimiento de las plantas. Por desgracia, estos no son sustentables y causan muchos problemas, dentro de los cuales el más grave es la contaminación de los suelos alterando el ciclo natural del nitrógeno.

Una de las formas de lograr un cultivo sin afectar este ciclo es la Lombricomposta, que es un método cada vez más popular de composteo pasivo y se reconoce como el composteo del futuro además el uso de desechos orgánicos en las comunidades rurales son una práctica antigua buscando con ello mejorar el contenido de materia orgánica del suelo para mantener la fertilidad del mismo.

Entre las principales ventajas del uso de la lombricomposta está:

- Aumenta la retención de agua.
- Aporta nutrientes, y facilita su absorción por la planta.
- Enriquece el suelo con microorganismos benéficos.
- Aporta materia orgánica y su carga bacteriana induce la humificación de la materia orgánica presente en el suelo.
- Mejora la estructura del suelo y su aireación.
- Aumenta la resistencia de la planta al ataque de plagas y enfermedades.
- Actúa como regulador del pH del suelo.
- Su aporte de ácidos húmicos y fúlvicos, propicia la formación de quelatos con sus propios nutrientes.

- Su aporte en Capacidad de Intercambio Catiónico a la solución del suelo, retiene más nutrientes evitando sus pérdidas por lixiviación



Figura 1. Lombricomposta



Figura 2 Composta casera

Por otro lado la composta es el material orgánico que se obtiene como producto de la acción microbiana controlada sobre residuos orgánicos tales como hojas, rastrojos, zacates, cáscaras, basuras orgánicas caseras, subproductos maderables (aserrín y virutas), ramas, estiércoles, y residuos industriales de origen orgánico; con estos residuos, en forma separada o bien mezclados, se forman pilas o montones, que por acción de los microorganismos dan origen a un material (materia orgánica) de gran utilidad para los suelos agrícolas ya que mejora la estructura y la fertilidad de estos.

El compostaje era practicado en la Antigüedad. Desde hace miles de años, los chinos han recogido y comportado todas las materias de sus jardines de sus campos y de sus casas, incluyendo materias fecales. En el Oriente Próximo, en las puertas de Jerusalén había lugares dispuestos para recoger las basuras urbanas: unos residuos se quemaban y con los otros se hacía compost. El descubrimiento, después de la Primera Guerra Mundial, de los abonos de síntesis populariza su utilización en la agricultura. En los últimos años se ha puesto de manifiesto que tales abonos químicos empobrecen la tierra a medio plazo.

En la actualidad los cultivos requieren mayor intensidad, lo que conlleva a un aumento en las dosis de fertilizantes. Entre las ventajas de la composta casera están:

- Da cuerpo a los suelos arenosos y ligeros
- Mejora el drenaje en los suelos arcillosos.
- Hortalizas, que se abonan con composta producen mejores cosechas de una mejor calidad con una buena resistencia a las plagas.

Para el presente proyecto se tomará en cuenta el diseño de experimentos como base para el análisis de los resultados.

### **13.3 Objetivos de la elaboración de las compostas:**

- ❖ Establecer la eficiencia en la obtención, uso y aplicación de las diferentes compostas.
- ❖ Aumentar la capacidad de retención de contenido nutricional.
- ❖ Transformación de materiales orgánicos biodegradables.
- ❖ Introducir una alternativa viable a los problemas causados por el mal manejo de desechos sólidos orgánicos al campo mexicano.

Con la finalidad de poder realizar una comparación efectiva se establece *la replicación* del experimento, teniendo 4 *unidades experimentales o prototipos* 2 de lombricomposta y 2 de composta casera.

### **Determinación de las variables por cada unidad experimental:**

Para la lombricomposta se plantean las siguientes variables:

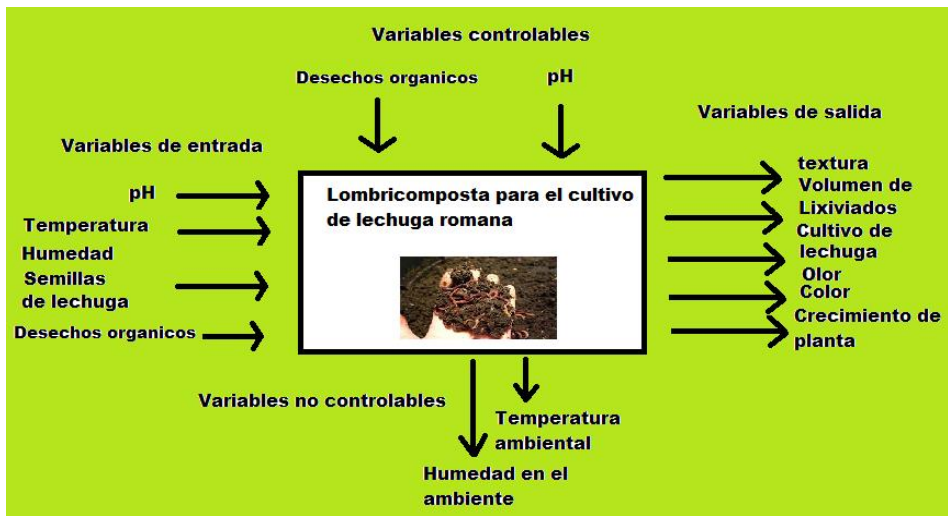


Figura 3. Diagrama de proceso

Para la composta casera se plantearon las siguientes variables:

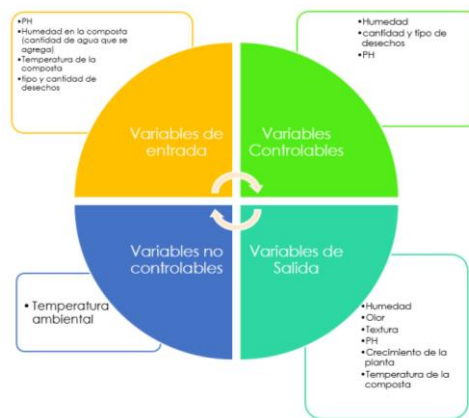


Figura 4. Diagrama de proceso: composta casera

Las variables en un proceso definen los productos que se obtendrán, muchas de estas se pueden manipular de manera que factores internos y externos al proceso no afecten su desarrollo. Respecto al proceso de lombricomposteo para cultivar lechuga romana y stevia, las variables se determinaron bajo las siguientes características.

Variables de entrada: Este tipo de variables son las que depende enteramente el proceso de lombricomposta. Cualquier falta de estas variables afectaría gravemente el producto final, llegando inclusive a no poder realizarse.

1. Desechos orgánicos: se refiere a la materia prima que se le agregaran a la composta para su descomposición.
2. Temperatura: La temperatura es un factor muy importante ya que si se alcanzan valores mayores a 65° centígrados muchos organismos micro bacterianos que ayudan a la descomposición morirían. Además de que las lombrices no soportan temperaturas mayores a 40°C. La temperatura es un factor que determina la velocidad de reproducción de las lombrices, así como la rapidez con que digieren el material orgánico.
3. Semillas: Dependiendo de la cantidad y calidad de las semillas, el cultivo germinara en mayor o menor cantidad, asumiendo que los demás factores sean los idóneos.
4. pH: se refiere al nivel de acidez que tiene los desechos orgánicos lo cual depende si provienen de fuentes cítricas o acidas, las cuales deben controlarse para tener la acidez óptima.
5. La humedad: Es importante un ambiente húmedo para prevenir la deshidratación de la lombriz y ayudarla a su desplazamiento en el sustrato. Por otro lado, la humedad excesiva puede producir condiciones anaeróbicas no aptas para las lombrices y la producción de lixiviados. El hecho de mezclar los materiales demasiado húmedos con otros más secos, es una práctica de control de humedad que mantienen al mínimo la producción de lixiviados.

Variables de salida: Este tipo de variables son los productos finales o subproductos diseñados para su obtención que se tendrán al final del proceso. Asumiendo que el proceso haya sido elaborado de la manera idónea.

1. Textura: El término textura refiere a la sensación que produce al tacto el roce con una determinada materia y en el cual el sentido del tacto es el principal decodificador de la

misma, ya que es el vehículo o encargado de producir la sensación que ostente la textura en cuestión: suavidad, dureza, rugosidad, entre otras.

2. Volumen de Lixiviados: Es un subproducto que se tiene por efecto de las lombrices. Siendo un concentrado de nutrientes que es el desecho de las lombrices.
3. Cultivo de la semilla: Una vez germinadas las semillas y con un crecimiento adecuado se espera tener un cultivo saludable de lechuga romana, así como de stevia.
4. El olor: se genera por una mezcla compleja de gases, vapores y polvo, de la lombricomposta donde la composición de la mezcla influye en el tipo de olor percibido por el receptor.
5. El Color: Algunos colores de la lombricomposta cambiarán un poco dependiendo del nivel de humedad y otros factores o sustancias que los representan naturalmente.
6. Crecimiento de la planta: Esta es la fase en la que el brote comienza a desarrollarse y la planta se encuentra en su etapa más delicada. Las primeras hojas empiezan a formarse y se comienza a desarrollar el sistema de la raíz. Durante esta etapa es importante que se mantenga el suelo húmedo, pero bien drenado.

Variables controlables: Durante el desarrollo del proceso, estas variables son controladas, a partir de técnicas y métodos de producción podemos alterarlas de tal forma a que nos ayuden a buscar el más alto rendimiento para la lombricomposta.

1. Desechos orgánicos: Dependiendo de la cantidad de materia orgánica depositada en la composta podremos controlar la cantidad de materia fertilizante final, así como la cantidad de nutrientes que podría tener la composta.
2. pH: Dependiendo del material orgánico agregado, si proviene de fuentes cítricas o ácidas, se puede controlar en cierta medida el pH de la composta.

Variables no controlables: Estas variables son las que afectan el proceso de la lombricomposta, y consideramos que al tener algunas variables controladas el efecto de las que se salen de control disminuye. Entre estas:

- ❖ Temperatura ambiental: La temperatura ambiental es algo que no podemos modificar en gran escala. Se podría hacer algo de manera local y específica, pero la composta no cuenta con un sistema de calefacción por lo que depende totalmente de la temperatura exterior.
- ❖ Humedad ambiental: De igual forma la humedad ambiental es algo que no podemos modificar en gran escala. Igualmente se podría implementar algo de manera local y específica, pero el diseño dispuesto en este proyecto no contempla ningún control de la humedad interior por lo que depende totalmente de la humedad exterior.

### 13.4 Diseño prototipo 1 y 2: Lombricomposta

El proceso de elaboración de lombricomposta permite la utilización de una amplia variedad de residuos orgánicos, sin embargo, es más conveniente limitarla a los Desechos de frutas y verduras de la preparación de los alimentos; es decir, únicamente los desechos de origen vegetal en crudo, cáscaras y tallos: Cáscara o trozos de melón, sandía, papaya, manzana, plátano; hojas de elote, lechuga, tallos de cilantro, perejil, jitomate, chile, etcétera, evitando los cítricos. Es importante que los residuos sean cortados en trozos pequeños (máximo de 2 cm).

Materiales:

- 2 tarjas de plástico
- Lombrices
- Desechos orgánicos (a excepción de cítricos)
- Lombrices californianas o rojas



Figura 5. Contenedor

1. En este proceso se sometió a una perforación inferior a una charola la cual se posicionará sobre la otra con el objetivo de separar mediante filtración el exceso de humedad y separar el lixiviado este se perforo con un desarmador caliente



Figura 6. Contenedor superior

2. La segunda charola se coloca en la parte inferior del compostero teniendo como funcionalidad la acumulación del lixiviado, este es un concentrado de nutrientes en forma líquida esta charola se coloca encima de la otra.



Figura 7. Montado del compostero

3. En la charola superior (perforada) se forma una pequeña cama de tierra, la cual alberga las lombrices y el alimento de las lombrices las cuales comenzaran con el proceso de composteo (esta fue tierra normal de jardín)





Figura 8. Agregado de lombrices

4. Se colocó la materia prima en el compostero los residuos colocados fueron cascara de plátano, hojas de lechuga, ralladura de zanahoria, así como trozos de pepino y cascara de melón



Figura 9. Introducción de material orgánico (desechos alimentarios)

5. Se cerró el compostero con una tapa la cual deja un pequeño flujo de aire.



Figura 10. Sellado del compostero

6. El residuo del compostero quedo de esta manera, las lombrices procesaron el alimento (materia prima) quedando de la siguiente manera:





Figura 11. Diseño final

### 13.5 Diseño prototipo 3 y 4: Composta casera

#### Materiales

Tabla 1. Materiales

Costal de yute	
Tierra común	
Tierra de hoja	
Excremento de vaca	

Caja de madera	
Desechos orgánicos	

### Procedimiento

1. Colocar el costal dentro de la caja de madera, el costal es una fibra natural que aportará beneficios al proceso de descomposición de la composta y también evitará las fugas de la tierra por las hendiduras de la caja.
2. Colocar marcas de medidas al costal a partir del fondo, las medidas deben de ser de cuatro centímetros aproximadamente.
3. Una vez teniendo las marcas comenzaremos por aplicar una capa de tierra teniendo como referencia la primera marca de cuatro centímetros.
4. Continuaremos por aplicar una segunda capa de excremento de vaca, tomando como referencia los siguientes cuatro centímetros.
5. Seguiremos colocando una capa de desechos orgánicos y así sucesivamente intercalando las capas hasta llegar a 25 centímetros de la caja aproximadamente.
6. Colocaremos por ultimo una capa de tierra de hoja, la capa de tierra de hoja debido a que también es una formación de desechos orgánicos contribuirá en el proceso de descomposición y aportará nutrientes.
7. Colocaremos una bolsa de plástico negra que aumentara la temperatura interna del compostaje.

- Colocaremos una tapa de metal negra que de igual manera que la bolsa ayudará a aumentar la temperatura para acelerar la proliferación de microorganismos que harán la labor de la descomposición de los desechos orgánicos.



Figura 11. Inicio prototipo



Figura 12. Preparación



Figura 13. Preparación compostaje



Figura 14. Finalización

Una vez terminado el procedimiento debemos esperar aproximadamente mes y medio para que la composta esté lista, podremos notar que ya está terminado el proceso cuando ya no se diferencie cada elemento de desecho orgánico que se colocó, la composta debe de ser supervisada para valorar la humedad y temperatura, debido a que si existe mucha humedad los desechos orgánicos solo se pudrirán y no podrán aportar ya nutrientes a la tierra.

### 13.6 Mediciones de prototipos 1 y 2

#### Temperatura y humedad de la lombricomposta

La temperatura y humedad de la lombricomposta se midió con un sensor de temperatura y humedad (DTH11) controlado por un microprocesador Arduino, cuyo uso y programación es muy fácil por lo que no nos llevó problema alguno utilizarlo.

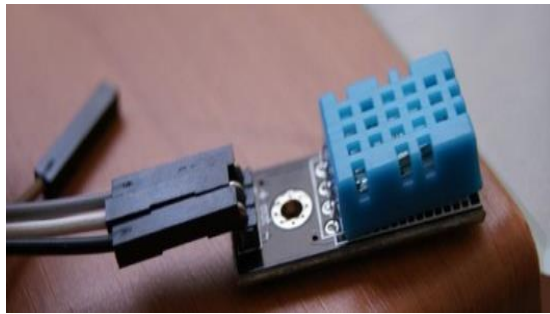


Figura 15. Sensor DTH11

### **Temperatura y humedad del ambiente**

Para las mediciones de temperatura y humedad del ambiente se pensó en usar el sensor DTH11, pero debíamos retirarlo de la lombricomposta y no deseábamos retirarlo de su entorno para no alterar las mediciones. Así que las mediciones las obtuvimos de la página web de meteorología de la NASA cuyas mediciones se entregan conforme a las latitudes y altitudes del lugar, en este caso la ciudad de Puebla. Estas mediciones son diarias y exactas.



## NASA Surface meteorology and Solar Energy - Available Tables



Latitude 19.455 / Longitude -98.769 was chosen.

## Geometry Information

Elevation: 1795 meters  
taken from the  
NASA GEOS-4  
model elevation

Northern boundary  
20

Center  
Latitude 19.5  
Longitude -98.5

Western boundary -99 Eastern boundary -98

Southern boundary  
19

## Meteorology (Temperature):

## Monthly Averaged Air Temperature At 10 m Above The Surface Of The Earth (°C)

Lat 19.455 Lon -98.769	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Average
22-year Average	12.3	14.1	16.7	18.9	19.7	19.0	18.2	18.3	17.6	15.9	14.3	12.5	16.5
Minimum	7.67	8.80	10.8	13.4	14.8	15.0	13.8	13.8	14.0	12.2	10.2	8.25	11.9
Maximum	17.7	19.8	22.8	24.7	24.9	23.3	22.9	23.2	21.8	20.2	19.0	17.6	21.5

[Parameter Definition](#)

## Average Daily Temperature Range (°C)

Lat 19.455 Lon -98.769	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
22-year Average	10.1	11.0	11.9	11.3	10.1 *	8.36	9.10	9.33	7.75	8.02	8.89	9.38

Figura 16. Página de la NASA meteorology

## Medición de pH

Para medir y controlar el pH se hizo lo siguiente:

- Coloca unos centímetros de jugo de repollo en una taza limpia y añade una o dos cucharadas de tierra. Espera durante treinta minutos y observa el color.
- El color púrpura o violeta indica que el pH es alrededor de 7, es decir, neutro.
- El color rosa indica que el suelo es ácido con un pH entre 1 y 7. Cuánto más ácido sea el suelo, más claro será el rosa.
- El color azul o verde indica que el suelo es alcalino, con un pH entre 8 y 14. Cuánto más alcalino sea el suelo, más claro será el color verde del jugo.

### 13.7 Comparación de las técnicas de medida por prototipo

Tabla 2. Técnicas de medida

Ventajas	Desventajas
Método sencillo	Fragilidad del componente
Medición rápida	Rango de temperatura: de 0° a 50° con 5% de precisión (pero solo mide por grados, no fracciones)
Económico	Que no presente los colores estandarizados.
Instrumento económico	No es adecuado para muestras de gran capacidad (se requiere un gran volumen).
Fácil realización	No es adecuado para muestras de gran capacidad.
Rango de humedad: de 20% al 80% con 5% de precisión	1 Muestra por segundo
Bajo consumo de energía	
Valores dados en °C y porcentaje (%)	

#### Crecimiento de la planta

La etapa de plántula se produce después de que la semilla germina, y dura entre una a dos semanas. Esta es la fase en la que el brote comienza a desarrollarse y la planta se encuentra en su etapa más delicada. Las primeras hojas empiezan a formarse y se comienza a desarrollar el sistema de la raíz. Durante esta etapa es importante que se mantenga el suelo húmedo, pero bien drenado. La lechuga romana cuenta con sistemas de raíces más profundas que las variedades de lechuga de hoja y necesita mucho espacio y humedad para un desarrollo saludable.

#### Color

Los colores de la lombricomposta son entre el negro a café oscuro, de acuerdo a lo experimentado.

El color del humus es café claro

### **Olor**

El olor es a tierra húmeda, el humus de lombriz está libre de olores.

Textura de la tierra es suave sin imperfecciones y húmeda.

### **13.8 Mediciones prototipos 3 y 4**

La composta como ya se mencionó antes debe ser supervisada en factores clave que nos indiquen aspectos específicos como humedad y temperatura. A continuación, se muestra una tabla de control de mediciones que se ha estado llevando a cabo diariamente, por métodos sencillos y caseros.

Tabla 3. Mediciones



Parámetros de medición	SEPTIEMBRE															OCTUBRE														
	18	19	20	21	22	23	24	25	26	27	28	29	30	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	
Temperatura °C	30	30	40	40	45	50	50	50	50	50	50	45	40	40	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	
Humedad %	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
Color																														
Olor	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	
Crecimiento (semilla)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
Parámetros de medición	OCTUBRE															NOVIEMBRE														
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	01	02	03	04	05	06	07	08	09	10	11	12	13	14	
Temperatura	29	25	22	20	30	29	28	25																						
Humedad	50	50	50	50	50	50	50	50																						
Color																														
Olor	M	M	M	M	M	M	M	P																						

Como se observa los primeros días la composta aumentó drásticamente de temperatura, alcanzando entre los 40°C a 60°C debido a que comenzaría el proceso de descomposición y además por la tapa y bolsa negras.

En los días posteriores la temperatura disminuyó considerablemente porque ya existía humedad entre los componentes, el costal de yute está sirviendo para mantener la humedad.

Los parámetros de coloración se miden, debido a que esta cambia según la humedad y absorción de los residuos orgánicos que tenga la tierra.

Anteriormente la humedad se ha medido bajo parámetros visuales, se toma en consideración la coloración de la bolsa de yute para saber si hay mucha poca o nada de humedad.

A continuación, se presentan los parámetros de medición visuales que se emplearon para la determinación de la humedad en el yute.



Figura 17. Escala comparativa visual de coloración del yute para mediciones de humedad

Cuando no existe nada de humedad el yute esta de su color normal



Figura 18. Yute seco

Cuando el yute esta húmedo y en las condiciones apropiadas para la composta esta de un color un poco más marrón.



Figura 19. Yute con la humedad idónea para composta

Si se aumenta la humedad en el yute, la coloración es marrón más intenso e incluso se nota la transferencia de la tierra que se cuela por sus fibras.



Figura 20. Yute con mucha humedad no apta para la composta

Cuando el yute está demasiado húmedo la coloración es café ya que la tierra ya es lodo y se cuela por medio de sus fibras.



Figura 20. Yute con exceso de humedad

## Temperatura

Para la medición de la temperatura optamos por la utilización de un termómetro digital casero.



Figura 21. Termómetro digital para medición de temperatura de la composta



Figura 22. Medición de temperatura con termómetro digital

## Color

La coloración es un factor indicador dentro del proceso de la composta ya que se ha observado que existe un cambio de coloración de la tierra conforme a la humedad y la absorción de desechos orgánicos, éste permite identificar cuando la composta ya esté lista para ser utilizada, momento en el cual toma un color más oscuro de la tierra inicial que se agregó a la composta.

A continuación, se muestra un comparativo de avance de la coloración



Figura 23. Color de la tierra antes de ser colocada en la composta y sin humedad



Figura 24. Coloración de la tierra una semana después de ser parte de la composta



Figura 25. Coloración de la tierra una semana después de ser parte de la composta y con la humedad suficiente



Figura 26. Coloración de la tierra 4 semanas después de formar parte de la composta y con la humedad adecuada

## Olor

Los parámetros de olor no pueden ser evidenciales debido a que se mide con el olfato y se puede tener una referencia si el olor va disminuyendo y empieza a ser más dominante el olor de la tierra y desechos orgánicos. En la tabla de control anterior se delimita la letra “M” para mucho, “P” para poco y “N” para nada.



Figura 27. Parámetros de olor de la composta

## Crecimiento de la semilla

Por el momento aún no se tienen parámetros de medición exacta debido a que la semilla recién ha sido plantada.

Pero se espera que aproximadamente en un mes la semilla alcance los 20 centímetros, 7 más del promedio, de ahí determinaremos si la composta fue un éxito.

### 13.9 Resultados

#### 14.9.1 Mediciones de prototipos 1 y 2

Tabla 4. Temperaturas

Semana	Temperatura del ambiente C°	Temperatura del ambiente C°
1	18	27
2	17	25
3	9	26
4	16	25
5	19	23
6	18	28
7		
8		

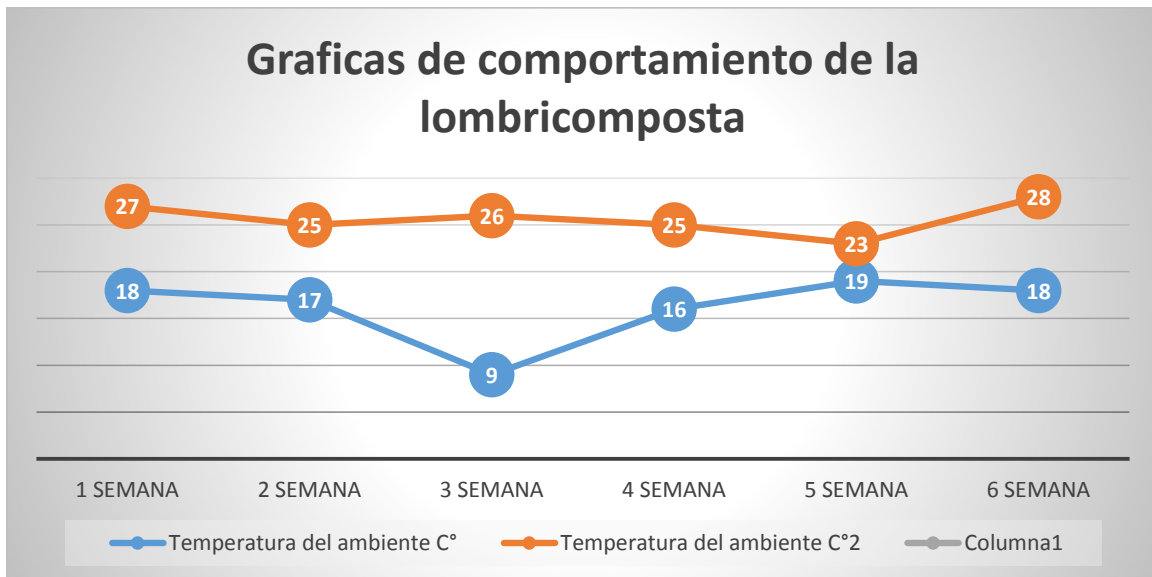


Figura 27. Comportamiento de la lombricomposta

Tabla 5.

Semana	Temperatura del compostero	Humedad del compostero	PH
1	18	80	9
2	19	92	9
3	17	80	7
4	22	82	5
5	25	77	6
6	24	79	8
7			
8			



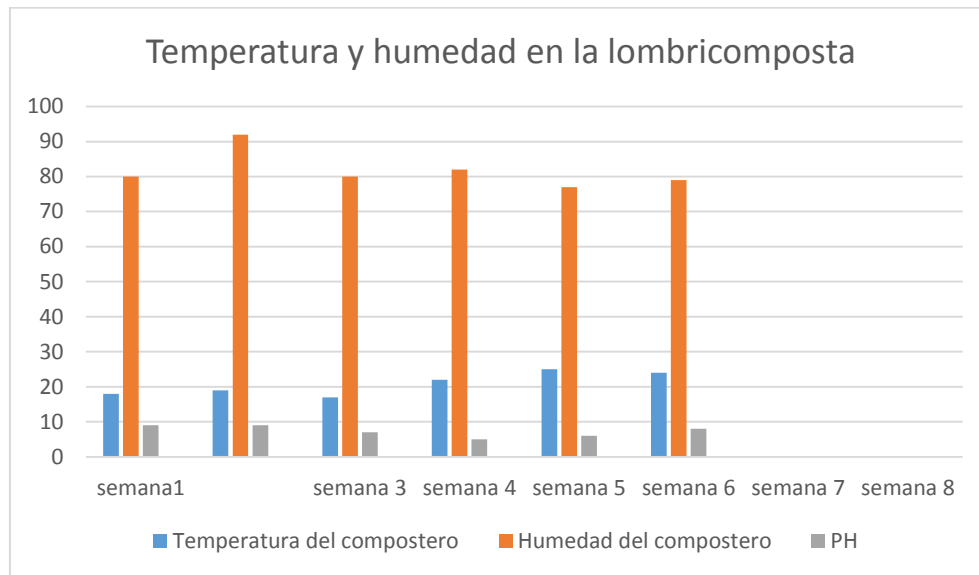


Figura 28.

## 14.9.2 Mediciones en composta casera

### Comportamiento de la Temperatura

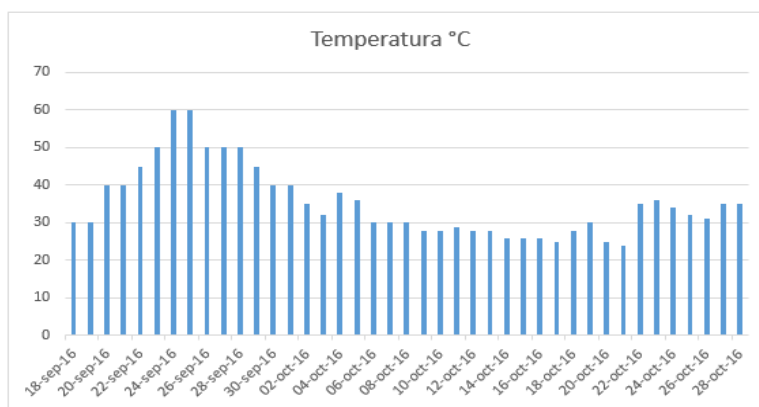


Figura 29.

## Comportamiento de la humedad

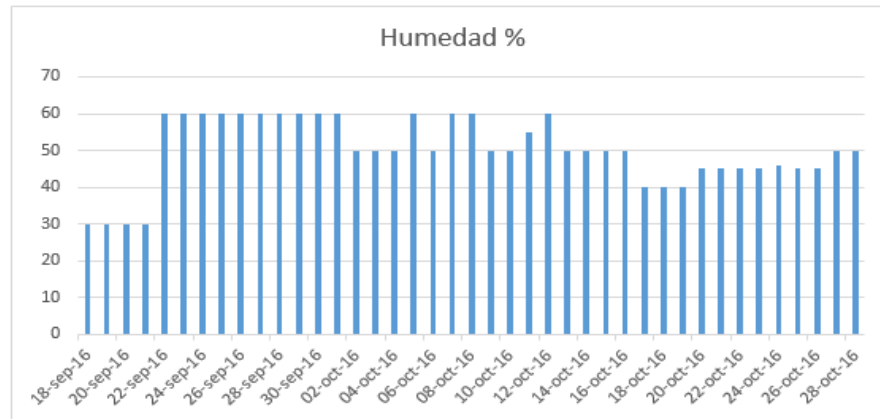


Figura 29.

### 13.9.3 Análisis de las observaciones prototipos 1 y 2

Con base a los resultados obtenidos, se busca un modelo matemático que pueda predecir con exactitud el comportamiento que tendría la lombricomposta y el cultivo de la lechuga y stevia, esto es importante ya que nos permitirá predecir el desempeño de nuestro proceso y mejorar los factores y variables que puedan estar involucrados.

El diseño del modelo matemático se basa en un diseño completamente aleatorizado, en el cual podemos especificar las características siguientes pertenecientes al modelo:

Interés: Un solo factor con varios niveles o tratamientos, ya que el fin de la lombricomposta es tener un cultivo sano de lechuga y de stevia, existiendo múltiples factores que lo podrían afectar como se muestra en las variables del diagrama de procesos.

Técnica estadística: Análisis de la varianza de un factor o una vía.

Objetivo: Comparar entre si múltiples variables.

Método: Descomposición de la variabilidad total de un experimento en componentes independientes.

Teóricamente es posible dividir esta variabilidad en dos partes, la originada por el factor de interés y la producida por los restantes factores que entran en juego, conocidos o no, controlables o no, que recibe el nombre de perturbación o error experimental.

### MODELO ESTADÍSTICO

$$y_{ij} = \mu + \tau_i + u_{ij} \quad , \quad i = 1, \dots, I; \quad j = 1, \dots, n_i$$

$y_{ij}$  : Variable aleatoria que representa la observación  $j$ -ésima del  $i$ -ésimo tratamiento (nivel  $i$ -ésimo del factor).

$\mu$  : Efecto constante, común a todos los niveles. Media global.

$\tau_i$  : Efecto del tratamiento  $i$ -ésimo. Es la parte de  $y_{ij}$  debida a la acción del nivel  $i$ -ésimo, que será común a todos los elementos sometidos a ese nivel del factor.

$u_{ij}$  : Variables aleatorias que engloban un conjunto de factores, cada uno de los cuales influye en la respuesta sólo en pequeña magnitud pero que de forma conjunta debe tenerse en cuenta. Deben verificar las siguientes condiciones:

1. La media sea cero:  $E[u_{ij}] = 0 \forall i, j$ .
2. La varianza sea constante:  $Var[u_{ij}] = \sigma^2 \forall i, j$
3. Independientes entre sí:  $E[u_{ij} u_{rk}] = 0$  si  $i \neq r$  ó  $j \neq k$ .
4. Distribución sea normal.

#### 13.9.5 Análisis de las observaciones prototipos 3 y 4

El objetivo general del estudio fue elaborar un modelo de superficie de respuesta que permitiera estimar condiciones en la producción de abono orgánico cuyo aporte de potasio fuera óptimo para la plantación de Stevia. En la producción de abono orgánico se utilizaron los desechos sólidos

generados en un hogar donde habitan 4 personas. La experiencia se planteó aplicando el diseño factorial fraccionado de Box-Behnken, utilizando los factores identificados y rangos de variación obtenidos preliminares los cuales fueron: olor (D), textura(C) y color(A), así como el periodo de crecimiento de la planta (B). La respuesta seleccionada como indicador de aporte de nutriente fue el potasio, principal nutriente en la Stevia Necesario para darle el sabor dulce característico. Ello permitió preparar una muestra.

El modelo de superficie consistió en la obtención de una función matemática que es de la forma:  $y = b_0 + b_i X_{ij} + b_{ij} X_i X_j + b_{1i} X_{1i} + e_{ij}$ , «  $i = 1,2,3,4$   $J = 1,2,3$  donde  $y$  corresponde a la respuesta analizada,  $X_{ij}$  corresponde al factor  $i$  nivel  $j$ , finalmente  $e_{ij}$  es el error aleatorio correspondiente a estas variables.

Se obtuvo la ecuación correspondiente al modelo de superficie para la respuesta analizada (cantidad de potasio,  $y_4$ ) y se agrupan de la siguiente manera:  $y_4 = 27,16 + 1,40A - 1,43B - 12,80C + 12,56D - 0,51AB - 3,36AC - 4,92AD + 3,63BC - 1,70BD - 5,47CD$ .

Con la ecuación obtenida se optimiza hasta obtener el punto máximo, de manera que se diseña un modelo que permite establecer las condiciones de mezcla inicial para preparar abono orgánico cuyo aporte de potasio se repite.

Para comprobar lo antes mencionado se presentan tablas de promedios de las variables analizadas:

- ✓ Promedio total de temperatura (D)=**35.3658**
- ✓ Promedio total de humedad (C)=**50.024**
- ✓ Promedio de la cantidad de desechos orgánicos (A)= **20%\***

\*La cantidad materia orgánica en los primeros días ira del 50%-70%, de la segunda a la quinta semana y dependiendo del proceso de descomposición se reducirá en un 20%, así que

se establece un 20% debido a que es la cantidad de materia orgánica que se encontrara en la composta en mayor tiempo. Una familia promedio de 4 personas produce a la semana 10 kg de residuos verdes y de cocina. El material dentro del recipiente tiende a compactarse hasta 400kg/m<sup>3</sup>.

✓ ¿Promedio de crecimiento de la planta (B)=?

\*Se delimita que aproximadamente el peso total de la composta sea de 25 kg y que el nutriente principal de estudio es el potasio se establece que el factor de conversión es 0,44, así que el contenido de P = 0.06 (constante de porcentaje de concentración de potasio en una composta casera) x 0.44 (factor de conversión del potasio) x 25 kg = 0.66 kg de Potasio en la composta. Si normalmente la stevia requiere un 3.45% potasio.

En bibliografía, sabemos que los valores adecuados para un buen compostaje son:

Tabla 6.

Fecha	Temperatura	Humedad %
18-sep-16	30	30
19-sep-16	30	30
20-sep-16	40	30
21-sep-16	40	30
22-sep-16	45	60
23-sep-16	50	60
24-sep-16	60	60
25-sep-16	60	60
26-sep-16	50	60
27-sep-16	50	60
28-sep-16	50	60
29-sep-16	45	60
30-sep-16	40	60
01-oct-16	40	60
02-oct-16	35	50
03-oct-16	32	50
04-oct-16	38	50
05-oct-16	36	60
06-oct-16	30	50
07-oct-16	30	60
08-oct-16	30	60
09-oct-16	28	50
10-oct-16	28	50
11-oct-16	29	55
12-oct-16	28	60
13-oct-16	28	50
14-oct-16	26	50
15-oct-16	26	50
16-oct-16	26	50
17-oct-16	25	40
18-oct-16	28	40
19-oct-16	30	40
20-oct-16	25	45
21-oct-16	24	45
22-oct-16	35	45
23-oct-16	36	45
24-oct-16	34	46
25-oct-16	32	45
26-oct-16	31	45
27-oct-16	35	50
28-oct-16	35	50
	<b>35.36585366</b>	<b>50.02439024</b>

Núm.	Factor considerado	Intervalo aceptable	Valor óptimo
1.	Composición inicial de la mezcla, relación C/N (%)	25 a 35 : 1	30:1
2.	Potencial de hidrógeno en la mezcla inicial, pH	6.5 a 8.5	7
3.	Contenido de humedad de la mezcla durante el compostaje (%)	40 - 60	60
4.	Contenido de oxígeno en el aire dentro de la cámara de compostaje (%)	> 5 %	Mínimo: 10 %
5.	Temperatura en la etapa termófila (°C)	Entre los 40 y 65	60
6.	Tamaño de partícula de los materiales (cm)	En general de 1 a 5 cm y para materiales leñosos 1 a 2 cm.	1- 2 cm

Figura 30. Niveles aceptables de los factores físicos y químicos para el compostaje, y los valores óptimos.

Con las condiciones establecidas en la elaboración de composta hemos llegado a los siguientes resultados preliminares

Tabla 6.

SIN COMPOSTA			CON COMPOSTA		
<u>LECTURA</u>	<u>No. DE DIA</u>	<u>MEDIDA EN cm</u>	<u>LECTURA</u>	<u>No DE DIA</u>	<u>MEDIDA EN cm</u>
<b>1</b>	0	0	<b>1</b>	0	0
<b>2</b>	3	0.5	<b>2</b>	3	1.1
<b>3</b>	6	0.6	<b>3</b>	6	1.5
<b>4</b>	9	0.85	<b>4</b>	9	1.7
<b>5</b>	12	1.15	<b>5</b>	12	2.1
<b>6</b>	15	1.415	<b>6</b>	15	2.3
<b>7</b>	18	1.68	<b>7</b>	18	2.92
<b>8</b>	21	1.945	<b>8</b>	21	3.34
<b>9</b>	24	2.21	<b>9</b>	24	3.76
<b>10</b>	27	2.475	<b>10</b>	27	4.18
<b>11</b>	30	2.74	<b>11</b>	30	4.6
<b>12</b>	33	3.005	<b>12</b>	33	5.02

Es importante remarcar que los resultados de la planta de stevia y lechuga con lombricomposta están aún en observación, por lo pronto podemos concluir que el proceso de compostaje es adecuado para la mejora en el crecimiento de las plantas y el modelo estadístico nos permitió determinar las variables que afectaban de manera directa la producción de composta y lombricomposta.

### 13.10 Conclusiones

Como puede observarse la implementación del aprendizaje basado en proyectos permitió obtener excelentes resultados en el desarrollo de las actividades propias de la materia, donde se comprobó que se lograron los objetivos propuestos en el programa de estudios, por otra parte la unión con el diseño de experimentos como herramienta estadística colaboró en dar orden a las mediciones de cada uno de los prototipos, también el diseño de experimentos (DoE), ayuda a investigar los efectos de las variables de entrada, sobre una variable de salida, al mismo tiempo.

Estos experimentos consisten en una serie de corridas, o pruebas, en las que se realizan cambios intencionales en las variables de entrada. En cada corrida se recolectan datos.

El DOE se *utiliza para identificar las condiciones del proceso y los componentes del producto que afectan la calidad, para luego determinar la configuración de factores que optimiza los resultados, de manera tal que, con base a los resultados obtenidos y mostrados anteriormente hemos determinado que las posibilidades de éxito de nuestra composta son muy favorables y cumple hasta ahora con todos los parámetros de metodologías planteados, por otra parte también se ha analizado por qué y que tan favorable seria plantar Stevia para corroborar los resultados de éxito llegando a las siguientes conclusiones*

En este proceso de experimentación de plantación con composta, se evaluaron distintos tipos de plantas para poder identificar de cual se obtendrán mejores resultados evaluando el costo-beneficio, al término de esta evaluación se determinó que la stevia es una planta de múltiple uso medicinal y uso como endulzante natural, se determina esta realizar la experimentación por las siguientes razones

- ✓ La Stevia requiere de múltiples nutrientes para poder adquirir ese sabor dulce tan distintivo.



- ✓ El terreno ideal para su cultivo es areno-arcillosa con una porción regular de humus y puede resistir suelos ligeramente alcalinos.
- ✓ Es resistente al frío pero no a la sequía el clima en su habitat natural es el subtropical, (El 35% de la superficie del estado presenta clima templado subhúmedo presente en la región central y sureste el 25% presenta clima cálido subhúmedo en la parte norte y sureste, el 19% presenta clima seco y semiseco hacia el sur y centro oeste, el 14% presenta clima cálido húmedo localiza en el norte y sureste, el 7% presenta clima templado húmedo en la región norte y una pequeña área hacia el sureste, también encontramos un pequeño porcentaje (0.2) de clima frío en la cumbre de los volcanes.)
- ✓ Su temperatura ideal esta entre los 15°C y los 30°C, con un límite inferior de -3, (La temperatura media anual del estado de Puebla es de 17.5°C, la temperatura máxima promedio es de 28.5°C y se presenta en los meses de abril y mayo, la temperatura mínima promedio es de 6.5°C durante el mes de enero y la precipitación media estatal es de 1 270 mm anuales, las lluvias se presentan en verano en los meses de junio a octubre.
- ✓ Esta plata se desarrolla mejor en temperaturas tibias entre los 5 ° C y 30 ° C.
- ✓ Requiere días con prolongada radiación solar, pero no mayores a los 43°, (En Puebla la temperatura máxima se estima a 28°C)
- ✓ Esta especie ha sido cultivada en distintas condiciones climatológicas obteniendo éxito en su plantación.
- ✓ Se cosecha en otoño, (La experimentación se llevara en el periodo de septiembre-diciembre del presente año)
- ✓ La floración de la Stevia es de aproximadamente un mes, (se cuenta con 3 meses para la experimentación)
- ✓ Transcurridos tres meses en el semillero los plantones estarán listos para su clasificación, (tiempo idóneo para la experimentación con composta)

Puebla cuenta con las condiciones climatológicas adecuadas para la experimentación con plantación de Stevia con composta, el tiempo que se requiere para realizar el comparativo y

resultados de las muestras es apto para el tiempo disponible, es una planta resistente con múltiples beneficios



Figura 31. Planta de Stevia

Las bajas temperaturas (invierno) retrasarán el compostaje, por lo que es conveniente iniciar una pila en primavera o verano.

De las observaciones obtenidas en la replicabilidad de cada una de las unidades experimentales se tienen las siguientes consideraciones, Tabla 8, donde fue posible determinar características de la composta cuando ya está madura y cuando aún no lo está.

Tabla 8. Replicabilidad.

Material	Composta inmadura	Composta madura
Color	Marrón poco obscuro	Marrón obscuro
Olor	Más o menos fuerte	Sin olor fuerte
Composición	Hay lombrices y partes del material que pueden ser identificados	NO hay lombrices y nada del material puede ser identificado.

Uso	Puede ser usado como cobertera para jardines, arbustos y árboles perennes	Se puede incorporar en la tierra
Cantidad	Usar poca cantidad ya que puede quemar las plantas	No hay riesgos puede realizar varias aplicaciones.

Tabla 9. Observaciones y sus posibles modificaciones.

Material	Causa	Solución
Mal olor	Falta oxígeno Demasiada agua Demasiado material verde El prototipo es muy grande o muy pequeño	Voltee el material. Agregue hojas secas , aserrín o paja. Agregue material color marrón Adecue tamaño
Centro muy seco	Falta de agua	Voltee y humedezca
Temperatura no sube	El prototipo es muy pequeño Falta materia verde Temperatura ambiente muy fria	Aisle los lados. Agregue cortes pequeños de pasto o restos de vegetales o frutas.
Pila muy húmeda	Excesiva lluvia Excesivo riego	Tapar con plástico, cuidar que se permita la aireación a través de hoyos. Agregue material seco Hojas, aserrín paja.
Vectores, moscas	Restos de cocina	Cubra los restos de desechos con tierra, composta vieja u hojas secas.

Como resultado de la replicabilidad y repetibilidad de los prototipos se obtuvieron las soluciones para las diferentes problemáticas observadas, en la tabla 9.

El presente proyecto se realizó en la materia de Dirección de proyectos de Bioenergía en el 10° cuatrimestre de la Universidad Tecnológica de Puebla. Con la dirección de la maestra Guillermina Sánchez López, la asesoría del Dr. José Dionicio Zacarías Flores, la bióloga Abigail Juárez Sánchez.



## Referencias

- Aliprantis, C. D. y Carmona, G. (2003). *Introduction to an Economic Problem: A models and modeling perspective*. En Lesh, R. y Doerr, H. (Eds.).
- Andreu, M<sup>a</sup> Á. González, J.A. Labrador, M<sup>a</sup> J. Quintanilla, I. Ruiz, T. (2004). Método del caso Ficha descriptiva y de necesidades. Valencia, España.
- Anderson, C. W. y Loynes, R. M. (1987). *The teaching of practical statistics*. New York: Wiley.
- Anderson, J. E. and Sungur E. A. (1999). *Community service statistics projects*. The American Statistician, 53, 132-136.
- Ausubel, D.P., Novak, J.D. and Hanesian, H. (1978). Educational psychology: a cognitive view. 2nd. ed. new york, holt rinehart and Winston.
- Ausubel, D.P, Novak, J.D. and Hanesian, H. (1980). Psicología educacional. Rio de janeiro, interamericana. Tradução para português, de Eva Nick et al., da segunda edição de educational psychology: a cognitive view.
- Ausubel, D.P., Novak, J.D. y Hanesian, H. (1983). Psicología educativa: un punto de vista cognoscitivo. México, editorial trillas. Traducción al español, de Mario Sandoval p., de la segunda edición de educational psychology: a cognitive view
- Batanero, C. y Díaz, C. (2011). *Estadística con proyectos*. Granada: Universidad de Granada.
- Espinoza J. (2009). Estrategias de enseñanza para la promoción de aprendizajes significativos, Ciencia ahora, 12 (34).
- Fernández A. (2006). Metodologías activas para la formación de competencias. Educatio sigloXXI, 24, 35-56.

Ferreiro R. “*Estrategias didácticas del aprendizaje cooperativo*”. México. Ed. Trillas. (2007).

Ferreira M. F. (2011). “*Implementación y evaluación de un modelo didáctico, basado en enfoques constructivistas, para la enseñanza de Estadística en el nivel superior*”. 3 de abril de 2016, de Universidad autónoma de baja california Sitio web: <http://iide.ens.uabc.mx/blogs/mce/files/2011/03/TesisMaestria-MFFerreira.pdf>

García C. J., Sánchez Q. C., Jiménez V., Gutiérrez T. M., (10 de octubre 2012). *Estilos de Aprendizaje y Estrategias de Aprendizaje: un estudio en discentes de postgrado. Estilos de aprendizaje*, 10, 17.

Sánchez L. Guillermina, Zacarías F. José, 2015, *Enseñanza de la estadística con la integración de 2 ideas didácticas: aprendizaje basado en proyectos (abp) y actividades reveladoras del pensamiento (mea). Una experiencia a nivel superior*” memorias del Congreso Internacional de Didáctica de la Matemática, International Mathematics Education Conference. Santa Marta Colombia





# 14

---

## El cálculo de áreas a través de la probabilidad

---

Giovanni Francisco Sanabria Brenes, Félix Núñez Vanegas.

gsanabria@itcr.ac.cr, fnunez@itcr.ac.cr.



### Resumen

Los currículum actuales de diversas carreras, a nivel universitario, exigen una transversalidad de los conocimientos, donde se hace necesario establecer conexiones entre distintos cursos con el fin de resolver problemas. Por otro lado, una potencialidad de la probabilidad es el de brindar un modelo para resolver problemas, incluso ajenos al azar, entre ellos, el cálculo de áreas. El presente trabajo brinda una propuesta para abordar el cálculo de áreas e introducir el concepto de integral definida por medio de la probabilidad y el uso de software. La propuesta vincula el cálculo diferencial, la probabilidad y el uso de software para resolver problemas de cálculo de áreas. La propuesta va dirigida a estudiantes universitarios con conocimientos básicos en probabilidad y cálculo diferencial.



**Palabras clave:** didáctica, probabilidad frecuencial, ley de los grandes números, cálculo de áreas, integral definida.

### **Abstract**

The current curricula of different careers, at the university level, require a transversality of knowledge, where it is necessary to establish connections between different courses in order to solve problems. On the other hand, a potentiality of probability is to provide a model to solve problems, even outside random, including the calculation of areas. The present work offers a proposal to approach the calculation of areas and introduce the concept of integral defined by means of the probability and the use of software. The proposal links differential calculus, probability and the use of software to solve area calculation problems. The proposal is aimed at university students with basic knowledge in probability and differential calculus.

**Keywords:** didactics, frequency probability, law of large numbers, calculation of areas, definite integral.

## **14.1 Introducción**

En una visión integral de educación matemática, los conceptos deben verse ligados unos con otros, y no en forma aislada, con el fin de ver su carácter de instrumento y carácter de objeto, de los cuales habla su Douady (1984). El ver un concepto como instrumento para resolver un problema es lo que le da sentido al concepto.

Por otro lado, en Núñez y Sanabria (2017), se indica que “el concepto puede intervenir en uno o varios marcos: geométrico, numérico y algebraico, entre otros. En cada marco el concepto se visualiza en términos de objetos y relaciones, formando significados del concepto en el marco”.

La ventaja de este juego de marcos, propuesto por Douady (1984), permite establecer correspondencias entre los significados que un mismo concepto adquiere en diferentes marcos, y contribuye a construir la diversidad semántica del concepto. Ella recomienda, que para haya un buen funcionamiento de los conocimientos de los alumnos, el docente debe proponer situaciones problemas en los que intervengan al menos dos cuadros.

Amén de lo anterior, hoy en día es importante tomar en cuenta en los diferentes procesos educativos, la tecnología. Recientemente, la Escuela de Matemática del Tecnológico de Costa Rica aprobó la carrera Enseñanza de la Matemática con Entornos Tecnológicos, aprobó el modelo pedagógico MTSK y TPACK para esta carrera. Específicamente, en la justificación que se da se indica que:

“Dadas las características de la carrera y su énfasis en el uso de la tecnología para la enseñanza de la matemática, el modelo, dado que el mismo se ajusta a los procesos de formación del plan de estudios aprobado, el cual considera un enfoque humanista y en el que los futuros docentes adquieran conocimientos, tanto en el área de la matemática, la pedagogía y la tecnología de manera integrada, según los principios establecidos en los modelos”.

Y más adelante, en el criterio número cuatro del mismo documento se insiste en el siguiente punto:

“Diseñar cursos específicos sobre el uso de la tecnología en la enseñanza y aprendizaje de la matemática, donde se aborde no solo el conocimiento tecnológico sino su relación con la matemática y con la pedagogía según el modelo TPACK”.

Este trabajo viene a aportar en esas líneas, al ver el cálculo de áreas como un problema de probabilidad, e introducir el concepto de integral definida ligado al concepto de probabilidad. A su vez, incorporamos a la tecnología en el ambiente de clase.

Por otro lado, la resolución de problemas debe estar presente en el proceso de enseñanza, y no solo al final para aplicar los conceptos adquiridos, sino también, al inicio para aprehender los conceptos. El profesor debe diseñar buenas situaciones problema para lograr estos objetivos.

Habíamos apuntado en Sanabria y Núñez (2016) que en el caso de la probabilidad, al revisar diversos libros de texto, aunque algunos plantean situaciones atractivas y contextualizadas, resolver un problema se reduce al cálculo de la probabilidad de un determinado evento. Esto da la sensación de que se busca calcular probabilidad sin ningún otro fin, más que el de calcular, y que después de abordar el estudio de un tópico de probabilidad, los problemas a resolver se reducen a calcular probabilidades.

En ese sentido entonces, es importante indicar cuáles problemas se ajustan bien en probabilidad. En Sanabria y Núñez (2016) se proponen tres tipos para la enseñanza de la probabilidad: Las situaciones centradas en el cálculo de probabilidades (ampliamente tratadas en los libros de texto), las situaciones sobre toma de decisiones y las situaciones ajenas al azar que utilizan la probabilidad como modelo.

Dentro de las situaciones ajenas al azar que utilizan la probabilidad como modelo, destacan los problemas de cálculo de áreas.

El presente trabajo brinda una propuesta que busca que los estudiantes apliquen el concepto de probabilidad y conocimientos de cálculo diferencial para poder resolver problemas de cálculo de área con ayuda de software (Excel o similar).

## **14.2 Sobre la propuesta**

### **14.2.1 Público meta**

La propuesta va dirigida a estudiantes que tiene conceptos básicos de probabilidad (probabilidad frecuencial, Ley de Laplace) y de cálculo diferencial (Cálculo de máximos y mínimos en un intervalo cerrado).

En el caso de Costa Rica los conceptos solicitados sobre probabilidad son vistos en secundaria según el Programa del Ministerio de Educación Pública. Por lo que, la propuesta puede ser usada como un trabajo extra clase que desarrollan los estudiantes que cursan Cálculo Diferencial después de haber visto el cálculo de máximos y mínimos en un intervalo cerrado. Además, creemos que facilitará en los estudiantes la comprensión del Cálculo Integral.

La propuesta puede ser fácilmente adaptada a estudiantes de secundaria eliminando algunas actividades de cálculo diferencial.

#### **14.2.2 Marco teórico**

Un requisito para utilizar la probabilidad como una herramienta para resolver un problema es que el problema posea una situación o experiencia azarosa.

Sin embargo, se puede presentar problemas desprovistos del azar, donde si aquel que lo resuelve quiere utilizar la probabilidad, debe recrear dentro del problema una situación azarosa, modelando las relaciones de la situación problema con los elementos teóricos de la probabilidad.

Este tipo de situaciones que utilizan la probabilidad como modelo requiere de un dominio más completo y complejo de la probabilidad, pues implican aplicar la probabilidad donde no se ve, construyendo una situación azarosa. En estas situaciones, aplicar el concepto adquirido de probabilidad, implica crear las condiciones necesarias para aplicarlo.

¿Cómo se aplica la probabilidad en este tipo de situaciones? Dado que la probabilidad es una medida relativa (medida de la posibilidad de que suceda de un evento), ésta puede ser utilizada para hallar ciertas medidas absolutas.

Más concretamente, de acuerdo a la Generalización de la ley de Laplace, dado un evento  $X$  de un espacio muestral equiprobable  $\Omega$ , con una medida asociada a  $\Omega$ , se tiene que la probabilidad de que ocurra  $X$  es:

$$P(X) = \frac{u(X)}{u(\Omega)}$$

Si  $P(X)$  y  $u(\Omega)$  son conocidos, entonces se puede hallar la medida de  $X$ .

Así, en ciertas situaciones que se requiera hallar una medida absoluta (por ejemplo la cantidad de objetos o un área), esta se puede ver como la medida de un evento de una situación azarosa construida.

Para comprender de qué estamos hablando, para el caso del cálculo de áreas, consideremos  $f$  una función continua no negativa en un intervalo  $[a, b]$ :

Como  $f$  es continua en  $[a, b]$ , entonces  $f$  alcanza su máximo y su mínimo. Sea  $M$  el máximo de  $f$  en  $[a, b]$ .

Considere la experiencia aleatoria: **Se eligen un punto azar del rectángulo  $R$  ( $[a, b] \times [0, M]$ )**, y considere el evento  $E$ : el punto se encuentra bajo la curva  $y=f(x)$ .

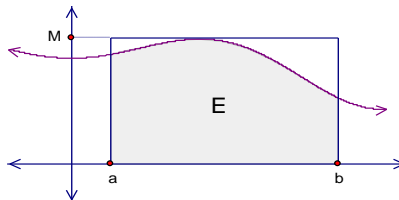


Gráfico 1

$$\text{Entonces } P(E) = \frac{\text{Area}(E)}{\text{Area}(R)} = \frac{\text{Area}(E)}{M(b-a)} \Rightarrow \text{Area}(E) = P(E) \cdot M(b-a)$$

Así, se define la integral definida para una función  $f$  continua no negativa en un intervalo  $[a, b]$  por

$$\int_a^b f(x) dx = P(E) \cdot M(b-a),$$

donde  $M$  es el máximo de  $f$  en  $[a,b]$ .

### 14.3 Una propuesta para introducir el cálculo de áreas utilizando probabilidad

Objetivo: Aplicar la probabilidad en el cálculo de áreas.

#### 14.3.1 Primera explicación y primer ejemplo

Recuerde que de acuerdo a la Generalización de la ley de Laplace, dado un evento  $X$  de un espacio muestral equiprobable  $\Omega$ , con  $u$  una medida asociada a  $\Omega$ , se tiene que la probabilidad de que ocurra  $X$  es:

$$P(X) = \frac{u(X)}{u(\Omega)}$$

Si  $P(X)$  y  $u(\Omega)$  son conocidos, entonces se puede hallar la medida de  $X$ .

Veamos la aplicación de esto en el siguiente ejemplo:

Ejemplo. Determine aproximadamente el área bajo la curva  $y = \sqrt{x}$  en el intervalo  $[0,1]$ .

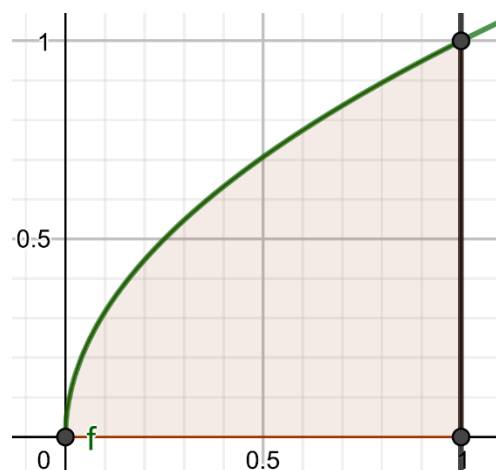


Gráfico 2

Veamos cómo se puede resolver esto: Sea  $A$  la región a la cual se le quiere determinar el área y  $\Omega$  el interior del cuadrado de vértices  $(0,0)$ ,  $(0,1)$ ,  $(1,0)$  y  $(1,1)$ :

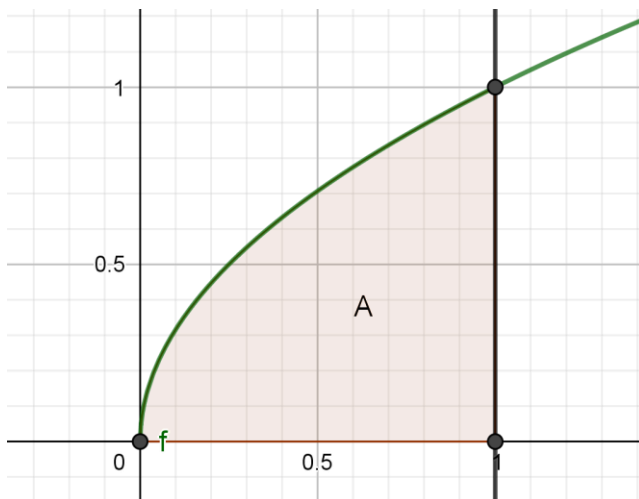


Gráfico 3

Ahora se necesita recrear una situación aleatoria: Se elige un punto  $Z$  al azar del interior del cuadrado. Note que la probabilidad de que  $Z$  esté en  $A$  es

$$P(Z \in A) = \frac{\text{Área}(A)}{\text{Área}(\Omega)} = \text{Área}(A)$$

Así, para determinar el área aproximada de  $A$ , basta hallar una aproximación a la probabilidad de que  $Z$  pertenezca a  $A$ . Para ello, utilizando Excel, debemos simular la elección al azar de un punto en el interior del cuadrado. Recuerde que un punto queda determinado de manera única por sus coordenadas.

Se escribe en Excel:

Tabla 7

	A	B	C
1	Coordenada X del punto Z	Coordenada Y del punto Z	¿El punto pertenece a A?
2	=Aleatorio()	=Aleatorio()	= SI(B2<RAIZ(A2); “SI”; “NO”)

Note que para saber si el punto está en A se compara la coordenada Y del punto con la imagen de la coordenada X del punto en la función  $f(x) = \sqrt{x}$ , por ejemplo:

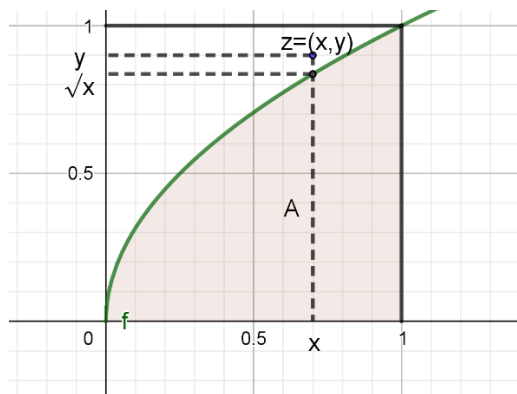


Gráfico 4

En este caso  $\sqrt{x} < y$  por lo tanto  $Z \notin A$ .

Luego, se arrastran las fórmulas de la fila 2 hasta la fila 1001, obteniendo 1000 simulaciones de la situación azarosa. Si se cuenta el número de simulaciones que tiene “SI” y se divide entre 1000, se obtiene una aproximación a  $P(A)$ , que constituye una aproximación de A.

### 14.3.2 Experimentación:

1. Se propuso a un grupo de 12 estudiantes de la carrera de Enseñanza de la Matemática con Entornos Tecnológico del Tecnológico de Costa Rica realizar la actividad anterior. Los jóvenes estaban cursando Cálculo y Análisis I. Dentro de los temas que se ven en el mismo está extremos



de una función continua en un intervalo cerrado. Indicamos que estos muchachos aún no han llevado el curso de cálculo integral. Se trabajó en un laboratorio y se usaron los paquetes de Excel y Geogebra. Con el fin de que se familiarizaran con la función ALEATORIO( ) de Excel, se les indicó que esta generaba números aleatorios entre 0 y 1. Luego se les pidió que en una celda generaran números aleatorios entre 2 y 3: Después de darles un par de minutos concluyeron que se debía escribir en la celda  $=2+(3-2)* \text{ALEATORIO}( )$ . Luego les pedimos que dieran una fórmula para generar números aleatorios entre dos números reales a y b cualesquiera. Escribieron la fórmula  $=a + (b-a)* \text{ALEATORIO}( )$ . Por otro lado, ya los estudiantes estaban familiarizados con el paquete GEOGEBRA. Se les pidió entonces trazar la gráfica de  $f(x) = \sqrt{x}$  y que sombrearan la región A, región a la cual se le quiere determinar el área y,  $\Omega$ , el interior del cuadrado de vértices (0,0), (0,1), (1,0) y (1,1). Luego se les preguntó cómo saber si dado un punto  $(x,y)$  en  $\Omega$ , éste quedaba en la curva. La respuesta de ellos fue que estaba en la curva si  $y = \sqrt{x}$ . Luego les preguntamos cómo saber si dicho punto se encontraba bajo la curva, nos indicaron que el punto estaba bajo la curva si  $y < \sqrt{x}$ . Enseguida les pedimos que simuláramos la situación en Excel tal como se indicó arriba. Después de realizar la simulación nos indicaron que una aproximación del área es 0,67.

Estaban sorprendidos porque el valor que obtuvieron, coincidía con el valor del área que indicaba Geogebra.

2. Luego se les propuso que estimaran el área bajo la curva  $y = (2 - x)e^x$  en el intervalo [0,2].

Sin la intervención nuestra, los muchachos graficaron la curva. Se dieron cuenta que x variaba de 0 a 2, por lo que escribieron en Excel  $0+2* \text{ALEATORIO}( )$ . Se dieron cuenta que y variaba de 0 al Máx de la función. Procedieron a calcularlo. Este máximo les dio e. Así que y variaba de 0 a e. En Excel colocaron lo siguiente:

Tabla 8

	A	B	C
1	Coordenada X del punto Z	Coordenada Y del punto Z	¿El punto pertenece a A?
2	=2*Aleatorio()	=EXP(1)*Aleatorio()	=SI(B2<(2-A2)EXP(A2); "SI"; "NO")

Calcularon la probabilidad frecuencial de que el punto perteneciera a la región A. Aproximadamente 0,80. Indicaron que el resultado estaba mal porque en Geogebra se marcaba como área 4,39. Se les dio tiempo para ver qué pasaba. Al fin un par de estudiantes concluyó que el área del rectángulo era de  $2e$ , por lo que:

$$\frac{Area(E)}{2e} \approx 0,80, \text{ de donde } Area(E) \approx 0,80 * 2e \approx 4,38.$$

La gráfica considerada en este ejercicio se da a continuación con el valor del área que indica Geogebra:

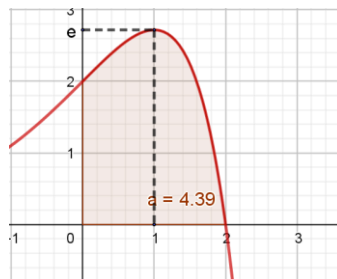


Gráfico 5

### 14.3.3 Actividad de cierre

Finalmente se les propuso la siguiente situación problema: Dada una función  $f$  continua no negativa en un intervalo  $[a,b]$ , ¿Cómo se puede hallar aproximadamente el área bajo la curva  $y=f(x)$  en el intervalo  $[a,b]$ ?

Los dejamos trabajar unos minutos. Al final concluyeron lo esperado: Como la función es continua en el intervalo cerrado  $[a,b]$ , la función alcanza su máximo absoluto. Indicaron que el área es aproximadamente  $\text{Área}(A) \approx \text{Máximo de la función} \cdot (b-a) \cdot P(A)$ .

Se les indicó al cierre de la actividad que lo que acabamos de descubrir era el concepto de integral definida para una función  $f$  continua no negativa en un intervalo  $[a, b]$  que se denota por

$$\int_a^b f(x)dx = P(E) \cdot M(b - a)$$

donde  $M$  es el máximo de  $f$  en  $[a,b]$ ,  $P(E)$  corresponde a la probabilidad de que un punto del rectángulo  $R$  ( $[a, b] \times [0, M]$ ), esté bajo la curva de  $y=f(x)$ .

#### 14.3.4 Actividades a realizar

Con el fin de reforzar lo establecido anteriormente, se propone desarrollar las siguientes situaciones problema:

(Similar al ejemplo). Halle aproximadamente el área bajo la curva  $y = \sqrt{x} + 1$  en  $[1,6]$ .

(Curva oculta). Determine aproximadamente el área de un círculo de radio 2.

(Polígono irregular). Halle aproximadamente el área del polígono ABCDE donde las coordenadas de los vértices  $A(1,1)$ ,  $B(3,3)$ ,  $C(5,7)$ ,  $D(7,4)$ ,  $E(10,2)$ .

(Debe hallar el máximo). Sea  $f(x) = (3 - x)e^x$ . Halle aproximadamente el área bajo la curva  $y = f(x)$  en  $[0,3]$ .

(Función por partes). Sea  $f(x) = \begin{cases} \frac{2}{x} + 2x & \text{si } x \geq 0 \\ \sqrt[3]{x^2 + 4x + 3} & \text{si } x < 0 \end{cases}$ . Halle aproximadamente el área

bajo la curva  $y = f(x)$  en  $[-3,2]$ .

(Entre una curva y una horizontal). Determine el área entre las curvas:

$$y = x \ln^2(x) + 5, y=2, x=1/2, x=2.$$

(Entre dos curvas). Determine el área entre las curvas.  $y = 1 - x^2, y = x^2$ .

#### 14.4 Conclusiones

Se abordó el problema del cálculo del área bajo una curva de un función continua y positiva en un intervalo cerrado desde un enfoque probabilístico. La ventaja de esto, en el juego de marcos, propuesto por Douady (1984), es que permite establecer correspondencias entre los significados que un mismo concepto adquiere en diferentes marcos, y contribuye a construir la diversidad semántica del concepto. Ella recomienda, que para haya un buen funcionamiento de los conocimientos de los alumnos, el docente debe proponer situaciones problemas en los que intervengan al menos dos cuadros.

La experimentación permitió llevar la propuesta al aula y permitió ver la posibilidad del abordaje de la misma con estudiantes que ya tienen conocimientos de cálculo diferencial.

Los muchachos aún no han llevado el curso de cálculo integral, por lo que la experimentación fue muy interesante, en el sentido de que hablamos del cálculo de áreas bajo una curva, sin hablar de sumas de Riemann. El uso de la tecnología facilitó el abordaje de la propuesta.

Los estudiantes manifestaron sentirse a gusto con lo trabajado en esta actividad.

Al aplicar la probabilidad al cálculo de áreas, mostramos una potencialidad de esta disciplina, ajena al azar: La de brindar un modelo para resolver problemas. De esta manera, el cálculo de

probabilidades no queda sólo ahí, si no que más bien se le da un uso práctico en la resolución de problemas.

## Referencias

- Brousseau, G. (1986). Fundamentos y Métodos de la Didáctica de las Matemáticas. Traducción al castellano del artículo “*Fondements et méthodes de la didactiques des mathématiques*” publicado en la revista *Recherches en Didactique des Mathématiques*, 7(2):33-115, y realizada por Julia Centeno, Begoña Melendo y Jesús MurilloCarvajal.
- Douady, R. (1984). *Relación enseñanza –aprendizaje*, Dialéctica instrumento – objeto, Juego de marcos. Cuadernos de Didáctica de las Matemáticas N° 3, IREM de Paris 7.
- Escuela de Matemática. (2017). *Modelo pedagógico de la Carrera Enseñanza de la Matemática con Entornos Tecnológicos*. Disponible en Escuela de Matemática del Tecnológico de Costa Rica.
- Sanabria, G; Núñez, F. (2016). *Probabilidad: Un modelo para resolver diversos problemas. Investigación en Educación Estadística y Probabilística*. Benemérita Universidad Autónoma de Puebla, Dirección de Fomento Editorial, Puebla, México.
- Núñez, F; Sanabria, G. (2017). *Un ejercicio sencillo de estadística descriptiva abordado desde un enfoque probabilístico*. Memorias del X CIEMAC. Cartago, Costa Rica.
- Vallecillos, A; Batanero, C (1997). Análisis del aprendizaje de conceptos clave en el contraste de hipótesis estadísticas mediante el estudio de casos. *Recherches en Didactique des Mathématiques*, 17(1), 29-48. 1997.





# 15

---

## **“Aprender haciendo”: Aplicación de tecnología en la enseñanza de los cuatro componentes de series de tiempo**

---

Carlos Medina Tello

cmedinatello@yahoo.com.mx



### **Resumen**

En la carrera de Ingeniería Industrial, una de las ofertadas en el Instituto Tecnológico de Zitácuaro, perteneciente al sistema Tecnológico Nacional de México, existe un especial interés por parte del profesorado que imparte la materia de Estadística Inferencial II, en que se implemente en medida de lo posible la tecnología con la que se dispone en el aula y el plantel en general, como el uso de dispositivos electrónicos (Laptop, Tablet, Smartphone) y la aplicación de Software especializado (Excel) en estos para la enseñanza de los temas propios de esta, tal es el caso del tema 2; Componentes de las series de tiempo, cuyo análisis se utiliza para detectar patrones de cambio o permanencia en la información estadística en intervalos o periodos regulares.



Siendo así, se se pretende dejar atrás las antiguas estrategias como la resolución de largos ejercicios a lápiz y papel, en donde los estudiantes deben estar utilizando una calculadora para realizar operaciones y hacer múltiples cálculos que requieren de mucho tiempo tanto para procesarlos como para irlos plasmando en el papel, además de crear tablas con una gran cantidad de datos ordenados para poder llegar al resultado esperado, obteniendo la información necesaria de las observaciones de los datos y su comportamiento. Agilizando de esta manera el procedimiento que se necesita para obtener los cuatro componentes de las series de tiempo, los cuales son: 1. Tendencia secular o variación secular 2. Fluctuación cíclica o variación cíclica 3. Variación estacional 4. Variación irregular; y haciendo a la vez más interesante para los alumnos el aprendizaje de este tema.

En el presente trabajo se muestra el estudio de los datos mediante cada uno de los componentes ya establecidos, en los que se proyectan patrones que siguen los datos en un intervalo de tiempo para obtener una estimación para el futuro, ya que el análisis de series de tiempo nos ayuda a aclarar la incertidumbre asociada con los acontecimientos futuros, basándonos en el uso de Excel y algunos complementos de este.

**Palabras clave:** Ingeniería Industrial, Estadística Inferencial, Tecnología, Software.

### **Abstract**

In the career of Industrial Engineering, one of the courses offered at the Technological Institute of Zitácuaro, belonging to the National Technological System of Mexico, there is a special interest on the part of the teaching staff who teach the subject of Inferential Statistics II, in which it is implemented as possible the technology with which it is available in the classroom and the campus in general, such as the use of electronic devices (Laptop, Tablet, Smartphone) and the application of specialized Software (Excel) in these for the teaching of the proper subjects of this, such is the case of topic 2; Components of the time series, whose analysis is used to detect patterns of change or permanence in the statistical information in intervals or regular periods.

This being so, it is intended to leave behind the old strategies such as the resolution of long exercises in pencil and paper, where students must be using a calculator to perform operations and make multiple calculations that require a lot of time both to process them and to shape them. on paper, in addition to creating tables with a large amount of data ordered to reach the expected result, obtaining the necessary information from the observations of the data and its behavior. Thus speeding up the procedure that is needed to obtain the four components of the time series, which are: 1. Secular trend or secular variation 2. Cyclical fluctuation or cyclical variation 3. Seasonal variation 4. Irregular variation; and making the learning of this topic more interesting for the students.

In the present work we show the study of the data through each one of the already established components, in which we project patterns that follow the data in a time interval to obtain an estimate for the future, since the analysis of series of Time helps us to clarify the uncertainty associated with future events, based on the use of Excel and some of its complements.

**Keywords:** Industrial Engineering, Inferential Statistics, Technology, Software.

## 15.1 Introducción

El papel que desarrolla la Estadística en áreas como la Ingeniería hoy en día es de suma importancia, principalmente porque al analizar datos recopilados en experimentos de cualquier tipo, se observa en la mayoría de las ocasiones que dichos datos están sujetos a situaciones que crean un tipo de incertidumbre o duda. La persona que realice dicho procedimiento debe tener la capacidad de tomar decisiones respecto de su objeto de análisis basándose en esos datos, para lo cual debe dotarse de herramientas adecuadas y un proceso bien elaborado que le permita actuar con seguridad y su conclusión sea confiable.

Esto, aunado a la importancia de la correcta aplicación de las TIC's en la educación ha generado una preocupación por implementar nuevos métodos de enseñanza que puedan ser más prácticos tanto para el profesor como para los estudiantes, aunque hay ciertas limitantes tales como el

presupuesto insuficiente que no alcanza para comprar licencias de un software estadístico profesional y una orientación asistencial para que sea empleado adecuadamente y sea útil en el perfil de egreso de los alumnos de Ingeniería Industrial.

En la asignatura de Estadística Inferencial II, el tema dos series de tiempo introduce al estudiante en los conceptos básicos de los modelos clásicos de series de tiempo, análisis de tendencias, análisis de variaciones cíclicas, medición de variaciones estacionales, aplicación de ajustes estacionales, pronósticos basados en factores de tendencia y estacionales. Para determinar el mejor pronóstico de la demanda de un bien con base en el análisis de la situación real, y los recursos requeridos para ello.

El análisis de series de tiempo es un método cuantitativo que utilizamos para determinar patrones de comportamiento en los datos recolectados a través del tiempo, este se utiliza para detectar patrones de cambio o permanencia en la información estadística en intervalos o periodos regulares. (Levin y Rubin, 2010)

En este trabajo se presenta una investigación destinada a amostrar una innovación en el proceso de enseñanza y aprendizaje de los cuatro componentes de las series de tiempo. La propuesta se ha aplicado en los grupos D y F de cuarto semestre en la materia de estadística inferencial II perteneciente a la carrera de Ing. Industrial del Instituto Tecnológico de Zitácuaro, ya que se considera necesario un cambio en los métodos de enseñanza, optando por otros más modernos. Mostrando como justificación, el fundamento teórico de nuestra proposición y el modo en el que la hemos desarrollado, aplicado y evaluado con los grupos piloto.

## **15.2 Marco Contextual**

La secuencia de enseñanza que está en etapa piloto implementada y evaluada en el semestre Enero-Junio del 2018, con grupos de estudiantes del 4to semestre D y F de la Carrera de Ingeniería Industrial del ITZ, y se han rescatado experiencias resultados de aprendizaje similares en los cursos anteriores. El aprendizaje de los estudiantes fue evaluado desde la visión de los indicadores

mediante datos distintos muy diferentes, buscando una coincidencia en los resultados de todos ellos. En concreto, se utilizó una tabla de datos similares y aleatorios como examen de prueba final individual escrita, coherente con la dinámica de clase y los objetivos de enseñanza de acuerdo al temario, como seguidamente se explicará.

En este artículo presentamos los resultados obtenidos con los grupos antes mencionados. Todos los estudiantes tenían un perfil similar, porque: a) el mismo método de enseñanza para la resolución de problemas del tema de series de tiempo con ayuda del programa Excel fue aplicado en ambos grupos; b) todos los alumnos pertenecientes a estos grupos han cursado y culminado anteriormente la materia de Estadística Inferencial I.

Los estudiantes recibieron la capacitación del método de enseñanza a cargo del profesor que contaba con amplia experiencia docente, según el progreso habitual: presentación formal de conceptos y propiedades, ejemplos, ejercicios de aplicación más o menos inmediata y uso de computadora para la automatización de cálculos.

Los estudiantes que siguieron el protocolo de enseñanza contaban con tecnologías que permitían alcanzar el objetivo del curso. Cada estudiante trabajaba de manera individual en el programa de actividades que se ha descrito en los apartados anteriores, proponía ideas para la resolución de los distintos problemas y tomaba nota del procedimiento desarrollado. Como ayuda para el reforzamiento del aprendizaje, se proporcionaron diversas herramientas externas que podían ser consultadas fuera del tiempo de clase, como videos tutoriales y enlaces aclarativos del tema.

Cada clase el profesor se daba a la tarea de elegir a un estudiante para que mostrara junto con él los pasos que se debían ir siguiendo en la resolución del problema, los alumnos expresaban la dominación de la técnica, la realización de los cálculos y la obtención de resultados ante el profesor y los estudiantes. Una vez finalizada la secuencia didáctica, cada alumno tenía la responsabilidad de entregar su trabajo hecho en clase para la evaluación del mismo. La implementación en el aula fue realizada por el Ing. Carlos Medina Tello, con la ayuda de los estudiantes más destacados y quienes desde un inicio dominaron el tema.

### 15.3 Antecedentes

La docencia que está centrada en el sujeto que aprende, tiene su origen en el siglo XVII, y es hasta principios del siglo XX cuando se plantea la idea de que el currículo y la enseñanza deben definitivamente centrarse en las necesidades, intereses y experiencias de los alumnos. Esta postura educativa, como ha sido señalado por diversos autores, estuvo representada por la nueva escuela europea y la educación progresista y democrática de John Dewey (Argudin Y, 2010; Díaz Barriga Arceo F, 2006; Tobón, T S, 2005).

“El uso de herramientas tecnológicas en los procesos de enseñanza y de aprendizaje condujo a pensar en la necesidad de replantear o construir nuevos marcos teóricos que pudieran explicar la relación entre el sujeto, el objeto de conocimiento y las herramientas utilizadas para acercarnos a ese objeto” (Vargas & Guzmán, 2012, pág. 90-91)

Es sabido que en la actualidad el uso de la tecnología ha aumentado de una manera muy significativa, y que se ha convertido en parte de nuestro día a día pues se aplica en la mayoría de las actividades que realizamos, por lo que es conveniente implementarlo también en el ámbito educativo, teniendo una especial preocupación por el área de ciencias básicas del Instituto Tecnológico de Zitácuaro y en particular en la estadística para que sea de ayuda en mejorar la comprensión de la misma, siendo aplicado en este proyecto para el tema los componentes de las series de tiempo.

Para lo cual es necesario que el docente maneje algún software con destreza y sea conocedor de las funciones del mismo, proponiendo una serie de actividades en Excel y complementos del mismo tales como son el Análisis de datos y Statgraphics para la enseñanza de la materia de Estadística Inferencial II.

Para la obtención de las fórmulas y ecuaciones utilizadas nos basamos en el libro “*Estadística para la administración y economía*” (Levin, 2010) de donde definimos principalmente los cuatro componentes de las series de tiempo, los cuales son los siguientes:

La tendencia secular o variación secular, es el primer componente de una serie de tiempo. En este el valor de la variable tiende a aumentar o disminuir en un periodo muy largo. El segundo tipo de variación observado en una serie de tiempo es la fluctuación cíclica, los movimientos cíclicos no siguen ningún patrón regular, sino que se mueven de manera un tanto impredecible. El ejemplo más común de fluctuación cíclica es el ciclo económico. El tercer tipo es la variación estacional, este tipo de variación implica patrones de cambio en el lapso de un año que tienden a repetirse anualmente. La variación irregular es el cuarto tipo de cambio que ocurre en el análisis de las series de tiempo, se encargan de describir los movimientos que existen cuando el valor de una variable puede ser completamente impredecible cambiando de manera aleatoria.

Por lo general en las series de tiempo se miden las variables independientes de tiempo en términos de semanas, meses o años, sin embargo para evitar dilemas es útil aplicar el proceso de codificación, con el que podemos convertir estas medidas tradicionales de tiempo a una forma que simplifica los cálculos.

Es importante recalcar que muchas tendencias pueden ser lineales, por lo que es necesario el uso de la ecuación de la recta, teniendo en cuenta las fórmulas:

<b>Ecuación para una línea recta</b>	
$Y = a + bX$ <div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>Variable dependiente</p> <p>Ordenada Y</p> </div> <div style="text-align: center;"> <p>Variable independiente</p> <p>Pendiente de la recta</p> </div> </div>	[12-1]

Figura 1.

<b>Ordenada Y de la recta de tendencia para valores de tiempo codificados</b>	
$a = \bar{Y}$	[15-4]

Figura 2.

<b>Pendiente de la línea de tendencia para valores de tiempo codificados</b>	
$b = \frac{\sum xY}{\sum x^2}$	[15-3]

Figura 3.

#### 15.4 Desarrollo de la secuencia

En esta sección se mostrara la interpretación de la captura de datos, manejo de fórmulas y salida de resultado por computadora (utilizando Excel) para que ser más eficiente a la hora de hacer los cálculos. Los cálculos de los factores se resumirán a continuación.

##### **1. Lo primero es capturar el problema que se realizara (15.7 del libro Estadística para la administración, Levin, 2010)**

Textualmente dice “Una compañía se especializa en la producción de equipo de recreación, para pronosticar las ventas con base a sus patrones de ventas históricas” (Ver Tabla 1).

##### **2. Posteriormente se obtienen todos los cálculos que se indican en el problema.**

En un problema de que se requiere sacar los cuatro componentes, la ocupación muestra las ventas históricas por trimestre durante cinco años. Necesitamos calcular el total móvil, la media móvil, media móvil centrada y el porcentaje respecto a la media móvil (Ver Tabla 2).

Tabla 1. Datos capturados e interpretados para la secuencia didáctica.

año	TRIMESTRE	OCUPACION
2005	I	16
	II	21
	III	9
	IV	18
2006	I	15
	II	20
	III	10
	IV	18
2007	I	17
	II	24
	III	13
	IV	22
2008	I	17
	II	25
	III	11
	IV	21
2009	I	18
	II	26
	III	14
	IV	25

Tabla 2. Cálculos de total móvil.



			PASO 1	PASO 2	PASO 3	PASO 4
1	2	3	4	5	6	7
AÑO	TRIMESTRE	OCUPACION	TOTAL MOVIL	MEDIA MOVIL	MEDIA MOVIL CENTRADA	% RESPECTO MEDIA MOVIL
2005	I	16				
	II	21				
	III	9	64	16	15.875	56.7
	IV	18	63	15.75	15.625	115.2
2006	I	15	62	15.5	15.625	96.0
	II	20	63	15.75	15.750	127.0
	III	10	63	15.75	16.000	62.5
	IV	18	65	16.25	16.750	107.5
2007	I	17	69	17.25	17.625	96.5
	II	24	72	18	18.500	129.7
	III	13	76	19	19.000	68.4
	IV	22	76	19	19.125	115.0
2008	I	17	77	19.25	19.000	89.5
	II	25	75	18.75	18.625	134.2
	III	11	74	18.5	18.625	59.1
	IV	21	75	18.75	18.875	111.3
2009	I	18	76	19	19.375	92.9
	II	26	79	19.75	20.250	128.4
	III	14	83	20.75		
	IV	25				

### 3. Se procede a calcular la pendiente, la ordenada y el pronóstico.

Primeramente se codifican los datos para después proceder a calcular lo que se necesita. Primero calcular la pendiente de los datos se debe sacar la sumatoria de la multiplicación de  $\sum xy$  dividirla entre la sumatoria de  $\sum x^2$ . Para la ordenada es la media de los valores de  $y$ . Por último el pronóstico es calculado mediante la fórmula de la pendiente  $\hat{Y} = a + bx$ .

Tabla 3. Cálculos para el pronóstico de ventas

PENDIENTE
0.297744361
ORDENADA
18
PRONOSTICO
25.44360902
22.31729323
21.72180451

#### 4. Calculo de índice estacional.

Se realizara una nueva tabla en donde el porcentaje de la media móvil se acomodara de manera vertical para la mejor interpretación de los datos. No se tomaran en cuenta los primeros dos y los últimos dos datos de estos porcentajes y se ordenaran de acuerdo a su trimestre y año. Se eliminara el porcentaje más alto y más bajo de los trimestres de todos los años quedándonos con 2 resultados. El primer cálculo que realizaremos será la media modificada de cada trimestre.

#### 15.5 Conclusiones.

El uso competente de las fórmulas relacionadas con series de tiempo es problemático para una proporción significativa de estudiantes ITZ, incluso tras recibir la secuencia didáctica, los estudiantes tienen dificultades para comprender la secuencia de las formulas a la hora de implementarlas, asumen de manera irreflexiva dichas formulas y se confunden, se requieren que al dejar el lápiz y el papel asuman una actitud más constructiva para su aprendizaje.

## Referencias

- Levin, R.I. y Rubin, D. S. (2010). *Estadística para administración y economía*. México: Pearson Educación.
- Alonso T., Javier. R, Abraham, C., Medina, C. (2017). “*Aprender Haciendo*”: *Modelos Predictivos en la enseñanza y aprendizaje de Estadística*. 05 de Marzo de 2018, de Manuscript Central Sitio web: <http://www.innovativejournal.in/index.php/sslej/article/view/1459>
- Sáez, A. J. (2012). *Apuntes de estadística para ingenieros*. Departamento de Estadística e Investigación Operativa Universidad de Jaén España: Creative Commons.
- Ronald, W., Raymond, M., Sharon, M. y Keying, Y. (2012). *Probabilidad y estadística para ingeniería y ciencias*. Novena edición. México: Pearson Educación.
- Vargas, V., y Guzmán, J. (2012). *Valor pragmático y epistémico de técnicas en la resolución de problemas verbales algebraicos en ambiente de hoja de cálculo*. Enseñanza de las ciencias, 89-107.

TENDENCIA RECIENTE EN LA EDUCACIÓN PROBABILÍSTICA Y ESTADÍSTICA

**Coordinado por José Dionicio Zacarías Flores.**

**Se terminó de imprimir en noviembre de 2019**

**en los talleres de A.L. Digital**

**El cuidado de la edición y la producción editorial son  
de Gladys Denisse Salgado Suárez.**

**El tiraje es de 1000 ejemplares.**

Este trabajo es la aportación de un grupo de investigadores nacionales e internacionales interesados por estudiar la problemática existente en el proceso de enseñanza y aprendizaje de las áreas de probabilidad y estadística. En esta edición se presentan estudios empíricos, innovaciones didácticas, recursos didácticos mediante el uso de tecnología digital, así como las principales ideas extraídas en una mesa de discusión que abordó, el tema principal del libro. “Tendencias recientes en la investigación de la enseñanza de la Probabilidad y la Estadística”.

---

