

Capítulo de Libro Interactivo para la Enseñanza de Tablas Estadísticas

Andrés Felipe Gómez Acevedo y Juan Camilo Rojas Velasco

Trabajo de Grado para optar al título de Licenciado en Matemáticas

Director

Tulia Esther Rivera Florez

Magister en Estadística

Universidad Industrial de Santander

Facultad de Ciencias

Escuela de Matemáticas

Licenciatura en Matemáticas

Bucaramanga

2026

Dedicatoria

A mi familia, especialmente a mis padres Emelina Velasco y Raúl Rojas, quienes me motivaron durante todo mi proceso de formación. Sus consejos y palabras de aliento son las que me han ayudado a perseverar.

A mi mejor amigo Roberth por su apoyo incondicional que me ha permitido entender que la verdadera fuerza nace del compromiso para responder a las expectativas de las personas a las que amamos.

Juan Camilo Rojas Velasco

Dedico este trabajo a mis familiares, en especial a mis padres Leonardo y Stella quienes con su apoyo incondicional y su confianza me acompañaron en cada etapa de este proceso; su esfuerzo y su ejemplo son la base de todo lo que he logrado.

A Miller, mi mejor amigo, por creer siempre en mí, incluso en los momentos en que yo mismo dudaba, y por estar presente de manera incondicional a lo largo de este camino.

A Johan, por haber sido una parte importante de mi formación académica y por aportar de distintas maneras, al profesional en el que me he convertido.

A Juan Camilo, mi compañero de tesis y compañero de las mil y un batallas, con quien compartí no sólo el desarrollo de este trabajo, sino también los momentos de duda, esfuerzo y satisfacción que solo quien vive un proceso así puede entender.

Y a todas las personas que confiaron en mí y me brindaron su apoyo cuando más lo necesité: sin ustedes, este camino habría sido mucho más difícil de recorrer.

Andrés Felipe Gómez Acevedo

Agradecimientos

En primer lugar agradecemos a Dios por permitirme culminar este primer título universitario y también por todos los beneficios recibidos durante este camino.

Agradecemos a la Universidad Industrial de Santander, a la Facultad de Ciencias y a la Escuela de Matemáticas, por brindarnos la formación académica y humana que hizo posible la culminación de este trabajo de grado.

A nuestra directora de tesis, Tulia Esther Rivera Florez, por depositar su confianza en este proyecto desde el inicio, por su acompañamiento constante, su paciencia y sus orientaciones oportunas, que fueron fundamentales para darle forma, sentido y rigor a este trabajo. Sin su guía, este proceso habría sido considerablemente más difícil de recorrer.

Finalmente, agradezco a todas las personas que, de forma directa o indirecta, hicieron parte de este proceso y contribuyeron a que este trabajo llegara a buen término.

Tabla de contenido

Introducción.....	12
1. Descripción del problema.....	13
2. Justificación.....	17
3. Antecedentes.....	21
4. Objetivos.....	28
4.1 Objetivo general.....	28
4.2 Objetivos específicos.....	28
5. Marco Teórico.....	29
5.1 Elementos de educación Matemática.....	29
5.2 Elementos de didáctica de la estadística.....	31
5.2.1 Tablas en estadística.....	31
5.2.2 Alfabetización estadística y transnumeración.....	32
5.2.3 Juicios de asociación.....	34
5.3 Recurso Didáctico Digital.....	35
5.3.1 Definición y Características de un Libro Interactivo Digital.....	35
5.3.2 El Libro Interactivo Digital (LID) como mediador de la interacción autónoma del estudiante en situaciones a-didácticas.....	36
5.4 Elementos teóricos de tablas en estadística.....	36
5.4.1 Tablas de frecuencia.....	37
5.4.1.1 Frecuencia absoluta.....	37

5.4.1.2 Frecuencia relativa.....	38
5.4.1.3 Frecuencia absoluta acumulada.....	38
5.4.1.4 Frecuencia relativa acumulada.....	39
5.4.2 Representación y organización de datos.....	39
5.4.3 Elementos Teóricos de Tablas de Contingencia.....	40
5.4.3.1 Porcentajes o frecuencias relativas al total (distribución conjunta).....	41
5.4.3.2 Porcentaje o frecuencias relativas por fila (distribución condicional por género).....	42
5.4.3.3 Porcentaje o frecuencias relativas por columna (distribución condicional por uso de redes).....	43
5.4.4 Asociación e independencia entre variables: el puente hacia la inferencia.....	44
5.4.5 La inferencia estadística como situación de validación.....	45
6. Metodología.....	48
7. Resultados.....	50
7.1 Diseño didáctico de la secuencia por capítulo.....	50
7.2 Enfoque didáctico del recurso.....	56
7.3 Estructura y organización del contenido.....	60
7.3.1 Sección I — Tablas de frecuencia.....	62
7.3.2 Sección II — Tablas de contingencia.....	68
7.3.3 Sección III — Prueba Chi-cuadrado.....	73
7.4 Revisión de estilo y contenido por panel de expertos sobre el capítulo de libro diseñado.....	83
7.4.1 Análisis de resultados.....	85

7.4.2 Fortalezas identificadas.....	91
7.4.3 Debilidades y ajustes implementados.....	91
8. Conclusiones.....	93
Referencias Bibliográficas.....	98

Lista de tablas

Tabla 1. Resumen de libros.....	21
Tabla 2. Tiempo de exposición a pantallas en estudiante.....	40
Tabla 3. Tabla de contingencia.....	41
Tabla 4. Tabla de contingencia con porcentajes sobre el total.....	42
Tabla 5. Tabla de contingencia con porcentajes por fila.....	43
Tabla 6. Tabla de contingencia con porcentajes por columna.....	44
Tabla 7. Diseño didáctico del LID según las fases de la Teoría de Situaciones Didácticas.....	52
Tabla 8. Estructura general del Libro Interactivo Digital (LID).....	60
Tabla 9. Instrumento de valoración para el contenido del capítulo diseñado.....	84
Tabla 10. Resultados de la validación por panel de expertos, por dimensión e ítem (n=6, escala 1-5; M= media; DE= desviación estándar).....	86

Lista de figuras

Figura 1. Ejemplo.....	22
Figura 2. Distribución X^2	46
Figura 3. Introducción construcción guiada sección I.....	63
Figura 5. Ejemplos dinámicos.....	65
Figura 6. Problemas.....	66
Figura 7. Situaciones.....	67
Figura 8. Introducción Capítulo II.....	69
Figura 9. Formas de leer una tabla de contingencia.....	70
Figura 10. Problemas.....	71
Figura Retos.....	72
Figura 12. Situaciones problema.....	73
Figura 13. Introducción.....	74
Figura 14. Construyendo lo esperado.....	75
Figura 15. Ejemplo guiado.....	76
Figura 16. Construyendo X^2	77
Figura 17. Interpretación de X^2	78
Figura 18. Introducción grados de libertad y valor crítico.....	79
Figura 19. Ejemplo guiado.....	80
Figura 20. Asociación vs causalidad.....	81
Figura 21. Aplicación libre.....	83

Lista de Apéndices

Los apéndices están disponibles en el repositorio Institucional

Apéndice A. Instrumento de validación - preguntas realizadas al panel

Apéndice B. Manual de Usuario

Resumen

Título: Capítulo de Libro Interactivo para la Enseñanza de Tablas Estadísticas

Autores: Andrés Felipe Gómez Acevedo y Juan Camilo Rojas Velasco

Palabras clave: Libro Interactivo Digital (LID), tablas de contingencia, educación estadística, Teoría de situaciones didácticas.

Descripción: La enseñanza de la estadística requiere superar el tratamiento centrado en la mecanización de procedimientos y favorecer procesos de interpretación y análisis crítico de los datos. El presente trabajo tuvo como objetivo diseñar un capítulo de Libro Interactivo Digital (LID) orientado al fortalecimiento de los procesos de lectura de tablas de contingencia en estudiantes de la Licenciatura en Matemáticas de la Universidad Industrial de Santander. La propuesta se enmarca en la modalidad de producción de material didáctico y se articuló la Teoría de Situaciones Didácticas y en los niveles progresivos de comprensión de datos propuestos por Curcio y Aoyama. El desarrollo del LID se realizó por medio de 5 fases: consulta bibliográfica, diseño y programación del prototipo, elaboración del instrumento de evaluación, revisión del capítulo por un panel de evaluadores y ajustes del recurso. El LID integra tres secciones orientadas al trabajo con tablas de frecuencia, tablas de contingencia y asociación entre variables mediante la prueba de independencia chi-cuadrado, incorporando situaciones de exploración, formulación, validación e institucionalización. Para valorar el contenido y estilo del capítulo se aplicó un instrumento a un panel de seis evaluadores. El panel de expertos valoraron favorablemente el recurso, con un promedio general de 4,56 sobre 5 y promedios por dimensión entre 4,46 y 4,67. Las observaciones permitieron realizar ajustes relacionados con la secuenciación de contenidos, la presentación de fórmulas, la navegación y la interacción del tutor virtual. El LID demostró pertinencia técnica y teórica; si bien la valoración se limitó a la validación de contenido por juicio de expertos y no permiten establecer efectos sobre el aprendizaje de estudiantes en un contexto de aula

* Trabajo de Grado

** Facultad de Ciencias. Escuela de Licenciatura en Matemáticas. Director: Tulia Esther Rivera Flórez. Magister en Estadística.

Abstract

Title: Interactive Book Chapter for Teaching Statistical Tables

Authors: Andrés Felipe Gómez Acevedo & Juan Camilo Rojas Velasco

Keywords: Interactive Digital Book (IDB), contingency tables, statistics education, Theory of Teaching Situations.

Description: Teaching statistics requires moving beyond an approach focused on the mechanization of procedures and instead promoting processes of interpretation and critical analysis of data. The objective of this study was to design a chapter for a Digital Interactive Book (DIB) aimed at strengthening students' ability to interpret contingency tables among students in the Bachelor of Science in Mathematics program at the Industrial University of Santander. The proposal falls within the framework of instructional material development and was based on the Theory of Instructional Situations and the progressive levels of data comprehension proposed by Curcio and Aoyama. The DIB was developed in five phases: literature review, prototype design and programming, development of the assessment instrument, review of the chapter by a panel of evaluators, and refinement of the resource. The DIB comprises three sections focused on working with frequency tables, contingency tables, and associations between variables using the chi-square test of independence, incorporating situations of exploration, formulation, validation, and institutionalization. To evaluate the chapter's content and style, an assessment instrument was administered to a panel of six evaluators. The panel of experts rated the resource favorably, with an overall average of 4.56 out of 5 and averages per dimension ranging from 4.46 to 4.67. Their comments allowed for adjustments related to the sequencing of content.

* Degree Work

** College of Sciences. Department of Mathematics. Advisor: Tulia Esther Rivera Flórez. Master's in Statistics.

Introducción

En la sociedad actual la capacidad de interpretar y analizar críticamente la información se ha convertido en una competencia esencial para la toma de decisiones informadas, por ello la enseñanza de la estadística resulta imprescindible para lograr una participación ciudadana activa. Sin embargo, tradicionalmente se ha visto reducida a la mecanización de algoritmos, dejando a un lado el desarrollo del sentido estadístico y la capacidad de realizar inferencias sobre la realidad, afectando la capacidad de dar juicios válidos a partir de los datos.

Debido a esta desvinculación entre la teoría y la práctica, el presente trabajo propone el diseño y desarrollo de un Libro Interactivo Digital (LID) enfocado en el uso e interpretación de Tablas de frecuencia y de contingencia. Fundamentado bajo el marco de la Teoría de las Situaciones Didácticas de Brousseau y enfocado mayormente en las situaciones a-didácticas que permiten al estudiante asumir la responsabilidad de su aprendizaje al interactuar con el LID para minimizar la intervención directa del docente y con ello mitigar una problemática muy común en nuestro medio como es la baja intensidad horaria asignada a los cursos de Estadística. El LID busca que el alumno interactúe directamente con el medio tecnológico para resolver problemas de transnumeración y asociación relacionadas con resúmenes de datos tabulares, mientras transita por los diferentes niveles de comprensión de Curcio y Aoyama que si bien fueron formulados para lectura de gráficos se pueden extender a este contexto de trabajo; se avanza hasta un nivel avanzado del uso de estos recursos el cual implica el uso de estadística inferencial, en particular la prueba Chi-cuadrado, permitiendo al estudiante comprender la complejidad de las relaciones bivariadas y su funcionalidad práctica en una sociedad que cada vez más adopta la toma de decisiones basada en datos como práctica común.

1. Descripción del problema

Por un lado, los desafíos en la enseñanza de la estadística trascienden la transmisión de contenidos puesto que implican abordar la ansiedad estudiantil y la complejidad de la abstracción matemática. Mientras que la didáctica en temas relacionados con los gráficos estadísticos ha avanzado, la interpretación de tablas permanece estancada en los métodos tradicionales donde se ven los datos de manera estática y en contextos alejados de la realidad lo cual es muy evidente en libros de texto tradicionalmente usados para la enseñanza de la estadística a todo nivel. Por lo anterior, surge la necesidad de aprovechar el desarrollo tecnológico para desarrollar recursos que logren combinar el razonamiento estadístico y el matemático bajo ambientes que favorezcan un aprendizaje activo. De manera empírica, varios estudios (Gutiérrez, 2021; Vásquez, 2014) sugieren que los recursos dinámicos son superiores a los textos estáticos al fomentar un aprendizaje real y mejorar el rendimiento académico.

En concordancia con lo anterior, Vásquez (2014) logra mostrar la influencia del libro interactivo multimedia (LIM) en la comprensión lectora de los alumnos a partir de los resultados obtenidos en la prueba final los cuales clasificaron como 70% en proceso y 30% de los estudiantes en el nivel destacado, con lo cual concluyeron que se obtuvo una mejora significativa con relación a los resultados de la prueba inicial. En concordancia, Gutiérrez (2021) presenta una propuesta de enseñanza a distancia flexible e integrada con la presencial por medio de un libro electrónico interactivo donde se presentan las simetrías del plano a través de una metodología de descubrimiento guiado de acuerdo con los niveles de Van Hiele. Se destaca que las TIC permiten a los docentes adaptar su enseñanza de acuerdo con las necesidades específicas de los estudiantes.

Por otra parte, aunque en los últimos años se ha enfatizado cada vez más en la enseñanza de la estadística, y poco a poco ha tomado un rol más importante debido a su valor para la toma de decisiones en una sociedad basada en datos aún no se ha llegado a consolidar como una disciplina autónoma que usa las matemáticas como lenguaje, esto se ve reflejado en las pruebas estandarizadas PISA 2022 donde se encuentran inmersas aplicaciones de la estadística en el apartado de Matemáticas. Las pruebas mencionadas anteriormente muestran bajos niveles de desempeño adquiridos por los estudiantes en matemáticas pero al no discriminar el componente de Probabilidad y Estadística es difícil identificar problemáticas particulares.

Complementando lo anterior, según el informe del rendimiento en la prueba PISA reportado por el Instituto Colombiano para la Evaluación de la Educación (ICFES, 2024), el 71% de los estudiantes colombianos se ubicaron debajo del nivel 2 de competencias en matemáticas y tan solo el 9,3% alcanzó niveles superiores; en contraste, el 46% de los estudiantes en los países de la Organización para la Cooperación y el Desarrollo Económico (OCDE) superan el nivel 2. Más aún, el panorama en niveles superiores, correspondiente a los niveles 5 y 6, en Colombia solo el 1% se ubican en estos niveles mientras que los estudiantes pertenecientes a la OCDE, ha mantenido cerca del 10% de sus estudiantes en niveles superiores.

Si bien es cierto que los resultados reportados por las pruebas PISA se presentan de manera general en el área de matemáticas, investigaciones como las de Batanero (2001), Méndez (2023) confirman que este bajo desempeño se traduce directamente en dificultades significativas para el desarrollo del pensamiento estadístico evidenciándose con mayor frecuencia en la educación superior y por consiguiente falencias en el desarrollo de habilidades relacionadas con el análisis de datos para el ámbito profesional.

Ante este panorama, una de las principales dificultades reside en la realización de análisis descriptivo de datos dentro de la cual se ubica la lectura y comprensión de tablas adicional a lo relacionado con el uso de representaciones gráficas; lo anterior, contrasta la suposición que la interpretación de estas representaciones es una tarea sencilla y no una actividad cognitiva compleja (Batanero, 2001).

Particularmente, la falta de preparación en el manejo de recursos estadísticos se torna problemático especialmente en la interpretación de tablas que son un aspecto clave del pensamiento estadístico, estas son un instrumento eficaz para comunicar, presentar y resumir información pero que requieren de un proceso denominado “transnumeración” que hace relación al proceso de obtener nueva información a partir del paso de un sistema de representación a otro (Arteaga et al., 2011), cuando el estudiante no logra realizar dicho proceso resulta en un avance lento hacia la comprensión de conceptos más elaborados.

Aunado a la complejidad cognitiva que implica el análisis tabular, el rápido crecimiento de recursos tecnológicos y análisis de datos ha llevado a replantear qué se entiende hoy por un análisis estadístico básico, para algunos las tablas de frecuencia simples pueden verse como un aspecto secundario en un análisis moderno y se opta por dar protagonismo al estudio de contextos multivariados y al uso de representaciones interactivas que faciliten la visualización de la información subyacente.

Bajo este contexto las tablas de contingencia emergen como un recurso básico en la enseñanza de la estadística descriptiva ya que estas tablas ofrecen la primera aproximación formal y necesaria para el análisis bivariado; en este sentido, los recursos didácticos deben orientarse hacia la enseñanza del análisis multivariado con el fin de preparar al estudiante para

comprender las interacciones complejas que caracterizan la mayoría de los problemas de datos actuales.

Desde el punto de vista de la Didáctica de la Estadística, el auge del análisis de datos en el contexto académico y social no está en relación directa con el volumen de investigación realizada especialmente en nuestro medio; hay escasez de recursos de apoyo innovadores y el predominio de un enfoque de enseñanza tradicional el cual se reduce a la repetición de ejercicios de libro de texto que suelen ser descontextualizados y ajenos al desarrollo de procesos cognitivos esenciales como la transnumeración (Wild y Pfannkuch, 1999, como se citó en Batanero, 2013). Este enfoque mecánico impide que los estudiantes aborden con éxito la complejidad del análisis multivariado ya que no desarrollan la capacidad de dar sentido a partir del cambio en las diferentes representaciones de los datos.

A la luz del poco énfasis que se hace a la comprensión de tablas y la necesidad de un recurso que fomente el aprendizaje activo, la exploración y la visualización para desarrollar pensamiento estadístico surgen varias inquietudes: ¿Qué recurso didáctico es más eficiente para un proceso de enseñanza de la estadística descriptiva a nivel de educación superior?, ¿qué características debe tener un libro de texto moderno que incorpore el desarrollo tecnológico a disposición?, ¿cómo incorporar habilidades como la exploración, visualización e interacción para un aprendizaje activo y autónomo?, y ¿qué actividades pueden contribuir a mejorar la comprensión del análisis de tablas y al desarrollo de un pensamiento estadístico en los estudiantes de Licenciatura en Matemáticas?

2. Justificación

La enseñanza de la estadística es un componente clave en la formación ciudadana ya que contribuye al desarrollo de competencias necesarias para resolver problemas a partir del análisis de datos (Hoyos, 2024); la estadística no debe considerarse como una rama más de las matemáticas sino como una herramienta para el desarrollo del pensamiento crítico que permite la interpretación de fenómenos en diferentes contextos, además de comprender y evaluar la información presente en diferentes medios de comunicación, contribuyendo a una participación activa de la sociedad.

Sin embargo, se enfrenta a retos persistentes tales como la baja comprensión de contenidos estadísticos aún en uno de los primeros temas que se aborda como son los análisis descriptivos de datos (Batanero, 2001). Esto se ve reflejado en muchos casos en un bajo rendimiento general en matemáticas, lo que enmarca la urgencia de fortalecer el pensamiento estadístico desde etapas tempranas pues estas dificultades afectan tanto la comprensión conceptual como la actitud de los estudiantes hacia la estadística descriptiva (Méndez, 2023).

Una de las principales dificultades identificadas: la comprensión de tablas y gráficos, lo que resulta contrario a la suposición de algunos docentes que su interpretación es una tarea sencilla, la investigación ha demostrado que se trata de una actividad cognitiva compleja que trasciende la simple lectura de valores. Batanero (2001), basándose en los niveles de comprensión propuestos por Curcio, argumenta que los estudiantes presentan un desarrollo limitado en su capacidad para interpretar la información visual y textual, lo que evidencia una habilidad deficiente para traducir información en diferentes representaciones.

Asimismo los análisis de tablas y gráficos constituyen elementos esenciales para la comprensión de datos puesto que el dominio de estas va más allá de una simple habilidad técnica ya que proporciona un potente instrumento para comunicar y resumir información de manera eficiente. De hecho, el proceso de obtener nueva información al cambiar de un sistema de representación a otro (por ejemplo, de un listado de datos a un gráfico) es un modo esencial del razonamiento estadístico denominado "transnumeración" (Arteaga et al., 2011).

Adicionalmente, se ha evidenciado una comprensión deficiente en conceptos fundamentales como las Medidas de Posición Central y de Dispersión debido a que en lugar de una aplicación flexible y significativa, los estudiantes tienden a emplear procedimientos mecánicos o procedimentales sin llegar a desarrollar una visión conceptual profunda (Batanero, 2001).

A pesar del auge de la estadística la investigación en su didáctica, y en particular sobre la comprensión de gráficos y tablas, no ha sido tan prolífera como en otras ramas de las matemáticas. Sumado a ello gran parte de los estudios iniciales fueron realizados desde la psicología cognitiva por autores como Cleveland & McGill (1984), así como Kosslyn (1989), sin embargo, la investigación en contextos escolares reales especialmente en las edades de 11 a 16 años es aún escasa; esta brecha investigativa entre lo cognitivo y el contexto real ha limitado la comprensión así como la identificación de dificultades que enfrentan los estudiantes en el desarrollo de conceptos estadísticos clave, lo que obstaculiza a su vez el desarrollo de una alfabetización estadística crítica igual que al desarrollo de la transnumeración, ambos elementos esenciales para la toma de decisiones basada en datos.

De este modo, resulta fundamental generar nuevos aportes tanto a nivel teórico como aplicado en la enseñanza de la estadística, atendiendo a las necesidades de educación básica como superior. Desde la perspectiva teórica, es necesario seguir profundizando en la descripción de las competencias requeridas, los niveles de comprensión y las concepciones erróneas de los estudiantes (Batanero, 2001). Complementariamente, desde la perspectiva aplicada, se necesita el diseño, experimentación y evaluación de métodos de enseñanza que promuevan la construcción autónoma del conocimiento y el trabajo con proyectos, que han demostrado ser beneficiosos en la comprensión de los conceptos (Batanero, 2013; Hoyos, 2024). De esta manera la creación de recursos didácticos innovadores, como libros de texto interactivos que fomenten la exploración y el autoaprendizaje es también una línea de trabajo prometedora para mejorar la formación estadística (Hoyos, 2024).

Por lo anterior, resulta necesaria la aplicación y desarrollo de recursos innovadores que promuevan un aprendizaje participativo, significativo y contextualizado, de manera que facilite la conexión entre teoría y práctica. Bajo esta consideración, un libro digital interactivo orientado a la enseñanza de Estadística descriptiva resulta ser una propuesta didáctica pertinente si este tipo de recurso integra múltiples lenguajes: textual, visual, audiovisual y manipulativo. Según Toapanta y Yugcha (2023) los recursos digitales estimulan la atención, la exploración y el aprendizaje autónomo, generando experiencias educativas más significativas y motivadoras. Esta propuesta está en concordancia con la evolución de la educación estadística que como subraya Batanero (2019) se ha orientado hacia la integración de proyectos y recursos digitales.

Adicionalmente, un texto interactivo digital reduciría la problemática de la baja intensidad horaria dispuesta a la enseñanza de la estadística tanto en nivel de educación básica

como a nivel de educación superior, pues permitiría un mayor acercamiento al desarrollo de competencias necesarias para enfrentar los desafíos actuales en el manejo y la interpretación de datos sin las limitaciones del enfoque tradicional. En adición, se debe pensar en el diseño de instrumentos y actividades evaluativas que fomenten la interpretación crítica y profunda de la información estadística, aspecto crucial para un aprendizaje sólido y funcional (Pallauta et al., 2023).

En este sentido, el Libro Interactivo Digital (LID) representa una evolución fundamental del material didáctico, ya que transforma la enseñanza de la estadística al integrar teoría, práctica y recursos multimedia, logrando que el proceso de enseñanza vaya mucho más allá del modelo pasivo donde el docente es el centro del aprendizaje, al estar centrado en el estudiante contribuye a la formación integral y al desarrollo de habilidades entorno al aprendizaje autónomo e independiente.

Esta propuesta aborda directamente las dificultades conceptuales documentadas por Batanero (2001), y su uso fomenta el desarrollo de competencias esenciales para el siglo XXI puesto que más allá de habilidades técnicas como el manejo de software y la visualización de datos, la interactividad permite a los estudiantes simular conceptos abstractos, fomentar el pensamiento crítico al cuestionar resultados y explorar cómo las decisiones influyen en la presentación de la información; asimismo al trabajar con datos reales el libro promueve la resolución de problemas a través de la exploración y la formulación de conjeturas, habilidades que se fortalecen con metodologías activas (Toapanta y Yugcha, 2023) mientras promueve de manera inherente la alfabetización digital, siendo esta una habilidad crucial para una participación informada en una sociedad basada en datos.

3. **Antecedentes**

Si bien existen estudios sobre el uso de recursos digitales e interactivos en la enseñanza de las matemáticas, los antecedentes sobre libros interactivos específicamente en estadística son escasos, lo cual evidencia un vacío en la literatura y justifica la pertinencia del presente trabajo. De acuerdo con Hoyos (2024) los libros de texto generales y educativos han tenido una evolución similar, partiendo de la civilización egipcia marcada por los papiros y pergaminos, tuvieron limitaciones para la producción de los mismos por la escasez de materia prima lo cual propició el descubrimiento del papel que revolucionó la creación de libros junto con los parámetros de escritura. Siglos después, el desarrollo de la tecnología abrió el camino para poder pasar del formato físico a diversos modos de almacenamiento que se fueron perfeccionando para mejorar la accesibilidad por medio del internet. Por otra parte, la autora presenta un análisis de diferentes libros de texto que pueden llamarse clásicos por su marcado uso en la enseñanza de cursos de estadística a nivel universitario; estos libros de texto, aunque son actualizados a través del tiempo conservan como característica general y principal recurso didáctico, la presentación de problemas resueltos y una gran cantidad de ejercicios propuestos (Tabla 1).

Tabla 1.

Resumen de libros

Libro	Año de edición	No aproximado de ejercicios y ejemplos resueltos
<i>Teoría y problemas de Probabilidad y Estadística</i> (Mac Graw Hill - primera edición)	1975	760
<i>Teoría y problemas de Probabilidad y Estadística</i> (Mac Graw Hill - cuarta edición)	2009	486
<i>Introducción a la probabilidad y estadística</i> (Mendenhall, Beaver & Beaver – décimo tercera edición)	2024	1.300

En particular, un texto clásico para la educación superior es el libro Jay L. Devore, Estadística para Ingenieros y Ciencias, presenta una estructura que menciona Hoyos (2024) la cual consta de definición, problemas resueltos y apartado de ejercicios propuestos, un recurso que ofrece es la implementación de una sección de resumen con los elementos teóricos y procedimentales del tema, allí se explica el paso a paso para la realización de lo que se quiere construir o bien desarrollar, por ejemplo en el subtítulo Tablas de contingencia mutuas (Figura 1). Cabe aclarar que en este texto aunque en el primer capítulo incluye un título que anuncia Métodos pictóricos y tabulares en estadística descriptiva no se desarrolla nada relacionado con tablas de frecuencias, no obstante en los capítulos finales se presentan las tablas de contingencia como insumo para realizar pruebas de inferencia estadística para evaluar bondad de ajuste e independencia.

Figura 1.

Ejemplo

14.3 Tablas de contingencia mutuas (o bidireccionales)

En las dos secciones previas, se estudiaron problemas inferenciales en los que los datos de cantidades se mostraba en una tabla rectangular de celdas. Cada tabla estaba formada por una fila y un número especificado de columnas, donde las columnas correspondían a categorías en las que la población se había dividido. Ahora se estudian problemas en los que los datos también están formados por cantidades o frecuencias, pero la tabla de información ahora tendrá I filas ($I \geq 2$) y J columnas, por tanto IJ celdas. Hay dos situaciones que por lo general se encuentran y en las que se muestran los datos:

1. Hay I poblaciones de interés, cada una correspondiente a una fila diferente de la tabla, y cada población está dividida en las mismas J categorías. Se toma una muestra de la i -ésima población ($i = 1, \dots, I$) y las cantidades se introducen en las celdas de la i -ésima fila de la tabla. Por ejemplo, los clientes de cada una de las $I = 3$ cadenas de tiendas departamentales podrían disponer de las mismas $J = 5$ categorías de pago: contado, cheque, tarjeta de crédito de la tienda, Visa y MasterCard.
2. Hay una sola población de interés, con cada individuo de la población clasificado con respecto a dos factores diferentes. Hay I categorías asociadas con el primer factor, y J categorías asociadas con el segundo factor. Se toma una sola muestra, y el número de individuos pertenecientes tanto a la categoría i del factor 1 como a la categoría j del factor 2 se introduce en la celda de la fila i , columna j ($i = 1, \dots, I; j = 1, \dots, J$). Como ejemplo, los clientes que hagan una compra podrían clasificarse de acuerdo con el departamento en el que hicieron la compra, con $I = 6$ departamentos, y de acuerdo con la forma de pago, con $J = 5$ como en (1) de líneas antes.

Denótese con n_{ij} el número de individuos de la(s) muestra(s) que caen en la (i, j) -ésima celda (fila i , columna j) de la tabla, es decir, la (i, j) cantidad de celda. La tabla que presente las n_{ij} se denomina **tabla de contingencia mutua**; un prototipo se muestra en la tabla 14.9.

Tabla 14.9 Una tabla de contingencia mutua

	1	2	...	j	...	J
1	n_{11}	n_{12}	...	n_{1j}	...	n_{1J}
2	n_{21}					$\vdots$
$\vdots$	$\vdots$					$\vdots$
i	n_{i1}	...		n_{ij}	...	
$\vdots$	$\vdots$					
I	n_{I1}	...				n_{IJ}

En situaciones del tipo 1, se desea investigar si las proporciones de las diferentes categorías son iguales para todas las poblaciones. La hipótesis nula expresa que las poblaciones son *homogéneas* con respecto a estas categorías. En situaciones del tipo 2, se investiga si las categorías de los dos factores se presentan independientemente una de otra en la población.

14.3 Tablas de contingencia mutuas (o bidireccionales) **589**

$$\begin{aligned}\hat{e}_{ij} &= \text{cantidad esperada estimada en celda } (i, j) = n_i \cdot \frac{n_{.j}}{n} \\ &= \frac{(\text{total de } i\text{-ésima fila})(\text{total de } j\text{-ésima columna})}{n} \quad (14.9)\end{aligned}$$

El estadístico de prueba también tiene la misma forma que en las situaciones del problema previo. El número de grados de libertad proviene de la regla empírica general. En cada fila de la tabla 14.9 hay $J - 1$ cantidades de celdas determinadas libremente (cada tamaño muestral n_i es fijo), de modo que hay un total de $I(J - 1)$ celdas determinadas libremente. Los parámetros $p_1, \dots, p_J$ se estiman, pero como $\sum p_i = 1$, sólo $J - 1$ de éstos son independientes. Por lo tanto, el grado de libertad es $gl = I(J - 1) - (J - 1) = (J - 1)(I - 1)$.

Hipótesis nula: $H_0: p_{1j} = p_{2j} = \dots = p_{ij} \quad j = 1, 2, \dots, J$

Hipótesis alternativa: $H_a: H_0$ no es verdadera

Valor de estadístico de prueba:

$$\chi^2 = \sum_{\text{todas las celdas}} \frac{(\text{observada} - \text{esperada estimada})^2}{\text{esperada estimada}} = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - \hat{e}_{ij})^2}{\hat{e}_{ij}}$$

Región de rechazo: $\chi^2 \geq \chi^2_{\alpha, (J-1)(I-1)}$

La información del valor P se puede obtener como se describe en la sección 14.1. La prueba se puede aplicar con seguridad mientras $\hat{e}_{ij} \geq 5$ para todas las celdas.

Ejemplo 14.13 Una compañía empaca un producto particular en latas de tres tamaños diferentes, cada uno con una línea de producción distinto. La mayor parte de las latas se apegan a especificaciones, pero un ingeniero de control de calidad ha identificado las siguientes razones de no

Tomado de *Probabilidad y Estadística para Ingeniería y Ciencias*, por J. Devore, 2008, Editorial Cengage Learning Editores.

El libro presenta en total alrededor de 1.200 problemas, 200 resueltos y 1.000 propuestos, que se distribuyen en 14 capítulos, de hecho este es uno de los principales recursos que el autor postula para facilitar la comprensión por parte de los estudiantes, el uso de datos reales y la disposición de un Manual de soluciones para el estudiante que desarrolla los ejercicios impares. Como novedad se incluye menos cálculos manuales para propiciar el uso de software y al comprar el texto se gana acceso a un sitio web con applets y el acceso a un sistema para gestión de archivos de clase, tareas y evaluaciones.

Por otro lado, en el libro *“Introduction to the Practice of Statistics”* de Moore, McCabe & Crai (2017). Las tablas de frecuencia y contingencia se presentan como herramientas para la exploración y el análisis de datos a problemas contextualizados, basados principalmente en investigaciones reales. Adicional a lo anterior, se proponen diferentes tipos de representaciones (gráfica y tabular) para responder a preguntas, realizar conjeturas o hipótesis con base en los datos. Si bien es cierto que el libro no cuenta con una sección dedicada para el análisis de tablas de frecuencia y contingencia también lo es que dada su estructura logra salirse del esquema tradicional (definición, ejemplo y ejercicios) así como la mecanización de algoritmos.

En conclusión, los libros de estadística tradicionales ofrecen una amplia cantidad de problemas/ejercicios así como ejemplos variados en concordancia con la sección que se esté estudiando sin embargo, debido a la misma naturaleza de la estadística el análisis de los datos varía dependiendo del contexto de los mismos así como la interpretación y apropiación del recurso acorde a un objetivo de análisis. Es por ello que un libro de estadística no debe ser estático o centrado en la memorización de algoritmos.

3.1 Plataformas y nuevos recursos en estadística

3.1.1 ALEKS – Mac Graw Hill

Esta novedosa plataforma es la opción que la editorial Mac Graw Hill está ofreciendo al público en sustitución de sus libros físicos. Antes de iniciar las sesiones exige que se realice una prueba de evaluación diagnóstica con el fin de poder determinar en cuál de las sesiones el estudiante debe ser situado para su aprendizaje, las preguntas de la prueba son de temas variados que van desde álgebra y pre cálculo hasta temas básicos en estadística. Una vez ubicado el estudiante en la sesión correspondiente a su nivel, el programa arroja de manera automática una definición del tema desde el cual se dará inicio con el fin de presentar la fórmula o procedimiento general para realizar los ejercicios de ese estilo para luego presentar problemas similares (ejercitación) y que el estudiante digite las respuestas.

Por otro lado, cuando el estudiante se equivoca en la realización de la actividad propuesta, la plataforma muestra la retroalimentación a la respuesta incorrecta y si el estudiante vuelve a equivocarse la plataforma muestra el procedimiento correcto para llegar a la solución del problema sin tratar de mostrar un paso a paso de la solución del mismo o guiar al estudiante en el desarrollo del problema.

3.1.2 StatKey

En esta novedosa plataforma se presentan varios temas para trabajar adoptando un enfoque de simulación, el alcance temático cubre desde estadística descriptiva y gráficos hasta pruebas de hipótesis por métodos de aleatorización; puntualmente, en cuanto a lectura de tablas lo más cercano que se presenta en esta plataforma es la opción para trabajo con dos variables

categorías, ya que presenta una tabla de datos donde va acompañado de un gráfico de barras pero se hace más énfasis en la lectura del diagrama de barras, sin embargo si se ignora el diagrama de barras se puede realizar la actividad con la tabla donde se enfocaría en la lectura de la misma.

3.1.3 Colección de applets de Rossman/Chance (2021)

Este repositorio de recursos variados ha venido ampliándose en el tiempo, hoy día cuenta con Applets para temas de análisis de datos, simulaciones de distribución de muestreo, probabilidad, inferencia estadística, métodos multivariados y otros temas de los catalogados como clásicos. Principalmente en inferencia estadística se presenta el subtema análisis de tablas de dos vías donde se hace énfasis en frecuencias y proporciones, conteniendo bases de datos predeterminadas o bien la posibilidad de cargar una propia.

Sin embargo, igual que en StatKey los Applets se enfatizan en el diagrama de barras más que en el uso y lectura de tablas, incluso se debe seleccionar la opción para que se pueda visualizar la tabla correspondiente.

Ahora bien, en cuanto a aplicación de las denominadas Tecnologías de la Información y la Comunicación (TIC) para la enseñanza en estadística y otras áreas de conocimiento se encontraron las siguientes referencias.

Por una parte, Vásquez (2014) en su tesis cuasi experimental con 60 alumnos de sexto grado de la institución educativa 10381 de Chota logra mostrar la influencia que introdujo el programa Libro Interactivo Multimedia (LIM) en la comprensión lectora de los alumnos ya que

los resultados obtenidos en la prueba final muestran una mejora significativa respecto a la inicial, permitiendo situar al 30% de los estudiantes en nivel destacado y 70% en proceso.

Por otra parte, Gutiérrez (2021) en su artículo presenta una propuesta de enseñanza a distancia flexible e integrada con la presencial por medio de un libro electrónico interactivo donde se presenta las simetrías del plano por medio de la metodología de descubrimiento guiado de acuerdo con los niveles de Van Hiele, los autores destacan que las herramientas TIC le permite a los docentes adaptar su enseñanza de acuerdo con las necesidades específicas de los estudiantes.

En concordancia, Madriz et al. (2024) promueven el uso de la tecnología para las disposiciones del programa de matemáticas como una herramienta que puede mejorar la enseñanza de estadística en la educación primaria y proponen un libro de estadística siguiendo los lineamientos del Programa de Estudio de Matemáticas de Costa Rica (PMCR).

4. Objetivos

4.1 Objetivo general

Diseñar un capítulo de libro interactivo digital (LID) orientado al fortalecimiento de los procesos de lectura de tablas de contingencia en estudiantes de la Licenciatura en Matemáticas de la Universidad Industrial de Santander.

4.2 Objetivos específicos

- Diseñar una secuencia didáctica de un capítulo del libro interactivo digital para la enseñanza de tablas de frecuencia y tablas de contingencia fundamentada en la Teoría de Situaciones Didácticas.
- Desarrollar el prototipo funcional de dicho capítulo mediante un modelo de inteligencia artificial, el lenguaje de programación Python y HTML que implemente la secuencia didáctica diseñada en el objetivo anterior.
- Evaluar el contenido del capítulo por parte de un panel formado por profesores de Estadística y estudiantes de Didáctica de la Probabilidad y Estadística.

5. Marco Teórico

5.1 Elementos de educación Matemática

La Teoría de las Situaciones Didácticas (TSD) desarrollada por Brousseau, distingue entre situaciones didácticas, a-didácticas y no didácticas (Espinosa y Bareño, 2018). Esta distinción permite comprender la relación entre docente, estudiante y conocimiento en distintos momentos. Para los fines de esta investigación, nos centraremos en la situación a-didáctica, la cual se caracteriza por llevar al estudiante a actuar de manera autónoma y sin una guía directa a partir de un problema contextualizado.

La construcción de un conocimiento matemático propuesto por Brousseau (1986, como se citó en Sadovsky, 2005), distingue diferentes momentos o fases según el tipo de interacción del estudiante con el medio (el problema, los materiales, compañeros, entre otros) y con el saber. Inicialmente formuló tres fases: acción, formulación y validación, entendidas como situaciones a-didácticas ya que en ellas el docente se retira intencionalmente para que el estudiante asuma la responsabilidad de producir el conocimiento. Posteriormente incorporó una cuarta fase, la institucionalización, de naturaleza propiamente didáctica en la que el profesor retoma un papel activo (Panizza, 2005; Sadovsky, 2005).

- Fase de acción: El estudiante interactúa directamente con el medio para resolver el problema planteado, tomando decisiones basadas en sus conocimientos previos y recibiendo retroacciones del propio medio, sin necesidad de verbalizar o justificar sus procedimientos.

- Fase de formulación: El estudiante debe comunicar a otros (compañeros, grupo) la estrategia o el procedimiento que empleó. Esto obliga a una reorganización cognitiva: describir, ordenar y hacer comprensible el propio razonamiento para que otro pueda replicarlo o discutirlo.
- Fase de validación: Los estudiantes someten sus formulaciones y procedimientos a un proceso de justificación y prueba, argumentando por qué su solución es correcta. No basta con que el procedimiento funcione; debe sostenerse mediante razonamientos matemáticos aceptables.
- Fase de institucionalización: El docente retoma las producciones, discusiones y conclusiones que surgieron en las fases anteriores y las convierte en saber oficial: las despersonaliza (deja de ser "la estrategia de tal estudiante") y las descontextualiza (deja de estar ligada solo a ese problema particular), vinculando explícitamente el conocimiento matemático que se pretendía enseñar.

Es aquí donde el estudiante recurre a sus conocimientos previos, genera hipótesis, y se aproxima a un comportamiento similar al de una comunidad científica al formular y contrastar posibles respuestas a través de la resolución del problema para finalmente dar paso a la institucionalización del conocimiento y así consolidar las ideas centrales y clarificar posibles errores (Chavarría, 2006, como se citó en Espinosa y Bareño, 2018).

Complementariamente, Espinosa y Bareño (2018) describe algunos elementos propios de una situación a-didáctica de la siguiente manera:

- Planteamiento del Problema: el docente plantea al estudiante un problema que se asemeja a situaciones de la vida real

- **Uso de Conocimiento Previo:** El estudiante debe abordar y resolver este problema haciendo uso de sus conocimientos previos
- **Autonomía:** El estudiante resuelve la situación sin la intervención directa del docente. La interacción del estudiante con el medio didáctico (problemas, materiales e instrumentos) es lo que determina el conocimiento que este utiliza para alcanzar una solución favorable
- **Rol Científico:** el estudiante genera hipótesis y conjeturas para resolver la situación
- **Institucionalización:** el docente brinda formalidad al nuevo conocimiento aparecido, acordando la notación y propiedades.

5.2 Elementos de didáctica de la estadística

Una vez que el estudiante ha logrado dominar la organización de una sola variable a través de la transnumeración en tablas de frecuencia, surge la necesidad de comprender fenómenos más complejos donde dos variables interactúan simultáneamente. Para esto, no basta con describir cuántos individuos pertenecen a una categoría ya que un verdadero sentido estadístico se alcanza cuando el alumno es capaz de cuestionar si existe una relación de dependencia entre las categorías de análisis. Por lo anterior, es necesaria la transición del análisis descriptivo al inferencial a través de las tablas frecuencia y contingencia respectivamente, haciendo uso de esta última y la prueba de hipótesis chi-cuadrado (X^2) como herramientas de validación para alcanzar un nivel de lectura crítica.

5.2.1 Tablas en estadística

En estadística, las Tablas de Contingencia (TC) permiten representar la distribución conjunta de dos variables categóricas, facilitando el análisis de posibles asociaciones entre ellas

(Batanero et al., 1994). La estructura de las tablas de contingencia (donde una celda puede derivar en frecuencias relativas respecto al total general, la fila o la columna) exige una habilidad rigurosa para distinguir la proporción buscada, lo cual es fundamental para el análisis bivariado y el desarrollo de un razonamiento estadístico avanzado. En un nivel de análisis superior las TC se usan para validar la independencia entre variables, este proceso se apoya en herramientas inferenciales como la prueba Chi-Cuadrado.

5.2.2 Alfabetización estadística y transnumeración

El proceso de enseñanza y aprendizaje de la estadística incluye el desarrollo de la alfabetización cuantitativa la cual consiste en entender, analizar y transmitir datos numéricos en diferentes situaciones (Batanero et al., 1994). Es aquí donde las tablas estadísticas se convierten en un recurso fundamental para el desarrollo de esta alfabetización cuantitativa ya que estas permiten organizar y presentar información de manera clara. Según Arteaga et al. (2011) las tablas cumplen la función de instrumentos de transnumeración, ya que permiten reconfigurar los datos brutos y generar nuevas inferencias mediante el cambio del sistema de representación.

Particularmente, Curcio (1989, citado en Batanero et al., 1994) propone una estructura jerárquica para evaluar los diferentes niveles del procesamiento cognitivo. Es importante señalar que, aunque los niveles fueron diseñados para la lectura de gráficas, es posible hacer la extensión y aplicación a la lectura de tablas como lo sostiene Arteaga et al. (2011). Acorde con lo anterior, la jerarquía se adapta de la siguiente manera:

- Leer los datos: lectura literal en la que se identifican valores o frecuencias explícitas dentro de las celdas de la tabla sin realizar inferencias .

- Leer dentro de los datos: Implica comparar y relacionar frecuencias entre filas y columnas, aplicando nociones matemáticas básicas para descubrir patrones, tendencias o asociaciones iniciales.
- Leer más allá de los datos: capacidad de formular predicciones o inferencias que trascienden la muestra representada, extendiendo la información disponible para generar conclusiones nuevas.

Cuando el propósito es el desarrollo de una cultura estadística, esta comprensión estructural debe complementarse con la valoración crítica de los datos, para ello se añade un nivel a los propuestos por Curcio (1989, citado en Batanero et al. 1994): “Leer detrás de los datos” (Friel et al., 2001, citado en Arteaga et al., 2011). Este nivel consiste en valorar críticamente el método de recolección de datos, su validez y fiabilidad, así como las posibilidades de extensión de las conclusiones. Nuevamente, si bien esta clasificación ha sido diseñada originalmente para gráficas podemos a partir de lo que sostiene Arteaga et al. (2011) extender, aplicar y adaptar cada subdivisión propuesta por Aoyama (2007) a la lectura de tablas en general. A continuación se presentan las subdivisiones adaptadas:

- Nivel racional/literal: Leen correctamente la tabla de datos, detectan tendencias y logran hacer predicciones pero sin tener en cuenta la fuente de información.
- Nivel crítico: se realiza una evaluación contextualizada de los datos presentados en las celdas, manteniendo una actitud reflexiva y cuestionando posibles sesgos en la información tabulada.

- Nivel hipotético: formula hipótesis propias y busca evidencia para proponer interpretaciones alternativas a las conclusiones iniciales de la tabla

5.2.3 Juicios de asociación

El dominio de juicios de asociación esenciales para el reconocimiento de relaciones entre variables y el ajuste de concepciones erróneas como la confusión entre correlación y causalidad son indispensables en un analista de datos (Batanero et al., 1994). Estos juicios están dados de la siguiente manera:

- Reconocimiento de relaciones entre variables: comprensión del tipo, dirección e intensidad de las asociaciones.
- Uso correcto de conceptos de proporción y probabilidad.
- Evitar concepciones erróneas: como la concepción causal (confundir correlación con causalidad), la concepción local (analizar datos aislados sin contexto) y la concepción determinista (asumir relaciones donde no hay variabilidad).

De esta manera, los juicios de asociación constituyen una base fundamental para el desarrollo de la alfabetización estadística pues permiten al estudiante ir más allá de la lectura literal de los datos y avanzar hacia una interpretación crítica y con sentido de los mismos. En particular su dominio resulta indispensable para alcanzar niveles superiores de comprensión ya que sin la capacidad de reconocer relaciones entre variables, cuestionar sesgos y evitar concepciones erróneas, difícilmente el estudiante podrá evaluar la validez de los datos, formular hipótesis propias o proponer interpretaciones alternativas con fundamento estadístico.

5.3 Recurso Didáctico Digital

Un recurso didáctico digital puede entenderse como cualquier herramienta mediada por la tecnología diseñada con una intención pedagógica explícita que permite al estudiante interactuar activamente con el contenido y a su vez construir conocimientos a partir de dicha interacción (Marqués, 2011). En el ámbito de la Didáctica de la Estadística Batanero (2019) señala que estos recursos han tomado relevancia al ofrecer entornos en los cuales el estudiante puede visualizar, representar y explorar los datos de manera dinámica. Es así que los recursos didácticos digitales se encuentran correlacionados con el proceso de transnumeración al permitir que el estudiante reconfigure la información y genere nuevas hipótesis a partir del cambio de representación.

5.3.1 Definición y Características de un Libro Interactivo Digital

Un Libro Interactivo Digital (LID) se define como un material didáctico que trasciende el formato estático que caracteriza un libro tradicional, al integrar múltiples formas de representar la información (textual, visual, audiovisual y manipulativo) en una plataforma digital (Toapanta y Yugcha, 2023). Entre las características principales de este tipo de recursos se enuncian:

- **Interactividad:** Permite al estudiante manipular datos, realizar simulaciones y recibir feedback inmediato, fomentando la exploración autónoma.
- **Multimedia:** Incorpora videos, animaciones y visualizaciones dinámicas que facilitan la comprensión de conceptos abstractos, como las distribuciones muestrales y las probabilidades condicionales.

- Adaptabilidad: Puede funcionar como un medio didáctico flexible que el estudiante utiliza a su propio ritmo.

5.3.2 El Libro Interactivo Digital (LID) como mediador de la interacción autónoma del estudiante en situaciones a-didácticas

Desde la perspectiva de la Teoría de las Situaciones Didácticas (TSD), el (LID) funciona como el recurso para situar al estudiante en una situación a-didáctica, al permitir la simulación y la manipulación dinámica de los datos. También se resalta que el LID reduce la necesidad de la intervención directa del docente lo cual resulta interesante en estadística dada la baja intensidad horaria y adicionalmente facilita:

- Formulación de conjeturas: Realiza cambios en las distribuciones de datos de las tablas de contingencia y observa el impacto inmediato.
- Contraste de hipótesis: Verifica sus interpretaciones sobre la transnumeración y las asociaciones bivariadas en tiempo real.
- Favorece la exploración y experimentación (soluciones no formales) lo cual supone un aprendizaje activo.

5.4 Elementos teóricos de tablas en estadística

El estudio de la estadística comienza con la necesidad de organizar la información para convertir los datos en objetos culturales con significado (Arteaga et al., 2011), donde las tablas de frecuencia son la una herramienta fundamental para cumplir este fin, lo que a su vez da paso hacia el desarrollo de la transnumeración al permitir que el estudiante logre “leer los datos” o

dicho de otra manera situarse en el Nivel 1 de comprensión (Curcio, 1987, citado en Arteaga et al., 2011). Sin embargo, la realidad demanda comprender cómo interactúan distintas variables, motivando de esta manera el tránsito hacia las tablas de contingencia, ya que con el análisis bivariado se facilita una lectura “dentro y más allá de los datos” (Niveles 2 y 3) sentando las bases para alcanzar la “lectura crítica” o Nivel 4 (Aoyama, 2007, citado en Arteaga et al, 2011).

De este modo, el recorrido desde las tablas de frecuencia hasta las tablas de contingencia no es solo una transición en términos de complejidad del recurso (pasar de “algo más fácil” a “algo más difícil”) sino un tránsito cognitivo que deja ver el desarrollo gradual de la alfabetización estadística, desde la organización de los datos hasta la interpretación crítica de relaciones entre variables.

5.4.1 Tablas de frecuencia

Los elementos teóricos que constituyen una tabla de frecuencia se describen y definen de la siguiente manera.

5.4.1.1 Frecuencia absoluta.

Es el número de veces que aparece un determinado valor (x_i) en el estudio. Se denota de la siguiente manera f_i

Propiedad: La suma de todas las frecuencias absolutas es igual al número total de datos (n)

$$\sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k = n$$

n = el número total de la muestra

5.4.1.2 Frecuencia relativa.

Representa la proporción de elementos que pertenecen a una clase respecto al total de datos y se denota de la siguiente manera f_r donde:

$$f_r = \frac{f_i}{n}$$

n = tamaño total de la muestra

f_i = frecuencia absoluta

Propiedad: la suma de todas las f_r es igual a 1:

$$\sum_{r=1}^k f_r = 1$$

Aclaración: si se desea conocer los porcentajes de (h_i) basta con multiplicar por 100.

5.4.1.3 Frecuencia absoluta acumulada.

Es la suma de las frecuencias absolutas de todos los valores inferiores o iguales al valor (x_i) considerado y se denota F_i donde:

$$F_i = \sum_{j=1}^i f_j = f_1 + f_2 + \dots + f_i$$

Propiedad: el último valor de la frecuencia absoluta acumulada (F_k) siempre es igual al total de datos (n)

5.4.1.4 Frecuencia relativa acumulada.

Es la suma de todas las frecuencias relativas (f_i) menores o iguales a un determinado valor (x_i) en el conjunto y se denota de la siguiente manera (F_r), donde:

$$F_r = \sum_{j=1}^r f_j = f_1 + f_2 + \dots + f_r$$

n = número total de datos de la muestra

Propiedad: el último valor de la frecuencia acumulada (F_k) es igual a 1

5.4.2 Representación y organización de datos

Para clarificar el tema anterior, se presenta a continuación un ejemplo de una tabla de frecuencia (Tabla 2). La tabla de frecuencia actúa como medio facilitador para el desarrollo del proceso de transnumeración al organizar los datos brutos y permitir que se extraigan significados iniciales, a través de esta representación se busca que el alumno logre "leer los datos" y alcance un Nivel 1 de comprensión. Los datos recolectados están listados uno a uno sin orden o edición alguna, a este formato que suele denominarse “datos libres”, para facilitar la visualización de esta muestra univariada un recurso es la tabulación que consiste en crear categorías o intervalos que agrupan los datos tal como se muestra a continuación.

Tabla 2.*Tiempo de exposición a pantallas en estudiante*

Intervalos	Frecuencia absoluta (f_i)	Frecuencia relativa (f_r)	Frecuencia absoluta acumulada (F_i)	Frecuencia absoluta acumulada (F_r)
0 - 30	4	$\frac{4}{17}$	4	$\frac{4}{17}$
31 - 60	5	$\frac{5}{17}$	9	$\frac{9}{17}$
61 - 90	8	$\frac{8}{17}$	17	$\frac{17}{17}$
Total	17	1		

5.4.3 Elementos Teóricos de Tablas de Contingencia

Las tablas de contingencia según Posada (2016) son un recurso para el análisis de datos bivariados o cruzados que cumplen con el objetivo de establecer relaciones entre las categorías y variables presentes en un conjunto. Este recurso marca una diferencia significativa respecto a las tablas de frecuencia pues incluso desde su estructura bidimensional permite el estudio de las relaciones entre categorías dando paso a que el estudiante tenga un soporte visual siendo este fundamental para el proceso de transnumeración.

Con el fin de visualizar y analizar de manera detallada los elementos estructurales que componen este tipo de representación bivariada, se propone el siguiente ejemplo práctico (Tabla 3). Este ejemplo parte de una muestra hipotética de 120 estudiantes de una institución educativa de Bucaramanga, en la cual se realiza un cruce de información entre dos variables. Por un lado, la primera categoría corresponde a “género”, subdividido en las categorías masculino y

femenino, mientras que por otro lado, la segunda se centra en “la preferencia del sabor de helado”, abarcando las categorías de vainilla, chocolate y fresa.

Tabla 3.

Tabla de contingencia

Sabor de helado Sexo	Vainilla	Chocolate	Fresa	Total
Masculino	14	34	12	60
Femenino	22	16	22	60
Total	36	50	34	120

Como se aprecia en la Tabla 3, su forma más básica contiene frecuencias absolutas conjuntas, es decir, el conteo de individuos que pertenecen simultáneamente a una categoría de cada variable. Esta representación permite al estudiante una lectura directa de los datos, como por ejemplo: ¿cuántos estudiantes hombres prefieren el sabor a vainilla? Sin embargo, la lectura simple de frecuencias absolutas no posibilita comparar grupos de tamaños distintos ni identificar patrones de asociación. Acorde con este escenario, surge la necesidad de transformar el tipo de representación, dando lugar a diferentes clases de tablas según los niveles (total, filas y columnas) que se describen a continuación.

5.4.3.1 Porcentajes o frecuencias relativas al total (distribución conjunta)

Para calcular la distribución conjunta se realiza el cociente de cada frecuencia absoluta con el total de los datos, es así que se obtiene la proporción de cada celda respecto al tamaño de muestra. Esta transformación constituye el primer acto de transnumeración: los conteos sin

procesar permiten un acercamiento hacia la lectura y formulación de hipótesis iniciales sobre las posibles asociaciones respecto a las categorías pero los valores en escala porcentual permiten evaluar la distribución de los datos de manera más efectiva. Un breve ejemplo de la distribución conjunta se visualiza en la Tabla 4.

Tabla 4.

Tabla de contingencia con porcentajes sobre el total

Sabor de helado Genero	Vainilla	Chocolate	Fresa	Total
Masculino	11,7%	28,3%	10,0%	50%
Femenino	18,3%	13,3%	18,3%	50%
Total	30,0%	41,7%	28,3%	100%

La Tabla 4 potencia que el estudiante logre leer entre los datos, al poder establecer relaciones como por ejemplo: el 28,3% de la muestra total corresponde a hombres que prefieren el helado de chocolate o que el chocolate parece ser el sabor más predilecto porque el 41,7% de la muestra lo eligió independientemente del género.

5.4.3.2 Porcentaje o frecuencias relativas por fila (distribución condicional por género)

Para determinar las proporciones dentro de cada fila, se divide cada frecuencia absoluta por el total de cada género, obteniendo así la distribución condicional. Esto nos lleva a un nuevo acto de transnumeración pues reestructura la misma información bajo una pregunta distinta: ¿cómo se distribuye el sabor de helado preferido dentro de cada categoría de género?(Tabla 5)

Tabla 5.*Tabla de contingencia con porcentajes por fila*

Sabor de helado Sexo	Vainilla	Chocolate	Fresa	Total
Masculino	23,3%	56,7%	20,0%	100%
Femenino	36,7%	26,7%	36,7%	100%

En la Tabla 5 se hace visible aquello que las frecuencias absolutas ocultaban: el 36,7% de las estudiantes mujeres prefieren por igual el sabor de helado a vainilla o fresa, frente al 23,3% y 20% de los hombres respectivamente. Esta comparación entre distribuciones condicionales, activa una lectura más allá de los datos pues el estudiante ya no solo describe sino que compara grupos y comienza a formular conjeturas e hipótesis sobre una posible asociación entre género y preferencia de sabor de helado.

5.4.3.3 Porcentaje o frecuencias relativas por columna (distribución condicional por uso de redes)

De manera similar, es posible calcular las proporciones dentro de cada columna, obteniendo la distribución condicional del género dado el sabor de helado preferido (Tabla 6). Este punto de vista responde a una pregunta distinta: entre quienes prefieren cada sabor de helado, ¿qué proporción corresponde a cada género?

Tabla 6.*Tabla de contingencia con porcentajes por columna*

Sabor de helado Sexo	Vainilla	Chocolate	Fresa	Total
Masculino	38,9%	68,0%	35,3%	50%
Femenino	61,1%	32,0%	64,7%	50%
Total	100%	100%	100%	100%

La elección entre la Tabla 5 (tabla por filas) y la Tabla 6 (Tabla por columnas) no es arbitraria ya que depende de la pregunta que se desee responder o bien el análisis que se requiera. Esta decisión forma en sí misma un elemento del pensamiento estadístico ya que el estudiante debe conocer criterios técnicos sobre cómo presentar la información de manera eficiente mientras que a su vez se sitúa en lo que Brousseau define como Fase de Formulación puesto que el estudiante debe seleccionar la representación más adecuada para comunicar su interpretación.

Es así que las distribuciones marginales (proporción de los totales de fila y columna sobre n) y las distribuciones condicionales (proporción de cada frecuencia conjunta entre su total por fila o por columna) constituyen el núcleo de la comprensión de tablas de contingencia en los niveles intermedios. El dominio de estas distribuciones es condición necesaria para acceder al nivel crítico que se aborda mediante la inferencia estadística

5.4.4 Asociación e independencia entre variables: el puente hacia la inferencia.

Antes de realizar la validación estadística formal, es necesario entender qué es la asociación entre variables de manera descriptiva. Intuitivamente, dos variables cualitativas se

consideran asociadas cuando la distribución de frecuencias o proporcione de una variable varía según las categorías de la otra, en caso contrario, se habla de independencia estadística.

Este momento de exploración descriptiva corresponde, en el marco de la Teoría de Situaciones Didácticas de Brousseau, a una situación de acción y formulación: el estudiante actúa sobre la tabla, produce sus propias representaciones y formula conjeturas sobre la posible relación entre las variables. La pregunta "¿será que estas variables están relacionadas?" es la que motiva el salto hacia la inferencia.

5.4.5 La inferencia estadística como situación de validación

La transición hacia la lectura crítica de los datos, requiere que el estudiante no sólo perciba una asociación en la tabla a nivel descriptivo sino que la valide mediante métodos inferenciales. Este proceso constituye, en términos de Brousseau, una situación de validación: el estudiante debe determinar si la asociación observada es estadísticamente significativa o puede atribuirse al azar.

A este respecto, el estadístico de prueba chi cuadrado en palabras de Quevedo (2011) “sirve para someter a prueba hipótesis referidas a distribuciones de frecuencias. En términos generales, esta prueba contrasta frecuencias observadas con las frecuencias esperadas de acuerdo con la hipótesis nula. Se usa para probar la asociación entre dos variables, también se puede usar para evaluar en qué medida se ajusta la distribución de frecuencias obtenida con los datos de una muestra, a una distribución teórica o esperada.”

Fórmula general del estadístico:

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Frecuencia observada en la fila i y columna j : O_{ij}

Frecuencia esperada en la fila i y columna j : E_{ij}

Número de filas: r

Número de columnas: c

Para hallar la frecuencia esperada E_{ij} se hace de la siguiente manera:

$$E_{ij} = \frac{(Total\ fila\ (i)) * (Total\ columna\ (j))}{Total\ global}$$

Una vez calculado el estadístico χ^2 se procede a ubicar este valor en la distribución indicada según los grados de libertad para así determinar si el valor observado es un valor poco probable de ocurrir y por ende no hay asociación entre variables. En la Figura 2 se visualizan diferentes distribuciones χ^2 así como propiedades.

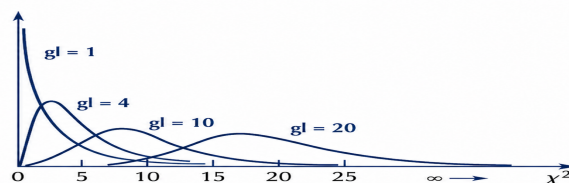
Figura 2.

Distribución χ^2

Propiedades de la distribución ji cuadrada

1. χ^2 es no negativo en valor; es cero o con valor positivo.
2. χ^2 no es simétrica; está sesgada a la derecha.
3. χ^2 está distribuida de modo que forma una familia de distribuciones, una distribución separada para cada diferente número de grados de libertad.

FIGURA 9.7
Varias distribuciones
ji cuadrada



Tomado de *Estadística elemental* (p.454), por R. Johnson y P. Kuby 11a edición, 2012, Editorial Cengage Learning.

De esta manera resulta fundamental la integración de esta prueba al Libro Interactivo Digital (LID) puesto que permite institucionalizar el conocimiento al automatizar el cálculo y centrar la atención del estudiante en la lógica del procedimiento que subyace a la interpretación de un *p-valor* y en la toma de decisiones sobre la hipótesis de independencia.

Cabe aclarar que la interpretación del *p-valor* y la decisión sobre la hipótesis (H_0) si se hace en forma conciente se constituye en un ejercicio muy exigente de interpretación crítica al demandar ir más allá de lo que la tabla muestra para cuestionarse si lo observado se ajusta a un comportamiento teórico y si puede generalizar a la población en estudio.

6. Metodología

Esta propuesta de investigación se enmarca en la modalidad de producción de material didáctico, para su desarrollo se organiza una planeación de cinco fases:

Actividad	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5
Consulta bibliográfica en temas como: Educación estadística, diseño, programación y práctica con herramientas para edición de textos interactivos en estadística (Python, html y modelo de IA Gpt - 4.0)	X				
Diseño y programación del prototipo tomando como base el material de clase utilizado por la directora del trabajo de		X			

grado					
Diseño de instrumento de evaluación para panel de expertos			X		
Revisión de estilo y contenido por panel de “expertos” sobre el capítulo de libro diseñado				X	
Ajuste y edición del capítulo de libro					X

7. Resultados

El producto de este trabajo es un capítulo de libro interactivo para la enseñanza de tablas estadísticas que contiene tres secciones. La primera se centra en la lectura, interpretación y construcción de tablas de frecuencia, la segunda sección aborda las tablas de contingencia de manera descriptiva y en la tercera sección se aborda la asociación entre variables dando paso a la prueba de independencia chi-cuadrado.

El libro interactivo digital de estadística descriptiva del dato a la decisión es un recurso educativo de acceso libre el cual fue desarrollado usando herramientas de software como Python, HTML, JavaScript y Cascading Style Sheets (CSS), en el siguiente link se tiene acceso al capítulo diseñado <https://g240935-debug.github.io/LID-UIS/>.

Este recurso está dirigido a estudiantes de licenciatura en matemáticas, sin embargo, el recorrido incluye situaciones que invitan a la exploración y el descubrimiento haciendo que el conocimiento se logre construir de manera progresiva independientemente del nivel educativo o la formación estadística previa. Este recurso no requiere instalación ni registro previo y puede ser utilizado desde cualquier dispositivo con acceso a internet, tiene como propósito central fortalecer la comprensión de tablas estadísticas mediante situaciones a-didácticas que promueven el aprendizaje autónomo teniendo en cuenta la TSD planteada por Brousseau y la adaptación a lectura de tablas de los niveles planteados por Curcio y Aoyama para Lectura de Gráficas.

7.1 Diseño didáctico de la secuencia por capítulo

Con el fin de evidenciar el cumplimiento del primer objetivo específico que se refería al diseño de una secuencia didáctica de un capítulo del libro interactivo digital fundamentada en la

Teoría de Situaciones Didácticas, la Tabla 7 sistematiza las situaciones de aprendizaje que componen el LID organizadas por sección y clasificadas según la fase de la TSD que predomina en cada una (Acción, Formulación, Validación e Institucionalización), así como el nivel de comprensión de datos de Curcio y Aoyama que se busca fortalecer durante todo el desarrollo del capítulo.

Como puede observarse, cada una de las tres secciones del LID recapitula un ciclo completo del esquema de Brousseau, en lugar de aplicarlo una única vez a la totalidad del libro. La Sección I inicia con una situación de exploración libre que genera la necesidad del conteo antes de introducir la notación formal de frecuencias o f_i , y cierra con actividades de validación donde el estudiante aplica lo construido de manera autónoma.

La Sección II reproduce el mismo principio planteando la situación del coordinador de Bienestar Estudiantil, en la que el estudiante enfrenta preguntas con respuesta no disponible a primera vista partiendo de una lista de datos sin procesar, esto lleva al estudiante a una situación a-didáctica ya que la necesidad de la tabla de contingencia surge del problema mismo, no de una indicación externa, y avanza progresivamente hacia una intuición pre-formal de independencia estadística que prepara el terreno para la Sección III.

La sección III se encuentra centrada en la prueba de chi-cuadrado, mantiene la misma lógica: el estudiante descubre por sí mismo la fórmula de la frecuencia esperada antes de que le sea revelada, y cuestiona la necesidad de elevar al cuadrado las diferencias antes de que dicho procedimiento se institucionalice. La progresión se mantiene hasta el cierre del capítulo pues en la situación final de aplicación libre, el tutor recibe primero la primera aproximación investigativa del estudiante, su contexto, pregunta y variables para luego guiarlo de manera

general. Se espera que el usuario sea capaz de retomar cualquier concepto trabajado a lo largo de todo el libro acorde con lo que la situación particular del estudiante requiera y cuestionando si sus propias decisiones responden efectivamente a la pregunta que él mismo planteó.

Adicional a lo anterior, es importante señalar que las páginas denominadas "Ejemplos Dinámicos", presentes en los Capítulos I y II, se excluyeron intencionalmente de esta clasificación ya que estas cumplen una función de refuerzo visual y de exploración libre de representaciones ya construidas previamente, pero no se ajustan a ninguna fase específica del ciclo Acción-Formulación-Validación-Institucionalización.

En conjunto, esta sistematización permite sostener que las decisiones de diseño del LID son el resultado de una aplicación consciente y verificable de la Teoría de Situaciones Didácticas a lo largo de la secuencia completa del recurso.

Tabla 7.

Diseño didáctico del LID según las fases de la Teoría de Situaciones Didácticas

Capítulo	Fase TSD	Situación	Nivel Curcio	Justificación
Sección I Tablas de Frecuencia	Acción	Exploración libre de datos (situación de apertura)	N1	Datos libres sin agrupar; genera la necesidad del conteo antes de nombrar f_i .
	Formulación	Construcción guiada $f_i \rightarrow f_r \rightarrow F_i \rightarrow F_r$	$N1 \rightarrow N2$	Cada columna se revela sólo tras evidencia de

Capítulo	Fase TSD	Situación	Nivel Curcio	Justificación
				comprensión de la columna anterior, no por avance automático.
	Institucionalización	Síntesis de la tabla completa	N2	Formalización explícita de los cuatro tipos de frecuencia.
	Validación	Problemas y Análisis (construcción autónoma)	N2→ N3	Aplicación independiente con verificación y justificación en lenguaje propio.
Sección II Tablas de Contingencia	Acción	Exploración:"El coordinador de Bienestar Estudiantil"	N1	Situación didáctica genética; preguntas complejas de responder con datos libres y sin procesar.
	Formulación	Actividad — Color favorito vs Personalidad	N2 →N3	Comparación de cifras absolutas de grupos de distinto tamaño antes de introducir el porcentaje relativo.

Capítulo	Fase TSD	Situación	Nivel Curcio	Justificación
	Institucionalización	Síntesis y Formas Parciales de lectura	N2 → N3	Formalización de frecuencia conjunta, condicional, marginal, y los sistemas de representación (% total/fila/columna).
	Validación	Problemas de Formulación y Validación	N2 → N3	Construcción completa de la tabla más justificación argumentada.
	Extensión hacia N4	Puente entre Asociación e Independencia y Exploración libre	N3 → N4	Intuición pre-formal de independencia estadística, previa al capítulo de chi-cuadrado.
Sección III Prueba Chi-cuadrado	Acción	Descubrimiento guiado de la fórmula de frecuencia esperada	N3 → N4	El estudiante propone su propia fórmula antes de que el sistema la revele; pistas progresivas.
	Formulación	Por qué elevar al cuadrado / cálculo de χ^2	N3 → N4	Cuestiona por qué una resta simple no basta, antes de institucionalizar $(O-E)^2$.

Capítulo	Fase TSD	Situación	Nivel Curcio	Justificación
	Validación	Grados de libertad, valor crítico y ejemplo completo	N3 → N4	Aplicación del procedimiento completo con verificación.
	Institucionalización	Síntesis Final	N4	Cierre formal del procedimiento y su interpretación.
	Integración (ciclo completo)	Situación libre final y Aplicación Libre (estudio propio)	N4	El tutor recopila primero la intención investigativa del estudiante (contexto, pregunta y variables) y lo guía con un andamiaje dialógico de alcance general, capaz de retomar cualquier concepto trabajado a lo largo del libro (frecuencias, asociación, independencia) según lo que su situación particular requiera, cuestionando si sus decisiones responden a su

Capítulo	Fase TSD	Situación	Nivel Curcio	Justificación
				propia pregunta de investigación. Al usar datos propios, el control sobre la corrección numérica es menor.

7.2 Enfoque didáctico del recurso

La estructura interna y el funcionamiento del Libro Interactivo Digital (LID), cuyos resultados de evaluación se analizan en este capítulo, se fundamentan operativamente en la noción de situación a-didáctica propuesta por Brousseau. Bajo este enfoque, cada actividad sitúa al estudiante ante un problema estadístico que exige la construcción activa y autónoma del conocimiento. En este entorno digital, el tutor virtual desempeña un rol fundamental al estar desarrollado a través de una estructura web integrada en Python y JavaScript (JS) y apoyado por el modelo de Inteligencia Artificial (IA) GPT-4o que le permiten actuar como el mediador didáctico central.

La integración de la IA permite que el sistema analice las respuestas y las conjeturas escritas por el usuario no para calificar de forma mecánica el acierto o error sino desde la devolución flexible y en tiempo real de las consecuencias de sus decisiones e interacciones, sin anticipar soluciones ni validar directamente sus procedimientos (Chavarría, 2006, como se citó en Espinosa y Bareño, 2018).

Por consiguiente, las fases de acción, formulación y validación se organizan secuencialmente durante el desarrollo de interfaz dinámica del LID gracias a la interactividad lograda con JS y a las retroalimentaciones ofrecidas por la IA, ofreciendo a su vez la institucionalización de conceptos analizados tales como: distribución conjunta, marginal, condicional y el estadístico chi-cuadrado, mediante la exploración e interacción con el medio. Con lo anterior, se aborda una de las mayores problemáticas de la educación tradicional como es el aprendizaje memorístico y mecánico que históricamente ha caracterizado la enseñanza de la disciplina (Batanero, 2001).

Adicional a lo anterior, se consideró la jerarquía de comprensión propuesta por Curcio (1989) citado por Batanero et al. (1994) y extendida por Arteaga et al. (2011) a la lectura de tablas, la cual guió la progresión cognitiva de los entornos interactivos programados en HTML, JavaScript y Python. Estos entornos fueron a su vez estructurados visualmente mediante Hojas de Estilo en Cascada (CSS). El diseño visual y el contraste logrados con CSS no constituyen elementos puramente estéticos, sino que reducen la carga visual extraña y, a su vez, guían con precisión la atención del estudiante hacia los patrones de asociación estadística.

A través del desarrollo secuencial de sus interfaces, el LID conduce al estudiante desde el nivel de leer los datos, mediante la identificación literal de frecuencias explícitas, transitando a su vez hacia el nivel de leer dentro de los datos al exigir la comparación, cálculo y manipulación dinámica de frecuencias entre filas y columnas para descubrir patrones de asociación bivariada. Posteriormente, el recurso promueve el nivel de leer más allá de los datos a través de la formulación de predicciones e inferencias informadas que trascienden la muestra analizada. Finalmente, se da paso al módulo desarrollado para la prueba Chi-cuadrado, adentrando al

estudiante en problemas y análisis de nivel superior correspondientes a leer detrás de los datos (Friel et al., 2001, como se citó en Arteaga et al., 2011), al articular las subdivisiones de este último nivel propuestas por Aoyama (2007, como se citó en Arteaga et al., 2011).

Esta progresión promueve juicios de asociación rigurosos (Batanero et al., 1994), estimulados intencionalmente a través de los scripts de interactividad en JavaScript, el procesamiento lógico en Python, la estilización visual en CSS y el procesamiento de lenguaje natural de la IA, para confrontar las concepciones erróneas más frecuentes en el análisis bivariado. Específicamente, las actividades confrontan de forma activa al estudiante para superar la concepción causal (confundir correlación con causalidad), la concepción local (analizar celdas o datos aislados ignorando el contexto global) y la concepción determinista (asumir relaciones perfectas y rígidas entre las variables categóricas), promoviendo una postura críticamente fundamentada ante la validez de las asociaciones observadas.

Ahora bien, desde el punto de vista técnico, el LID fue construido mediante la integración de cuatro tecnologías complementarias: HTML como lenguaje de marcado para la estructuración del contenido, CSS para la organización visual de los elementos de la interfaz, JavaScript (JS) como motor de la interactividad, y Python con Flask micromarco de trabajo (*microframework*), como núcleo del servidor backend encargado del procesamiento lógico y la gestión de las sesiones de aprendizaje. Esta arquitectura de capas permitió separar con claridad las responsabilidades del sistema: mientras HTML y CSS definen lo que el estudiante ve y cómo lo percibe, JavaScript gestiona las acciones que el estudiante ejecuta en tiempo real como el cambio de sistema de representación, el llenado de tablas o la navegación entre páginas y Python

administra el estado de cada sesión, enruta los mensajes al modelo de inteligencia artificial y registra los avances del estudiante.

La conexión entre el LID y el modelo GPT-4o se realizó mediante la API de OpenAI, consumida desde el servidor Flask a través de solicitudes HTTP (Protocolo de Transferencia de Hipertexto). Cada vez que el estudiante envía un mensaje al tutor virtual, JavaScript captura la entrada y la transmite al servidor; Python la recibe, la incorpora al historial conversacional de la sesión activa y llama al modelo con un conjunto de instrucciones pedagógicas denominado system prompt, diseñado específicamente para cada capítulo y fase del LID.

El system prompt mencionado anteriormente controla el rol del tutor virtual como mediador didáctico de la siguiente manera: define el protocolo secuencial de fases, prohíbe la entrega directa de respuestas, establece las condiciones bajo las cuales el modelo puede institucionalizar un concepto y especifica la notación matemática que se debe usar, adicional a ello se controla el tono con el cual el tutor responde al estudiante así como responder desde las consecuencias de las decisiones del estudiante, evitando tanto la validación halagadora como la corrección punitiva. La respuesta generada por el modelo retorna al servidor, que la devuelve al frontend, donde JavaScript la renderiza en el panel de chat del estudiante.

Con el propósito de proteger la clave de acceso a la API y mantener el historial conversacional de forma centralizada, el servidor fue desplegado en la plataforma Render, de modo que la comunicación entre el navegador del estudiante y el modelo de IA nunca se produce de forma directa sino siempre a través de este intermediario seguro.

Para distinguir las sesiones y avances de diferentes estudiantes en el LID, cada conversación tiene un identificador único generado aleatoriamente y almacenado en el navegador del estudiante, lo que permite retomar el avance en cualquier momento independientemente de si la página en la que se encontraba contaba o no con tutor de IA, garantizando así la aplicación coherente de los *prompts* a lo largo de toda la interacción.

7.3 Estructura y organización del contenido

El LID está organizado por secciones entre las cuales se abordan los temas de tablas de frecuencia, tablas de contingencia y la prueba chi-cuadrado. Con fines expositivos, se asignó una numeración secuencial a cada sección del LID (Tabla 8); esta numeración no aparece en la interfaz del recurso y corresponde únicamente a una convención adoptada para efectos de organización y referencia dentro del presente documento.

Tabla 8.

Estructura general del Libro Interactivo Digital (LID)

Capítulo	N.º	Contenido
Sección 1 Tablas de Frecuencia	1.1	Construcción guiada
	1.2	Síntesis (Tabla de frecuencias completa)
	1.3	Ejemplos Dinámicos (Explora la Tabla de Frecuencias)
	1.4	Problemas (Completa la tabla de frecuencias)
	1.5	Análisis (Construye y analiza tu propia tabla)

Capítulo	N.º	Contenido
Sección 2 Tablas de Contingencia	2.1	Actividad (El coordinador de Bienestar Estudiantil necesita responder...)
	2.2	Síntesis (Lo que construiste en esta sesión)
	2.3	Formas Parciales (Formas de leer una tabla de contingencia)
	2.4	Actividad (Color favorito $\times$ Personalidad)
	2.5	Asociación e Independencia (¿Están relacionadas estas variables?)
	2.6	Ejemplos Dinámicos
	2.7	Formulación
	2.8	Validación
	2.9	Exploración libre
Sección 3 Inferencia con Chi-cuadrado	3.1	Introducción
	3.2	Acción (descubrimiento guiado de la fórmula de frecuencia esperada)
	3.3	Ejemplo (aplicación de la fórmula paso a paso)
	3.4	Formulación (cálculo del estadístico χ^2)
	3.5	Síntesis

Capítulo	N.º	Contenido
	3.6	Validación (grados de libertad y valor crítico)
	3.7	Ejemplo
	3.8	Reflexión (asociación vs causalidad)
	3.9	Síntesis Final
	3.10	Aplicación (situación libre final de evaluación)
	3.11	Aplicación Libre (tu propio estudio estadístico)

Las secciones presentadas en la tabla anterior describen el recorrido que realizará el estudiante a lo largo del LID. En cada una encontrará situaciones concretas que van desde la exploración inicial de datos sin procesar hasta la aplicación de herramientas de análisis inferencial, pasando por ejercicios de construcción, ejemplos dinámicos y espacios de reflexión.

7.3.1 Sección I — Tablas de frecuencia

La primera sección introduce al estudiante en la construcción de la distribución de frecuencias a partir de una situación de exploración deliberadamente incómoda: antes de presentar cualquier tabla, se exhiben las cuarenta respuestas de una encuesta en su forma de dato libre (sin ordenar, sin resumir), y se le pide al estudiante que extraiga conclusiones sobre el comportamiento de la variable a partir de esa lista. Esta situación busca que el estudiante experimente por sí mismo la dificultad de interpretar información no procesada antes de que se le ofrezca alguna herramienta de resumen, de modo que la necesidad de organizar los datos surja de la interacción directa con el medio y no de una instrucción externa. Es así como el agrupamiento

y el cálculo de frecuencias emergen como una respuesta natural a un problema real y no como un procedimiento impuesto, lo que favorece una comprensión más profunda de su utilidad estadística en contextos concretos.

Figura 3.

Introducción construcción guiada sección I

The screenshot shows a digital interface for a statistics textbook. At the top, it says 'CAPÍTULO I · CONSTRUCCIÓN GUIADA' and 'f_i - Exploración'. Below this, a section titled 'SITUACIÓN' contains the text: 'Antes de construir la tabla... Se encuestó a 40 estudiantes sobre su bebida favorita al momento de estudiar. Estas son las respuestas **libres** tal como se recolectaron, sin ningún procesamiento.' Below the text is a list of 40 numbered items, each representing a student's favorite beverage. The list is as follows:

1	Café Negro
2	Jugo Natural
3	Café Negro
4	Té / Aromática
5	Café Negro
6	Café Negro
7	Jugo Natural
8	Té / Aromática
9	Café Negro
10	Café Negro
11	Bebida Energizante
12	Café Negro
13	Jugo Natural
14	Café Negro
15	Bebida Energizante
16	Té / Aromática
17	Té / Aromática
18	Jugo Natural
19	Té / Aromática
20	Té / Aromática
21	Café Negro
22	Jugo Natural
23	Café Negro
24	Té / Aromática
25	Bebida Energizante
26	Café Negro
27	Café Negro
28	Café Negro
29	Jugo Natural
30	Café Negro
31	Jugo Natural
32	Bebida Energizante
33	Té / Aromática
34	Café Negro
35	Jugo Natural
36	Café Negro
37	Café Negro
38	Té / Aromática
39	Café Negro
40	Té / Aromática

Below the list, there is a section titled 'ANTES DE CONTINUAR' with the question: '¿Qué puedes concluir sobre el comportamiento de la variable *Bebida favorita al momento de estudiar*, mirando los datos tal como están?'. There is a text input field for the answer. At the bottom, there are navigation buttons: 'Continuar →', '← Presentación', and 'Síntesis → →'.

Nota: Lista de cuarenta respuestas en formato de dato libre, antes de cualquier organización tabular.

La transición del dato bruto a la representación organizada se presenta de manera dinámica, cuando el estudiante se enfrenta a la lista desordenada (Figura 3), experimenta la necesidad de la transnumeración donde a partir de esto un tutor virtual guía la construcción progresiva de las cuatro frecuencias: Absoluta, Relativa, Absoluta Acumulada y Relativa Acumulada.

El tutor virtual institucionaliza cada una de las frecuencias mencionadas anteriormente de manera secuencial, dando paso entre ellas sólo después de que el estudiante haya completado el

nivel de razonamiento requerido o bien agotando el límite de preguntas al tutor sobre un mismo concepto. Tras cada institucionalización, el tutor propone un ejemplo de uso natural y una situación comunicativa distinta como contarle un dato a un compañero, redactarlo para un periódico universitario, reportarlo a un coordinador administrativo, sintetizar en una diapositiva, exigiendo en cada caso un nivel de lectura de Curcio distinto y creciente, de modo que el estudiante no solo calcule la frecuencia sino que reconozca para qué sirve y en qué contexto resulta apropiado la aplicación de cada uno de los conceptos estadísticos, como se visualiza en la Figura 4.

Figura 4.

Construcción guiada sección I

LIBRO INTERACTIVO DIGITAL Y TABLAS DE CONTINGENCIA - UIS - 2026

CAPÍTULO I · CONSTRUCCIÓN GUIADA · f_i activa 02

SITUACIÓN · BUCARAMANGA

Construyendo la tabla paso a paso

El tutor te guiará para construir la tabla de frecuencias completa. Cada columna aparecerá a medida que el tutor la introduzca: primero f_i , luego f_r , después F_i y finalmente F_r .

DISTRIBUCIÓN DE LA VARIABLE BEBIDA FAVORITA AL MOMENTO DE ESTUDIAR - F_1 (N = 40)	
Bebida	f_i
Café Negro	18
Té / Aromática	10
Jugo Natural	8
Bebida Energizante	4
Total	40

f_i ✓ f_r F_i F_r

Fase 1: Frecuencia Absoluta (f_i)
El tutor te guía para entender el conteo directo de cada categoría.

Tutor Virtual UIS
En línea

Observa la tabla inicial con f_i . Responde las preguntas del tutor.

¡Hola! Vamos a comenzar. Se encuestaron a 40 jóvenes de la UIS sobre su bebida preferida para estudiar. La tabla inicial ya tiene los datos de frecuencia absoluta (f_i). Ahora, ¿puedes identificar cuántos prefieren 'Café Negro'?

Escribe tu respuesta aquí... **Enviar** ↑

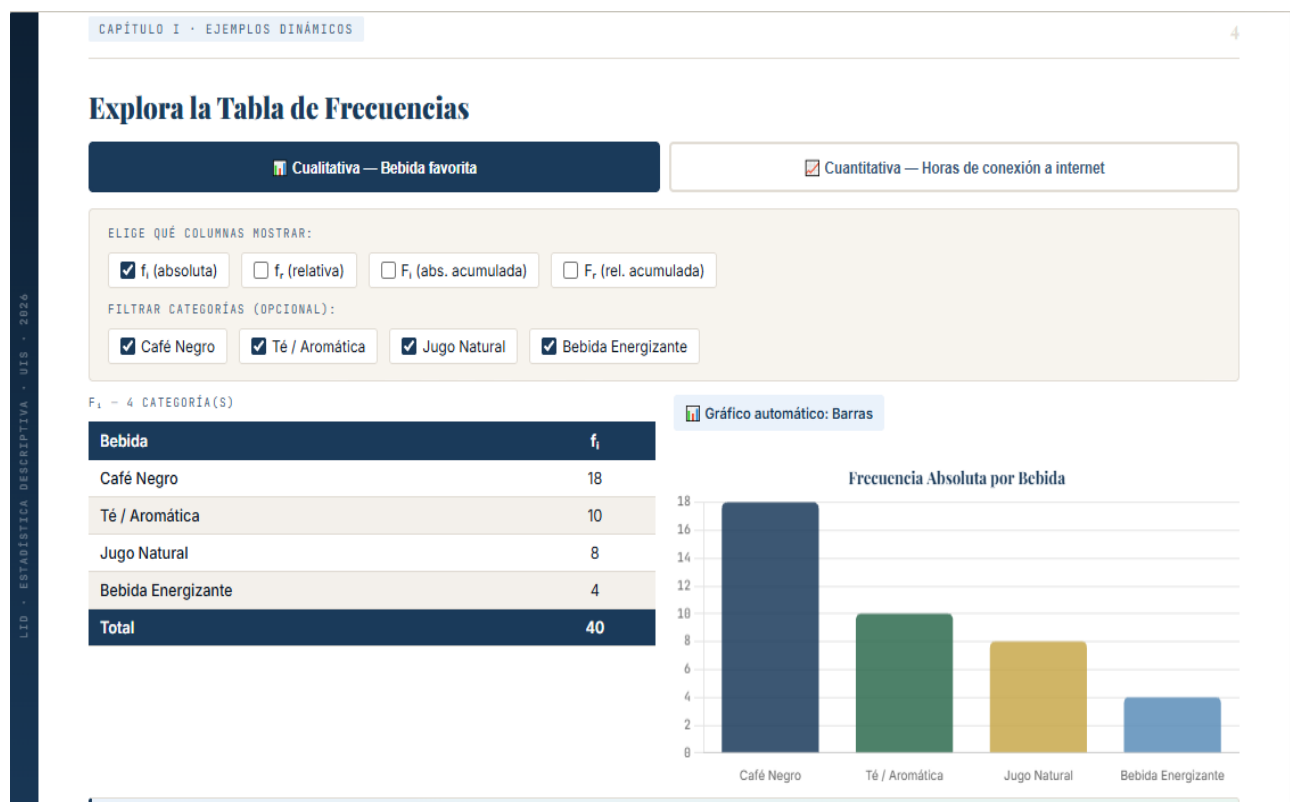
← Presentación Sintesis →

Nota: El tutor virtual acompaña la formalización de las frecuencias tras la institucionalización de cada concepto.

La sección se completa con ejemplos dinámicos donde el estudiante alterna entre una variable cualitativa y una cuantitativa (figura 5), seleccionando columnas de la tabla para observar cómo se actualiza automáticamente el gráfico correspondiente a barras, circular o histograma, según corresponda acorde al tipo de variable y una de las distribuciones de frecuencias permitiendo establecer la diferencia de comportamiento e interpretación de los gráficos en base al contexto.

Figura 5.

Ejemplos dinámicos



Nota: Ejemplos dinámicos con selección de variable y actualización automática del gráfico correspondiente.

Una vez terminado el proceso en esta página, se propone la realización de la síntesis sobre lo que se trabajó con el tutor virtual dando paso a la página de problemas de aplicación con tablas de frecuencia incompletas (figura 6) que el estudiante completa y verifica bajo su propio criterio, al momento de verificar el programa le informa al estudiante el número de celdas que faltan por completar. También se presenta en la misma página las fórmulas clave, una pista para poder empezar con la solución del mismo y el indicador del progreso respecto al diligenciamiento de la tabla.

Figura 6.

Problemas

SECCIÓN I • PROBLEMAS

Completa la tabla de frecuencias

Problema 1
Problema 2
Problema 3
Problema 4

PROBLEMA 2

Se registró el medio de transporte que usan 80 jóvenes de la UIS para llegar a la universidad: A pie (30), Bus (25), Bicicleta (15), Moto (7), Carro particular (3). Completa la tabla de frecuencias completa.

Categoría	f_i	f_r .00	F_i	F_r .00
A pie	30	ej: 0.2	30	ej: 0.2
Bus	25	ej: 0.2	?	0.69
Bicicleta	15	ej: 0.2	70	ej: 0.2
Moto	7	0.09	77	ej: 0.2
Carro particular	3	0.04	?	1.00
Total	80	1.00	80	1.00

FÓRMULAS CLAVE

$$f_r = f_i / n$$

$$F_i = F_{(i-1)} + f_i$$

$$F_r = F_i / n$$

$$n = \sum f_i$$

PISTA

F_i se acumula fila a fila: $F_1 = f_1$, $F_2 = F_1 + f_2$, y así sucesivamente.

PROGRESO

0 / 8 celdas completadas

← Ejemplos Dinámicos

Problema libre → →

Nota: Tabla de frecuencias incompleta con indicador de progreso y panel de fórmulas clave.

Ahora bien, luego de que el estudiante tenga interiorizados y reforzados los conceptos se proponen situaciones con un nivel de inferencia mayor pues el estudiante debe responder a las preguntas haciendo uso de manera implícita de las fórmulas (figura 7), adicional a lo anterior se debe construir la tabla correspondiente a los datos del contexto logrando de esta manera no solo realizar cálculos automatizados sino también el planteamiento de estrategias o contrastes de patrones para llenar más rápido la tabla haciendo uso de reglas teóricas y observación.

Figura 7.

Situaciones

Situación 1

Situación 2

Situación 3

SITUACIÓN 1

En una tienda escolar se vendieron 120 refrigerios durante una semana. Las categorías disponibles son: Jugo de fruta, Sándwich, Frutas frescas, Empanada, Yogur. Los datos de ventas son: Jugo de fruta (38), Sándwich (30), Frutas frescas (22), Empanada (18), Yogur (12).

DATOS DISPONIBLES (N = 120):

Jugo de fruta: 38

Sándwich: 30

Frutas frescas: 22

Empanada: 18

Yogur: 12

Construye la tabla — arrastra las filas o escríbelas en el orden que consideres necesario:

	Categoría	f_i	f_r	F_i	F_r
⋮	Jugo de fruta	38	?	?	?
⋮	Sándwich	38	?	?	?
⋮	Frutas frescas	22	?	?	?
⋮	Empanada	18	?	?	?
⋮	Yogur	12	?	?	?

✓ Verificar tabla completa

↶ Reiniciar

Preguntas de análisis

Responde después de completar tu tabla. Cada pregunta requiere interpretar los datos — no basta con leer un número.

F_r — FRECUENCIA RELATIVA

Pregunta 1: ¿Qué proporción del total de refrigerios representa la combinación de Sándwich y Yogur juntos? Explica por qué esta frecuencia relativa combinada es útil para el tendero al momento de planeear su inventario.

Escribe aquí tu análisis...

F_i — ACUMULADA ABSOLUTA

Pregunta 2: Si ordenas las categorías de mayor a menor venta, ¿cuántos refrigerios acumula la tabla hasta incluir los tres productos más vendidos (F_i acumulada)? ¿Qué conclusión sacas sobre la concentración de las ventas?

Escribe aquí tu análisis...

F_r — ACUMULADA RELATIVA

Pregunta 3: Construye la tabla ordenando las categorías de menor a mayor venta. ¿En qué posición acumulada se supera el 50% de los refrigerios vendidos ($F_r > 0.50$)? ¿Cómo cambia esta interpretación respecto al orden de mayor a menor?

Escribe aquí tu análisis...

Nota: Situación de construcción completa de la tabla a partir de un contexto con información implícita.

7.3.2 Sección II — Tablas de contingencia

La segunda sección del recurso didáctico traslada al estudiante del análisis descriptivo univariado al bivariado. Iniciando con una breve introducción sobre lo que se encontrará en el desarrollo de la misma, dando paso a una situación donde se presenta una tabla sobre la preferencia de actividades extracurriculares según el género en jóvenes de la UIS, por medio de esta, el tutor virtual guía la lectura e interpretación de la tabla de contingencia pasando por frecuencia conjunta, frecuencia marginal y frecuencia condicionada las cuales son institucionalizadas por el tutor virtual cuando el estudiante ha realizado dicho razonamiento. Cabe destacar que el tutor siempre realiza cuestionamientos por medio de los cuales confirma que realmente el estudiante está solucionando de manera consciente el problema. Al igual que en la sección I se plantean situaciones de la vida cotidiana buscando mostrar contextos en los cuales se aplica dicho conocimiento, permitiendo que en la siguiente página se pueda realizar la síntesis de lo trabajado con el tutor.

Figura 8.*Introducción Capítulo II*

CAPÍTULO II · ACTIVIDAD
Institucionalización
06

SITUACIÓN

Actividades Extracurriculares en jóvenes
Se encuestaron 100 jóvenes de Bucaramanga sobre su actividad extracurricular preferida. Los datos se cruzaron con el género para construir la siguiente tabla.

Género \ Actividad	Deportes	Danza	Música	Total Marginal
Hombres	30	5	15	50
Mujeres	10	25	15	50
Total Marginal	40	30	30	100

¿Cómo cambia lo que ves al cambiar la representación?

Ver como Diagrama de Barras

Representación tabular activa

Tutor Virtual UIS
En línea

¡Genial! Vamos a empezar con un pequeño ejercicio. Imagina que tenemos una tabla que muestra las preferencias de actividades extracurriculares según el género en jóvenes de la UIS.

Quiero que me digas cuántas 'Mujeres que prefieren Danza' hay en esta tabla. Tómate tu tiempo para pensarlo.

25

¡Perfecto! Ahora, cuéntame, ¿cómo cruzaste la información en la tabla para llegar a ese número 25? Es importante entender cómo se obtiene esa información.

Escribe tu respuesta aquí...

Enviar ↑

← Presentación
Síntesis → →

Nota: Presentación de la tabla de contingencia sobre preferencias extracurriculares en jóvenes de la UIS.

Una vez realizada la síntesis el estudiante encontrará un apartado sobre las formas de leer una tabla de contingencia donde se muestra un ejemplo para dar idea de las posibles preguntas que se pueden responder de acuerdo con el tipo de representación, de igual manera funcionan los ejemplos dinámicos donde se hace énfasis en los cambios que hay en las gráficas en cada una de las representaciones (Figura 9).

Figura 9.*Formas de leer una tabla de contingencia*

Nota: Visualización de las formas de leer una tabla de contingencia con actualización dinámica del gráfico según el tipo de representación.

Los problemas que se plantean sobre el tema de tablas de contingencia presentan un contexto junto con una pregunta orientadora a resolver, para ello el estudiante debe elegir el tipo de representación que permite responder la pregunta y completar la información de la tabla correspondiente al problema dando la posibilidad de verificar la tabla e incluso consultar al tutor sobre los datos propuestos por el estudiante, lo cual envía la información al tutor virtual para que guíe al estudiante al óptimo desarrollo del problema (Figura 10).

Figura 10.*Problemas*

PROBLEMA 1 DE 3

Se encuestaron 90 estudiantes de la UIS sobre su facultad (Ciencias / Ingenierías / Humanidades) y su modalidad de estudio preferida (Individual / Grupal / Mixta). La tabla tiene algunas celdas ocultas.

PREGUNTA QUE GUÍA ESTA ACTIVIDAD:
¿Dentro de los estudiantes de Ingenierías, cuál es la modalidad más frecuente? ¿Ese porcentaje cambia si lo calculas respecto al total general?

1. Escoge el sistema de representación que mejor responde la pregunta:

Frecuencias absolutas % sobre el total % por fila % por columna

FÓRMULAS CLAVE

Abs.	% total	% fila	% col.
f_{ij}	f_{ij}/N	$f_{ij}/f_{i\cdot}$	$f_{ij}/f_{\cdot j}$

2. Completa las celdas vacías con el sistema que escogiste:

↓ / →	Individual	Grupal	Mixta	Total
Ciencias	14	8	?	30
Ingenierías	?	18	?	40
Humanidades	6	?	4	20
Total	30	36	24	90

3. Justifica tu elección:

Explica por qué escogiste ese sistema y qué revela sobre los datos...

[Verificar tabla](#) [Consultar tutor](#)

Nota: Problema con pregunta orientadora, selector de tipo de representación y opción de verificación de la tabla.

Los retos funcionan de manera similar a los problemas, solo que con la variante de que se le brinda al estudiante una información inicial, aumentando el nivel de exigencia ya que además de elegir la representación que debe usar para responder a la pregunta, también tiene que interpretar los datos para completar la tabla. La actividad se plantea en tres retos con contextos distintos, y en cada uno el estudiante no solo debe acertar en las celdas, sino justificar por escrito la elección del sistema de representación, lo que exige una argumentación explícita más allá del cálculo. Como apoyo, se mantiene visible un panel de fórmulas clave, y el estudiante cuenta con dos formas de retroalimentación independientes: la verificación inmediata de sus celdas y la posibilidad de consultar al tutor virtual para profundizar en su razonamiento (Figura 11).

Figura 11.*Retos*

RETO 1 DE 3

Se encuestaron 90 estudiantes de la UIS sobre su medio de transporte (Bus / Bicicleta / A pie) y su puntualidad (Siempre / A veces / Nunca).

- 30 estudiantes usan Bus.
- De quienes usan Bicicleta, la mitad llega Siempre a tiempo.
- 10 estudiantes van A pie y llegan A veces.
- Solo 3 estudiantes van A pie y Nunca llegan a tiempo.
- En total, 24 estudiantes llegan Siempre a tiempo.

CONSTRUYE LA TABLA QUE RESPONDA A:
¿Qué medio de transporte se asocia con mejor puntualidad?

1. Escoge el sistema de representación adecuado:

Frecuencias absolutas

% sobre el total

% por fila

% por columna

FÓRMULAS CLAVE

Abs.
 f_{ij}

% total
 f_{ij}/N

% fila
 $f_{ij}/f_{i\cdot}$

% col
 $f_{ij}/f_{\cdot j}$

2. Construye la tabla completa (todas las celdas vacías):

↓ / →	Siempre	A veces	Nunca	Total fila
Bus	?	?	?	30
Bicicleta	?	?	?	25
A pie	?	?	?	15
Total col	24	32	14	90

3. Preguntas de análisis:

% POR FILA

Pregunta 1: Calcula el porcentaje de puntualidad "Siempre" dentro de cada medio de transporte. ¿Cuál tiene la proporción más alta? ¿Por qué no es suficiente comparar los conteos absolutos?

Escribe aquí tu análisis...

% TOTAL

Pregunta 2: Del total de 90 estudiantes, ¿qué porcentaje va en Bus Y llega Siempre? ¿Ese dato por sí solo te permite concluir que el bus es el mejor medio para ser puntual?

Escribe aquí tu análisis...

N4 • CAUSALIDAD

Pregunta 3: ¿Podría haber una variable que explique tanto la elección del transporte como la puntualidad, sin que una cause la otra directamente? Propón al menos una.

Escribe aquí tu análisis...

✓ Verificar tabla

Consultar tutor

Nota: Reto de construcción con panel de fórmulas, verificación de celdas y consulta al tutor virtual.

Para finalizar se plantean dos situaciones en las cuales se propone una afirmación a explorar pidiendo al estudiante que complete la tabla con datos de tal manera que pueda reflejar la afirmación presentada al inicio, esta información es enviada al tutor para ser debatida buscando la manera de reafirmar lo aprendido durante el capítulo.

Figura 12.

Situaciones problema

Situación 1

Situación 2

SITUACIÓN 1

En la UIS se quiere estudiar si existe relación entre el **turno de clase preferido** (Mañana / Tarde / Noche) y el **rendimiento académico** (Alto / Bajo). Se encuestaron **90 estudiantes**.

AFIRMACIÓN A EXPLORAR:
"Los estudiantes que prefieren el turno de mañana tienden a tener mejor rendimiento académico."

Distribuye los 90 encuestados como creas que refleja la afirmación. Arrastra o escribe en cada celda:

MAÑANA
38 encuestados

TARDE
35 encuestados

NOCHE
25 encuestados

Turno / Categoría	Alto	Bajo	Total fila
Mañana	?	?	30
Tarde	?	?	35
Noche	?	?	25
Total col	0	0	90

Total ingresado: 0 / 90

Enviar al tutor para discutir

RECUERDA

% fila
 f_{22}/f_{2+}

% col
 f_{22}/f_{+2}

¿Asociación?
¿Cómo saberlo?

Tutor Virtual UIS

En línea

😊 Distribuye los datos y envíalos al tutor. Él te hará preguntas para que descubras si tu distribución realmente prueba la afirmación.

¡Perfecto! Tienes la libertad de distribuir los datos como prefieras dentro de la tabla de contingencia.

¿Cuál es la *afirmación* que te gustaría explorar o demostrar con tu *distribución* de datos?

Escribe tu razonamiento o pregunta... Enviar ↑

Nota: Situación de exploración libre con afirmación inicial y envío de la tabla construida al tutor para su discusión.

7.3.3 Sección III — Prueba Chi-cuadrado

Se establece al inicio un conflicto cognitivo entre dos personajes que analizan la misma tabla de datos y llegan a conclusiones opuestas sobre si existe o no una asociación entre las

variables, planteando la necesidad de una herramienta que permita determinar con mayor certeza dicha relación. A través de tres preguntas abiertas, se obliga al estudiante a expresar su propio razonamiento antes de enviarlo al tutor, quien no discute ni contradice su postura sino que profundiza progresivamente mediante nuevas preguntas calibradas según el nivel de razonamiento que el estudiante demostró en sus respuestas, evitando intencionalmente nombrar la prueba de chi-cuadrado en esta etapa para sostener la necesidad conceptual antes de resolverla formalmente.

Esta dinámica corresponde a la fase de Acción a-didáctica de Brousseau, en la que el estudiante enfrenta el conflicto sin ninguna herramienta formal todavía (Figura 13).

Figura 13.

Introducción

CAPÍTULO III · INTRODUCCIÓN
¿Podemos estar seguros?

¿Puede la tabla mentir?

Has aprendido a construir tablas de contingencia y a calcular porcentajes por fila, columna y total. Pero hay algo que esas herramientas **no pueden responder**:

Ana estudió la misma tabla y dijo:
"Hay una relación clara entre el transporte y la puntualidad."

Carlos analizó la misma tabla y dijo:
"La diferencia podría ser puro azar de la muestra."

¿Quién tiene razón? ¿Cómo se podría decidir con certeza?

La tabla que analizaron (90 estudiantes):

Transporte ↓ / Puntualidad →	Siempre	A veces	Nunca	Total
Bus	10	14	6	30
Bicicleta	12	8	5	25
A pie	2	10	3	15
Total	24	32	14	90

Antes de continuar — escribe tu intuición

PREGUNTA 1
Mirando la tabla, ¿crees que el medio de transporte y la puntualidad están relacionados? ¿Por qué?

porque se ve la puntualidad con transporte

PREGUNTA 2
¿Podría ser que esa diferencia aparezca solo porque encuestaron a *ese grupo específico* de 90 estudiantes, y con otro grupo los números fueran distintos?

si

PREGUNTA 3
¿Qué necesitarías saber o calcular para estar *seguro* de que la relación es real y no producto del azar?

no sé

Enviar al tutor

Nota: conflicto cognitivo inicial entre dos posturas opuestas sobre la misma tabla, con espacio de respuesta abierta antes de la intervención del tutor.

Con lo anterior es necesario que el estudiante construya el concepto de “frecuencia esperada” de manera empírica. Para ello se le pide al estudiante que simule cómo se vería la tabla si no hubiera relación, respetando los totales de fila y columna (Figura 14).

Figura 14.

Construyendo lo esperado

CAPÍTULO III · ACCIÓN
Construye lo esperado
15

SITUACIÓN A-DIDÁCTICA

Usa la misma tabla de 90 estudiantes. Los **totales de fila y columna están fijos** — no puedes cambiarlos. Tu tarea: distribuir los 90 casos en las celdas internas como si el transporte y la puntualidad **no tuvieran ninguna relación** entre sí.

PREGUNTA GUÍA:
Si no hay ninguna relación entre las dos variables, ¿cómo se repartirían los estudiantes en cada celda?

Marginales fijos (no editables):

Bus: 30
Bicicleta: 25
A pie: 15
Siempre: 24
A veces: 32
Nunca: 14
N = 90

↓/-	Siempre	A veces	Nunca	Total fila
Bus	?	?	?	30
Bicicleta	?	?	?	25
A pie	?	?	?	15
Total col	24	32	14	90

PREGUNTA DE INTERPRETACIÓN (N3)

Imagina que en la realidad los datos observados se parecieran mucho a esta tabla "sin relación" que construiste. ¿Qué dirías sobre la conexión entre el medio de transporte y la puntualidad? ¿Y si fueran muy distintos?

Tu interpretación...

Nota: Actividad de construcción empírica de la tabla esperada bajo el supuesto de independencia.

Se hace la transición a un nuevo caso de estudio para formalizar el cálculo paso a paso dando a entender que se toma la tabla observada para compararla con la tabla esperada pidiendo al estudiante que identifique en qué celda hay más estudiantes de los esperados apoyándose con el tutor (Figura 15).

Figura 15.*Ejemplo guiado*

CAPÍTULO III · EJEMPLO ● Construyendo E_{ij} 16

EJEMPLO GUIADO - FRECUENCIAS ESPERADAS
Se encuestaron 60 estudiantes sobre su **programa** (Matemáticas / Estadística) y su **uso de software** (R / Python / SPSS). ¿Cómo calcular las frecuencias que *esperaríamos* si no hubiera ninguna relación entre programa y software?

← Paso anterior Paso 1 / 5 Siguiente paso →

Paso 1 — La tabla observada (O_{ij})
Estos son los datos que recogiste: cuántos estudiantes de cada programa usan cada software. Estos valores se llaman **frecuencias observadas O_{ij}** .

Ver tabla como:
☐ Observada (O_{ij})
☐ Esperada (E_{ij})
☒ Diferencia

↓/→	R	Python	SPSS	Total
Matemáticas	20	8	2	30
Estadística	6	12	12	30
Total	26	20	14	60

INTERPRETACIÓN (N3)
Activa la vista "Diferencia". ¿En qué celda hay más estudiantes de los esperados? ¿Qué te dice eso sobre la relación entre el programa y el software preferido?

estadística

Nota: Ejemplo guiado de comparación entre tabla observada y tabla esperada, con apoyo del tutor virtual.

Una vez calculada la diferencia entre lo observado y lo esperado en cada celda se evidencia que hay valores tanto positivos como negativos. Al sumarlos tal cual, el resultado es cero (o muy cercano a cero), no porque la tabla se acerque a la independencia, sino porque los signos se cancelan entre sí y ocultan el verdadero tamaño del alejamiento. Ante este panorama, se le pide al estudiante proponer con sus propias palabras una operación matemática que convierta cualquier diferencia en un valor positivo. El sistema no acepta que el estudiante simplemente nombre la operación correcta sino que el tutor virtual le exige argumentar por qué esa propuesta resuelve efectivamente el problema de la cancelación de signos, ofreciendo hasta tres pistas progresivamente más concretas si la justificación inicial resulta insuficiente. Solo una vez que el estudiante logra sustentar su propuesta se revela formalmente la operación y se avanza

a la construcción del estadístico de prueba X^2 , cuyo resultado se discute posteriormente con el tutor virtual (Figura 16). Seguido a lo anterior se define lo ya encontrado y adicional a ello se abre un espacio para que el estudiante exprese con sus palabras qué es una frecuencia esperada y que nos dice de X^2 (Figura 17).

Figura 16.

Construyendo X^2

DESCUBRIMIENTO — ¿CÓMO MEDIR EL ALEJAMIENTO TOTAL?

Ya tienes las diferencias $O_{ij} - E_{ij}$ de cada celda (ejemplo de 60 estudiantes, Programa $\times$ Software). Parece natural sumarlas todas para medir cuánto se aleja la tabla de la independencia. Antes de ver ninguna fórmula, vamos a probarlo.

Paso 1 — Suma las diferencias tal cual

Estas son las diferencias $O_{ij} - E_{ij}$ de las 6 celdas. Súmalas todas (con su signo) y escribe el resultado:

Celda	O_{ij}	E_{ij}	$O_{ij} - E_{ij}$
Matemáticas/R	20	13.00	+7.00
Matemáticas/Python	8	10.00	-2.00
Matemáticas/SPSS	2	7.00	-5.00
Estadística/R	6	13.00	-7.00
Estadística/Python	12	10.00	+2.00
Estadística/SPSS	12	7.00	+5.00

Suma de las diferencias (con signo) =

Exacto: la suma da **cero** (o casi). Aunque la tabla claramente se aleja de lo esperado, las diferencias positivas y negativas *se cancelan*. Ese es el problema que hay que resolver. 🐞

Paso 2 — El problema de los signos

La suma dio (aprox.) **cero**, aunque la tabla claramente se aleja de lo esperado. Las diferencias positivas y negativas *se cancelan* entre sí. Necesitamos una forma de medir el tamaño del alejamiento sin que el signo lo borre.

¿Qué operación matemática convierte cualquier número (positivo o negativo) en un valor positivo que crece con la magnitud? Escribe tu idea:

Nota: Construcción del estadístico X^2 a partir de la propuesta argumentada del estudiante sobre el tratamiento de las diferencias.

Figura 17.Interpretación de χ^2

Frecuencias Esperadas (E_{ij})

$$E_{ij} = (f_{i \cdot} \times f_{\cdot j}) / N$$

Representan lo que *ocurriría* si las dos variables fueran completamente independientes. Si $O_{ij} \approx E_{ij}$ en todas las celdas, no hay evidencia de asociación.

⚠ Condición: $E_{ij} \geq 5$ en todas las celdas para que la prueba sea válida.

Estadístico χ^2

$$\chi^2 = \sum (O_{ij} - E_{ij})^2 / E_{ij}$$

Mide el *alejamiento global* entre lo observado y lo esperado. $\chi^2 = 0$ significa independencia perfecta. Cuanto mayor sea χ^2 , mayor evidencia de asociación.

Contribución por celda

$$(O_{ij} - E_{ij})^2 / E_{ij}$$

Las celdas con mayor contribución son las que más se alejan de la independencia. Son la "firma" de la asociación.

Tabla resumen del ejemplo trabajado

Celda	O_{ij}	E_{ij}	$(O-E)^2/E$
Matemáticas/R	20	13.00	3.7692
Matemáticas/Python	8	10.00	0.4000
Matemáticas/SPSS	2	7.00	3.5714
Estadística/R	6	13.00	3.7692
Estadística/Python	12	10.00	0.4000
Estadística/SPSS	12	7.00	3.5714
χ^2 total			15.4813

Explica con tus propias palabras qué significa E_{ij} y qué nos dice χ^2 . El tutor cuestionará hasta que tu explicación sea sólida.

Explica con tus propias palabras qué son las frecuencias esperadas y qué nos dice el estadístico χ^2 cuando es grande vs. cuando es pequeño...

[Enviar al tutor de consolidación](#)

[Continuar: gl y valor crítico →](#)

Nota: Captura del espacio de síntesis intermedia: el estudiante verbaliza con sus propias palabras el significado de la frecuencia esperada y del estadístico χ^2 .

En la fase de validación, el propósito es que el estudiante comprenda el significado de los grados de libertad y su impacto al valor crítico, habilitando al estudiante para efectuar inferencias estadísticas rigurosas e incluso concluir el contraste de hipótesis.

Figura 18.**Introducción grados de libertad y valor crítico**

CAPÍTULO III · VALIDACIÓN ● ¿Cuándo es suficiente?

19

ACTIVIDAD – GL Y VALOR CRÍTICO

El mismo valor de χ^2 puede ser evidencia fuerte o débil dependiendo del tamaño de la tabla. Aquí explorarás por qué los **grados de libertad** importan.

Antes de revelar: los tres casos tienen el mismo $\chi^2 = 6.5$ pero distinto tamaño de tabla. Conjetura: ¿en cuál crees que hay *más* evidencia de asociación? Anótalo en la reflexión de abajo y luego pulsa "Revelar valores críticos".

CASO 1
Tabla 2×2 (4 celdas)
gl = 1
 $\chi^2 = 6.5$
 $\chi^2_{crítico} = ?$
¿Se rechaza H_0 ?

CASO 2
Tabla 2×4 (8 celdas)
gl = 3
 $\chi^2 = 6.5$
 $\chi^2_{crítico} = ?$
¿Se rechaza H_0 ?

CASO 3
Tabla 3×3 (9 celdas)
gl = 5
 $\chi^2 = 6.5$
 $\chi^2_{crítico} = ?$
¿Se rechaza H_0 ?

Revelar valores críticos

REFLEXIÓN
Con el mismo $\chi^2 = 6.5$ pero distintos grados de libertad, ¿en cuál de los tres casos hay más evidencia de asociación? ¿Por qué crees que el tamaño de la tabla afecta la interpretación?

Ordénalos de más a menos evidencia y explica...

Tabla de valores críticos ($\alpha = 0.05$)

gl	1	2	3	4	5	6
$\chi^2_{crítico}$	3.841	5.991	7.815	9.488	11.07	12.59

APLICA AL EJEMPLO
En el ejemplo del Cap III (Programa × Software, 2×3), calcula los grados de libertad. Luego ubica el χ^2 que calculaste en pág. 18 y compara con el valor crítico. ¿Qué concluyes?

gl = (filas-1)(col-1) = ? χ^2 calculado = ? Valor crítico = ?

¿Rechazas H_0 ? ¿Qué concluyes sobre la relación programa-software?

Verificar gl y valor crítico **Consultar tutor**

Nota: Introducción a los grados de libertad y el valor crítico como herramientas para completar el contraste de hipótesis.

Se plantea un ejemplo guiado sobre la posible asociación entre el nivel inglés y la participación en intercambios con el cual se muestra el paso a paso en los cuales se pueden realizar preguntas al tutor. En cada uno de los pasos se muestra el planteamiento de hipótesis, verificación del supuesto, el cálculo del estadístico de prueba, la elección de los grados de libertad, la realización de la comparación con los valores críticos y la conclusión estadística después de todo el proceso de comprobación (figura 19).

Figura 19.

Ejemplo guiado

CAPÍTULO III · EJEMPLO
Ciclo completo
20

EJEMPLO GUIADO – DE LA TABLA A LA CONCLUSIÓN

Se encuestaron 75 estudiantes sobre su nivel de inglés (Básico / Avanzado) y su participación en intercambios (Sí / No). ¿Existe asociación estadística entre estas variables?

← Paso anterior
Paso 3 / 8
Siguiente paso →

Paso 3 — Contribución de cada celda

$(O_{ij} - E_{ij})^2 / E_{ij}$:

Celda	O	E	Contrib.
Básico/Sí	3	8.80	3.8227
Básico/No	27	21.20	1.5868
Intermedio/Sí	7	7.33	0.0151
Intermedio/No	18	17.67	0.0063
Avanzado/Sí	12	5.87	6.4120
Avanzado/No	8	14.13	2.6616

Preguntarle al tutor sobre este paso

Tutor Virtual UIS
En línea

Dado que el valor de χ^2 calculado (14.5046) es mayor que el valor crítico (5.991), se rechaza H_0 .

Ahora, reflexionando un poco más allá de los números, ¿qué podrías decir sobre la relación entre el nivel de inglés y la participación en el intercambio? ¿Crees que hay algún otro factor que podría estar influyendo en esta relación?

¿Tienes alguna duda?
Enviar ↑

← gl y valor crítico

¿Asociación = causalidad? → →

Nota: Ejemplo guiado paso a paso del proceso completo de la prueba Chi-cuadrado sobre la asociación entre nivel de inglés y participación en intercambios.

Buscando marcar la diferencia entre asociación y causalidad se presentan tres paradojas clásicas de correlación como lo son zapatos y crimen, helado y ahogamientos, internet y

esperanza de vida, en las que el estudiante debe identificar qué variable oculta podría explicar cada asociación, obligándolo a construir una abstracción de la relación entre variables más allá del simple cálculo matemático. El tutor virtual no válida la primera respuesta que coincida con el ejemplo del libro: si el estudiante solo lo repite, le exige proponer un caso propio del contexto colombiano. A esto se suman tres preguntas adicionales de mayor profundidad sobre cuándo sería irresponsable concluir causalidad, cómo afectaría al estadístico de prueba un aumento considerable en el tamaño de la muestra, y qué papel juega el diseño del estudio (experimental u observacional) en la posibilidad de establecer causalidad, potenciando así el pensamiento crítico al cuestionar, en distintos contextos, si los valores matemáticos realmente reflejan una relación causal entre las variables o si sólo evidencian una asociación estadística que aún debe interpretarse (Figura 20).

Figura 20.

Asociación vs causalidad

CAPÍTULO III · REFLEXIÓN

N4 · Pensamiento crítico

21

¿ASOCIACIÓN IMPLICA CAUSALIDAD?

Rechazar H_0 significa que hay evidencia estadística de asociación. Pero eso **no significa que una variable cause la otra**. Analiza los tres casos reales a continuación.

Caso 1 — Zapatos y crimen

Un estudio encontró asociación significativa entre el número de pares de zapatos que posee una persona y su probabilidad de cometer un delito.

¿Qué variable oculta podría explicar esta asociación?

Variable oculta posible...

Caso 2 — Helado y ahogamientos

Los datos muestran que las ventas de helado y los ahogamientos en piscinas están asociados (χ^2 significativo).

¿Qué variable oculta podría explicar esta asociación?

Variable oculta posible...

Caso 3 — Internet y esperanza de vida

Los países con mayor penetración de internet tienen mayor esperanza de vida (asociación muy significativa).

¿Qué variable oculta podría explicar esta asociación?

Variable oculta posible...

LID... ESTADÍSTICA DESCRIPTIVA - UIS - 2020

PREGUNTA N4 · LIMITACIONES

¿Cuándo sería estadística y éticamente irresponsable concluir causalidad a partir de un χ^2 significativo? Da un ejemplo propio del contexto colombiano.

Tu análisis...

PREGUNTA N4 · TAMAÑO MUESTRAL

Si repitieras el estudio de puntualidad con 9000 estudiantes (manteniendo las mismas proporciones), ¿qué le pasaría a χ^2 ? ¿Eso lo haría más o menos útil para tomar decisiones?

Tu análisis...

PREGUNTA N4 · DISEÑO DE ESTUDIO

¿Qué tipo de estudio adicional (diferente a una encuesta) te permitiría ir más allá de la asociación estadística y explorar si existe una relación causal?

Tu análisis...

Enviar al tutor

Nota: Actividad de análisis crítico sobre correlaciones falsas o engañosas y distinción entre asociación estadística y causalidad.

Por último, se abre un espacio de aplicación libre organizado en cuatro etapas: primero, el estudiante define su problema estadístico describiendo el contexto de sus datos, la pregunta que desea responder y las variables involucradas; segundo, carga su propia base de datos en formato CSV o Excel; tercero, explora sus datos mediante una vista previa y la construcción de gráficas con las variables que él mismo selecciona, interpretando el patrón que observa; y cuarto, sin que el tutor se lo indique debe decidir cuál de las herramientas trabajadas a lo largo del libro corresponde a su problema, y justificar dicha elección antes de iniciar la conversación con el tutor virtual, quien a partir de ese momento acompaña el análisis retroalimentando las decisiones e interpretaciones del estudiante sobre sus propios datos (Figura 21).

Figura 21.*Aplicación libre*

CAPÍTULO III · APLICACIÓN LIBRE ● Tu propio estudio estadístico 24

1 Define tu problema estadístico

Antes de subir cualquier dato, reflexiona sobre lo que quieres investigar. El análisis estadístico comienza con una pregunta, no con una tabla.

¿De dónde vienen tus datos? (describe el contexto real)

Ej: encuesta a 80 estudiantes de mi colegio sobre hábitos de estudio y rendimiento académico...

¿Qué pregunta estadística quieres responder?

Ej: ¿Existe relación entre el tiempo de estudio diario y el rendimiento académico?

¿Qué variables tiene tu tabla y de qué tipo crees que son?

Ej: Género (categórica), Horas de estudio (numérica), Rendimiento (categórica: Alto/Medio/Bajo)...

Continuar --> Subir mi tabla

← Situación libre

El tutor aparecerá aquí cuando inicies el análisis en el paso 4.

Nota: Espacio de aplicación libre con carga de datos propios, construcción de gráficas y análisis acompañado por el tutor virtual.

7.4 Revisión de estilo y contenido por panel de expertos sobre el capítulo de libro diseñado

Con el propósito de evaluar la validez del contenido y estilo del recurso diseñado, se estructuró un proceso de evaluación mediante el juicio de un panel de expertos conformado por profesores de Estadística, licenciados en matemáticas y estudiantes de Didáctica de la probabilidad y Estadística de la Universidad Industrial de Santander (UIS). Para guiar esta validación de manera sistemática, se diseñó un instrumento tipo formulario cuyas preguntas se derivan directamente de las diferentes dimensiones que componen este trabajo. En la tabla 9

presentada a continuación se detallan las dimensiones evaluadas articulando lo planteado en el marco teórico.

Tabla 9.

Instrumento de valoración para el contenido del capítulo diseñado

Dimensión evaluada	Fundamento teórico	Ítems	Indicador
Rigurosidad y pertinencia del contenido estadístico	Batanero (2001); Quevedo (2011)	4-8	Precisión conceptual, notacional y procedimental de los elementos estadísticos presentados
Coherencia con la Teoría de Situaciones Didácticas	Brousseau (citado en Espinosa & Bareño, 2018)	9-13	Grado en que las actividades preservan las condiciones de una situación a-didáctica: autonomía, devolución de consecuencias e institucionalización diferida
Progresión de los niveles de comprensión	Curcio (1989, citado en Batanero et al., 1994); Aoyama (2007,	14-18	Grado en que el capítulo conduce gradualmente al estudiante desde la lectura literal hasta la lectura crítica de la tabla de contingencia

	citado en Arteaga et al., 2011)		
Juicios de asociación y prevención de errores conceptuales	Batanero et al. (1994)	19-21	Capacidad del recurso para evitar las concepciones causal, local y determinista en el análisis bivariado
Mediación del tutor virtual (IA)	Marquès (2011); Batanero (2019)	22-24	Calidad pedagógica de la retroalimentación entregada por el recurso como medio didáctico
Usabilidad e interfaz	Toapanta & Yugcha (2023)	25-27	Claridad de navegación y adecuación del lenguaje al público objetivo
Valoración global		28-32	Pertinencia general del recurso y triangulación cualitativa mediante preguntas abiertas

7.4.1 Análisis de resultados

El instrumento de validación fue aplicado a un panel de seis evaluadores conformado por estudiantes egresados del programa de Licenciatura en Matemáticas y estudiantes de Didáctica de la Probabilidad y Estadística durante el semestre 2026-I. Resulta pertinente señalar que la totalidad del panel (6 de 6) respondió negativamente a la pregunta sobre el uso previo de algún

recurso digital interactivo para la enseñanza o el aprendizaje de la estadística. Si bien no fue un criterio de selección de los evaluadores, este resultado sugiere que este tipo de recursos didácticos e interactivos resulta todavía escaso en el contexto de formación de los participantes lo cual refuerza la pertinencia del aporte que representa el desarrollo del LID. La Tabla 10 presenta las medias y desviaciones estándar obtenidas por ítem, organizadas según las siete dimensiones de la Tabla de operacionalización del instrumento.

Tabla 10.

Resultados de la validación por panel de expertos, por dimensión e ítem (n=6, escala 1-5; M= media; DE= desviación estándar)

Dimensión	Ítem	Descripción breve	M	DE
Rigurosidad y pertinencia del contenido estadístico	3	Definición rigurosa de frecuencias	4.33	0.52
	4	Consistencia de la notación matemática	4.83	0.41
	5	Pertinencia de ejemplos y datos	4.50	0.84
	6	Precisión de la prueba chi-cuadrado	4.33	0.82
	Promedio de la dimensión		4.50	
Coherencia con la TSD	8	Autonomía en problema contextualizado	4.83	0.41

Dimensión	Ítem	Descripción breve	M	DE
	9	El tutor evita dar la respuesta directa	4.17	0.75
	10	Institucionalización diferida	4.50	0.55
	11	Formulación de hipótesis propias	4.33	0.52
	Promedio de la dimensión		4.46	
Progresión niveles Curcio/Aoyama	13	Lee los datos	4.50	0.55
	14	Lee dentro de los datos	4.83	0.41
	15	Lee más allá de los datos	4.67	0.52
	16	Lee detrás de los datos	4.33	0.82
	17	Progresión gradual entre niveles	4.67	0.52
	Promedio de la dimensión		4.60	
Juicios de asociación	18	Distingue correlación de causalidad	4.33	0.52
	19	Evita análisis aislado de datos	4.67	0.52
	20	Cuestiona relaciones deterministas	4.50	0.55
	Promedio de la dimensión		4.50	
Mediación del tutor virtual (IA)	21	Preguntas claras y pertinentes	4.67	0.52
	22	Tono adecuado ante errores	4.67	0.52

Dimensión	Ítem	Descripción breve	M	DE
	23	Retroalimentación genuina	4.67	0.52
	<i>Promedio de la dimensión</i>		4.67	
Usabilidad e interfaz	24	Navegación clara e intuitiva	4.33	0.82
	25	Visualización correcta de tablas/gráficos	4.83	0.41
	26	Lenguaje apropiado al público objetivo	4.83	0.41
	<i>Promedio de la dimensión</i>		4.67	
Valoración global	27	Aporte pertinente en general	4.50	0.55
	<i>Promedio de la dimensión</i>		4.50	
Promedio general del instrumento			4.56	

Como se observa en la Tabla 10 las siete dimensiones evaluadas se agrupan en un rango relativamente pequeño que está entre 4.46 y 4.67 (sobre una escala de 5.00), demostrando una valoración favorable del panel hacia el capítulo. Por un lado, las dimensiones de Mediación del tutor virtual y Usabilidad e interfaz obtuvieron una media de 4.67 siendo estas las más altas de

todas las medias por dimensión indicando no sólo una valoración favorable sino un consenso entre evaluadores dado los bajos valores en cuanto a desviación estándar.

Por otro lado, la dimensión de Coherencia con la TSD, aunque también valorada favorablemente (4.46), registró la media más baja de todo el instrumento en el ítem que evalúa si el tutor evita dar la respuesta directa (4.17, DE = 0.75). Este resultado se explica en gran medida por un hallazgo cualitativo clave de uno de los evaluadores quien señaló que el recuadro de "Fórmulas Clave" en la actividad bivariada de Color $\times$ Personalidad se presentaba de manera anticipada. Según el experto, este recurso "revela las herramientas de resolución antes de tiempo" e "interrumpe el proceso en el que el alumno debe construir su propia conjetura".

En cuanto a la variabilidad de las respuestas, las mayores desviaciones estándar se concentraron en cuatro ítems específicos entre 0.82 y 0.84: la pertinencia de los ejemplos (ítem 5), la precisión de la prueba Chi-cuadrado (ítem 6), la lectura crítica "detrás de los datos" (ítem 16) y la claridad en la navegación (ítem 24)

La dispersión en el ítem 5 evidenciaría un desacuerdo en el panel respecto al tipo de contextos seleccionados para las actividades pero dado el bajo valor no es relevante. Mientras que algunos evaluadores valoraron positivamente el uso de ejemplos basados en situaciones cotidianas y subjetivas de los estudiantes, como la relación entre el color favorito y la personalidad por su potencial motivador, otro docente mostró mayor cautela sugiriendo ejemplos con variables de mayor control y dificultad para no interferir con la comprensión del concepto matemático.

Del mismo modo, la variabilidad en el ítem 6 refleja la tensión natural entre la rigurosidad formal de la prueba Chi-cuadrado y el enfoque intuitivo-computacional que se adoptó en el LID para facilitar su visualización y comprensión a nivel de pregrado.

Por su parte, en el ítem 16 responde a una diferencia de criterios sobre el alcance del recurso: algunos expertos consideraron que el andamiaje del tutor virtual y las preguntas abiertas son suficientes para movilizar la lectura "detrás de los datos", mientras que otros evaluadores consideraron que las actividades todavía priorizan un nivel de interpretación intermedio ("leer entre los datos") y que se requiere mayor énfasis en cuestionar los sesgos del origen de la información.

Siguiendo, la dispersión en el ítem de navegación (ítem 24) sí se vincula directamente a un aspecto técnico concreto reportado en las observaciones cualitativas, donde algunos expertos manifestaron confusión ocasional sobre "cómo avanzar a la siguiente pantalla". Esto se resolvió optimizando las transiciones visuales y los botones de control en la hoja de estilos (CSS).

Finalmente respecto a la valoración global, cinco de los seis evaluadores (83.3%) recomendarían el capítulo "con ajustes", y uno (16.7%) lo recomendaría sin reservas; ningún evaluador emitió una recomendación negativa. Entre las fortalezas señaladas con mayor frecuencia se destacan la variedad de representaciones (tabular y gráfica), el acompañamiento del tutor virtual, y el rigor matemático combinado con un formato interactivo. Como principal oportunidad de mejora, además del ajuste ya discutido sobre el momento de presentación de fórmulas clave, uno de los evaluadores señaló una inconsistencia notacional puntual en el uso del símbolo de sumatoria aplicado sobre n , en lugar de expresar la igualdad $\sum f_i = n$.

7.4.2 Fortalezas identificadas

El Tutor virtual como mediador: El panel destacó la integración del tutor con IA como un recurso para generar conocimiento válido y confiable. Señalan que "fomenta la reflexión profunda de los contenidos y su aplicación" y promueve la construcción activa del conocimiento al guiar sin dar respuestas explícitas. Visualización Dinámica: La disposición de tener diferentes representaciones (tablas y gráficos) fue calificada positivamente, argumentando que "contribuye a la formación compleja de la imagen mental de un objeto matemático" y capta la atención del estudiante de forma interactiva.

7.4.3 Debilidades y ajustes implementados

Respecto al nivel de dificultad de los problemas planteados, se recibió la observación: "El nivel de los problemas planteados es medio o básico; sería interesante incorporar un problema opcional con mayor dificultad al final de cada tema." Frente a esta observación es pertinente aclarar que el libro ya contaba desde su diseño original con una situación libre final en la cual el estudiante debe aplicar de manera integrada todos los conocimientos adquiridos a lo largo del capítulo. Si bien esto no reproduce la sugerencia puntual de un reto opcional al cierre de cada tema, sí atiende la necesidad de fondo señalada por el evaluador: ofrecer un espacio de mayor exigencia cognitiva más allá de los problemas de aplicación directa.

En cuanto a la notación matemática, uno de los evaluadores identificó una inconsistencia puntual en la página 5, donde se aplicaba el símbolo de sumatoria sobre "n" donde esta representa el valor de la constante tamaño de la muestra, en lugar de expresar correctamente la

igualdad $\Sigma f_i = n$. Esta observación fue verificada directamente sobre el recurso y la notación fue corregida en consecuencia.

Referente al diseño de la interfaz y secuenciación de los contenidos, atendiendo a una observación sobre la página 9, se quitó el recuadro "FÓRMULAS CLAVE" el cual se presentaba antes de las preguntas de indagación ya que la ubicación inicial de las fórmulas interrumpía o limitaba el proceso de descubrimiento autónomo así como el proceso de formulación de conjeturas e hipótesis propias del estudiante. Se decidió reestructurar la interfaz asegurando un proceso de acompañamiento que respete la fase de exploración del alumno antes de la formalización teórica.

La interacción con la IA también tuvo la siguiente observación: "En algunas ocasiones, al responder una pregunta el tutor reitera lo mencionado en preguntas pasadas, lo que puede llegar a ser redundante." Esta observación responde a una decisión de diseño intencional: el tutor virtual no avanza a la formalización de un concepto hasta verificar que el estudiante demuestre comprensión genuina, incluso si esto implica profundizar más de una vez sobre la misma idea antes de continuar. Este comportamiento es consistente con el principio de institucionalización diferida de la Teoría de Situaciones Didácticas donde el saber formal solo se nombra después de que el estudiante lo ha construido por sí mismo.

Adicionalmente y como complemento al recurso principal se desarrolló un manual de usuario con la finalidad de que el estudiante comprenda mejor el funcionamiento del LID, particularmente la interactividad propia del tutor virtual y las distintas herramientas presentes en el mismo (Anexo 2)

8. Conclusiones

La construcción y lectura de resúmenes tabulares no es un tema que goce de protagonismo en los libros de texto tradicionales de Estadística a nivel de Educación superior. Tras la realización de este proyecto podemos afirmar que además del componente algorítmico surgen dos componentes a nivel cognitivo a los cuales el profesor debe atender por un lado la transnumeración como proceso mental que permite dar sentido a las representaciones obtenidas cuando se transforman los datos a otros formatos y por el otro la capacidad de análisis crítico mezclando elementos técnicos y del contexto. Para diseñar una actividad de clase eficiente, además nos resultó muy útil vincular la teoría de las situaciones didácticas a través de la cual pudimos imprimir un enfoque tanto experimental como activo al recurso diseñado siguiendo un proceso gradual de avance.

Los resultados obtenidos a través de la validación por panel de expertos permiten afirmar que el Libro Interactivo Digital (LID) para la enseñanza de tablas de contingencia constituye un recurso didáctico pertinente, coherente con la Teoría de Situaciones Didácticas y con los niveles progresivos de comprensión de datos propuestos por Curcio y Aoyama. El promedio general obtenido (4.56/5) y la reducida dispersión entre las siete dimensiones evaluadas (4.46-4.67) indican que la valoración favorable no se concentra en un aspecto aislado del diseño, por ejemplo solo en la interfaz o solo en el contenido matemático, sino que se sostiene de manera consistente a través de las dimensiones pedagógicas, tecnológicas y disciplinares que fueron objeto de evaluación. Esto permite concluir que el objetivo de validar, mediante juicio de expertos, la pertinencia didáctica y disciplinar del capítulo de Tablas de Contingencia del LID fue alcanzado satisfactoriamente.

Un hallazgo particularmente significativo de este proceso fue la coincidencia entre las observaciones del panel externo y decisiones de diseño que ya habían sido revisadas internamente, como fue el caso del recuadro de fórmulas clave eliminado de la fase de exploración de una de las actividades. Que un evaluador, sin conocimiento del proceso interno de diseño, identificara de manera independiente la misma tensión pedagógica (la anticipación de herramientas de resolución antes de que el estudiante formule su propia conjetura) constituye una forma de triangulación que fortalece la validez del criterio de diseño adoptado, más allá de lo que un único instrumento o una única perspectiva podrían ofrecer por sí solos. De manera similar, la observación sobre la aparente redundancia en las intervenciones del tutor virtual, aunque en principio señalada como debilidad, permitió reafirmar y hacer explícito un principio de diseño central del recurso: la institucionalización diferida del saber, que impide formalizar un concepto hasta que el estudiante demuestre comprensión genuina de él.

El LID transforma la enseñanza tradicional de la estadística en un proceso de exploración reflexiva, enfocando al estudiante en la interpretación de los datos estadísticos y en la construcción de tablas estadísticas. Esto impacta directamente en la calidad del pensamiento estadístico, promoviendo ciudadanos y futuros licenciados capaces de argumentar críticamente frente a fenómenos del mundo real. Este recurso sirve además como un ejemplo práctico de cómo la inteligencia artificial puede ser adaptada al servicio de teorías didácticas robustas, cambiando la percepción de la IA hacia una herramienta potenciadora de situaciones a-didácticas de aprendizaje y no un simple sustituto de la explicación magistral. En términos prácticos, uno de sus mayores aportes es su capacidad para mitigar la baja intensidad horaria presencial característica de los cursos de estadística: al contar con un tutor virtual disponible de manera no

simultánea, así el aula de clase puede reservarse para debates profundos o proyectos de investigación, en lugar de la transmisión unidireccional de contenidos.

Es necesario señalar, que el instrumento aplicado fue diseñado y ejecutado como un proceso de validación de contenido por juicio de expertos, orientado a evaluar la calidad, rigurosidad y coherencia teórica del diseño del recurso, y no como una evaluación de impacto o efectividad del aprendizaje en estudiantes reales. En este sentido, los resultados presentados dan cuenta de la pertinencia del diseño desde la perspectiva de especialistas en el área, pero no permiten, por sí mismos establecer conclusiones sobre el efecto del LID en el aprendizaje efectivo de las tablas de contingencia. Esta distinción resulta relevante tanto para la interpretación correcta del alcance de este trabajo como para la identificación de las líneas de trabajo futuro que se plantean más adelante en la que se sugiere hacer una implementación en aula.

En cuanto a las limitaciones del recurso, estas recaen principalmente en su dependencia tecnológica: el LID necesita conexión a internet para su óptimo funcionamiento y el tutor virtual requiere del pago a plataformas de interfaz de programación de aplicaciones (API) para operar. El recurso pierde su funcionalidad de retroalimentación en tiempo real y de sostenimiento de entornos a-didácticos autónomos si se implementa en lugares que carezcan de conectividad o de equipos de cómputo con la capacidad de ejecutar interfaces web. A esto se suma una limitación de naturaleza distinta: la integración de un modelo de lenguaje de gran escala como GPT-4o para la calificación y retroalimentación deja un margen de respuestas inconsistentes o impredecibles. Aunque se diseñaron directrices estrictas basadas en la Teoría de las Situaciones Didácticas, existe el riesgo latente de que la IA incurra en el efecto de dar pistas excesivas que resuelvan el

problema por el estudiante o presente variaciones sutiles en el rigor matemático al evaluar. El control sobre estas interacciones requiere un refinamiento constante de los prompts y del backend del programa.

Finalmente, cabe reconocer que el panel de evaluadores, aunque calificado en su formación disciplinar y didáctica, estuvo compuesto por un número reducido de participantes ($n=6$), lo cual es coherente con la naturaleza cualitativa de un proceso de validación por juicio de expertos, pero limita la posibilidad de generalizar estadísticamente los resultados obtenidos. Es pertinente destacar que la totalidad del panel manifestó no haber utilizado previamente un recurso digital interactivo similar para la enseñanza de la estadística lo que sin haber sido un criterio de selección, sugiere la escasez de este tipo de herramientas en el contexto de formación de los participantes y refuerza la pertinencia del aporte que representa el desarrollo del LID. Como líneas de trabajo futuro se identifican, en primer lugar, la implementación y evaluación del capítulo diseñado en un curso real de estadística que permitirá no sólo contrastar la valoración experta obtenida en este trabajo con evidencia empírica sobre el proceso de aprendizaje, sino también comparar la mediación del LID frente a mediaciones tradicionales como los libros de texto. En segundo lugar, se propone incluir una evaluación diagnóstica objetiva al inicio del recurso, que permita identificar el nivel de lectura de datos en el que se encuentra cada estudiante y situarlo en problemas que partan de dicho nivel para avanzar de manera progresiva e individualizada. Finalmente, se identifican como líneas adicionales la extensión de este mismo proceso de validación por el panel de expertos a los demás capítulos del LID, y la incorporación de los ajustes de usabilidad sugeridos que no fueron abordados en la presente versión del recurso.

A través de este trabajo de grado, valoramos que el programa de pregrado en Licenciatura en Matemáticas incluya asignaturas electivas de programación, las cuales nos permitieron desarrollar el presente trabajo. La culminación de este proceso no pretende ser el final de una investigación, sino una invitación abierta a toda la comunidad académica, especialmente a los docentes de matemáticas en formación a sumarse activamente al diseño y desarrollo de recursos educativos innovadores. Con el desarrollo tecnológico de las últimas décadas, hoy se cuenta con múltiples recursos tecnológicos y didácticos a partir de los cuales el educador matemático no necesita ser un experto en los diferentes campos de la programación para la creación de nuevos recursos interactivos. Al dar el paso de consumidores de tecnología a creadores de medios didácticos, estaremos dotando a la educación matemática de herramientas contextualizadas, dinámicas y profundamente humanas, capaces de transformar las aulas de nuestra región y del país.

Referencias Bibliográficas

- Arteaga, P., Batanero, C., Cañadas, G., y Contreras, J. M. (2011). Las tablas y gráficos estadísticos como objetos culturales. *Números*, 76(1), 55-67.
<https://mdc.ulpgc.es/files/original/9e064e13e0ac4e50158ce316fe200251f555653f.pdf>
- Batanero, C. (2001). *Didáctica de la estadística*. Grupo de investigación en Educación Estadística, Universidad de Granada.
- Batanero, C., y Díaz, C. (2011). *Estadística con proyectos*. Departamento de Didáctica de la Matemática, Universidad de Granada.
- Batanero, C. (2013). Sentido estadístico: componentes y desarrollo [Conferencia]. *I Jornadas Virtuales de Didáctica de la Estadística, la Probabilidad y la Combinatoria*, Granada, España.
https://www.academia.edu/26342430/Sentido_estad%C3%ADstico_Componentes_y_desarrollo
- Batanero, C. (2019). Treinta años de investigación en educación estocástica: Reflexiones y desafíos. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*.
www.ugr.es/local/fqm126/civeest.html
- Batanero, C., Godino, J. D., Green, D., Holmes, P., & Vallecillos, A. (1994). Errores y dificultades en la comprensión de los conceptos estadísticos elementales. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.

https://www.researchgate.net/publication/246088808_Errors_and_difficulties_in_understanding_elementary_statistical_concepts

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387), 531-554.

https://www.researchgate.net/publication/6062457_Graphical_Perception_and_Graphical_Methods_for_Analyzing_Scientific_Data

Espinosa, A. J., y Bareño, D. M. S. (2018). Enseñando matemáticas con situaciones a-didácticas. *Revista Boletín REDIPE*, 7(12), 133-143.

<https://revista.redipe.org/index.php/1/article/view/652/603>

Gutiérrez, A. (2021). Diseño de libros electrónicos para la enseñanza a distancia a estudiantes con diferentes capacidades matemáticas [Presentación de conferencia]. *V COBISEMAT*

Hoyos Castellanos, M. J. (2024). *Análisis de datos usando gráficos estadísticos desde un enfoque exploratorio* [Trabajo de grado de pregrado, Universidad Industrial de Santander].

<https://noesis.uis.edu.co/entities/publication/89c24121-9054-4b98-a3d6-0e5177aff18b>

Instituto Colombiano para la Evaluación de la Educación, ICFES. . (2024). *Informe nacional de resultados para Colombia 2022: Programa para la Evaluación Internacional de Alumnos (PISA)*.

<https://www.mineducacion.gov.co/porta/micrositios-preescolar-basica-y-media/Pruebas-Internacionales/Pruebas-PISA-2022/>

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3(3), 185-225.

https://www.academia.edu/3117937/Understanding_charts_and_graphs

Madriz, C., Ortega, R., Valdés, Z., y Meza, L. (2024). La enseñanza de los gráficos estadísticos: reseña de un libro de texto interactivo para tercer año de la educación costarricense.

Revista Digital Matemática, Educación e Internet, 25, 1-23.

<https://revistas.tec.ac.cr/index.php/matematica/article/view/7237/7094>

Marquès, P. (2011). *Los medios didácticos*. Departamento de Pedagogía Aplicada, Facultad de Educación, UAB.

<http://peremarques.net/medios.htm>

Méndez Marín, M. Á. (2023). *Dificultades en el aprendizaje de la estadística*. [Trabajo de grado de posgrado, Universidad Externado de Colombia].

<https://bdigital.uexternado.edu.co/server/api/core/bitstreams/655acb77-c0bb-44c8-8b48-9a99f378b284/content>

Moore, D. S., McCabe, G. P., & Craig, B. A. (2017). *Introduction to the practice of statistics* (5.^a ed.). W. H. Freeman.

https://www.academia.edu/34554431/Introduction_to_the_Practice_of_Statistics

O'Bannon, B. W., Skolits, G. J., & Lubke, J. K. (2017). The influence of digital interactive textbook instruction on student learning preferences, outcomes, and motivation. *Journal of Research on Technology in Education*, 49(3-4), 103-116.

<https://doi.org/10.1080/15391523.2017.1303798>

- Pallauta, J. D., Batanero, C., y Gea, M. M. (2023). Un instrumento para evaluar la comprensión de tablas estadísticas en educación secundaria. *Enseñanza de las Ciencias*, 41(3), 89-112.
<https://ensciencias.uab.cat/article/view/v41-n3-pallauta-batanero-gea>
- Panizza, M. (2005). Conceptos básicos de la teoría de situaciones didácticas. En H. Alagia, A. Bressan y P. Sadovsky (Eds.), *Reflexiones teóricas para la educación matemática* (pp. 13–65). Libros del Zorzal.
- Posada Hernández, G. J. (2016). *Elementos básicos de estadística descriptiva para el análisis de datos*. Fondo Editorial Luis Amigó.
https://www.funlam.edu.co/uploads/fondoeditorial/120_Ebook-elementos_basicos.pdf
- Sadovsky, P. (2005). La teoría de las situaciones didácticas: un marco para pensar y actuar la enseñanza de la matemática. En H. Alagia, A. Bressan y P. Sadovsky (Eds.), *Reflexiones teóricas para la educación matemática* (pp. 13–65). Libros del Zorzal.
- Toapanta Quishpe, S. M., y Yugcha Oña, E. F. (2023). *Recursos digitales para el aprendizaje activo*. [Trabajo de grado de pregrado, Universidad Técnica de Cotopaxi].
- Vásquez, J. (2014). *El programa libros interactivos multimedia en la comprensión lectora de los alumnos de sexto grado de la IE 10381 de Chota, 2014* [Trabajo de grado de posgrado, Universidad Nacional de Cajamarca].
<https://repositorio.unc.edu.pe/bitstream/handle/20.500.14074/1030/Tesis%20lim%20enero%2022.pdf?sequence=1&isAllowed=y>

Apéndice A Instrumento de validación - preguntas realizadas al panel

Sección	N.º	Tipo	Pregunta
Sección 1 Datos del evaluador	1	<i>Selección</i>	Rol actual
	2	<i>Si/No + texto</i>	¿Ha usado previamente algún recurso digital interactivo para enseñar o aprender estadística? Si la respuesta es Sí, escriba el nombre del recurso en "Otros".
Sección 2 Rigurosidad y pertinencia del contenido estadístico	3	<i>Likert 1-5</i>	Los conceptos de frecuencia conjunta, marginal y condicional están definidos con rigurosidad estadística.
	4	<i>Likert 1-5</i>	La notación matemática empleada es consistente a lo largo del capítulo.
	5	<i>Likert 1-5</i>	Los ejemplos y datos utilizados son apropiados para el nivel de Licenciatura en Matemáticas.
	6	<i>Likert 1-5</i>	La prueba de independencia chi-cuadrado se presenta con precisión conceptual y procedimental.
	7	<i>Abierta</i>	¿Identificó algún error conceptual, matemático o de notación? Descríbalo con la mayor precisión posible (página o situación específica).
Sección 3 Coherencia con la Teoría de Situaciones Didácticas	8	<i>Likert 1-5</i>	Las actividades sitúan al estudiante en un problema contextualizado que exige acción autónoma sin intervención directa de un docente.
	9	<i>Likert 1-5</i>	El tutor virtual evita dar la respuesta directa y en su lugar devuelve consecuencias de las decisiones del estudiante.
	10	<i>Likert 1-5</i>	La institucionalización de cada concepto (el momento en que se nombra formalmente) ocurre después de que el estudiante ha explorado y argumentado, no antes.
	11	<i>Likert 1-5</i>	Las situaciones permiten que el estudiante formule hipótesis y conjeturas propias antes de recibir retroalimentación.
	12	<i>Abierta</i>	¿Identificó algún momento en que el recurso "revele" la respuesta antes de que el estudiante la construya por sí mismo? Descríbalo.
Sección 4 Progresión de los niveles de comprensión (Curcio/Aoyama)	13	<i>Likert 1-5</i>	El capítulo logra que el estudiante lea los datos (identificación literal de frecuencias).
	14	<i>Likert 1-5</i>	El capítulo logra que el estudiante lea dentro de los datos (comparación de frecuencias, patrones de asociación).
	15	<i>Likert 1-5</i>	El capítulo logra que el estudiante lea más allá de los datos (predicciones que trascienden la muestra).
	16	<i>Likert 1-5</i>	El capítulo logra que el estudiante lea detrás de los datos (nivel crítico: cuestionar validez, sesgos, variables ocultas).
	17	<i>Likert 1-5</i>	La progresión entre estos niveles es gradual y evita saltos abruptos de dificultad.
Sección 5 Juicios de asociación y prevención de errores conceptuales	18	<i>Likert 1-5</i>	El recurso ayuda a distinguir correctamente entre correlación y causalidad.
	19	<i>Likert 1-5</i>	El recurso evita que el estudiante analice datos de forma aislada, promoviendo lectura contextualizada.
	20	<i>Likert 1-5</i>	El recurso cuestiona adecuadamente la tentación de asumir relaciones deterministas entre variables.

Sección	N.º	Tipo	Pregunta
Sección 6 Mediación del tutor virtual (IA)	21	<i>Likert 1-5</i>	Las preguntas del tutor virtual son claras y pertinentes al momento de la actividad.
	22	<i>Likert 1-5</i>	El tutor mantiene un tono adecuado (ni punitivo ni excesivamente validador) frente a los errores del estudiante.
	23	<i>Likert 1-5</i>	La retroalimentación del tutor virtual contribuye genuinamente a la comprensión, más allá de indicar si la respuesta es correcta o incorrecta.
Sección 7 Usabilidad e interfaz	24	<i>Likert 1-5</i>	La navegación entre páginas y actividades es clara e intuitiva.
	25	<i>Likert 1-5</i>	Las tablas, gráficos y elementos interactivos se visualizan correctamente.
	26	<i>Likert 1-5</i>	El lenguaje utilizado es apropiado para el público objetivo (estudiantes de Licenciatura en Matemáticas).
Sección 8 Valoración global y recomendaciones	27	<i>Likert 1-5</i>	En términos generales, este recurso constituye un aporte pertinente para la enseñanza de tablas de contingencia.
	28	<i>Sí/No, con ajustes</i>	Recomendaría el uso de este capítulo en un curso real de estadística.
	29	<i>Abierta</i>	¿Cuáles considera que son las tres mayores fortalezas del capítulo?
	30	<i>Abierta</i>	¿Cuáles considera que son las tres mayores debilidades u oportunidades de mejora?
	31	<i>Abierta</i>	¿Considera que la disposición de diferentes tipos de representación (tabular y gráfica) aporta significativamente al desarrollo de las ideas centrales del capítulo?
	32	<i>Abierta</i>	¿Algún comentario adicional?
	33	<i>Abierta</i>	¿Qué ajustes sugiere para mejorar la experiencia, claridad y usabilidad de este recurso?

Nota. El instrumento fue administrado mediante formulario en línea. Las preguntas de escala Likert se calificaron de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo).

Apéndice B *Manual de Usuario*

L I D · E S T A D Í S T I C A

Manual de Usuario

*Guía para navegar el Libro Interactivo Digital y trabajar con
el tutor virtual*

Licenciatura en
Matemáticas Universidad
Industrial de Santander
(UIS)
2026

1. Introducción

El Libro Interactivo Digital (LID) de Estadística Descriptiva es una herramienta de aprendizaje pensada para construir, paso a paso y con la guía de un tutor virtual, los conceptos fundamentales de tablas de frecuencia, tablas de contingencia y la prueba de independencia chi-cuadrado.

A diferencia de un libro de texto tradicional, este material no entrega las definiciones de entrada: te propone situaciones, preguntas y retos, y te acompaña mientras tú mismo llegas a las conclusiones. El tutor virtual no busca decirte si "acertaste" o "fallaste" de inmediato — su función es ayudarte a razonar, devolverte las consecuencias de tus propias respuestas, y guiarte hacia el concepto correcto a través del diálogo.

Este manual está pensado tanto para estudiantes que usarán el LID como parte de su proceso de aprendizaje, como para docentes o evaluadores que necesiten comprender su funcionamiento.

2. Antes de empezar

2.1 Requisitos

- Un navegador web actualizado: Google Chrome, Mozilla Firefox o Microsoft Edge (evita versiones muy antiguas).
- Conexión a internet estable — el tutor virtual necesita comunicarse con un servidor para generar sus respuestas.
- No es necesario instalar nada: el LID funciona directamente en el navegador.

Importante

Evita usar el modo incógnito o privado del navegador si quieres que tu progreso se conserve entre sesiones. En modo incógnito, todo lo que hiciste se borra al cerrar la ventana.

3. Navegación general

El libro se organiza en páginas que avanzas de una en una, siguiendo una secuencia pensada para que cada concepto se apoye en el anterior.

3.1 Botones de navegación

- En la parte inferior de cada página encontrarás dos botones: "← [nombre de la página anterior]" a la izquierda, y "[nombre de la página siguiente] →" a la derecha.
- Estos botones te permiten avanzar o retroceder en cualquier momento.

3.2 Puntos de navegación

En la parte inferior de la pantalla hay una fila de puntos pequeños — cada uno representa una página del libro. El punto de color dorado indica en qué página estás actualmente. Puedes hacer clic en cualquier punto para saltar directamente a esa página

3.3 Indicador de fase

En algunas páginas con tutor, verás junto al título un pequeño punto de color acompañado de una etiqueta (por ejemplo, "Exploración", "Institucionalización"). Este indicador te muestra en qué momento de la actividad te encuentras — no requiere ninguna acción de tu parte, es solo una referencia visual.

4. El tutor virtual: cómo interactuar con él

Esta es la parte más importante del libro, y la que más se diferencia de un material de estudio convencional. El tutor virtual no es un buscador de respuestas correctas: es un compañero de razonamiento que dialoga contigo.

4.1 Cómo escribir buenas respuestas

- Responde con tus propias palabras, explicando tu razonamiento — no solo con un número suelto.
- Si la pregunta pide un cálculo, incluye el resultado Y una breve explicación de cómo lo obtuviste.
- No te preocupes por "decir lo que el tutor quiere escuchar" — el objetivo es que expreses tu propio pensamiento, incluso si no estás seguro.
- Si una respuesta te parece incompleta al releerla, complétala antes de enviarla: el tutor identificará respuestas muy cortas o ambiguas y te pedirá que las desarrolles.

4.2 Las tres formas en que aparece el tutor

A. Preguntas iniciales

En algunas actividades, antes de que el tutor aparezca, encontrarás dos o tres preguntas abiertas para responder primero (por ejemplo, tu primera impresión sobre los datos). Escribe tus respuestas en los recuadros de texto y pulsa "Enviar al tutor →". El tutor iniciará la conversación retomando justamente lo que escribiste.

B. Diálogo continuo

En la mayoría de páginas, verás una ventana de chat con un campo de texto en la parte inferior. Escribe tu respuesta y pulsa "Enviar" (o la tecla Enter). El tutor responderá en unos segundos.

C. Momentos de descubrimiento guiado

En ciertos puntos clave del libro (por ejemplo, al descubrir la fórmula de la frecuencia esperada en el capítulo de chi-cuadrado), el tutor no te dará la respuesta directamente. Te hará una pregunta, y si tu primer intento no es suficiente, te dará una pista progresivamente más

concreta en tu siguiente intento. Esto es intencional: forma parte del diseño pedagógico del libro. Tras algunos intentos, si aún no llegas por tu cuenta, el sistema revelará la fórmula formalmente — nunca te dejará atascado indefinidamente.

¿Por qué el tutor no dice "correcto" o "incorrecto"?

Es intencional. En vez de un veredicto inmediato, el tutor suele devolverte una pregunta o la consecuencia matemática de tu respuesta, para que seas tú quien evalúe si tiene sentido. Esta forma de aprender (construyendo el conocimiento en vez de recibirlo ya hecho) es el corazón del diseño de este libro.

4.3 Si el tutor no parece avanzar

- Revisa que tu respuesta incluya un dato concreto (un número, una comparación) y no solo una idea general.
- Si te pide "qué te haría falta saber" o algo similar, intenta identificar específicamente qué información falta, no solo decir que "falta algo".
- Ten paciencia con las preguntas que se sienten repetidas: muchas veces el tutor está profundizando en un mismo punto antes de dejarte avanzar al siguiente — es parte del proceso, no un error.

4.4 Lectura en voz alta

Junto al nombre del tutor encontrarás un pequeño ícono de altavoz. Actívalo para que las respuestas del tutor se lean en voz alta automáticamente — útil si prefieres escuchar mientras lees.

5. El aviso "¿Seguro que quieres continuar?"

Si intentas avanzar a la siguiente página sin haber completado la actividad (por ejemplo, sin haber interactuado con el tutor, o sin haber terminado de construir un concepto), aparecerá un aviso como este:

Aviso de continuidad

Estás a punto de avanzar sin haber completado esta situación de aprendizaje. La construcción del concepto que trabajamos aquí depende de que la resuelvas antes de continuar — saltarla ahora puede dejarte sin la base que necesitas más adelante. Aun así, tú decides.

Este aviso NUNCA te bloquea. Tienes dos opciones:

- "Volver a la actividad" — te quedas en la página para terminarla.

- "Continuar de todos modos" — avanzas igual; el sistema simplemente deja constancia de que decidiste saltarla, sin impedirte seguir.

La decisión siempre es tuya. El aviso existe para que sea una decisión consciente, no un obstáculo.

6. Herramientas interactivas del libro

6.1 Tablas para completar

- Las celdas vacías (marcadas con "?") son las que debes completar tú.
- Usa el botón "Verificar tabla" para revisar tu trabajo: las celdas correctas se resaltan en verde, las incorrectas en rojo.
- El botón "Limpiar" reinicia todas tus respuestas en esa tabla si quieres empezar de nuevo.
- Un contador de progreso ("X / Y celdas completadas") te indica cuánto llevas.

6.2 Selector de sistema de representación

En las páginas de tablas de contingencia encontrarás pestañas como "Frecuencias absolutas", "% sobre el total", "% por fila" y "% por columna". Cada una cambia la forma en que se interpreta la tabla — selecciona la que la actividad te pida antes de completar las celdas.

6.3 Arrastrar y soltar

En algunas actividades (como la construcción de frecuencias acumuladas), puedes arrastrar las filas de la tabla para reordenar las categorías antes de calcular los valores. Simplemente haz clic sostenido sobre una fila y suéltala en la posición deseada.

6.4 Alternar entre decimal y porcentaje

En las columnas de proporción (f_r , F_r) verás un pequeño botón junto al encabezado (".00" o "%"). Al hacer clic, toda la columna cambia de formato — útil para comparar ambas representaciones del mismo dato. Si ya habías escrito una respuesta, se convierte automáticamente al nuevo formato sin que pierdas tu trabajo.

6.5 La microcalculadora

En la esquina superior de cada página encontrarás un pequeño botón con un ícono de calculadora. Al hacer clic, se abre una calculadora básica (suma, resta, multiplicación, división, porcentaje y cuadrado) para que hagas cálculos rápidos sin salir del libro ni perder tu lugar.

7. Tu progreso se guarda solo

- No necesitas guardar manualmente: cada respuesta que escribes y cada conversación con el tutor se conserva automáticamente.
- El progreso se guarda en tu propio dispositivo y navegador — si usas otro computador o celular, no verás el mismo avance (cada dispositivo lleva su propio registro).
- Si recargas la página o cierras y vuelves a abrir el navegador (sin usar modo

incógnito), tu trabajo seguirá ahí.

8. Estructura del capítulo

El LID se organiza en tres secciones, cada una con varias páginas de actividad. Este mapa te ayuda a ubicarte:

Sección I · Tablas de Frecuencia	
Presentación	Introducción a la sección y contexto del problema
Construcción guiada	El tutor te acompaña a construir f_i , f_r , F_i y F_r paso a paso
Síntesis	Repaso de la tabla completa
Ejemplos dinámicos	Explora ejemplos adicionales de forma libre
Problemas	Completa tablas de frecuencia por tu cuenta
Análisis	Construye y analiza tu propia tabla, con acompañamiento del tutor
Sección II · Tablas de Contingencia	
Presentación	Introducción a la sección
Actividad	Construcción guiada de una tabla de contingencia con el tutor
Síntesis	Repaso de lo construido
Formas Parciales	Explora las distintas formas de leer una tabla (% total, por fila, por columna)
Asociación Independencia	Puente hacia la idea de independencia estadística
Ejemplos dinámicos	Explora ejemplos adicionales
Formulación	Retos de construcción de tablas (Reto 1, 2 y 3)
Validación	Retos de validación con preguntas de análisis
Exploración libre	Diálogo abierto con el tutor sobre independencia
Sección III · Prueba Chi-cuadrado	
Presentación	Introducción a la prueba de independencia
Introducción	Primer contacto con el problema

Acción	Descubrimiento guiado de la fórmula de frecuencia esperada
Ejemplo	Ejemplo completo paso a paso
Formulación	Descubrimiento de por qué se eleva al cuadrado, y cálculo de χ^2
Síntesis	Verbalización de lo aprendido
Validación	Grados de libertad y valor crítico
Ejemplo	Segundo ejemplo completo
Reflexión	Asociación vs. causalidad
Síntesis Final	Cierre del capítulo
Aplicación Libre / Aplicación	Situación final y estudio con tus propios datos

9. Solución de problemas comunes

Situación	Qué hacer
La página se ve rota o vacía al cargar	Recarga la página con Ctrl+Shift+R (Windows/Linux) o Cmd+Shift+R (Mac). Esto fuerza una recarga completa sin usar archivos guardados en caché.
El tutor no responde o tarda mucho	Revisa tu conexión a internet. Si el mensaje "Escribiendo..." no cambia después de 15-20 segundos, recarga la página — tu conversación anterior se conserva.
Escribí una respuesta y el tutor repite la misma pregunta	Intenta ser más específico: incluye el número o el razonamiento completo, no solo una palabra suelta. El tutor está diseñado para pedir más detalles antes de avanzar.
No veo mi progreso de una sesión anterior	Verifica que estés usando el mismo navegador y el mismo dispositivo — el progreso no se comparte entre computador y celular, ni entre navegadores distintos en el mismo equipo.
Aparece un aviso preguntando si quiero continuar	No es un error: es un recordatorio de que aún no completaste la actividad de esa página. Puedes elegir "Volver a la actividad" o "Continuar de todos modos" — la decisión es tuya.

Perdí mi progreso por completo	Esto puede ocurrir si navegas en modo incógnito/privado, o si borraste manualmente los datos de navegación de tu navegador. Evita el modo incógnito si quieres conservar tu avance.
---------------------------------------	---

10. Preguntas frecuentes

¿Puedo volver a una página anterior sin perder mi avance?

Sí. Puedes navegar libremente hacia atrás y hacia adelante; tu progreso en cada página se conserva.

¿Qué pasa si me equivoco al responder al tutor?

Nada grave. El tutor está diseñado justamente para trabajar con respuestas incompletas o incorrectas, devolviéndote la consecuencia de tu razonamiento para que lo ajustes tú mismo.

¿El tutor recuerda lo que hablamos antes en la misma página?

Sí, dentro de una misma actividad el tutor mantiene el hilo completo de la conversación, incluso si recargas la página.

¿Necesito completar el libro en una sola sesión?

No. Puedes cerrar el libro y continuar después desde donde lo dejaste, siempre que uses el mismo dispositivo y navegador.