Revista de Educación Estadística
Facultad de Ciencias Básicas
Universidad Católica del Maule

Revista de Educación Estadística
https://revistaeduest.ucm.cl/
Vol. 5, pp. 1-19, ene. 2026 - dic. 2026
ISSN 2810-6164

DOI: https://doi.org/10.29035/redes.5.1.2

ESTADÍSTICA PARA MARKETING: EFECTIVIDAD DE LA ENSEÑANZA CONTEXTUALIZADA EN EDUCACIÓN SUPERIOR

Estatística para marketing: eficácia do ensino contextualizado no Ensino Superior

Statistics for marketing: effectiveness of contextualized teaching in Higher Education

Miguel Alejandro Guzmán Méndez[1]
Universidad del Valle de Atemajac (La Piedad, México)
Rey Josué Pérez Gómez[2]
Instituto Tecnológico de La Piedad (La Piedad, México)

Resumen
La enseñanza de estadística en programas de ciencias sociales enfrenta desafíos vinculados a actitudes negativas y a la percepción de irrelevancia profesional de la disciplina. El objetivo de este estudio fue evaluar si una propuesta didáctica que emplea el análisis del comportamiento del consumidor como contexto significativo mejora el desempeño académico en estadística de estudiantes de licenciatura en marketing, en comparación con la enseñanza tradicional. Se adoptó un diseño cuasi-experimental de grupos de comparación no equivalentes con medición post-test, comparando dos grupos de 40 estudiantes que cursaron la misma asignatura en periodos consecutivos. El análisis combinó pruebas estadísticas inferenciales sobre el desempeño académico con un análisis cualitativo de productos estudiantiles. Los resultados muestran un desempeño promedio considerablemente más alto en el grupo experimental y la ausencia de reprobación, con un tamaño del efecto amplio. El análisis cualitativo sugiere un desarrollo de competencias interpretativas y de aplicación contextualizada de conceptos estadísticos. Se concluye que la contextualización significativa constituye una estrategia pedagógica prometedora, si bien las limitaciones del diseño obligan a interpretar los hallazgos como evidencia sugestiva y no concluyente.
Palabras clave: Educación estadística, Contextualización, Comportamiento del consumidor, Educación Superior, Enseñanza de la estadística.

*Autor de correspondencia: ingalexmendez23@gmail.com (M.A. Guzmán Méndez)
[1] https://orcid.org/0000-0001-5167-0387 (ingalexmendez23@gmail.com).
[2] https://orcid.org/0009-0006-9525-2009 (ing.josue.perez@gmail.com).

Resumo
O ensino de estatística em programas de ciências sociais enfrenta desafios vinculados a atitudes negativas e à percepção de irrelevância profissional da disciplina. O objetivo deste estudo foi avaliar se uma proposta didática que emprega a análise do comportamento do consumidor como contexto significativo melhora o desempenho acadêmico em estatística de estudantes de graduação em marketing, em comparação com o ensino tradicional. Adotou-se um delineamento quase-experimental de grupos de comparação não equivalentes com medição pós-teste, comparando dois grupos de 40 estudantes que cursaram a mesma disciplina em períodos consecutivos. A análise combinou testes estatísticos inferenciais sobre o desempenho acadêmico com uma análise qualitativa de produtos estudantis. Os resultados mostram um desempenho médio consideravelmente mais alto no grupo experimental e a ausência de reprovação, com um tamanho de efeito amplo. A análise qualitativa sugere o desenvolvimento de competências interpretativas e de aplicação contextualizada de conceitos estatísticos. Conclui-se que a contextualização significativa constitui uma estratégia pedagógica promissora, embora as limitações do delineamento obriguem a interpretar os achados como evidência sugestiva e não conclusiva.
Palavras-chave: Educação estatística, Contextualização, Comportamento do consumidor, Ensino Superior, Ensino da estatística.

Abstract
Teaching statistics in social science programs faces challenges linked to negative attitudes and the perception that the discipline is professionally irrelevant. The aim of this study was to evaluate whether a teaching proposal that uses consumer behavior analysis as a meaningful context improves the academic performance in statistics of undergraduate marketing students, compared with traditional instruction. A quasi-experimental design with non-equivalent comparison groups and post-test measurement was adopted, comparing two groups of 40 students who took the same course in consecutive terms. The analysis combined inferential statistical tests on academic performance with a qualitative analysis of student products. Results show a considerably higher average performance in the experimental group and the absence of failing grades, with a wide effect size. The qualitative analysis suggests the development of interpretive and contextualized application competencies regarding statistical concepts. It is concluded that meaningful contextualization is a promising pedagogical strategy, although the limitations of the design require interpreting the findings as suggestive rather than conclusive evidence.
Keywords: statistics education, contextualization, consumer behavior, higher education, statistics teaching.

Recibido: 14/12/2025 - *Aceptado*: 27/06/2026

1. INTRODUCCIÓN

La enseñanza de la estadística en programas de ciencias sociales y administrativas constituye un desafío pedagógico persistente en la educación superior. Diversas investigaciones han documentado que los estudiantes de estas disciplinas exhiben con frecuencia actitudes desfavorables hacia la estadística, asociadas a experiencias previas adversas y a un escepticismo

sobre su utilidad profesional (Gal et al., 1997; Rodríguez-Santero y Gil-Flores, 2019). Estos factores afectivos pueden operar como barreras para el aprendizaje, incluso cuando los estudiantes poseen las capacidades cognitivas necesarias.

El problema resulta particularmente relevante en México, donde la reprobación en cursos de estadística dentro de programas de administración y marketing constituye un obstáculo recurrente señalado por el profesorado y observado en la propia experiencia institucional de la que parte este trabajo. Esta situación afecta el progreso académico individual, contribuye al rezago estudiantil y, en casos extremos, a la deserción universitaria. De manera más amplia, limita el desarrollo de competencias de razonamiento cuantitativo que resultan cada vez más críticas en el contexto profesional contemporáneo, en el que la toma de decisiones basada en datos se ha convertido en un imperativo organizacional (Batanero et al., 2013).

La investigación en educación matemática ha identificado que las actitudes hacia la estadística poseen un carácter multidimensional. Por una parte, las percepciones de irrelevancia profesional, en las que el estudiante no logra conectar conceptos estadísticos abstractos con aplicaciones concretas de su carrera futura (Gal et al., 1997). Por otra, las creencias sobre la propia capacidad, en las que el estudiante se percibe como poco apto para las matemáticas y anticipa el fracaso antes de intentar comprender el contenido (Murtonen et al., 2008). A ello se suman experiencias previas desfavorables que condicionan la disposición con que el estudiante enfrenta la asignatura.

Desde una perspectiva constructivista, el aprendizaje significativo ocurre cuando el estudiante puede conectar el conocimiento nuevo con estructuras cognitivas y experiencias previas (Brown et al., 1989). En el ámbito de la educación estadística, esto sugiere que la presentación de conceptos en contextos abstractos o artificiales dificulta el anclaje cognitivo y favorece un aprendizaje superficial. Garfield y Ben-Zvi (2008) sostienen específicamente que los estudiantes aprenden estadística de forma más efectiva cuando los conceptos se sitúan en contextos que reconocen como significativos y profesionalmente relevantes.

El presente estudio responde a esta problemática mediante la implementación y evaluación de una propuesta didáctica que utiliza el análisis del comportamiento del consumidor como contexto significativo para enseñar estadística a estudiantes de licenciatura en marketing. El análisis del comportamiento del consumidor es el campo, central en la formación en marketing, que estudia cómo las personas seleccionan, adquieren y utilizan productos y servicios; en términos de datos, se traduce en el examen de variables como frecuencia de compra, gasto, preferencias y

satisfacción. La elección de este contexto no es arbitraria: todos los estudiantes son consumidores y poseen conocimiento experiencial del fenómeno; además, el análisis de consumidores resulta nuclear para los roles profesionales en marketing, investigación de mercados y estrategia comercial, y permite explorar conceptos estadísticos desde técnicas descriptivas básicas hasta el modelado predictivo sin simplificación artificial.

La pregunta que orienta esta investigación es la siguiente: ¿la implementación de una propuesta didáctica basada en la contextualización significativa mediante el análisis del comportamiento del consumidor mejora el desempeño académico en estadística de los estudiantes de licenciatura en marketing, en comparación con la enseñanza tradicional? Para responderla se diseñó un estudio cuasi-experimental que comparó dos grupos de 40 estudiantes cada uno en la Universidad del Valle de Atemajac (UNIVA): un grupo control (enero-abril 2023) que recibió enseñanza tradicional y un grupo experimental (enero-abril 2024) que participó en la intervención contextualizada. La evaluación combinó el análisis cuantitativo del desempeño académico mediante pruebas estadísticas inferenciales con el análisis cualitativo de productos estudiantiles.

2. MARCO TEÓRICO

2.1. Actitudes hacia la Estadística en Estudiantes de Ciencias Sociales

La investigación sobre las actitudes hacia la estadística ha tenido un desarrollo sustancial en las últimas décadas. Para ofrecer una visión articulada del campo, a continuación, se caracteriza primero el constructo y su medición, y después se reseñan dos aportes empíricos, atendiendo en cada caso al colectivo estudiado, a las dimensiones consideradas y a las conclusiones principales.

El instrumento más difundido para evaluar estas actitudes, el Survey of Attitudes Toward Statistics (SATS-36), distingue seis dimensiones interrelacionadas: afecto, competencia cognitiva, valor, dificultad, interés y esfuerzo (Rodríguez-Santero y Gil-Flores, 2019). La dimensión de valor —entendida como la percepción de utilidad profesional de la disciplina— resulta especialmente pertinente para el presente estudio, pues sitúa la relevancia percibida en el centro de cualquier intervención que aspire a modificar la disposición del estudiante hacia la asignatura.

Estos autores trabajaron con estudiantes de Ciencias de la Educación y obtuvieron evidencias de validez estructural para la versión española del instrumento. Su aporte es tanto instrumental como conceptual: las actitudes hacia la estadística se construyen socialmente a través de experiencias

educativas previas y de discursos culturales sobre las matemáticas, por lo que su medición requiere instrumentos validados. De ello se desprende que las intervenciones pedagógicas efectivas deben atender no solo los contenidos conceptuales, sino también las dimensiones afectivas y socioculturales.

Murtonen et al. (2008) examinaron a estudiantes universitarios de educación y psicología inscritos en cursos de métodos de investigación cuantitativa y encontraron que las dificultades experimentadas se vinculan tanto con la percepción de que esas competencias no serán necesarias en la vida laboral como con las creencias sobre la propia capacidad para afrontar la materia. Su trabajo sugiere que las intervenciones capaces de hacer visible la utilidad profesional de la disciplina y de construir confianza de forma gradual pueden resultar particularmente efectivas. En una revisión posterior del campo, Murtonen y Balloo (2019) sitúan estas dificultades dentro del desarrollo más amplio del pensamiento científico en la educación superior.

En conjunto, estos trabajos comparten la idea de que las actitudes hacia la estadística no son preferencias superficiales, sino disposiciones con consecuencias tangibles para el aprendizaje. Convergen, además, en que la percepción de valor profesional y la confianza en las propias capacidades son componentes determinantes, lo cual fundamenta una propuesta didáctica orientada precisamente a hacer visible la relevancia profesional de la disciplina.

2.2. Aprendizaje Contextualizado y Significativo en Educación Estadística

La teoría del aprendizaje situado sostiene que el conocimiento no es una entidad abstracta separable del contexto en el que se adquiere y utiliza (Brown et al., 1989). En educación estadística, esta perspectiva ha dado origen a un movimiento pedagógico que enfatiza el aprendizaje mediante la investigación estadística auténtica, con datos reales en contextos significativos (Garfield y Ben-Zvi, 2008).

Wild y Pfannkuch (1999) desarrollaron el concepto de razonamiento estadístico situado, al argumentar que el pensamiento estadístico no es un conjunto de procedimientos abstractos, sino una forma de razonar sobre fenómenos reales con apoyo en datos. Desde esta perspectiva, enseñar estadística de manera efectiva exige situar los conceptos en dominios donde resulte natural formular preguntas estadísticas, recopilar datos y emplear el análisis para informar decisiones.

A partir de la literatura revisada, en este trabajo se adoptan tres criterios para juzgar la pertinencia

didáctica de un contexto de enseñanza: familiaridad experiencial, es decir, que el estudiante posea conocimiento previo del dominio; relevancia profesional, esto es, una conexión clara con sus aspiraciones futuras; y riqueza conceptual, que el contexto permita explorar conceptos estadísticos de manera auténtica sin simplificación excesiva. Estos tres criterios sintetizan las orientaciones que recorren la caracterización del campo realizada por Zieffler et al. (2018) y las propuestas de enseñanza basada en la investigación estadística auténtica, y son los que guiaron la selección del contexto en el presente estudio.

La literatura ofrece, además, propuestas concretas de contextualización en distintos dominios. Rossman et al. (2006) describen un curso introductorio posterior al cálculo, dirigido a futuros profesores de secundaria, en el que los conceptos se construyen a partir del análisis de datos reales en lugar de la deducción formal. Burrill y Biehler (2011), por su parte, identifican las ideas estadísticas fundamentales que deben estructurar la enseñanza y subrayan la necesidad de anclarlas en contextos con significado para el aprendiz, tanto en el currículo escolar como en la formación del profesorado.

2.3. Énfasis en el Razonamiento Conceptual sobre el Cálculo Mecánico

Un debate persistente en educación estadística concierne al equilibrio apropiado entre el cálculo manual y el razonamiento conceptual. Tradicionalmente, los cursos de estadística dedicaban un tiempo sustancial a que los estudiantes calcularan estadísticos a mano mediante fórmulas. Este enfoque ha sido cuestionado de manera fundamental ante la disponibilidad generalizada de software estadístico (Cobb y Moore, 1997; Garfield y Ben-Zvi, 2008).

Las directrices curriculares contemporáneas (GAISE, 2016) recomiendan invertir la distribución tradicional del tiempo didáctico: dedicar el tiempo mínimo al cálculo mecánico —solo el suficiente para comprender la lógica subyacente— y el tiempo máximo a la interpretación, el razonamiento y la toma de decisiones estadísticamente informadas. Esta recomendación se fundamenta en que, en contextos profesionales reales, ningún profesional calcula manualmente una desviación estándar o una prueba de hipótesis, pero sí necesita interpretar reportes estadísticos, evaluar la validez de los análisis y decidir con base en evidencia cuantitativa (Gal et al., 1997).

Diversas experiencias de rediseño curricular han respaldado esta aproximación. Garfield et al. (2012) describen y evalúan un curso introductorio de nivel terciario organizado en torno al

modelado y la simulación con apoyo de software, en lugar del cálculo mediante fórmulas, y reportan avances en el razonamiento estadístico de los estudiantes sin que la reducción del cálculo manual comprometiera la comprensión conceptual. Esta experiencia resulta directamente pertinente para el presente estudio, dado que una de las diferencias entre las condiciones comparadas radica en el papel asignado a la herramienta de cálculo, aspecto que se discute de manera explícita más adelante.

2.4. Errores Conceptuales Persistentes en el Razonamiento Estadístico

La investigación ha identificado errores conceptuales que los estudiantes cometen de manera sistemática, incluso después de la instrucción formal (Kahneman, 2011; Tversky y Kahneman, 1974). No se trata de simples olvidos, sino de malentendidos estructurales sobre conceptos fundamentales que requieren intervención pedagógica específica. Entre el conjunto de errores documentados en la literatura, este estudio se concentra en cuatro: la confusión entre correlación y causalidad, la malinterpretación del valor p, la omisión de la variabilidad y la equiparación entre significancia estadística e importancia práctica. La elección de estos cuatro responde a dos criterios: son los errores que la propuesta didáctica abordó de forma deliberada en el aula y son, además, los que la literatura identifica como más consecuentes para la toma de decisiones profesionales en contextos de marketing. No se pretende, por tanto, ofrecer un inventario exhaustivo, sino delimitar el foco del análisis cualitativo posterior.

La confusión entre correlación y causalidad es el error más documentado y de mayores consecuencias. El estudiante observa que dos variables están correlacionadas y concluye de manera automática que una causa a la otra. Kahneman (2011) sostiene que esta tendencia refleja un sesgo cognitivo profundo. Castro Sotos et al. (2007), en su revisión sistemática de la evidencia empírica disponible, documentaron que, incluso después de una instrucción explícita, una proporción considerable de estudiantes universitarios persiste en este error.

La malinterpretación del valor p constituye un segundo error frecuente, dada la complejidad técnica de su definición correcta. La misma revisión de Castro Sotos et al. (2007) muestra que los estudiantes intercambian con frecuencia los términos de la probabilidad condicional e interpretan, de manera incorrecta, el valor p como la probabilidad de que la hipótesis nula sea verdadera. Esta confusión persiste tras la instrucción explícita y se observa incluso entre investigadores

profesionales (Wasserstein et al., 2019).

La omisión de la variabilidad es un tercer error: los estudiantes tienden a resumir distribuciones empleando únicamente el promedio e ignoran que la variabilidad es igualmente informativa (delMas et al., 2007). Por último, la equiparación entre significancia estadística e importancia práctica lleva a asumir que un resultado estadísticamente significativo es, por ello, relevante en términos prácticos, sin reconocer que con muestras suficientemente grandes diferencias triviales pueden alcanzar significancia estadística (Lin et al., 2013).

3. METODOLOGÍA

3.1. Diseño del Estudio

Este estudio adoptó un diseño cuasi-experimental de grupos de comparación no equivalentes con medición post-test (Campbell y Stanley, 1963; Shadish et al., 2002). Los diseños cuasi-experimentales resultan apropiados cuando la asignación aleatoria no es posible por razones éticas, prácticas o institucionales; en contextos educativos reales, los estudiantes se inscriben de manera natural en secciones específicas de los cursos, lo que impide la aleatorización. Campbell y Stanley (1963) identificaron las amenazas principales a la validez interna en estos diseños —historia, maduración, selección, mortalidad e instrumentación—, y Cook y Campbell (1979) argumentan que un diseño cuasi-experimental bien ejecutado puede generar evidencia interpretable cuando se documenta la equivalencia de los grupos en variables observables y se consideran de manera sistemática las explicaciones alternativas. En una revisión contemporánea de estos diseños, Reichardt (2019) insiste en que la credibilidad de las estimaciones depende menos del diseño considerado en abstracto que de la explicitación y el descarte razonado de las hipótesis rivales. Estas consideraciones orientaron tanto el diseño del estudio como la interpretación cautelosa de sus resultados.

Se compararon dos grupos de estudiantes que cursaron la misma asignatura de Estadística Aplicada en periodos académicos consecutivos: un grupo control que recibió enseñanza tradicional (enero-abril 2023) y un grupo experimental que participó en la intervención contextualizada (enero-abril 2024). El diseño es cuasi-experimental porque los grupos no fueron asignados de manera aleatoria, sino que están constituidos por estudiantes inscritos naturalmente en periodos distintos. Dado que ambos grupos pertenecían al mismo programa académico, al mismo nivel curricular (quinto

semestre de la Licenciatura en Marketing) y a la misma institución, se asume una equivalencia razonable en sus características académicas y demográficas observables, sin desconocer las limitaciones que esta condición impone.

3.2. Contexto y participantes

El estudio se realizó en la Universidad del Valle de Atemajac (UNIVA), campus La Piedad, Michoacán, México, una institución privada de educación superior. La asignatura evaluada fue Estadística Aplicada, materia obligatoria del quinto semestre de la Licenciatura en Marketing, impartida durante un cuatrimestre de 16 semanas, con tres horas semanales y un total de 48 horas. El grupo control estuvo conformado por 40 estudiantes que cursaron la asignatura durante el cuatrimestre enero-abril 2023. Sus características demográficas fueron: edad promedio de 21.3 años (rango 19-24), 62.5 % mujeres (n = 25) y 37.5 % hombres (n = 15), todos estudiantes de tiempo completo. La preparación matemática previa era heterogénea: aproximadamente el 60 % reportó haber cursado solo álgebra básica en preparatoria y el 40 % cursos adicionales de trigonometría o cálculo diferencial.

El grupo experimental estuvo conformado por 40 estudiantes que cursaron la asignatura durante el cuatrimestre enero-abril 2024. Sus características demográficas fueron similares: edad promedio de 21.1 años (rango 19-23), 60 % mujeres (n = 24) y 40 % hombres (n = 16), todos estudiantes de tiempo completo. La preparación matemática previa reportada fue comparable a la del grupo control. La similitud en las características observables fortalece la inferencia de que las diferencias en el desempeño pueden vincularse a la intervención pedagógica; no obstante, se reconoce la posible existencia de diferencias en variables no observadas, como la motivación intrínseca o las experiencias previas específicas con la estadística.

3.3. Descripción de las Intervenciones

3.3.1. Enseñanza tradicional (grupo control)

El grupo control recibió enseñanza de la estadística conforme a un enfoque pedagógico convencional. La presentación de los contenidos partía de definiciones formales y fórmulas matemáticas antes de considerar las aplicaciones; por ejemplo, el coeficiente de correlación de

Pearson se introducía mostrando primero la fórmula algebraica y calculando ejemplos numéricos. Las actividades enfatizaban el cálculo manual de estadísticos con apoyo de calculadora, y las evaluaciones asignaban una puntuación significativa a la precisión aritmética. Los ejemplos y ejercicios empleaban conjuntos de datos pequeños y genéricos, diseñados sobre todo para facilitar el cálculo manual, sin un contexto profesional rico. El curso seguía la estructura modular secuencial del libro de texto, con módulos relativamente independientes entre sí.

3.3.2. Enseñanza contextualizada (grupo experimental)

La intervención contextualizada se estructuró en cinco módulos temáticos, cada uno con el análisis del comportamiento del consumidor como contexto motivador y aplicación concreta. El Módulo 1, Estadística descriptiva: caracterización de poblaciones de clientes (tres semanas), partió de una pregunta auténtica sobre la caracterización de los clientes de una cadena de cafeterías e introdujo las medidas de tendencia central y dispersión como herramientas para responder preguntas empresariales, con énfasis en la interpretación contextualizada. El Módulo 2, Asociación entre variables categóricas (dos semanas), introdujo la prueba chi-cuadrada a partir de una pregunta sobre diferencias de satisfacción entre canales de venta y dedicó tiempo sustancial a la interpretación correcta del valor p. El Módulo 3, Correlación y causalidad (cuatro semanas), el más extenso, abordó la distinción entre correlación y causalidad mediante ejemplos de correlaciones espurias y casos empresariales. El Módulo 4, Segmentación de mercados (cuatro semanas), introdujo el análisis de conglomerados a partir del problema de la heterogeneidad de clientes. El Módulo 5, Visualización de datos y comunicación (tres semanas), se centró en la comunicación efectiva de hallazgos estadísticos. A lo largo de los cinco módulos operaron tres principios: la contextualización auténtica, el cálculo asistido por software y el abordaje explícito de los errores conceptuales.

Es necesario explicitar que las dos condiciones difieren en más de un elemento. Además del contexto de los ejemplos (genérico frente a comportamiento del consumidor), las condiciones difieren en la herramienta de cálculo: el grupo control empleó calculadora para el cálculo manual, mientras que el grupo experimental utilizó una hoja de cálculo (Excel) para las computaciones, lo que liberó tiempo de clase hacia la interpretación. Esta diferencia no es menor y se retoma de manera explícita en la sección de resultados y discusión y en las limitaciones, ya que constituye

una variable que el diseño no aísla.

3.4. Instrumentos y procedimientos de evaluación

Ambos grupos fueron evaluados mediante el mismo sistema de calificación, compuesto por cuatro componentes con idéntica ponderación en las dos condiciones: examen escrito (30 %), proyecto integrador (30 %), trabajos en clase (20 %) y trabajos independientes (20 %). La calificación final, en escala de 0 a 100, se calculó como el promedio ponderado de los cuatro componentes; la calificación mínima aprobatoria, conforme al reglamento institucional, es de 70 puntos.

Respecto de la equivalencia de las evaluaciones, es preciso ser transparente sobre su alcance. Los instrumentos específicos —los reactivos concretos del examen y las consignas del proyecto— no fueron idénticos entre los grupos, debido a la diferencia temporal y a la naturaleza contextualizada de la intervención. La equivalencia que se sostiene es funcional y se estableció controlando tres aspectos: el tipo de competencias evaluadas (comprensión conceptual, capacidad de aplicación e interpretación de resultados), el nivel de complejidad cognitiva requerido según la taxonomía de Bloom, y el número y la estructura de reactivos por nivel cognitivo. En concreto, los exámenes de ambos grupos contenían la misma cantidad de reactivos en cada nivel taxonómico (recordar, comprender, aplicar y analizar), de modo que la diferencia entre instrumentos residía en el contexto de los enunciados y no en su demanda cognitiva. No obstante, se reconoce explícitamente que esta equivalencia funcional no elimina la posibilidad de que el grupo experimental estuviera mejor preparado para el formato específico de evaluación que recibió; esta es una limitación que se retoma más adelante.

En cuanto al análisis cualitativo de los proyectos, este se apoyó en un conjunto de criterios definidos a priori, aunque no en una rúbrica formalmente validada. Los criterios fueron tres: la integración de varias técnicas estadísticas para responder a una pregunta de investigación coherente, la interpretación contextualizada de los resultados —entendida como la traducción de los valores numéricos a implicaciones empresariales— y la presencia o ausencia de los errores conceptuales abordados en el curso, en particular la inferencia de causalidad a partir de correlación. Cada proyecto se valoró de manera holística respecto de estos tres criterios. Se reconoce que la valoración la realizó únicamente el docente-investigador, sin la participación de un segundo evaluador independiente, lo que constituye una limitación que se discute en la sección

correspondiente.

3.5. Análisis de datos

Para el análisis cuantitativo se calcularon estadísticos descriptivos —media, mediana, desviación estándar, valores mínimo y máximo— de las calificaciones finales de ambos grupos. La significancia estadística de las diferencias de medias se evaluó mediante la prueba t de Student para muestras independientes, previa verificación del supuesto de normalidad con la prueba de Shapiro-Wilk. El tamaño del efecto se estimó mediante la d de Cohen. Las tasas de aprobación se compararon con la prueba chi-cuadrada. El nivel de significancia se estableció en $\alpha = 0.05$ y los análisis se realizaron con el software R, versión 4.3.

El análisis cualitativo consistió en un análisis temático de los proyectos finales del grupo experimental, orientado a identificar patrones en la calidad de la interpretación contextualizada y en el manejo de los errores conceptuales. El análisis fue de carácter inductivo, si bien estuvo informado por la literatura sobre errores conceptuales y por los tres criterios descritos en el apartado anterior.

3.6. Consideraciones éticas

El estudio se desarrolló como parte de las actividades académicas regulares del curso. Todos los estudiantes fueron evaluados de manera justa conforme a los criterios explicitados en el programa oficial. Los datos de calificaciones se anonimizaron para el análisis, con el fin de proteger la confidencialidad de los participantes.

3.7. Limitaciones metodológicas

Este diseño cuasi-experimental presenta limitaciones inherentes que deben considerarse al interpretar los hallazgos. La ausencia de aleatorización admite la posibilidad de diferencias preexistentes en variables no observadas. La diferencia temporal entre los grupos permite la confusión con efectos asociados al periodo, como el incremento de la experiencia docente. Los instrumentos de evaluación, aunque funcionalmente equivalentes, no fueron idénticos. La condición experimental difiere de la condición control en más de un factor —contexto de los ejemplos y herramienta de cálculo—, por lo que el diseño no permite aislar la contribución

específica de cada uno. Por último, el primer autor fue a la vez implementador de la intervención y evaluador, lo que introduce un posible sesgo de expectativas. En conjunto, estas limitaciones implican que los hallazgos deben interpretarse como evidencia sugestiva y no como demostración concluyente de una relación causal.

4. RESULTADOS Y DISCUSIÓN

Esta sección presenta de manera conjunta los resultados del estudio y su discusión, conforme al formato de la revista. Se organiza en torno a tres ejes: el desempeño cuantitativo de los grupos, el análisis cualitativo de los productos estudiantiles y la interpretación de los hallazgos a la luz de la literatura, incluidas las amenazas a la validez que condicionan su lectura.

4.1. Desempeño académico de los grupos

La Tabla 1 presenta los estadísticos descriptivos de las calificaciones finales de ambos grupos. El grupo experimental obtuvo una calificación promedio superior a la del grupo control, con una diferencia de 12.90 puntos. La mediana del grupo experimental también fue superior, lo que indica que la diferencia no es atribuible a valores atípicos. La desviación estándar fue menor en el grupo experimental, lo que sugiere una menor dispersión del desempeño, y el valor mínimo de este grupo fue notablemente más alto, lo que apunta a que la mejora se concentró también en los estudiantes de rendimiento más bajo.

Tabla 1

Estadísticos descriptivos de las calificaciones finales por grupo

Estadístico	Grupo control (n = 40)	Grupo experimental (n = 40)	Diferencia
Media	79.18	92.08	+12.90
Mediana	79.50	92.50	+13.00
Desviación estándar	7.54	4.25	-3.29
Mínimo	66	78	+12
Máximo	92	98	+6
Aprobados (≥ 70)	35	40	+5
Tasa de aprobación	87.5 %	100 %	+12.5 %

Nota: Calificaciones en escala de 0 a 100. Calificación mínima aprobatoria: 70 puntos.

La prueba de Shapiro-Wilk indicó que ambas distribuciones eran aproximadamente normales (control: $W = 0.97$, $p = 0.38$; experimental: $W = 0.96$, $p = 0.21$), lo que permitió aplicar la prueba

t de Student para muestras independientes. Esta prueba reveló una diferencia estadísticamente significativa entre los grupos, $t(78) = 9.426$, $p < 0.001$ (bilateral). El tamaño del efecto, estimado mediante la d de Cohen, fue $d = 2.11$, valor que se clasifica como amplio según los criterios convencionales (Cohen, 1988). El intervalo de confianza del 95 % para la diferencia de medias fue [10.18, 15.62].

Respecto de las tasas de aprobación, el grupo control registró un 87.5 % (35 de 40) y el grupo experimental un 100 % (40 de 40). Dado que una de las frecuencias esperadas es inferior a 5 (2.5 reprobados por grupo), la prueba chi-cuadrada no resulta apropiada para esta comparación y se aplicó en su lugar la prueba exacta de Fisher. El resultado, $p = 0.055$ (bilateral), indica que la diferencia favorece al grupo experimental pero no alcanza el nivel de significancia convencional de 0.05. Este resultado se reporta tal cual y contrasta con la magnitud observada en la comparación de medias; su lectura debe ser prudente, ya que con cinco reprobados en un solo grupo la potencia estadística para detectar diferencias en esta variable es reducida. La Tabla 2 detalla las frecuencias.

Tabla 2

Frecuencias de aprobación por grupo

Resultado	Grupo control	Grupo experimental	Total
Aprobados (≥ 70)	35 (87.5 %)	40 (100 %)	75
Reprobados (< 70)	5 (12.5 %)	0 (0 %)	5
Total	40	40	80

Nota: Prueba exacta de Fisher, $p = 0.055$ (bilateral). No se aplicó la prueba chi-cuadrada porque la frecuencia esperada mínima (2.5) es inferior a 5.

4.2. Análisis cualitativo de los productos estudiantiles

El análisis temático de los proyectos finales del grupo experimental, realizado conforme a los tres criterios descritos en la metodología, permitió identificar algunos patrones recurrentes. A título ilustrativo se describen dos proyectos. En el primero, un equipo analizó las preferencias de consumo en la cafetería universitaria a partir de una encuesta a 120 usuarios; el equipo integró estadística descriptiva y prueba chi-cuadrada e interpretó los resultados en términos de gestión de inventario y de personal. De manera relevante para los criterios del estudio, al examinar la correlación entre gasto y satisfacción, el equipo evitó la inferencia causal y reconoció de forma explícita que la correlación observada no permitía concluir una relación de causa y efecto. En el segundo proyecto, otro equipo realizó un análisis de segmentación con datos simulados de clientes

de una tienda en línea, identificó tres segmentos y propuso estrategias diferenciadas, con visualizaciones orientadas a una audiencia no técnica.

A partir de estos y otros proyectos se observaron tres patrones: la integración de varias técnicas estadísticas para responder a preguntas coherentes, la interpretación de los resultados en términos empresariales más allá del reporte de cifras, y la ausencia, en la mayoría de los trabajos, del error de inferir causalidad a partir de correlación. Conviene precisar el alcance de esta lectura. La caracterización anterior corresponde a la valoración de un único evaluador —el docente-investigador— sobre la base de criterios definidos a priori, pero no de una rúbrica validada ni de una segunda codificación independiente. Por consiguiente, estos patrones se presentan como observaciones cualitativas que ilustran posibles efectos de la intervención, y no como mediciones cuya fiabilidad haya sido establecida formalmente. Para evitar una interpretación sesgada, se ha optado por describir las conductas observables de los estudiantes —por ejemplo, la formulación explícita de salvedades sobre la causalidad— en lugar de atribuirles, sin un procedimiento de validación, competencias de orden superior.

4.3. Interpretación de los hallazgos y amenazas a la validez

Los resultados cuantitativos y cualitativos apuntan, de manera convergente, hacia una asociación entre la propuesta didáctica contextualizada y un mejor desempeño académico. La diferencia de medias y el tamaño del efecto observado son compatibles con la literatura que ha defendido la utilidad de los contextos auténticos en educación estadística (Garfield y Ben-Zvi, 2008; Zieffler et al., 2018). Ahora bien, la magnitud del efecto obtenida ($d = 2.11$) es considerablemente mayor que la que suele reportarse en intervenciones educativas comparables. Antes que interpretar este dato como evidencia de una potencia pedagógica excepcional, resulta más prudente considerarlo una señal de que el valor obtenido podría estar influido por las amenazas a la validez propias del diseño, que se examinan a continuación.

La amenaza más relevante para la interpretación de este estudio es que la condición experimental no aísla una sola variable. El grupo experimental difiere del control en, al menos, dos aspectos simultáneos: el contexto de los ejemplos —comportamiento del consumidor frente a datos genéricos— y la herramienta de cálculo —hoja de cálculo frente a calculadora—. La literatura ofrece razones para suponer que el segundo factor, por sí solo, puede tener un efecto considerable:

el uso de software libera tiempo cognitivo y de clase que puede redirigirse al razonamiento, y reduce la carga aritmética que constituye una de las barreras del aprendizaje (Cobb y Moore, 1997; Garfield et al., 2012). En consecuencia, no es posible determinar, con los datos disponibles, qué proporción de la mejora observada se debe a la contextualización del marketing y qué proporción al cambio de herramienta de cálculo. El presente estudio, por tanto, evalúa el efecto de un paquete de intervención y no el de la contextualización como variable aislada; discernir entre ambos componentes exige un diseño factorial que los manipule de manera independiente, lo cual se plantea como una línea de investigación futura.

Una segunda amenaza es la no identidad de los instrumentos de evaluación. Aunque se controló el nivel de complejidad cognitiva y la estructura de reactivos por nivel taxonómico, la diferencia en el contexto de los enunciados pudo favorecer al grupo experimental, más familiarizado con problemas situados en escenarios de marketing. La diferencia temporal entre los grupos y el doble papel del primer autor como docente y evaluador operan en la misma dirección: ambos podrían inflar la diferencia observada con independencia del mérito intrínseco de la propuesta. Estas consideraciones no anulan los hallazgos, pero sí delimitan su alcance.

En cuanto a la interpretación del análisis cualitativo, conviene ser igualmente cauteloso. Los proyectos del grupo experimental muestran indicios de una aplicación contextualizada de los conceptos y de una menor incidencia del error de causalidad. Sin embargo, dado que no se dispuso de una rúbrica validada ni de un segundo codificador, y dado que no se analizaron de manera comparable los productos del grupo control, estas observaciones deben leerse como evidencia ilustrativa y exploratoria. Permiten formular la hipótesis de que la intervención favorece la interpretación situada, pero no la confirman.

Finalmente, conviene situar el estudio en relación con la literatura internacional. La dirección de los hallazgos es coherente con la noción de razonamiento estadístico situado de Wild y Pfannkuch (1999) y con el énfasis de Rodríguez-Santero y Gil-Flores (2019) en el peso de las dimensiones afectivas y de la percepción de valor dentro de las actitudes hacia la estadística. Asimismo, el estudio aporta evidencia desde un contexto latinoamericano, menos representado en la investigación en educación estadística, lo que contribuye a una base empírica geográficamente más diversa. Esta contribución, no obstante, debe entenderse en el marco de las limitaciones ya señaladas.

5. CONCLUSIONES

Este estudio cuasi-experimental evaluó una propuesta didáctica que utiliza el análisis del comportamiento del consumidor como contexto significativo para enseñar estadística a estudiantes de licenciatura en marketing. La pregunta de investigación indagaba si este enfoque contextualizado mejora el desempeño académico en comparación con la enseñanza tradicional.

En respuesta directa a esa pregunta, la evidencia disponible es compatible con una respuesta afirmativa, aunque condicionada. El grupo que recibió la propuesta contextualizada mostró un desempeño promedio más alto, una menor dispersión de las calificaciones y la ausencia de reprobación, con un tamaño del efecto amplio; el análisis cualitativo de sus proyectos sugiere, además, una aplicación situada de los conceptos estadísticos. No obstante, esta respuesta afirmativa debe matizarse en dos sentidos. En primer lugar, el diseño no permite atribuir la mejora exclusivamente a la contextualización del marketing, ya que la intervención modificó simultáneamente la herramienta de cálculo; lo que el estudio respalda es la efectividad de un paquete pedagógico, no la de la contextualización como factor aislado. En segundo lugar, las amenazas a la validez —ausencia de aleatorización, diferencia temporal, instrumentos no idénticos y doble papel del investigador— implican que la magnitud del efecto observado podría estar sobrestimada. Por todo ello, los hallazgos se interpretan como evidencia sugestiva que justifica nuevas investigaciones, y no como una demostración concluyente.

Desde el punto de vista de la práctica docente, y con la prudencia que imponen las limitaciones anteriores, el estudio sugiere que vincular los contenidos estadísticos con dominios profesionalmente relevantes, redistribuir el tiempo del cálculo hacia la interpretación y abordar de forma explícita los errores conceptuales más persistentes constituyen orientaciones razonables para los docentes de estadística en contextos similares.

Varias líneas de investigación futura se derivan de este trabajo. La más prioritaria es un diseño factorial que manipule de manera independiente la contextualización y la herramienta de cálculo, con el fin de aislar la contribución de cada componente. Resultan igualmente necesarios estudios con asignación aleatoria, con instrumentos de medición actitudinal validados —como la escala SATS— aplicados antes y después de la intervención, con evaluación independiente de los productos estudiantiles mediante rúbricas y doble codificación, y con seguimiento longitudinal que permita valorar la retención del aprendizaje. La incorporación de una medición actitudinal

explícita es especialmente pertinente, ya que permitiría contrastar la hipótesis de que la reducción de las barreras afectivas media el efecto observado, hipótesis que el presente estudio no estuvo en condiciones de evaluar.

En síntesis, este estudio representa un esfuerzo por hacer la estadística más accesible al conectarla con experiencias que los estudiantes reconocen y valoran, sin renunciar al rigor conceptual. Sus resultados, leídos con la cautela que el diseño exige, ofrecen una base razonable para seguir investigando el papel de la contextualización significativa en la educación estadística.

REFERENCIAS BIBLIOGRÁFICAS

Batanero, C., Díaz, C., Contreras, J.M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números. Revista de Didáctica de las Matemáticas, 83*, 7-18.

Brown, J.S., Collins, A. y Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42. https://doi.org/10.3102/0013189X018001032

Burrill, G. y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A Joint ICMI/IASE Study* (pp. 57-69). Springer. https://doi.org/10.1007/978-94-007-1131-0_10

Campbell, D.T. y Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Castro Sotos, A.E., Vanhoof, S., Van den Noortgate, W. y Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review, 2*(2), 98-113. https://doi.org/10.1016/j.edurev.2007.04.001

Cobb, G. W. y Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*(9), 801-823. https://doi.org/10.1080/00029890.1997.11990723

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2.ª ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Cook, T.D. y Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

delMas, R., Garfield, J., Ooms, A. y Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58. https://doi.org/10.52041/serj.v6i2.483

Gal, I., Ginsburg, L. y Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. En I. Gal y J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37-51). IOS Press.

GAISE. (2016). *Guidelines for assessment and instruction in statistics education: College report 2016*. American Statistical Association.

Garfield, J. y Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer. https://doi.org/10.1007/978-1-4020-8383-9

Garfield, J., delMas, R. y Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM, 44*(7), 883-898. https://doi.org/10.1007/s11858-012-0447-5

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Lin, M., Lucas, H. C. y Shmueli, G. (2013). Research commentary: Too big to fail: Large samples and the p-value problem. *Information Systems Research, 24*(4), 906-917. https://doi.org/10.1287/isre.2013.0480

Murtonen, M. y Balloo, K. (Eds.). (2019). *Redefining scientific thinking for higher education: Higher-order thinking, evidence-based reasoning and research skills*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-24215-2

Murtonen, M., Olkinuora, E., Tynjälä, P. y Lehtinen, E. (2008). "Do I need research skills in working life?": University students' motivation and difficulties in quantitative methods courses. *Higher Education, 56*(5), 599-612. https://doi.org/10.1007/s10734-008-9113-9

Reichardt, C. S. (2019). *Quasi-experimentation: A guide to design and analysis*. Guilford Press.

Rodríguez-Santero, J. y Gil-Flores, J. (2019). Actitudes hacia la estadística en estudiantes de Ciencias de la Educación. Propiedades psicométricas de la versión española del Survey of Attitudes Toward Statistics (SATS-36). *Revista Electrónica de Investigación y Evaluación Educativa, 25*(1), 3. https://doi.org/10.7203/relieve.25.1.12676

Rossman, A., Medina, E. y Chance, B. (2006). A post-calculus introduction to statistics for future secondary teachers. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics (ICOTS-7)* (pp.1-6). International Statistical Institute.

Shadish, W.R., Cook, T.D. y Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Tversky, A. y Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Wasserstein, R. L., Schirm, A. L. y Lazar, N. A. (2019). Moving to a world beyond "p < 0.05". *The American Statistician, 73*(sup1), 1-19. https://doi.org/10.1080/00031305.2019.1583913

Wild, C.J. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-248. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Zieffler, A., Garfield, J. y Fry, E. (2018). What is statistics education? En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 37-70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2

Cómo citar:

Guzmán Méndez, M.A. y Pérez Gómez, R J. (2026). Estadística para marketing: efectividad de la enseñanza contextualizada en Educación Superior. *Revista de Educación Estadística, 5*, 1-19. https://doi.org/10.29035/redes.5.1.2



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.