PROBABILIDAD Y ESTADÍSTICA:

UNA MIRADA DIDÁCTICA AL ANÁLISIS DE DATOS Y EL RAZONAMIENTO DEL AZAR

AUTORÍA

JOSÉ LUIS SAQUINAULA BRITO

EDICIONES RISEI

COLECCIÓN:
Educación en acción: Praxis, currículo y subjetividades

UNEMI POSGRADOS

Probabilidad y Estadística: una mirada didáctica al análisis de datos y el razonamiento del azar

Autores

Saquinaula Brito José Luis
Universidad Estatal de Milagro
jsaquinaulab@unemi.edu.ec
https://orcid.org/0000-0003-2080-2548

Editorial: *Ediciones RISEI*
Colección Educación en acción: Praxis, currículo y subjetividades
Título del libro: Probabilidad y Estadística: una mirada didáctica al análisis de datos y el razonamiento del azar
Autoría: Saquinaula Brito José Luis
Edición: Primera edición
Año: 2025
ISBN digital: 978-9942-596-47-5
DOI: https://doi.org/10.63624/risei.book-978-9942-596-47-5

Coordinación editorial: Jorge Maza-Córdova y Tomás Fontaines-Ruiz
Corrección de estilo: Unidad de Redacción y Estilo
Diagramación y diseño: Unidad de Diseño
Revisión por pares: Sistema doble ciego de revisión externa

Machala – Ecuador, diciembre de 2025

Este libro fue diagramado en InDesign.
Disponible en: https://editorial.risei.org/
Contacto: info@risei.org

Prólogo

La estadística y la probabilidad suelen presentarse como materias difíciles, llenas de fórmulas y procedimientos que muchos aprenden sin llegar a comprender del todo. Sin embargo, la realidad actual nos recuerda cada día que pensar con datos es una habilidad indispensable para interpretar información, tomar decisiones informadas y participar de manera crítica en la vida social. Este libro surge de la necesidad de acercar estos conceptos de un modo claro y significativo, mostrando que la estadística no es un lenguaje reservado para expertos, sino una herramienta para comprender mejor el mundo.

A lo largo de sus capítulos, el texto propone un recorrido que conecta los contenidos con situaciones reales, ejemplos cotidianos y preguntas auténticas que dan sentido a cada técnica. Más que explicar procedimientos, busca revelar la lógica que hay detrás de ellos y cómo pueden ayudarnos a describir, comparar, predecir y argumentar a partir de datos. De este modo, cada noción estadística se convierte en un punto de partida para pensar, no en un requisito que se aprende de memoria.

El enfoque didáctico que atraviesa la obra coloca en el centro la variabilidad, la incertidumbre y la necesidad de formar un pensamiento crítico. Las simulaciones, los ejemplos contextualizados y las interpretaciones guiadas están presentes no para reemplazar el razonamiento, sino para enriquecerlo. Aprender estadística implica aprender a preguntar, dudar, interpretar y justificar; es un ejercicio intelectual que invita a mirar con atención las historias que se esconden detrás de los datos.

Este libro está dirigido a docentes, estudiantes e incluso lectores que, sin formación especializada, desean comprender mejor la información que los rodea. La intención es acompañar al lector en la construcción de una mirada más reflexiva frente a los datos y al azar, fortaleciendo su capacidad para interpretar la complejidad del mundo actual. Si estas páginas logran despertar curiosidad, claridad conceptual y una actitud crítica frente a la información, habrán cumplido su misión fundamental.

Contenido

índice de tabla y figuras

Capítulo I

Índice de figuras

Capítulo II

Índice de figuras

Capítulo III

Índice de tablas

Índice de figuras

Capítulo IV

Índice de tablas

Índice de figuras

Capítulo I

Comprender la estadística desde la experiencia: fundamentos y representaciones

Introducción

Hablar de estadística en educación suele remitirnos, casi de inmediato, a tablas, números y gráficos. Sin embargo, detrás de cada dato hay una historia: un estudiante que asistió o faltó, un grupo que aprendió con mayor o menor ritmo, una realidad social que se mueve con matices que ningún promedio puede capturar por completo. Este capítulo nace justamente de esa idea: los datos no son solo cifras; son representaciones de experiencias humanas que merecen ser leídas con atención y sensibilidad.

A medida que analizamos información educativa o social, descubrimos que lo importante no es únicamente "cuánto" ocurre un fenómeno, sino cómo se comporta, qué tan uniforme es, dónde aparecen tensiones y en qué lugares se esconden diferencias que pueden pasar inadvertidas. Por eso, la estadística descriptiva no es un ejercicio frío, sino una forma de mirar. A veces los valores se agrupan y nos hablan de estabilidad; en otras ocasiones se

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

dispersan y revelan desigualdades, excepciones o procesos que están tomando rumbos inesperados. En este recorrido, media y mediana nos dan una primera impresión, pero la dispersión, los valores atípicos y la variabilidad aportan la profundidad que permite interpretar el panorama completo.

En tiempos donde la información circula con rapidez y se simplifica sin miramientos, aprender a leer datos con criterio se vuelve una habilidad esencial. No basta aceptar un promedio como verdad absoluta; hay que preguntarse qué esconde, qué muestra y qué transforma. Estas preguntas no son técnicas, son intelectuales y éticas: nos obligan a mirar más allá del número y reconocer las realidades diversas que conviven dentro de cualquier conjunto de datos.

El capítulo que sigue se propone acompañar al lector en esa tarea. No pretende convertir la estadística en un conjunto de fórmulas, sino en una herramienta para pensar con mayor claridad. A lo largo de las secciones, encontraremos ejemplos, comparaciones, casos reales y situaciones que permiten entender cómo los datos adquieren sentido cuando se interpretan con cuidado. La intención es que cada lector, sea docente, investigador o estudiante, pueda acercarse a la estadística como un modo de comprender lo que ocurre a su alrededor y no solo como un contenido que debe memorizarse.

Con esta mirada abierta y reflexiva iniciamos el capítulo, invitando a detenerse, observar, comparar y, sobre todo, interpretar. Porque en educación y en la vida misma entender los datos es, al final, una forma de entender a las personas y las historias que dan origen a esos datos.

La estadística como herramienta para describir la realidad

La estadística se ha convertido en un recurso indispensable para interpretar fenómenos sociales, educativos, científicos y cotidianos. Su presencia en la toma de decisiones, en los medios de comunicación, en las instituciones y en la vida diaria ha transformado la manera en que las personas comprenden su entorno. Sin embargo, detrás de cada dato y de cada representación existe un proceso que merece ser examinado con detenimiento. Los números no son meras copias de la realidad; son construcciones que dependen de decisiones conceptuales, metodológicas y contextuales. Comprender este trasfondo permite desarrollar un pensamiento más reflexivo sobre lo que significan los datos y sobre cómo influyen en la manera en que narramos el mundo.

El propósito de este epígrafe es profundizar en esta dimensión menos visible de la estadística, articulando tres aspectos esenciales. En primer lugar, se examina cómo se construyen los

datos y qué factores intervienen en su configuración. En segundo lugar, se aborda el papel de las representaciones estadísticas, entendidas como narraciones visuales que no solo muestran información, sino que orientan la interpretación del fenómeno. Finalmente, se explora la relación entre lo individual y lo colectivo, subrayando las tensiones que emergen cuando se sintetizan experiencias diversas en cifras agregadas. La intención no es cuestionar el valor de la estadística, sino mostrar su complejidad y destacar la necesidad de un pensamiento cuidadoso al utilizarla para describir la realidad.

La construcción de los datos: decisiones que configuran la realidad

Los datos son frecuentemente presentados como entidades objetivas, como si fueran una reproducción fiel de la realidad observable. Sin embargo, toda producción de datos implica una serie de decisiones técnicas y conceptuales que influyen en lo que finalmente se recoge y registra. Garfield y Ben-Zvi (2008) señalan que un dato no es solo un número; es el resultado de una operación de medición que presupone categorías, instrumentos y criterios que no son ajenos al contexto cultural y educativo en el que se generan. Por ello, comprender cómo se construyen los datos es un paso fundamental para interpretar adecuadamente lo que representan.

La primera decisión aparece en la definición del fenómeno a estudiar. Para analizar los hábitos de estudio de un grupo de estudiantes, por ejemplo, es necesario decidir qué se entiende por "estudiar": ¿leer?, ¿resolver ejercicios?, ¿preparar trabajos?, ¿participar en tutorías? Cada elección delimita un tipo de información y excluye otras formas de actividad. Batanero y Borovcnik (2016) advierten que la delimitación conceptual del objeto de estudio determina la naturaleza de los datos y, por tanto, la perspectiva desde la cual se describirá el fenómeno. Un "dato" sobre horas de estudio solo representa una parte de la experiencia educativa, no su totalidad.

La segunda decisión recae en la forma de medición. En muchos contextos escolares se emplean cuestionarios donde los estudiantes estiman cuántas horas dedican a una actividad. Sin embargo, la percepción del tiempo es subjetiva y puede variar considerablemente. Los registros automáticos, como plataformas virtuales que contabilizan minutos de conexión, proporcionan información más precisa, pero también introducen nuevos problemas: ¿todo el tiempo de conexión implica estudio activo?, ¿qué ocurre con las actividades no registradas digitalmente?

Estos interrogantes muestran que la medición nunca es completamente transparente. Cada instrumento captura un aspecto, pero deja fuera otros.

La tercera decisión se relaciona con la selección de la muestra. Los resultados obtenidos dependen en gran medida de quiénes participan en la recolección de datos. Wild y Pfannkuch (1999) explican que la representatividad es un aspecto crucial del pensamiento estadístico. En el aula, esto se evidencia cuando se analizan encuestas a partir de la participación voluntaria: quienes responden suelen ser estudiantes más comprometidos o con más disponibilidad, lo que tiende a sesgar los resultados.

Estas decisiones no son fallas; son parte del proceso natural de producción de datos. No obstante, comprenderlas permite reconocer que los datos no "hablan" solos: son interpretaciones codificadas de la realidad. Este reconocimiento es esencial para evitar conclusiones simplificadas, especialmente en fenómenos sociales complejos como el aprendizaje, la convivencia escolar o la participación estudiantil.

Cuando la institución analiza las calificaciones finales de un grupo de 8 estudiantes (Figura 1), observa que el promedio alcanza 6,81, mientras que la mediana se sitúa en 6,65.

Figura 1.
Estadísticos descriptivos de la calificación final sin recuperación.

	Estudiante	Tarea	Proyecto	Examen_final	Recupera...	Promocio...	Nota_final_sin_rec
1	S1	9	8	7		1	7.7
2	S2	7	8	6	8	1	6.8
3	S3	6	7	5	7	1	5.8
4	S4	9	9	9		1	9.0
5	S5	8	6	5		0	5.9
6	S6	7	7	6		1	6.5
7	S7	5	6	4	7	1	4.8
8	S8	8	8	8		1	8.0

Nota. La figura presenta los estadísticos descriptivos calculados en Jamovi a partir del rendimiento final de 8 estudiantes.

Como se muestra en la Figura 2, estas cifras muestran que la mayoría de los estudiantes se mueve en un rango de desempeño intermedio, sin grandes distancias entre la media y el valor central. Aun así, el comportamiento de las notas evidencia diferencias importantes: el estudiante con menor puntuación obtiene 4,80, mientras que el de mayor rendimiento llega a 9,00.

Esa amplitud refleja que algunos avanzan con mayor seguridad, mientras otros requieren un acompañamiento más cercano. Si bien estas cifras no capturan la complejidad completa del aprendizaje, sí ofrecen una primera lectura útil para que el equipo docente identifique tendencias, reconozca necesidades

y ajuste sus estrategias de enseñanza con miras a fortalecer el proceso formativo.

Figura 2.
Estadísticos descriptivos del puntaje del proyecto

Descriptivas

Descriptivas

	Proyecto
N	8
Perdidos	0
Media	7.38
Mediana	7.50
Desviación típica	1.06
Mínimo	6
Máximo	9

Nota. La figura resume los valores descriptivos obtenidos en la variable "Proyecto", incluyendo la media, mediana, desviación típica y los puntajes mínimo y máximo registrados en el grupo.

Esa amplitud refleja que algunos avanzan con mayor seguridad, mientras otros requieren un acompañamiento más cercano. Si bien estas cifras no capturan la complejidad completa del aprendizaje, sí ofrecen una primera lectura útil para que el equipo docente identifique tendencias, reconozca necesidades y ajuste sus estrategias de enseñanza con miras a fortalecer el proceso formativo.

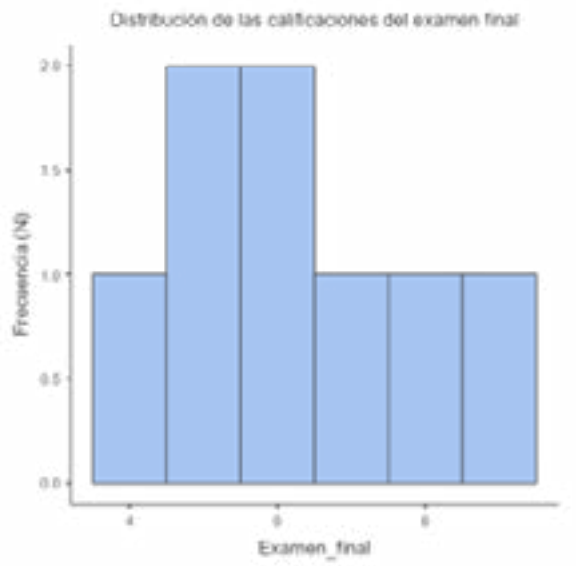
Representar para comprender: narrativas estadísticas y sus límites

Las representaciones estadísticas (gráficos, tablas, diagramas y resúmenes numéricos) ocupan un papel central en la comprensión de los fenómenos. Una tabla bien organizada o un gráfico claro puede revelar patrones, tendencias y relaciones que no son visibles en los datos en bruto. Sin embargo, representan más que una simple traducción; construyen narraciones que orientan la interpretación del fenómeno. Curcio (1989) enfatiza que la comprensión de los gráficos depende tanto del diseño como de las capacidades lectoras del observador, y ambos aspectos pueden modificar sustancialmente el sentido atribuido a los datos.

La selección del tipo de representación ya es, en sí misma, una elección cargada de significado. Un gráfico de líneas sugiere continuidad temporal; un diagrama de barras destaca comparaciones entre categorías; un histograma permite visualizar la forma de distribución; un boxplot revela variabilidad y valores atípicos. En el aula, muchos estudiantes interpretan cada gráfico como si mostrara la "verdad" de los datos, sin notar que cada representación enfoca un aspecto y deja otros en segundo plano.

Un ejemplo ilustrativo aparece cuando se comparan histogramas de diferentes tamaños de intervalo: la misma distribución puede verse dispersa o concentrada según cómo se definan los rangos, generando conclusiones distintas. Por ejemplo, en la Figura 3 se observa que las calificaciones del examen final tienden a concentrarse en un rango intermedio, principalmente entre 5 y 7 puntos, donde se ubica la mayoría del grupo. Solo aparecen dos casos que se apartan de esa franja: un estudiante con un puntaje alto de 9 y otro con una nota baja de 4. Esta distribución sugiere un nivel de rendimiento relativamente uniforme, aunque con diferencias puntuales que permiten identificar tanto un desempeño destacado como una dificultad aislada dentro del grupo.

Figura 3.
Histograma de la calificación final sin recuperación.



Nota. El gráfico muestra la distribución de las calificaciones finales sin recuperación obtenidas por ocho estudiantes, generada mediante el software Jamovi.

Las elecciones estéticas también influyen en la interpretación. El uso de colores intensos, ejes truncados o escalas desiguales puede exagerar diferencias mínimas o minimizar patrones

importantes. Esto se observa con frecuencia en medios de comunicación y redes sociales. Un gráfico que muestra el aumento de un indicador de salud puede parecer alarmante si el eje inicia cerca del valor máximo, aun cuando el cambio real sea pequeño. Enseñar a identificar estos efectos ayuda a los estudiantes a desarrollar una lectura más cuidadosa de las representaciones visuales.

La narrativa que construye un gráfico también depende del orden de los elementos. Una tabla puede mostrar los datos ordenados alfabéticamente o según valor numérico, dando énfasis a diferentes aspectos. En un gráfico de barras, el orden puede sugerir tendencias inexistentes.

Por ejemplo, cuando las calificaciones se representan respetando el orden original de los estudiantes en la tabla, la gráfica ofrece una imagen mucho menos lineal y, en apariencia, más caótica. Las barras suben y bajan sin seguir una secuencia reconocible, lo que refleja con mayor fidelidad la variabilidad real del grupo. En este caso, el gráfico no sugiere ninguna tendencia general de mejora o deterioro, sino que muestra simplemente las diferencias individuales de cada estudiante (Figura 4).

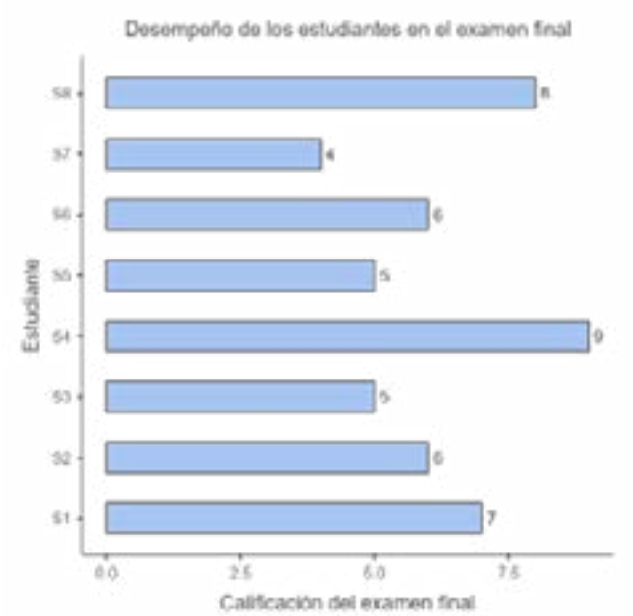
Esta representación resulta más transparente porque evita imponer una estructura visual que no está presente en los datos. Ver el gráfico sin ordenar permite comprender el rendimiento desde una perspectiva más abierta, donde lo relevante no es la forma global del dibujo, sino las particularidades de cada caso. Al contrastarlo con la versión ordenada, se hace evidente cómo pequeñas decisiones de presentación pueden modificar la narrativa visual sin modificar los valores, recordándonos que toda representación gráfica implica una interpretación y no únicamente una descripción literal de los datos.

Apoyo didáctico: En contextos educativos, la lectura de gráficos se convierte en una actividad fundamental para desarrollar el pensamiento estadístico. Garfield y Ben-Zvi (2008) señalan que los estudiantes deben aprender no solo a "leer" datos, sino a "leer entre los datos": preguntarse qué se destaca, qué se omite y qué decisiones gráficas influyen en lo que se percibe. Actividades como comparar diferentes representaciones del mismo conjunto de datos permiten observar cómo cambia la interpretación según el modelo elegido. Esta experiencia resulta reveladora, pues los estudiantes descubren que la representación no es neutra.

Otro elemento clave es el uso de medidas de resumen. La media, la mediana, el rango o la desviación estándar son herramientas poderosas para sintetizar información, pero también pueden simplificar en exceso el fenómeno. Una media puede ocultar desigualdades internas, mientras que un rango no informa

sobre la distribución real de los valores. Cuando se presentan únicamente los resúmenes numéricos, se corre el riesgo de transmitir una imagen parcial. Wild y Pfannkuch (1999) destacan que el pensamiento estadístico requiere ir y venir entre lo global y lo particular, evitando quedarse solo con uno de esos niveles.

Figura 4.
Distribución de las calificaciones finales sin recuperación



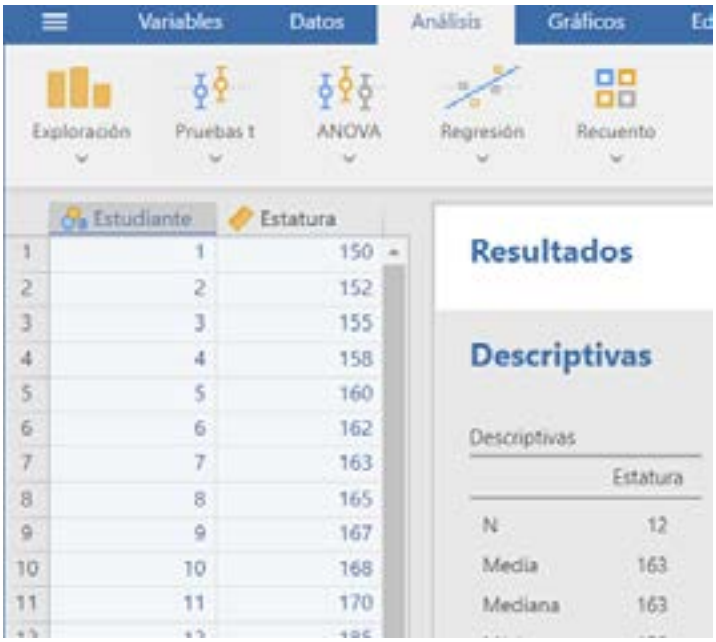
Nota. La figura presenta la distribución de las calificaciones finales sin recuperación obtenidas por ocho estudiantes. El gráfico fue generado en el software Jamovi y permite observar, de manera comparativa, el desempeño individual de cada estudiante en el examen final.

Otro elemento clave es el uso de medidas de resumen. La media, la mediana, el rango o la desviación estándar son herramientas poderosas para sintetizar información, pero también pueden simplificar en exceso el fenómeno. Una media puede ocultar desigualdades internas, mientras que un rango no informa sobre la distribución real de los valores. Cuando se presentan únicamente los resúmenes numéricos, se corre el riesgo de transmitir una imagen parcial. Wild y Pfannkuch (1999) destacan que el pensamiento estadístico requiere ir y venir entre lo global y lo particular, evitando quedarse solo con uno de esos niveles.

Ejemplo: en una clase de primer año de bachillerato, el docente decidió registrar la estatura de sus 12 estudiantes con el fin de analizar la variabilidad del grupo y trabajar conceptos básicos de estadística. Las estaturas, medidas en centímetros, fueron las siguientes: 150, 152, 155, 158, 160, 162, 163, 165, 167, 168, 170 y 185 (Figura 5). Con esta información desea describir la distribución

del grupo, identificar medidas de tendencia central y reconocer posibles diferencias entre estudiantes. A partir de estos datos, el objetivo es que los alumnos comprendan cómo se comporta una variable cuantitativa continua en un conjunto real de personas y qué conclusiones pueden extraerse de su análisis.

Figura 5.
Distribución de las estaturas registradas en el grupo de estudio



Nota. La figura muestra la tabla descriptiva generada en Jamovi a partir de las estaturas registradas para los doce participantes del grupo. Se incluyen medidas de tendencia central y dispersión que permiten observar la variabilidad del conjunto, así como el valor del percentil 25, útil para interpretar la distribución de los datos.

La tabla de descriptivos muestra que se trabajó con las estaturas de 12 estudiantes, cuya media y mediana coinciden en 163 cm. Este dato sugiere que el centro de la distribución está claramente ubicado alrededor de ese valor y que no hay una asimetría marcada en la parte central del conjunto. El mínimo registrado es de 150 cm y el máximo de 185 cm, lo que indica un rango amplio de estaturas dentro del grupo. Además, el percentil 25 se sitúa en 157 cm, de modo que una cuarta parte de los estudiantes mide menos de esa cifra, mientras que el resto se concentra por encima. En conjunto, la tabla permite apreciar que, aunque la mayoría de las estaturas se agrupa en torno a valores medios, existe al menos un estudiante significativamente más alto, lo que aporta variabilidad al grupo y ayuda a discutir cómo los valores extremos influyen en el análisis estadístico.

En definitiva, representar datos significa construir una narrativa visual o numérica que guía la interpretación. Reconocer esta característica es fundamental para evitar lecturas superficiales y para apreciar que cada representación ilumina aspectos distintos de la realidad.

Enseñar esta dimensión narrativa permite que los estudiantes comprendan que los gráficos no solo muestran; también cuentan historias que deben ser interpretadas con cuidado.

De lo individual a lo colectivo: tensiones y efectos de las agregaciones estadísticas

Uno de los desafíos más significativos de la estadística consiste en traducir experiencias individuales en descripciones colectivas. Esta transición entre lo particular y lo general constituye un proceso complejo en el que se sintetizan múltiples realidades para obtener indicadores globales. Sin embargo, esta síntesis puede generar tensiones importantes, especialmente cuando los resúmenes estadísticos ocultan variaciones internas que resultan relevantes para comprender el fenómeno.

Gal (2002) advierte que la interpretación de estadísticas agregadas requiere habilidades que permitan distinguir entre patrones colectivos y comportamientos individuales. En el aula, este desafío se evidencia cuando los estudiantes analizan encuestas que mencionan promedios, porcentajes o medianas sin mostrar la distribución completa.

Por ejemplo: Supón que se aplica una encuesta rápida a 20 estudiantes de un colegio para saber si leen al menos un libro al mes. Además, se registra: Curso: (8.º / 9.º), Género (Mujer / Hombre) y si Lee libro al mes (Sí / No) (Figura 6). Como se podrá comprobar "el 70 por ciento de los estudiantes lee al menos un libro al mes" ofrece una visión general del grupo, pero no revela si existen diferencias significativas entre subgrupos, como curso, género, acceso a recursos o motivación lectora.

Los resultados muestran que, de los 20 estudiantes encuestados, 14 afirman leer al menos un libro al mes, lo que representa el 70 % del grupo, mientras que 6 (30 %) señalan que no tienen este hábito de lectura mensual (Figura 7).

A primera vista, la cifra de quienes sí leen sugiere una práctica relativamente extendida de lectura en el estudiante, pero al mismo tiempo evidencia que casi un tercio permanece al margen de esa rutina. Esta información permite sostener el enunciado de que "el 70 por ciento de los estudiantes lee al menos un libro al mes", aunque sigue siendo un dato global: no indica si este comportamiento se distribuye de la misma manera entre cursos, géneros u otras características, por lo que todavía no es posible saber si existen subgrupos con mayor o menor participación en la lectura.

Figura 6.
Distribución registrada de la estatura y características del grupo de estudio

	Estudiante	C...	Género	Lectura de al menos un libro al
1	1	8.º	Mujer	Sí
2	2	8.º	Mujer	Sí
3	3	8.º	Mujer	No
4	4	8.º	Mujer	Sí
5	5	8.º	Hombre	No
6	6	8.º	Hombre	No
7	7	8.º	Hombre	Sí
8	8	8.º	Hombre	No
9	9	9.º	Mujer	Sí
10	10	9.º	Mujer	Sí
11	11	9.º	Mujer	Sí
12	12	9.º	Mujer	Sí
13	13	9.º	Hombre	Sí
14	14	9.º	Hombre	Sí
15	15	9.º	Hombre	No
16	16	9.º	Hombre	Sí
17	17	9.º	Mujer	No
18	18	9.º	Hombre	Sí

Nota. La figura muestra la base de datos utilizada para el análisis descriptivo del grupo de estudiantes generada en Jamovi.

Figura 7.
Distribución de estudiantes según la lectura de al menos un libro al mes

Recuento

Frequencias de Lectura de al menos un libro al mes

Lectura de al menos un libro al mes	Recuentos	% del Total	% Acumulado
Sí	14	70.0 %	70.0 %
No	6	30.0 %	100.0 %

Nota. La tabla resume las respuestas de los 20 estudiantes encuestados sobre si leen al menos un libro al mes.

Estas tensiones también se observan en el análisis del rendimiento académico. Cuando se presentan los resultados de una evaluación mediante una media y una desviación estándar, se obtiene una descripción global del grupo.

Sin embargo, esta representación puede ocultar la existencia de estudiantes que requieren apoyo específico. Dos grupos con la misma media pueden tener distribuciones muy diferentes: en uno, todos los estudiantes pueden situarse cerca del promedio; en el otro, puede haber brechas considerables entre quienes obtienen puntajes altos y bajos. En contextos educativos, estas diferencias son esenciales para planificar estrategias pedagógicas personalizadas.

La cuestión de las agregaciones cobra especial relevancia cuando se analizan fenómenos sociales complejos, como la distribución de recursos, las condiciones de vida o la participación ciudadana. Ben-Zvi y Makar (2016) sostienen que los estudiantes deben aprender a examinar estos indicadores desde una perspectiva que considere tanto los patrones globales como las dinámicas internas. Por ejemplo, un estudio sobre el acceso a internet puede mostrar que el "80 por ciento de los hogares posee conexión", pero ese número no informa sobre la calidad del servicio, la disponibilidad de dispositivos o el uso efectivo de plataformas digitales. La agregación oculta diferencias que pueden ser cruciales para comprender desigualdades educativas.

Las tensiones entre lo individual y lo colectivo se vuelven aún más visibles cuando se consideran fenómenos donde existen valores extremos. Un ejemplo habitual en clase consiste en analizar ingresos familiares. La media puede verse fuertemente influida por pocos valores muy altos, mientras que la mediana puede ofrecer una descripción más fiel del comportamiento típico. En este caso, la media no representa adecuadamente al grupo. Estas situaciones muestran que la elección del indicador no solo depende de criterios técnicos, sino también del propósito de la descripción.

Apoyo didáctico: En la práctica docente, trabajar con distribuciones completas permite que los estudiantes comprendan la riqueza y complejidad de los datos. Actividades donde se comparan diferentes grupos, donde se identifican segmentos con comportamientos divergentes o donde se analizan valores atípicos ayudan a reconocer que la realidad social es diversa y que las estadísticas deben interpretarse con matices. La comprensión profunda de estas tensiones favorece una mirada más sensible hacia las diferencias dentro de los grupos.

Por otra parte, la agregación también tiene efectos en la toma de decisiones. Cuando las instituciones educativas utilizan indicadores globales para establecer políticas, existe el riesgo de ignorar las particularidades de ciertos estudiantes o contextos. La estadística puede contribuir a decisiones informadas, pero solo si se reconoce la complejidad de los datos y se evita la tentación

de reducir los fenómenos a cifras únicas. En el aula, esto se puede trabajar mediante estudios de caso donde los estudiantes analicen cómo varían las conclusiones según se consideren o no los detalles internos de la distribución.

Tipos de datos y variables en contextos educativos y sociales

Comprender los tipos de datos y las variables que los generan es esencial para analizar cualquier fenómeno educativo o social con seriedad intelectual. Aunque la clasificación de datos suele presentarse como un procedimiento técnico, su alcance es mucho más profundo: permite reconocer el modo en que los fenómenos se conceptualizan, se traducen en información y finalmente se interpretan. En palabras de Garfield y Ben-Zvi (2008), la estadística no opera sobre una copia exacta de la realidad, sino sobre una versión conceptualizada de ella. Esta idea resulta decisiva en educación, donde los fenómenos incluyen emociones, percepciones, comportamientos, aprendizajes y contextos que rara vez son simples de reducir a números.

La literatura en educación estadística ha mostrado que los datos no son entidades neutrales. Wild y Pfannkuch (1999) explican que todo análisis estadístico comienza con decisiones invisibles: qué observar, cómo medir, cómo registrar y qué considerar relevante. Por ello, reflexionar sobre los tipos de datos y variables no solo fortalece la comprensión técnica de la estadística, sino que también promueve un pensamiento crítico indispensable para interpretar fenómenos educativos y sociales en toda su complejidad.

A continuación, se desarrollan tres dimensiones fundamentales: la naturaleza del dato, la clasificación cualitativa y cuantitativa, y la construcción de variables que permiten comprender y representar la realidad desde múltiples ángulos.

Comprender la naturaleza del dato: significados, decisiones y contextos

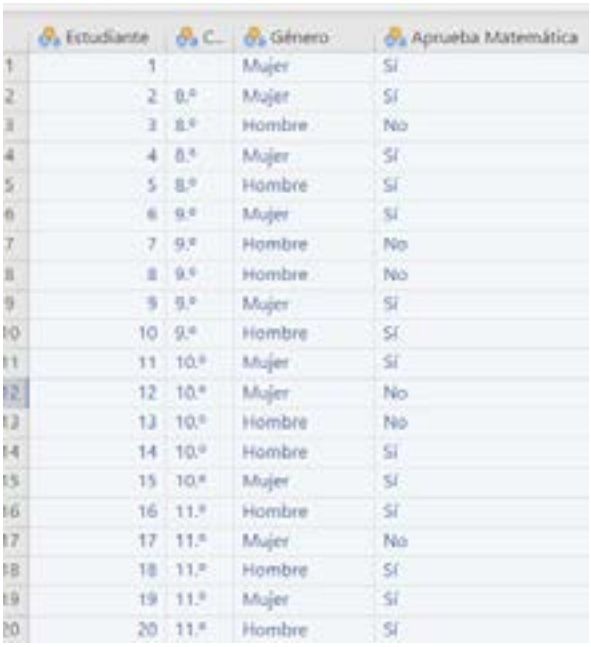
Cada dato que aparece en un gráfico, una tabla o un informe escolar representa una serie de decisiones previas: qué se considera relevante, qué se omite, qué se clasifica y qué se cuantifica. Esta dimensión, a menudo invisible, configura la forma en que comprendemos los fenómenos. Según Gal (2002), interpretar datos sin preguntarse por su origen conduce a conclusiones apresuradas, pues la información nunca está desligada del contexto que la produce.

En educación, este asunto se hace evidente en la evaluación del aprendizaje. Cuando un docente asigna una calificación numérica, esa cifra parece clara, objetiva y comparable.

Sin embargo, detrás de ella existe un entramado metodológico: criterios de evaluación, ponderaciones, tipos de tareas, carácter de las actividades y nivel de complejidad de los ítems. Batanero y Díaz (2011) señalan que incluso en contextos aparentemente cuantitativos, como la evaluación escolar, intervienen componentes cualitativos que moldean el dato final.

Un ejemplo puede estar referido a que un informe puede indicar que "el 90 por ciento de los estudiantes aprueba Matemática", lo cual parece dar una imagen positiva (Figura 8),es decir los datos del grupo pueden dar la impresión de que el rendimiento en Matemática es homogéneo, pues la mayoría de estudiantes aprueba la asignatura.

Figura 8.
Registro de las características del grupo de estudio utilizadas para el análisis



	Estudiante	C...	Género	Aprueba Matemática
1	1		Mujer	Sí
2	2	8.°	Mujer	Sí
3	3	8.°	Hombre	No
4	4	8.°	Mujer	Sí
5	5	8.°	Hombre	Sí
6	6	9.°	Mujer	Sí
7	7	9.°	Hombre	No
8	8	9.°	Hombre	No
9	9	9.°	Mujer	Sí
10	10	9.°	Hombre	Sí
11	11	10.°	Mujer	Sí
12	12	10.°	Mujer	No
13	13	10.°	Hombre	No
14	14	10.°	Hombre	Sí
15	15	10.°	Mujer	Sí
16	16	11.°	Hombre	Sí
17	17	11.°	Mujer	No
18	18	11.°	Hombre	Sí
19	19	11.°	Mujer	Sí
20	20	11.°	Hombre	Sí

Nota. La figura muestra la base de datos ingresada en Jamovi con información del curso, género y desempeño académico de los 20 estudiantes encuestados.

Sin embargo, cuando se revisa con detenimiento la distribución por cursos (Figura 9), tal como se muestra en la tabla, aparecen matices que el valor global por sí solo no revela. Aunque en cada curso se observa un predominio de estudiantes que aprueban, la proporción de aprobados y no aprobados no es idéntica entre niveles.

Figura 9.
Distribución de estudiantes que aprueban Matemática según el curso

Tablas de Contingencia

Tablas de Contingencia

	Aprueba Matemática		
Curso	Sí	No	Total
8.º	4	1	5
9.º	3	2	5
10.º	3	2	5
11.º	4	1	5
Total	14	6	20

Nota. La tabla resume cuántos estudiantes aprobaron y reprobaron Matemática en cada curso.

En 8.º y 11.º, por ejemplo, la aprobación es claramente mayoritaria, mientras que en 9.º y 10.º el número de estudiantes que no superan la materia es más visible. Esto demuestra que la cifra total funciona como un indicador sintético, pero también puede ocultar diferencias internas importantes. Para comprender el fenómeno educativo con mayor precisión, no basta con quedarse en el porcentaje general: es necesario observar los subgrupos, contrastar patrones y situar el dato dentro de un contexto más amplio.

Otros ejemplos pueden describir situaciones similares como, por ejemplo: supongamos que un investigador desea medir la "participación en clase". Podría registrarla como número de intervenciones orales (cuantitativo), como categoría (alta, media, baja), como nivel de iniciativa (ordinal) o mediante notas de observación (cualitativo descriptivo). Cada modalidad produce datos distintos, y por lo tanto, interpretaciones distintas. Wild y Pfannkuch (1999) afirman que estas decisiones forman parte del pensamiento estadístico y deben enseñarse explícitamente para que los estudiantes comprendan el carácter construido del dato.

Por ello, comprender la naturaleza del dato implica reconocer que todo registro es fruto de una decisión. Formar estudiantes capaces de analizar críticamente información requiere enseñarles que los datos no son simples productos de medición, sino interpretaciones codificadas de la realidad.

Datos cualitativos y cuantitativos: matices, tensiones y posibilidades analíticas

La distinción entre datos cualitativos y cuantitativos es fundamental, pero también es frecuente caer en simplificaciones. En educación y sociedad, estas categorías conviven y se complementan, y su interacción permite captar la complejidad de los fenómenos. Garfield y Ben-Zvi (2008) indican que cada tipo de dato abre y cierra determinadas posibilidades analíticas, por lo que la elección del tipo de dato nunca es trivial.

Datos cualitativos: comprender significados y categorías

Los datos cualitativos permiten registrar aspectos que no se pueden expresar numéricamente: percepciones, categorías, experiencias, narraciones o clasificaciones no ordinales. Estos datos son indispensables para comprender dimensiones humanas como la motivación, el clima escolar, la convivencia o la percepción docente. Por eso, en educación y ciencias sociales, los datos cualitativos no constituyen un complemento, sino una parte esencial de la representación estadística.

Entre los datos cualitativos, los nominales identifican categorías sin orden natural. Ejemplos frecuentes son:

- tipo de recurso digital utilizado,
- asignatura favorita,
- rol asumido en el trabajo grupal,
- área de interés profesional.

Esta información ayuda a comprender la diversidad del grupo, aunque no permite establecer jerarquías.

En cambio, los ordinales sí establecen un orden, aunque sin distancias cuantificables entre categorías:

- niveles de motivación,
- percepción del clima escolar,
- calidad del aprendizaje según el propio estudiante,
- niveles de logro.

Gal (2002) advierte que una dificultad habitual aparece cuando los datos ordinales se convierten artificialmente en números para realizar promedios. Esa práctica puede dar una falsa sensación de precisión. Por ejemplo, asignar valores del 1 al 5 a percepciones sobre satisfacción puede ser útil para análisis exploratorios, pero promediar esas respuestas no implica que las distancias entre categorías sean iguales.

Datos cuantitativos: magnitudes, decisiones y límites

Los datos cuantitativos permiten medir cantidades y realizar operaciones matemáticas. No obstante, que sean numéricos no significa que sean precisos o neutros. Como señalan Garfield y Ben-Zvi (2008), todo dato cuantitativo está mediado por un instrumento y por un criterio de medición.

Los **datos discretos** representan conteos enteros: número de tareas realizadas, de errores cometidos o de libros leídos.

Los **datos continuos** pueden tomar muchos valores dentro de un intervalo: tiempo de estudio, temperatura, calificaciones con decimales.

Sin embargo, en educación, muchos datos que parecen continuos provienen de escalas discretas. Una calificación de 8,7 da la sensación de exactitud, pero su significado depende de la estructura de la evaluación. Wild y Pfannkuch (1999) explican que la precisión numérica no es sinónimo de precisión conceptual.

Tensiones entre ambos tipos de datos

En fenómenos educativos y sociales, lo cualitativo y lo cuantitativo se entrelazan. Investigar la motivación, por ejemplo, requiere articular:

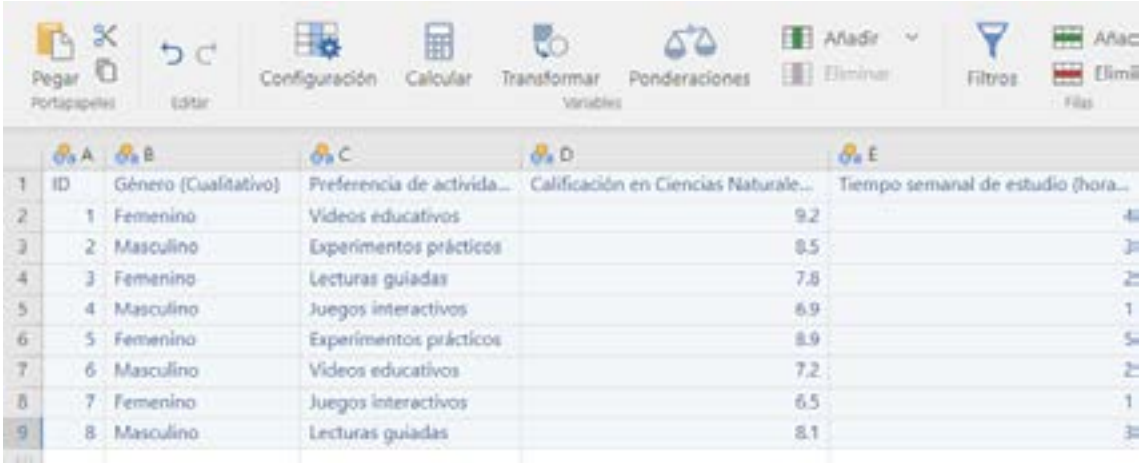
- datos cuantitativos (tiempo dedicado, tareas completadas),
- datos cualitativos (razones, percepciones, emociones),
- y datos ordinales (nivel de interés, disposición para aprender).

Ben-Zvi y Makar (2016) sostienen que los análisis más sólidos provienen de integrar diferentes tipos de datos, pues permiten captar matices que un único enfoque no logra mostrar. Esta visión integrada resulta especialmente relevante en educación, donde las variables contienen dimensiones afectivas, cognitivas, sociales y contextuales.

Un ejemplo de la utilización de estas variables (Figura 10) puede estar referido a la necesidad que tenga un docente de comprender por qué algunos estudiantes muestran mejores resultados en Ciencias Naturales que otros, a pesar de recibir las mismas explicaciones en clase.

Para tener una mirada más clara, el docente puede decidir recopilar información sencilla pero significativa: qué tipo de actividades disfrutan más, cuánto tiempo dedican a estudiar en casa y cómo les fue en la última evaluación. Con esa pequeña muestra, se puede descubrir si las preferencias de aprendizaje y el esfuerzo individual se reflejan en el rendimiento académico, y, sobre todo, qué tipo de actividades podría potenciar en sus clases. A partir de los datos registrados, se espera analizar los patrones que surgen y plantear recomendaciones que permitan reforzar el proceso de enseñanza y aprendizaje con estrategias más acordes a las necesidades reales del grupo.

Figura 10.
Datos cualitativos y cuantitativos registrados



	A	B	C	D	E
1	ID	Género (Cualitativo)	Preferencia de actividad...	Calificación en Ciencias Naturale...	Tiempo semanal de estudio (hora...
2	1	Femenino	Videos educativos	9.2	4
3	2	Masculino	Experimentos prácticos	8.5	3
4	3	Femenino	Lecturas guiadas	7.8	2
5	4	Masculino	Juegos interactivos	6.9	1
6	5	Femenino	Experimentos prácticos	8.9	5
7	6	Masculino	Videos educativos	7.2	2
8	7	Femenino	Juegos interactivos	6.5	1
9	8	Masculino	Lecturas guiadas	8.1	3

Nota. La figura presenta la información recopilada por el docente para comprender mejor las características y el desempeño de los estudiantes.

Apoyo didáctico: Para comprender mejor por qué algunos estudiantes avanzan con mayor facilidad en Ciencias Naturales, el docente puede trabajar con un conjunto de variables que le permitan observar tanto aspectos descriptivos del grupo como indicadores de rendimiento. Por un lado, incorporar variables cualitativas, como el género y la actividad de aprendizaje que cada estudiante prefiere, porque estos datos ofrecen pistas sobre los estilos de participación, la motivación y las formas en que cada niño se involucra con los contenidos. Por otro lado, incluir variables cuantitativas, como la calificación obtenida y las horas semanales de estudio, ya que aportan medidas objetivas que permiten contrastar el desempeño real y el nivel de dedicación fuera del aula.

Con esta combinación, el docente puede identificar si existe alguna relación entre la manera en que los estudiantes aprenden y los resultados que alcanzan, y así determinar qué tipo de estrategias podría fortalecer en sus clases. El análisis de esta información pretende, en última instancia, orientar decisiones pedagógicas más ajustadas a las necesidades y características del grupo.

Variables educativas y sociales: construcción, interpretación y efectos

Las variables constituyen el eje estructural de cualquier análisis estadístico. Sin ellas, los datos carecen de forma y las preguntas de investigación quedan en un plano abstracto. En contextos educativos y sociales, definir adecuadamente una variable implica un ejercicio de interpretación que va mucho más allá de escoger un nombre o un formato de registro. Como señalan Garfield y

Ben-Zvi (2008), toda variable refleja una elección conceptual que delimita qué aspectos del fenómeno serán observados, cuáles permanecerán invisibles y qué sentido adquieren los valores que posteriormente se analizarán. Por ello, comprender la construcción de las variables es comprender también la manera en que representamos la realidad.

Esta dimensión es particularmente relevante en educación, donde se estudian fenómenos complejos que integran dimensiones cognitivas, afectivas, sociales y culturales. Variables como "aprendizaje", "participación", "motivación", "clima escolar", "bienestar emocional" o "competencia digital" no son objetos simples ni directamente observables. Cada una puede abordarse desde múltiples perspectivas, y la manera en que se elige operacionalizarlas condiciona profundamente la interpretación posterior. Wild y Pfannkuch (1999) destacan que la calidad de un análisis estadístico depende en gran parte de la solidez conceptual con la que se definan las variables iniciales, pues estas actúan como filtros que dan forma a la realidad analizada.

Definir una variable implica tomar posición respecto de lo que se considera importante en un fenómeno. En el estudio del rendimiento académico, por ejemplo, existen múltiples formas de conceptualizarlo:

1. Como puntaje en pruebas estandarizadas.

Ofrece datos comparables, pero captura solo una parte del aprendizaje.

2. Como promedio de calificaciones.

Integra actividades diversas, aunque puede verse influido por criterios poco homogéneos.

3. Como logro en estándares curriculares.

Se centra en desempeños específicos, pero exige sistemas de evaluación consistentes.

4. Como desarrollo de competencias.

Refleja procesos, pero requiere instrumentos más complejos y subjetivos.

5. Como autopercepción del propio aprendizaje.

Incorpora la voz del estudiante, aunque no mide directamente habilidades objetivas.

Cada definición origina una variable distinta y conduce a "versiones" diferentes del mismo fenómeno. Batanero y Díaz (2011) explican que la elección de una variable no es neutra: implica escoger un ángulo de análisis y renunciar a otros posibles. Esta renuncia es inevitable, pero debe ser consciente y argumentada.

Supongamos que una institución evalúa la efectividad de su nuevo programa de Matemática. Si la variable seleccionada es "puntaje en pruebas estandarizadas", se obtendrá una lectura

centrada en habilidades específicas. Si la variable elegida es "confianza matemática", la lectura será emocional y motivacional. Si la variable utilizada es "uso de estrategias de resolución de problemas", la evaluación se centrará en prácticas cognitivas. Una misma intervención puede parecer exitosa o limitada según la variable escogida.

Este ejemplo evidencia por qué el acto de definir variables constituye una parte esencial del análisis y no un simple paso preliminar.

Variables observables y variables latentes

En educación y ciencias sociales es frecuente distinguir entre variables observables y variables latentes.

Variables observables

Son aquellas que pueden registrarse directamente mediante instrumentos concretos, como:

- número de horas de estudio;
- asistencia diaria;
- cantidad de intervenciones orales;
- puntaje obtenido en una prueba;
- frecuencia de uso de una plataforma digital.

Estas variables permiten un registro más objetivo, pero no necesariamente capturan la totalidad del fenómeno. Por ejemplo, medir la "participación" solo como número de intervenciones puede invisibilizar a estudiantes que participan mediante la escucha activa, el trabajo colaborativo o la escritura.

Variables latentes

Las variables latentes representan constructos no directamente observables, tales como:

- motivación;
- sentido de pertenencia;
- ansiedad matemática;
- clima escolar;
- liderazgo estudiantil;
- percepción de competencia.

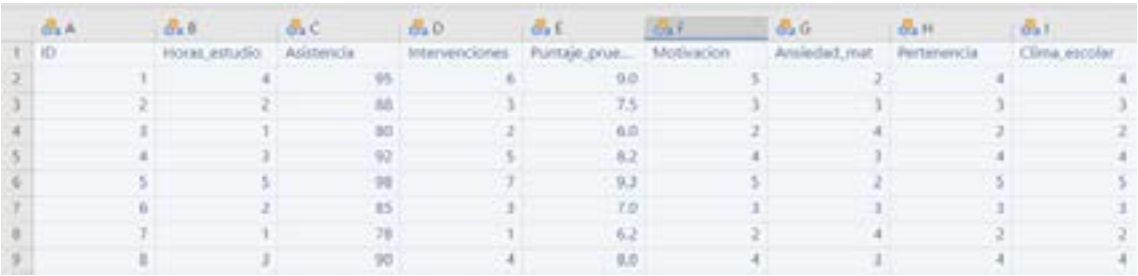
Gal (2002) recuerda que estas variables requieren procedimientos indirectos de medición implica asumir modelos teóricos y criterios interpretativos. Su complejidad no las vuelve menos valiosas; al contrario, permiten estudiar dimensiones profundas de la experiencia educativa. No obstante, requieren interpretaciones cuidadosas, pues pueden variar según el instrumento utilizado y el contexto de aplicación.

Por ejemplo, con el propósito de comprender mejor cómo se relacionan ciertos hábitos académicos con factores más profundos del aprendizaje, una docente decidió recopilar información de ocho estudiantes de quinto año.

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

Para ello registró varias variables observables, como las horas de estudio semanal, el porcentaje de asistencia, el número de intervenciones orales y el puntaje obtenido en una prueba reciente (Figura 11). Junto con ello, aplicó una escala tipo Likert para evaluar variables latentes vinculadas al funcionamiento emocional y social del grupo, específicamente la motivación, la ansiedad matemática, el sentido de pertenencia y la percepción del clima escolar.

Figura 11.
Registro de variables observables y latentes en el grupo de estudiantes



	A	B	C	D	E	F	G	H	I
1	ID	Horas_estudio	Asistencia	Intervenciones	Puntaje_prue...	Motivacion	Ansiedad_mat	Pertenencia	Clima_escolar
2	1	4	95	6	9.0	5	2	4	4
3	2	2	88	3	7.5	3	3	3	3
4	3	1	80	2	6.0	2	4	2	2
5	4	3	92	5	8.2	4	3	4	4
6	5	5	98	7	9.3	5	2	5	5
7	6	2	85	3	7.0	3	3	3	3
8	7	1	78	1	6.2	2	4	2	2
9	8	3	90	4	8.0	4	3	4	4

Nota. La figura muestra la matriz de datos ingresada en Jamovi, donde se combinan variables observables, como horas de estudio, asistencia, intervenciones y puntaje en la prueba, junto con variables latentes evaluadas mediante escala Likert, entre ellas motivación, ansiedad matemática, sentido de pertenencia y clima escolar.

El interés del docente será identificar si existen patrones entre estos dos tipos de variables, por ejemplo, si los estudiantes con mayor motivación tienden a estudiar más, si la ansiedad matemática se refleja en menores puntajes o si un clima escolar más favorable se asocia con mejores intervenciones en clase. Con la información registrada, se busca analizar estas relaciones para fundamentar decisiones pedagógicas que ayuden a fortalecer el aprendizaje y el bienestar del grupo.

Operacionalizar: transformar un concepto en una medida

Operacionalizar significa traducir un concepto abstracto en indicadores concretos y medibles. Este proceso exige una comprensión clara del constructo teórico.

Para comprender de manera más precisa la competencia digital docente (Figura 12), un investigador puede comenzar identificando aquellos aspectos que son visibles en la práctica educativa.

Entre ellos se encuentran el nivel de alfabetización tecnológica, la variedad de herramientas digitales que el docente emplea y la frecuencia con la que las utiliza en sus clases.

Figura 12.
Operacionalización de la competencia digital docente

Dimensión	Indicador	Descripción del indicador	Tipo de variable	Ejemplo de ítem (escala Likert 1–5)
Alfabetización tecnológica	Nivel de alfabetización tecnológica	Grado de familiaridad con conceptos básicos y uso inicial de recursos digitales.	Ordinal	"Me siento capaz de manejar funciones básicas de dispositivos y aplicaciones digitales".
Uso de herramientas digitales	Variedad de herramientas digitales empleadas	Número y diversidad de aplicaciones, plataformas o recursos digitales utilizados en la práctica docente.	Cuantitativa discreta	"En mis clases uso al menos tres tipos diferentes de herramientas digitales".
Frecuencia de uso	Frecuencia de uso de herramientas digitales	Regularidad con la que el docente incorpora recursos digitales en su labor diaria.	Ordinal	"Utilizo herramientas digitales para preparar o desarrollar mis clases varias veces por semana".
Adaptación a entornos virtuales	Capacidad de adaptación a entornos virtuales	Facilidad para desenvolverse en modalidades virtuales y para resolver imprevistos tecnológicos.	Ordinal	"Me adapto con rapidez a nuevas plataformas o cambios en los entornos virtuales de aprendizaje".
Confianza digital	Confianza percibida en el manejo de plataformas	Nivel de seguridad personal al interactuar con sistemas digitales educativos.	Ordinal	"Siento confianza al utilizar plataformas de gestión del aprendizaje con mis estudiantes".

Nota. La figura detalla las dimensiones, indicadores y tipos de variables empleados para operacionalizar la competencia digital docente. Se incluyen además ejemplos de ítems formulados en escala Likert de 1 a 5.

Sin embargo, también intervienen dimensiones menos evidentes que requieren una mirada más profunda. La capacidad de adaptación a entornos virtuales, la confianza que el docente percibe al manejar plataformas digitales y su habilidad para integrar esos recursos con intención pedagógica constituyen componentes esenciales de la competencia digital.

Apoyo didáctico: Estos elementos permiten observar cómo el profesorado incorpora la tecnología en su rutina pedagógica y hasta qué punto domina funciones básicas y aplicaciones que facilitan el desarrollo de actividades de enseñanza y aprendizaje, así como entender no solo el uso instrumental de las herramientas, sino también cómo el docente las incorpora de manera estratégica para enriquecer los procesos educativos y responder a las demandas de los entornos digitales actuales.

Cada uno de estos aspectos representa dimensiones diferentes de una misma variable. Garfield y Ben-Zvi (2008) sostienen que la operacionalización debe ser coherente con los objetivos del estudio y con la naturaleza del fenómeno. Una definición excesivamente estrecha puede producir análisis incompletos; una excesivamente amplia puede dificultar la interpretación.

La Figura 13 reúne los puntajes asignados a cada docente en las distintas dimensiones que conforman la competencia digital, y permite observar de manera conjunta cómo se distribuyen los niveles de alfabetización tecnológica, la variedad y frecuencia de uso de herramientas digitales, así como la adaptación, la confianza y la integración pedagógica

Figura 13.
Puntuaciones registradas en las dimensiones de la competencia digital docente

	A	B	C	D	E	F	G
1	ID	Alfabetizacio...	Variedad_herr...	Frecuencia_uso	Adaptacion_v...	Confianza_di...	Integracion_pedag...
2	1	4	3	4	4	4	5
3	2	3	2	3	3	3	3
4	3	2	1	2	2	2	2
5	4	4	4	4	5	4	4
6	5	5	5	5	5	5	5
7	6	3	3	3	3	3	3
8	7	2	2	2	2	2	2
9	8	4	4	4	4	4	4

Nota. La figura muestra la matriz de datos ingresada en Jamovi correspondiente a las seis dimensiones evaluadas de la competencia digital docente.

Las puntuaciones recogidas en las diferentes dimensiones de la competencia digital docente muestran que no se trata solo de "saber usar" tecnología, sino de integrar ese saber en la práctica cotidiana,

se observan perfiles más consolidados, con valores altos y relativamente equilibrados en alfabetización tecnológica, frecuencia de uso e integración pedagógica, junto a otros más irregulares, donde la confianza o la adaptación a entornos virtuales quedan rezagadas.

El contexto como parte integral de la variable

Entender una variable sin considerar el contexto en el que surge es una de las formas más comunes de distorsionar la realidad que se intenta analizar. Una medida tan simple como el "tiempo dedicado al estudio" puede adquirir significados radicalmente distintos según el entorno social y educativo en el que se observe. En contextos urbanos, suele interpretarse como una señal de esfuerzo o disciplina; en zonas rurales, en cambio, puede revelar carencias materiales que obligan al estudiante a invertir más tiempo para lograr los mismos resultados. También puede ser expresión de dinámicas familiares complejas o de niveles elevados de presión académica. Del mismo modo, el "logro académico" no representa un estándar universal, pues depende de los criterios de evaluación de cada institución, la disponibilidad de recursos, las prácticas docentes y las expectativas socioculturales que influyen en lo que cada comunidad considera un buen desempeño.

Esta variabilidad en el significado de las variables ha sido ampliamente discutida en la literatura contemporánea sobre educación y análisis de datos. Ben-Zvi y Makar (2016) advierten que interpretar variables sin atender al contexto produce conclusiones incompletas y, en muchos casos, desconectadas de la experiencia real de los estudiantes. Considerar el contexto no solo enriquece la interpretación estadística, sino que también permite tomar decisiones pedagógicas más justas y pertinentes, evitando generalizaciones que invisibilizan las múltiples realidades que conviven dentro del sistema educativo.

Efectos de una definición inadecuada de la variable

La manera en que se define una variable determina, en gran medida, la calidad y la profundidad del análisis que se puede realizar con ella. Cuando una variable está mal construida o parte de una definición limitada, los resultados se vuelven incompletos y, en ocasiones, abiertamente contradictorios. Por ejemplo, reducir el "abandono escolar" únicamente a la inasistencia prolongada deja fuera factores estructurales como la inseguridad, el trabajo infantil o la migración, todos ellos ampliamente documentados en estudios internacionales (UNESCO, 2021).

Algo similar ocurre cuando se pretende medir el "bienestar estudiantil" solo con escalas cuantitativas, lo que puede ocultar experiencias subjetivas, relaciones familiares frágiles o tensiones emocionales que requieren una lectura cualitativa más profunda.

Incluso en el ámbito disciplinar, definir la "competencia matemática" únicamente mediante pruebas de cálculo conduce a ignorar dimensiones clave como el razonamiento, la argumentación o la modelización.

Diversos investigadores han insistido en que la definición de una variable no es un trámite conceptual, sino una decisión metodológica con efectos directos sobre la interpretación de los datos. Hernández-Sampieri y Mendoza (2018) destacan que una variable mal delimitada genera mediciones inconsistentes que comprometen la validez de los resultados. Del mismo modo, Messick (1995) advierte que la validez de un constructo depende no solo de los instrumentos que lo miden, sino también de la claridad con que se define su estructura conceptual. En este sentido, reflexionar sobre la construcción de las variables no es un lujo académico, sino una condición esencial para producir análisis estadísticos responsables y decisiones pedagógicas bien fundamentadas.

De manera general, la elección de las variables en educación no es un detalle técnico, sino una decisión que orienta qué se considera valioso y, por tanto, qué tipo de prácticas y políticas se impulsan. Cuando una institución decide medir solo rendimiento numérico, termina reforzando lógicas centradas en pruebas y resultados estandarizados; si incorpora variables de desarrollo socioemocional, abre espacio a tutorías, acompañamiento y cuidado; y si valora la participación, fomenta actividades colaborativas y voces estudiantiles más visibles.

Tabulación y representaciones gráficas: leer y comunicar información

La tabulación y las representaciones gráficas constituyen un componente fundamental del razonamiento estadístico. No se trata únicamente de ordenar datos o de embellecer información, sino de transformar registros dispersos en narrativas comprensibles. Tanto en educación como en los estudios sociales, los gráficos y las tablas no solo muestran cifras: construyen significados, revelan patrones y permiten que las personas entiendan fenómenos que, sin estas herramientas, serían invisibles o incomprensibles. Como afirman Garfield y Ben-Zvi (2008), la representación gráfica no es un añadido decorativo de la estadística, sino la vía por la cual los estudiantes aprenden a pensar con datos.

Al analizar cómo los datos se organizan y representan, aparecen dos dimensiones esenciales: la posición teórica y la posición didáctica, que examina cómo estas herramientas pueden favorecer aprendizajes significativos en las aulas

La articulación entre ambas perspectivas permite comprender que tabular y graficar no son actividades mecánicas, sino procesos que exigen interpretación, reflexión y comunicación.

Tabular para organizar y revelar patrones

La tabulación constituye la primera estructura formal que permite convertir un conjunto disperso de registros en una organización significativa. Wild y Pfannkuch (1999) explican que el pensamiento estadístico inicia con la capacidad de "imponer estructura" sobre datos desorganizados; en otras palabras, construir una forma que permita ver lo que antes era invisible. Las tablas revelan frecuencias, distribuciones, agrupamientos y comportamientos inusuales, y actúan como un puente entre los datos brutos y las interpretaciones posteriores.

Por ejemplo, un docente quiso comprender por qué algunos estudiantes estaban teniendo dificultades en el último bloque de Matemática, para ello, les pidió que anotaran de manera anónima cuántos ejercicios de la guía habían logrado resolver sin ayuda. Al finalizar la clase, tenía una lista desordenada de números:

3, 5, 4, 2, 10, 4, 3, 4, 2, 3, 12, 4, 5. A primera vista, esa serie de valores no decía mucho. Podría parecer simplemente una mezcla irregular de cantidades sin patrón evidente. Sin embargo, cuando decidió organizar los datos en una tabla de frecuencias (Figura 14), la situación cambió por completo.

Figura 14.
Distribución de ejercicios resueltos sin ayuda por los estudiantes

Frequencias de Ejercicios

Ejercicios	Recuentos	% del Total	% Acumulado
2	2	15.4 %	15.4 %
3	3	23.1 %	38.5 %
4	4	30.8 %	69.2 %
5	2	15.4 %	84.6 %
10	1	7.7 %	92.3 %
12	1	7.7 %	100.0 %

Nota. La figura muestra la frecuencia con la que los estudiantes resolvieron ejercicios de manera autónoma.

En este sentido llama la atención es que la mayoría de los estudiantes se mueve entre 2 y 4 ejercicios resueltos; de hecho, casi siete de cada diez están en ese rango. Esto da la sensación de que la guía les resultó manejable, pero todavía exigente:hacen algunos ejercicios por su cuenta, pero no tantos como para

pensar que todos dominan el tema con soltura. También se ve un pequeño grupo que llega a 5 ejercicios, lo cual refuerza la idea de un rendimiento bastante parejo, sin demasiadas diferencias dentro del conjunto principal.Por otra parte, destaca los dos valores más altos, 10 y 12 ejercicios. Son casos aislados, pero no por eso menos importantes. Más bien, funcionan como señales claras de que hay estudiantes que están muy por encima del resto, ya sea porque entienden el contenido con mayor facilidad o porque han tenido más práctica previa. Estos contrastes no se notaban en la lista de números desordenados, pero la tabla los pone en evidencia de inmediato

En contextos educativos, tabular implica tomar decisiones conceptuales: seleccionar variables, definir categorías, establecer niveles de detalle y ordenar la información según criterios analíticos. Un docente que analiza la evolución del rendimiento en Matemática puede tabular por unidades de aprendizaje, por niveles de logro o por frecuencia de actividades. Cada tabulación construye una lectura distinta del mismo fenómeno. Lo mismo ocurre en estudios sociales: tabular el acceso a internet por territorio, edad o nivel socioeconómico genera interpretaciones específicas, y todas están mediadas por las elecciones iniciales del investigador.

Desde una perspectiva teórica, este proceso muestra que la tabla no es neutral: es una construcción que destaca unos aspectos y atenúa otros. Batanero y Díaz (2011) señalan que toda tabulación conlleva un posicionamiento conceptual, pues las decisiones sobre la organización del dato determinan las posibilidades de interpretación. En consecuencia, tabular implica comprender el fenómeno y, simultáneamente, delimitar su lectura.

La investigación en educación estadística coincide en que muchos estudiantes reproducen tablas sin comprender las implicaciones de cada decisión. Garfield y Ben-Zvi (2008) indican que la enseñanza debería enfatizar que tabular no es llenar casillas, sino estructurar información para pensar con ella. Por eso, trabajar con datos reales permite que los estudiantes descubran patrones y comprendan el sentido de la organización.

Representaciones gráficas: visualizar, interpretar y construir narrativas

Las representaciones gráficas constituyen una ampliación visual del razonamiento estadístico. Curcio (1989) estableció que comprender un gráfico implica leer datos, leer entre los datos y leer más allá de los datos, es decir, interpretar relaciones, detectar tendencias y formular conclusiones.

En este sentido, los gráficos son formas de narración visual que permiten transformar lo cuantitativo en relaciones descriptivas más accesibles.

- Los distintos tipos de gráficos ofrecen perspectivas diferenciadas:
- Gráficos de barras: permiten comparar categorías y visualizar contrastes inmediatos.
- Gráficos de líneas: describen trayectorias temporales y facilitan analizar cambios y ritmos.
- Histogramas: muestran la forma de la distribución y revelan simetrías, sesgos o concentraciones.
- Diagramas de caja: permiten observar variabilidad, dispersión y valores atípicos.
- Gráficos de dispersión: exponen relaciones entre variables, especialmente al analizar asociaciones.
- Gráficos de sectores: muestran proporciones, aunque requieren interpretaciones cautelosas.

Desde una posición teórica, Garfield y Ben-Zvi (2008) destacan que las representaciones visuales realizan una transformación conceptual del fenómeno. Un histograma no solo reproduce frecuencias; sintetiza la estructura completa de la variable. Un diagrama de caja no solo ordena datos; visibiliza desigualdades internas que no se perciben en promedios o valores individuales.Por ejemplo, para conocer mejor la realidad de cada grupo (Figura 15)., un docente comparó las notas de los cursos A y B utilizando un diagrama de caja.

Aunque ambos cursos tenían promedios similares el gráfico mostró diferencias importantes (Figura 16) en el curso A las calificaciones estaban más agrupadas, mientras que en el curso B había una dispersión mucho mayor, con estudiantes muy avanzados y otros con claras dificultades. Esta visualización le permitió reconocer que, aunque los dos cursos parecían similares en los números generales, sus necesidades internas eran distintas y necesitaban estrategias de apoyo diferenciadas.

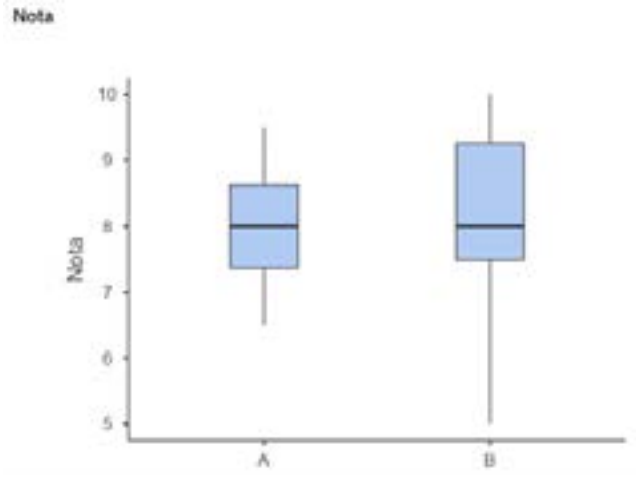
La literatura crítica observa que los gráficos pueden inducir interpretaciones erróneas cuando las escalas, colores o categorías no se eligen con criterio. Gal (2002) sostiene que la lectura crítica de gráficos se ha vuelto un componente indispensable de la alfabetización estadística contemporánea, especialmente en un entorno saturado de información visual. Comprender un gráfico implica analizar no solo lo que se muestra, sino también las decisiones detrás de su diseño.

Figura 15
Registros de ejercicios resueltos sin ayuda por los estudiantes en los cursos A y B

	ID	Curso	Nota
1	1	A	6.5
2	2	A	7.0
3	3	A	7.5
4	4	A	8.0
5	5	A	8.5
6	6	A	9.0
7	7	A	9.5
8	8	A	8.0
9	9	B	5.0
10	10	B	6.0
11	11	B	8.0
12	12	B	8.0
13	13	B	10.0
14	14	B	10.0
15	15	B	8.0
16	16	B	9.0

Nota. La figura muestra la matriz de datos ingresada en Jamovi con las calificaciones obtenidas por los estudiantes de ambos cursos.

Figura 16
Diagrama de caja de las notas obtenidas por los estudiantes de los cursos A y B



Nota. El diagrama muestra la distribución de las calificaciones registradas en ambos cursos.

La literatura crítica observa que los gráficos pueden inducir interpretaciones erróneas cuando las escalas, colores o categorías no se eligen con criterio. Gal (2002) sostiene que la lectura crítica de gráficos se ha vuelto un componente indispensable de la alfabetización estadística contemporánea, especialmente

en un entorno saturado de información visual. Comprender un gráfico implica analizar no solo lo que se muestra, sino también las decisiones detrás de su diseño.

En educación, enseñar a representar datos implica no solo aprender a construir gráficos, sino comprender qué tipo de gráfico responde mejor a cada pregunta de investigación. Wild y Pfannkuch (1999) explican que representar visualmente permite que los estudiantes avancen desde la observación hacia la explicación. Al producir gráficos propios, deben justificar decisiones, argumentar sus interpretaciones y construir una comunicación basada en datos.

La mirada investigativa: evidencias, dificultades y aportes al razonamiento estadístico

La investigación en educación estadística ha dedicado especial atención a comprender cómo los estudiantes leen, interpretan y producen tablas y representaciones gráficas, y cuáles son los procesos cognitivos implicados en estas tareas. Desde sus primeros desarrollos, este campo ha mostrado que la capacidad de trabajar con representaciones no se adquiere de manera espontánea; por el contrario, demanda instrucción explícita, experiencias progresivas y oportunidades para interactuar críticamente con datos reales. Garfield y Ben-Zvi (2008) enfatizan que el razonamiento estadístico se construye mediante ciclos de interpretación, discusión y comunicación, donde las tablas y los gráficos cumplen un papel central. La investigación confirma que estas representaciones son algo más que herramientas visuales: son mediadores cognitivos que orientan la comprensión, estructuran el pensamiento y permiten construir argumentos basados en datos.

Investigaciones sobre cómo se comprende una representación: niveles, procesos y vacíos persistentes

Uno de los aportes más influyentes en este campo proviene del trabajo de Curcio (1989), quien estableció tres niveles de comprensión gráfica: leer los datos, leer entre los datos y leer más allá de los datos. Este modelo ha sido validado por múltiples estudios posteriores y sigue siendo un referente para analizar la evolución del pensamiento estadístico. "Leer los datos" implica identificar valores explícitos; "leer entre los datos" exige comparar, inferir y detectar tendencias; y "leer más allá" supone interpretar patrones, anticipar comportamientos o justificar conclusiones. La investigación muestra que muchos estudiantes permanecen anclados en el primer nivel, incluso en grados superiores, lo que evidencia una brecha entre la exposición a los gráficos y el desarrollo efectivo del razonamiento.

Wild y Pfannkuch (1999) ampliaron esta mirada al mostrar que los estudiantes suelen centrarse en detalles superficiales, como puntos aislados o segmentos llamativos, mientras ignoran la estructura global de la distribución. Esta atención fragmentada dificulta que adviertan variabilidad, tendencias o relaciones estadísticas más complejas.

Estudios recientes confirman que la comprensión de un gráfico depende de procesos cognitivos simultáneos: decodificación visual, reconocimiento de patrones, articulación entre la representación y la variable conceptual, y contextualización del fenómeno. No basta con identificar puntos o barras; es imprescindible comprender qué significan en relación con la pregunta que motivó el análisis.

Dificultades documentadas: errores recurrentes y obstáculos conceptuales

La literatura investigativa identifica una serie de dificultades recurrentes que enfrentan los estudiantes al trabajar con tablas y gráficos. Estas dificultades no se limitan a errores técnicos, sino que revelan obstáculos más profundos del razonamiento estadístico.

Entre los desafíos más frecuentes se encuentran:

1. Interpretación literal y no relacional del gráfico

Los estudiantes tienden a describir lo que "ven" sin relacionarlo con el fenómeno. Garfield y Ben-Zvi (2008) documentan casos donde los alumnos interpretan una línea ascendente como "sube porque sí", sin conectar el cambio con la variable estudiada.

2. Confusión entre forma visual y magnitud

Especialmente con gráficos de barras o áreas, algunos estudiantes creen que una barra más ancha representa "más" aunque no sea más alta. Esto muestra una lectura centrada en rasgos perceptuales y no en las variables (Curcio, 1989).

3. Dificultad para interpretar variabilidad

La variabilidad suele ser uno de los conceptos más difíciles. En diagramas de caja, por ejemplo, muchos estudiantes no identifican valores atípicos o interpretan la mediana como un "promedio más exacto" (Garfield & Ben-Zvi, 2008).

4. Problemas para leer escalas no uniformes

Cuando los ejes cambian la escala, el gráfico puede parecer engañoso. La investigación muestra que los estudiantes rara vez detectan estas decisiones de diseño, lo que afecta la interpretación crítica (Gal, 2002).

5. Sobrerreliance en valores individuales

Los estudiantes suelen fijarse en el valor máximo o mínimo sin considerar patrones globales. La mirada queda fragmentada y no se construyen conclusiones integradas.

6. Falta de conexión entre tabla y gráfico

Muchos estudiantes no vinculan la tabla con la representación visual correspondiente. La investigación señala que estas dos herramientas se trabajan de forma aislada, cuando en realidad están profundamente conectadas (Batanero & Díaz, 2011).

Estas dificultades muestran que la comprensión de gráficos no depende únicamente de verlos o reproducirlos, sino de desarrollar una mirada analítica que permita reconocer estructuras, relaciones y significados.

Los estudios en didáctica de la estadística han identificado prácticas que favorecen el desarrollo de la comprensión gráfica:

1. Trabajar con datos reales y significativos:

Ben-Zvi y Makar (2016) muestran que los estudiantes desarrollan mayor sensibilidad hacia los patrones y tendencias cuando los datos provienen de fenómenos familiares: asistencia, hábitos digitales, rendimiento, percepciones. La cercanía favorece la formulación de preguntas y la interpretación crítica.

2. Combinar varias representaciones del mismo fenómeno

La investigación indica que presentar un mismo conjunto de datos en tabla, histograma, gráfico de líneas y diagrama de caja permite a los estudiantes comprender diferentes facetas de la distribución. Esta triangulación desarrolla una comprensión más robusta (Garfield & Ben-Zvi, 2008).

3. Promover la explicación oral y escrita

Explicar un gráfico obliga a estructurar el pensamiento. Los estudios señalan que cuando los estudiantes justifican por qué eligieron un tipo de gráfico o cómo interpretan la tendencia, la comprensión se profundiza.

4. Enseñar explícitamente la lectura crítica

Gal (2002) sostiene que la educación moderna debe abordar la manipulación visual de datos. Enseñar a detectar escalas truncadas, gráficos engañosos o narrativas sesgadas fortalece la alfabetización estadística.

5. Guiar la construcción de gráficos, no solo su lectura

Construir gráficos requiere tomar decisiones: elegir la escala, seleccionar el tipo de gráfico, organizar los datos. La investigación confirma que este proceso es más formativo que leer gráficos ya hechos (Wild & Pfannkuch, 1999).

Trabajar con tablas y representaciones gráficas no solo ayuda a "embellecer" los datos, sino que se convierte en una vía privilegiada para desarrollar componentes centrales del razonamiento estadístico. La tabulación permite estructurar el fenómeno, ordenar información dispersa y hacer visibles patrones que de otro modo pasarían desapercibidos, mientras que los gráficos facilitan la comprensión de la variabilidad, la forma de la distribución

y la presencia de valores extremos. Esta organización visual de la información vuelve más accesible la comparación entre grupos, distribuciones y tendencias, y crea un soporte concreto para la construcción de argumentos basados en datos. En esa línea, autores como Garfield y Ben-Zvi (2008) destacan que el trabajo sistemático con representaciones tabulares y gráficas favorece que el estudiante pase de "ver números" a interpretar evidencias, lo que fortalece la capacidad de justificar conclusiones de manera fundamentada.

Al mismo tiempo, la lectura crítica de gráficos reales abre la puerta a una ciudadanía más informada y reflexiva. Gal (2002) subraya que la alfabetización estadística implica no solo comprender procedimientos, sino también interpretar mensajes que circulan en medios, informes y debates públicos. Desde esta perspectiva, las representaciones actúan como un puente entre datos, contexto y narrativa: permiten conectar valores numéricos con significados sociales, educativos y culturales, y cuestionar qué se muestra, qué se oculta y qué se da por supuesto. En síntesis, la investigación en educación estadística coincide en que el uso intencional de tablas y gráficos es uno de los caminos más sólidos para desarrollar un pensamiento estadístico auténtico, basado en comprensión, análisis crítico y comunicación clara de la información (Garfield & Ben-Zvi, 2008; Gal, 2002).

Medidas de tendencia central y dispersión: interpretar la información numérica

La interpretación de datos numéricos requiere comprender cómo se organizan, cómo se concentran y qué tan alejados se encuentran los valores respecto a un punto de referencia. En este proceso, las medidas de tendencia central y dispersión constituyen herramientas esenciales para sintetizar la información y revelar patrones que no son visibles a simple vista. Garfield y Ben-Zvi (2008) destacan que estas medidas permiten entender la estructura de un fenómeno, identificar comportamientos típicos y evaluar la variabilidad que caracteriza a los datos. En educación y en las ciencias sociales, esta comprensión resulta especialmente relevante, pues muchos fenómenos presentan distribuciones heterogéneas que requieren análisis cuidadoso para evitar interpretaciones simplistas.

Comprender estas medidas no es únicamente un ejercicio técnico: implica desarrollar la capacidad de argumentar con datos, contextualizar valores numéricos y reconocer que los resúmenes estadísticos deben interpretarse a la luz del fenómeno estudiado. En esta sección se amplía la reflexión sobre estas medidas, integrando ejemplos concretos y aportes de la investigación reciente.

Comprender el "centro": media, mediana y moda como representaciones del valor típico

Las medidas de tendencia central permiten sintetizar un conjunto de datos en un solo valor, pero cada una expresa una forma distinta de interpretar qué significa "lo característico" de un grupo. La media suele entenderse como un punto de equilibrio: suma todos los valores y los distribuye de manera uniforme. Sin embargo, cuando existen diferencias marcadas dentro del conjunto, este promedio puede ofrecer una visión distorsionada. La mediana, en cambio, señala el valor que divide a la población en dos partes iguales y se mantiene estable incluso frente a valores extremos. La moda aporta otra mirada complementaria, pues identifica el valor que aparece con mayor frecuencia y resulta especialmente útil cuando se analizan categorías o preferencias. Estas distinciones no son simplemente técnicas; responden a preguntas sobre cómo se comportan realmente los datos y qué tan homogéneo o diverso es un fenómeno.

En el campo educativo y social, elegir la medida adecuada es fundamental para comprender la realidad sin perder matices. Prodromou (2019) señala que los estudiantes desarrollan un pensamiento estadístico más profundo cuando trabajan con datos auténticos, porque pueden observar directamente cómo la media se desplaza cuando aparecen valores atípicos, cómo la mediana captura la estabilidad del grupo y cómo la moda revela patrones de comportamiento que de otro modo pasarían desapercibidos.

Del mismo modo, Konold et al. (2017) señalan destacan que enseñar estas diferencias ayuda a superar la idea de que la estadística se reduce a algoritmos, promoviendo en cambio una lectura crítica y contextualizada de la información. En conjunto, estos aportes muestran que comprender las medidas centrales implica mucho más que calcular números: implica aprender a interpretar el sentido de los datos y a reconocer qué dicen sobre el fenómeno que se estudia.

Por ejemplo, en un centro de salud comunitario se implementó recientemente un taller educativo para pacientes con hipertensión, con el propósito de mejorar sus hábitos diarios y ayudarles a controlar la presión arterial. Después de un mes, el equipo médico quiso comprobar si existían diferencias reales entre quienes participaron en el taller y quienes continuaron con la atención habitual.

Para ello, registraron la presión arterial sistólica de un grupo de pacientes en ambos escenarios y organizaron los datos (Figura 17) con el fin de comparar la distribución de valores, identificar posibles patrones y evaluar si el taller estaba teniendo un impacto clínicamente significativo.

El análisis de estos resultados permitirá orientar decisiones sobre la continuidad del programa y sobre la necesidad de ajustes en las intervenciones educativas que se ofrecen a la comunidad.

Figura 17.
Registros de presión arterial sistólica (PAS) de los pacientes

	ID	Grupo	PAS
1	1	Taller	128
2	2	Taller	130
3	3	Taller	125
4	4	Taller	132
5	5	Taller	127
6	6	Taller	129
7	7	Taller	135
8	8	Taller	126
9	9	Control	142
10	10	Control	150
11	11	Control	138
12	12	Control	146
13	13	Control	141
14	14	Control	149
15	15	Control	144
16	16	Control	152

Nota. La figura muestra la matriz de datos utilizada en el análisis.

Figura 18.
Estadísticos descriptivos de la presión arterial sistólica en los grupos Taller y Control

Descriptivas

Descriptivas

	Grupo	PAS
N	Taller	8
	Control	8
Media	Taller	129
	Control	145
Mediana	Taller	129
	Control	145
Moda	Taller	125*
	Control	138*

* Existe más de una moda, sólo se reporta la primera

Nota. La tabla presenta los valores de N, media, mediana y moda de la presión arterial sistólica (PAS) en ambos grupos.

Los resultados descriptivos (Figura 17) muestran un patrón claro: aunque ambos grupos tienen el mismo número de participantes, las medidas de tendencia central son consistentemente más bajas en el grupo que asistió al taller. Tanto la media como la mediana se sitúan en 129 mmHg en este grupo, mientras que en el grupo control alcanzan alrededor de 145 mmHg. Incluso los valores más frecuentes (moda) siguen esta misma dirección.

Prodromou (2019) mostró que trabajar con datos reales y dinámicos permite que los estudiantes entiendan cómo valores extremos afectan la media, cómo la mediana refleja estabilidad y cómo la moda evidencia patrones de preferencia o comportamiento. Estos elementos convierten la enseñanza de las medidas centrales en una oportunidad para promover pensamiento crítico sobre los datos.

La dispersión como clave interpretativa: variación, estabilidad y desigualdad en los datos

En el análisis estadístico, comprender únicamente los valores centrales de un conjunto de datos suele resultar insuficiente para interpretar el comportamiento real de un fenómeno. La dispersión se convierte, por ello, en una clave interpretativa fundamental, ya que permite evaluar el grado en que los valores se alejan del centro y, con ello, reconocer patrones de variación, niveles de estabilidad y formas de desigualdad presentes en los datos. Por ejemplo, si conectamos la idea de variación con los datos de presión arterial sistólica (PAS) de los dos grupos (Taller y Control), podemos imaginar la siguiente situación: en el grupo Taller, la mayoría de los valores se concentran cerca de los 129 mmHg (que coinciden en la media y la mediana), con registros que oscilan en un rango relativamente estrecho alrededor de ese valor.

Esto indicaría una variación baja y, por tanto, un comportamiento más estable del indicador en las personas que participaron en la intervención educativa. En cambio, en el grupo Control, donde la media y la mediana se sitúan en 145 mmHg, los valores podrían estar mucho más dispersos, con algunos participantes ligeramente por encima de 140 mmHg y otros con cifras cercanas o superiores a 160 mmHg. Aunque ambos grupos tienen el mismo tamaño muestral (n = 8), la mayor variación en el grupo Control sugeriría heterogeneidad estructural y posibles influencias externas no controladas, como diferencias en el acceso a controles médicos o en la adherencia a tratamientos.

La dispersión también ofrece una ventana interpretativa sobre la estabilidad. Un conjunto con baja desviación estándar o bajo rango intercuartílico suele interpretarse como estable, en el sentido de que sus valores no difieren abruptamente entre sí.

Esto es especialmente relevante cuando se evalúan procesos educativos, donde la estabilidad puede sugerir que los aprendizajes son relativamente homogéneos, o en estudios clínicos, donde una baja dispersión podría indicar respuestas similares al tratamiento. En contraste, valores con amplia dispersión alertan sobre contextos inestables, brechas internas o procesos que requieren mayor atención o diferenciación metodológica.

Figura 19.
Medidas descriptivas de la presión arterial sistólica (PAS) en los grupos Taller y Control.

Descriptivas

Descriptivas

	Grupo	PAS
N	Taller	8
	Control	8
Perdidos	Taller	0
	Control	0
Media	Taller	129
	Control	145
Mediana	Taller	129
	Control	145
Moda	Taller	125*
	Control	138*
Desviación típica	Taller	3.30
	Control	4.06
RIC	Taller	3.75
	Control	7.50
Mínimo	Taller	125
	Control	138
Máximo	Taller	135
	Control	152

* Existe más de una moda, solo se reporta la primera

Nota. La figura muestra el número de participantes y las principales medidas de tendencia central y dispersión de la presión arterial sistólica (PAS) por grupo.

Finalmente, la dispersión constituye un indicador clave para analizar la desigualdad. Mientras medidas como la media o la mediana describen tendencias generales, la dispersión revela la distancia entre grupos o individuos. En investigaciones sobre rendimiento académico, ingreso económico o acceso a servicios, una alta dispersión puede reflejar inequidades estructurales que se ocultan tras un valor central aparentemente adecuado.

De manera general, la dispersión no es un elemento técnico aislado, sino un recurso interpretativo que aporta profundidad analítica. Permite reconocer cuánto cambian los datos, cuán estable es un proceso y qué desigualdades emergen dentro de un conjunto.

Integrar esta mirada favorece análisis más rigurosos y decisiones mejor fundamentadas, especialmente en contextos educativos, sociales o de investigación aplicada.

Comprender los valores atípicos: impacto interpretativo, decisiones analíticas y sentido del dato

El análisis estadístico contemporáneo exige observar no solo lo que ocurre "en el centro", sino también cómo ciertos valores inusuales (valores atípicos u outliers) influyen en la comprensión de un fenómeno. Lejos de ser simples anomalías, estos valores pueden constituir señales importantes sobre desigualdades, comportamientos excepcionales o transformaciones emergentes del contexto. Bakker y Wagner (2019) subrayan que su interpretación adecuada depende del propósito analítico: un valor extremo puede enriquecer el análisis o distorsionarlo, según cómo se contextualice.

Las medidas de tendencia central y dispersión reaccionan de forma distinta frente a los valores atípicos y la media por su parte es muy sensible a los valores extremos, Un solo valor inusualmente alto puede arrastrar el promedio y alterar por completo la interpretación. La mediana, en cambio, permanece estable y puede convertirse en una alternativa más adecuada cuando el conjunto presenta desigualdades estructurales, como ocurre con ingresos o acceso a dispositivos tecnológicos.

Las medidas de dispersión también se ven afectadas: la desviación estándar aumenta de manera notable ante un valor extremo, y el rango puede volverse engañoso si un solo dato inflado define la amplitud del conjunto. En este sentido, el recorrido intercuartílico (RIC) se vuelve especialmente útil porque se concentra en el 50% de datos centrales, ignorando los extremos y ofreciendo una visión más robusta del comportamiento general.

Consideremos el siguiente ejemplo sin valores extremos (Figura 18), las horas de estudio se distribuyen de manera homogénea, con una media y mediana de 5.50, una desviación típica baja (1.05) y un rango acotado entre 4 y 7 horas.

Sin embargo, al incorporar el valor atípico de 20 horas (Figura 19), la media aumenta abruptamente a 7.57 y la desviación típica sube a 5.56, indicando una dispersión mucho mayor.

En contraste, la mediana cambia mínimamente (de 5.50 a 6) y el RIC solo aumenta ligeramente, lo que confirma su estabilidad frente a valores extremos. Este comportamiento evidencia que la media, el rango y la desviación estándar son altamente sensibles a los outliers, mientras que la mediana y el RIC ofrecen una representación más robusta del comportamiento central del grupo.

Figura 20.
Medidas descriptivas de las horas de estudio sin valores atípicos.

Descriptivas

Descriptivas

	Horas_estudio
N	6
Perdidos	2
Media	5.50
Mediana	5.50
Moda	5.00ª
Desviación típica	1.05
RIC	1.00
Mínimo	4
Máximo	7

	A	Horas_estudio
1	1	4
2	2	5
3	3	5
4	4	6
5	5	6
6	6	7

Nota. La figura muestra la base de datos y las medidas descriptivas correspondientes a las horas de estudio registradas por seis estudiantes

Figura 21.
Medidas descriptivas de las horas de estudio sin valores atípicos.

Descriptivas

Descriptivas

	Horas_estudio
N	7
Perdidos	1
Media	7.57
Mediana	6
Moda	5.00ª
Desviación típica	5.56
RIC	1.50
Mínimo	4
Máximo	20

	A	Horas_estudio
1	1	4
2	2	5
3	3	5
4	4	6
5	5	6
6	6	7
7	7	20

Nota. La figura muestra la base de datos y las medidas descriptivas correspondientes a las horas de estudio registradas por siete estudiantes

En educación y ciencias sociales, los valores atípicos suelen ser "puntos de tensión interpretativa", según Konold et al. (2017) señalan, pues pueden evidenciar situaciones de riesgo, inequidades o dinámicas que requieren atención. Un estudiante que obtiene una calificación considerablemente inferior al resto no es un error estadístico: puede señalar dificultades de aprendizaje, problemas emocionales o falta de recursos. En estudios sociales, un ingreso extremadamente bajo o alto puede reflejar desigualdades profundas o transformaciones socioeconómicas en curso.

Las investigaciones recientes destacan además el papel de la tecnología en la comprensión de los valores atípicos. Prodromou (2019) demuestra que la manipulación dinámica de datos permite visualizar instantáneamente cómo se altera la media o la desviación estándar cuando aparece un outlier. Bakker y Akkerman (2019), por su parte, recuerdan que en entornos de big data los valores extremos pueden indicar errores, pero también fenómenos emergentes que requieren análisis crítico.

Conclusiones

Al recorrer este capítulo se hizo evidente que la estadística descriptiva no es solo un conjunto de técnicas, sino una forma de aproximarse a la realidad con una mirada más atenta y consciente. Las medidas de centro, la dispersión y la identificación de valores atípicos permiten interpretar fenómenos educativos y sociales con mayor precisión, evitando quedarse en afirmaciones simplificadas que pierden de vista la complejidad de los datos. Comprender cómo se comportan los valores dentro de un conjunto ayuda a reconocer patrones, detectar desigualdades y tomar decisiones mejor fundamentadas.

También se vuelve claro que las representaciones estadísticas son herramientas que acompañan el pensamiento crítico. Un promedio, por sí solo, nunca cuenta toda la historia; solo cuando se contrasta con la variabilidad, la forma de la distribución y la presencia de casos inusuales es posible obtener una lectura completa. Esta perspectiva invita a analizar con cuidado cualquier afirmación sustentada en datos, especialmente en ámbitos como la educación o la salud, donde cada número representa situaciones humanas que requieren sensibilidad e interpretación contextualizada.

Finalmente, el capítulo subraya la importancia de formar a estudiantes y profesionales capaces de dialogar con los datos, no solo de reproducir cálculos. En un entorno saturado de información, la capacidad de interpretar tablas, gráficos y estadísticas se transforma en una competencia esencial para participar de manera crítica en la vida académica y social. Aprender a mirar más allá del valor central, a preguntarse por lo que los números muestran y por lo que ocultan, es un paso fundamental para construir una comprensión más honesta y profunda de los fenómenos que estudiamos.

Referencias

Bakker, A., & Akkerman, S. (2019). The interdisciplinary character of data science. Harvard Data Science Review, 1(1).

Batanero, C., & Borovcnik, M. (2016). Statistics and probability in high school. Springer. https://doi.org/10.1007/978-3-319-23986-6

Batanero, C., & Díaz, C. (2011). Estadística y probabilidad en educación secundaria. Editorial Universidad de Granada.

Ben-Zvi, D., & Aridor-Berger, K. (2016). Students making sense of data in project-based learning. Journal of Mathematical Behavior, 43, 1–14. https://doi.org/10.1016/j.jmathb.2016.04.001

Ben-Zvi, D., & Makar, K. (2016). The role of context in developing young students' statistical thinking. Springer.

Curcio, F. R. (1989). Developing graph comprehension. National Council of Teachers of Mathematics.

Gal, I. (2002). Adult's statistical literacy: Meanings, components, responsibilities. International Statistical Review, 70(1), 1–25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. Springer.

Hernández-Sampieri, R., & Mendoza, C. (2018). Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta. McGraw-Hill.

Konold, C., Finzer, W., & Kreetong, K. (2017). Modeling as a core component of structuring data. Statistics Education Research Journal, 16(2), 191–212. https://doi.org/10.52041/serj.v16i2.190

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Pfannkuch, M., & Ben-Zvi, D. (1999). Innovations in statistical education: Learning analytics and data modeling. Springer.

Prodromou, T. (2019). Data modelling and statistical reasoning with dynamic data. Springer. https://doi.org/10.1007/978-3-030-03633-1

UNESCO. (2021). Informe de seguimiento de la educación en el mundo 2021/2022: No dejar a nadie atrás. UNESCO Publishing.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Capítulo II

Pensar el azar: fundamentos didácticos y conceptuales de la probabilidad

Introducción

El estudio de la probabilidad invita a mirar el mundo desde una perspectiva distinta, una en la que la incertidumbre no es un obstáculo, sino una puerta para comprender cómo se comportan los fenómenos que no siguen un patrón fijo. En la vida cotidiana convivimos con el azar sin pensarlo demasiado: cuando anticipamos el clima, participamos en un juego, tomamos una decisión con información incompleta o interpretamos una noticia basada en datos. Sin embargo, esas experiencias suelen construir intuiciones que, aunque útiles, no siempre coinciden con la manera en que la probabilidad describe matemáticamente lo posible. Este capítulo propone adentrarse en ese territorio fronterizo entre intuición y formalización, explorando cómo las personas interpretan la incertidumbre, cómo se estructura un experimento aleatorio y cómo pueden enseñarse las ideas fundamentales del azar de forma significativa.

Desde una perspectiva didáctica, pensar el azar implica reconocer que las ideas probabilísticas no se desarrollan de manera espontánea. Requieren experiencias que permitan observar la variabilidad, discutir

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

expectativas, contrastar lo que creemos que debería ocurrir con lo que realmente sucede y elaborar modelos que ayuden a organizar esa complejidad. A través de ejemplos, simulaciones, representaciones y análisis de situaciones reales, este capítulo busca mostrar cómo la probabilidad puede enseñarse no como una colección de reglas aisladas, sino como un modo de razonamiento que nos permite interpretar, predecir y tomar decisiones en contextos donde no existe certeza absoluta.

En conjunto, el contenido de este capítulo invita a construir una comprensión más profunda y humana del azar. No se trata únicamente de aprender a calcular probabilidades, sino de desarrollar una sensibilidad para reconocer patrones en la incertidumbre, cuestionar intuiciones iniciales y valorar el papel que juega el razonamiento probabilístico en la vida diaria. Pensar el azar, en este sentido, es también aprender a pensar con apertura, cautela y sentido crítico frente a un mundo que rara vez se comporta de manera totalmente predecible.

De la intuición del azar al concepto formal de probabilidad

Comprender la probabilidad comienza mucho antes de encontrarse con definiciones formales: nace en la manera en que las personas interpretan lo inesperado, anticipan resultados o explican por qué ciertos eventos suceden y otros no. Desde edades tempranas, todos desarrollamos ideas intuitivas sobre el azar basadas en experiencias cotidianas, conversaciones, juegos y observaciones informales; sin embargo, estas intuiciones no siempre coinciden con la lógica matemática que sustenta el concepto de probabilidad. Este epígrafe propone un tránsito reflexivo desde esas primeras percepciones, a veces imprecisas o cargadas de sesgos, hacia una comprensión más rigurosa y estructurada del azar. Se busca mostrar que la probabilidad no surge de memorizar reglas, sino de reconstruir el pensamiento inicial, confrontarlo con fenómenos reales y reconocer que detrás de la incertidumbre existen patrones y modelos capaces de describir, con sorprendente precisión, el comportamiento de lo aleatorio.

La intuición del azar: creencias, expectativas y razonamientos espontáneos

La manera en que niños, jóvenes e incluso adultos conciben el azar está profundamente anclada en la experiencia cotidiana. Antes de llegar al aula, las personas ya han desarrollado una serie de creencias implícitas sobre cómo "deberían" comportarse los fenómenos inciertos. Estas ideas conforman lo que la literatura denomina intuiciones del azar, una categoría que recoge percepciones previas, razonamientos espontáneos y explicaciones culturalmente transmitidas. Aunque estas intuiciones no forman parte del lenguaje matemático formal, sí constituyen el punto de partida desde el cual se construye el pensamiento probabilístico escolar.

Los estudios recientes han confirmado que estas intuiciones no desaparecen con el tiempo; por el contrario, evolucionan, se complejizan y continúan influyendo en la interpretación de los fenómenos aleatorios incluso en la adultez. Retamal y Alsina (2022), en un estudio con estudiantes chilenos de 8 a 14 años, muestran que la mayoría espera que los resultados aleatorios se "compensen" rápidamente, especialmente en secuencias de lanzamientos de monedas o dados.

A esta expectativa de equilibrio inmediato se suman otras creencias igual de arraigadas, que suelen aparecer con fuerza cuando las personas intentan explicar situaciones marcadas por la incertidumbre. En el ámbito educativo latinoamericano, Alsina y Salcedo (2020) han mostrado que muchos estudiantes interpretan el azar desde supuestos cargados de causalidad o de intención, como si los resultados no fueran producto de un proceso aleatorio sino de algún tipo de fuerza oculta. Así, cuando un dado arroja varios valores altos seguidos, no es raro que algunos lo atribuyan a que el dado está "cargado" o a que un jugador trae mala suerte, aun cuando el objeto sea perfectamente justo.

Estas explicaciones, por más equivocadas que sean desde la perspectiva matemática, no deben verse únicamente como fallos conceptuales. Más bien reflejan un esfuerzo por dotar de sentido a la variabilidad natural del azar, un intento por encontrar estabilidad donde no siempre la hay. Esta necesidad de organizar lo incierto es una reacción profundamente humana y se hace visible tanto en situaciones cotidianas como en el aprendizaje formal de la probabilidad. Por eso, más que corregir de manera directa estas ideas, la tarea educativa consiste en acompañar a los estudiantes para que puedan reconocer por qué surgen, qué función cumplen y cómo es posible reemplazarlas por una comprensión más rigurosa del carácter aleatorio de los fenómenos.

Apoyo didáctico: En este escenario, el papel del docente no consiste en "corregir" de inmediato estas ideas, sino en comprenderlas como parte de un sistema de pensamiento más amplio. La literatura contemporánea insiste en que las intuiciones son recursos cognitivos útiles cuando se integran adecuadamente a actividades de exploración, discusión y reflexión.

Un ejemplo concreto puede ilustrar esta transición. Si en un grupo de estudiantes se lanza una moneda diez veces y aparece una secuencia de cinco caras consecutivas, es frecuente escuchar afirmaciones como: "Ahora tiene que salir sello" o "Va a equilibrarse". Cuando el docente repite el experimento cien o doscientas veces o utiliza una simulación digital que permite miles de repeticiones los estudiantes observan que la proporción se acerca a 50 %, pero que no existe una

compensación inmediata tras una racha. Esta observación empírica genera una tensión cognitiva que impulsa la reorganización conceptual.

Este tipo de experiencias pone en evidencia la importancia de la intuición como punto de partida del pensamiento probabilístico. Lejos de ser un obstáculo, constituye una base fértil para desarrollar discusiones, experimentos y análisis que conduzcan progresivamente hacia la comprensión formal del azar. En palabras de Fischbein y Malkiel (2019), "el pensamiento intuitivo no desaparece con el aprendizaje; se transforma y continúa cohabitando con las ideas formales, influyendo en la toma de decisiones incluso cuando la teoría es conocida".

En síntesis, comprender las intuiciones del azar implica reconocer la complejidad cognitiva, emocional y cultural desde la que los estudiantes interpretan los fenómenos inciertos. La enseñanza de la probabilidad no puede limitarse a transmitir reglas; debe partir de estas creencias iniciales, generando un diálogo entre la intuición y el conocimiento matemático que permita construir, de manera gradual y sólida, una nueva forma de pensar lo indeterminado.

La construcción del concepto formal de probabilidad: del experimento al modelo matemático

Transitar desde la intuición hacia el concepto formal de probabilidad implica un cambio epistemológico profundo. No se trata únicamente de aprender definiciones, sino de comprender que el azar posee regularidades que solo emergen cuando los fenómenos se analizan de manera sistemática. Para que este tránsito ocurra de forma significativa, la literatura reciente propone tres pilares: experimentación, simulación y formalización progresiva.

a) La experimentación como puente entre intuición y pensamiento frecuencial

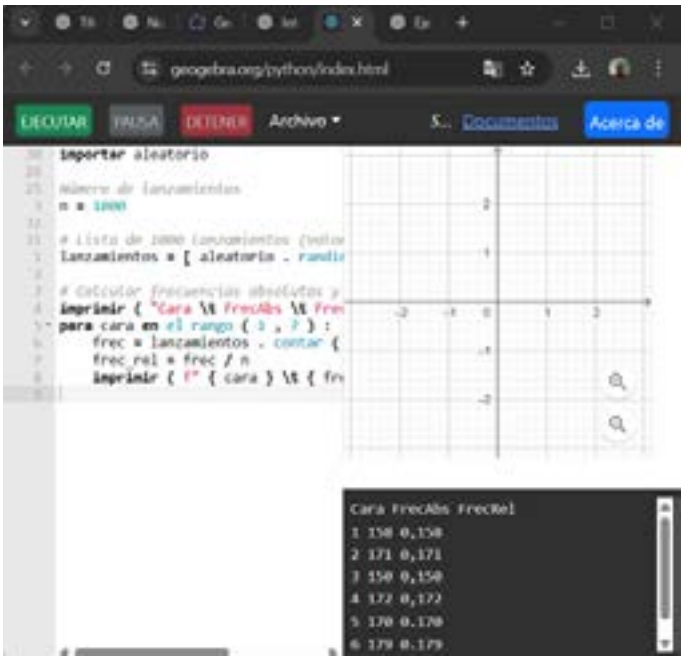
El trabajo con experimentos repetidos permite que los estudiantes observen la estabilización de las frecuencias relativas. Chance et al. (2016) señalan que esta aproximación simulation-based favorece la comprensión de que la probabilidad no es una predicción individual, sino una descripción colectiva del comportamiento del azar. Cuando los estudiantes realizan solo unas pocas repeticiones, sus expectativas intuitivas dominan la interpretación; pero cuando aumentan el número de ensayos, comienzan a advertir que la variabilidad empieza a "ordenarse".

La integración de herramientas digitales ha sido clave en esta transformación pedagógica. Prodromou (2019) demuestra que las simulaciones permiten visualizar la convergencia de frecuencias, comparar modelos y experimentar con escenarios

que serían imposibles de reproducir manualmente. Este enfoque, además, reduce el sesgo perceptual y permite trabajar con volúmenes de datos suficientes para que los patrones probabilísticos sean evidentes.

Por ejemplo, un docente de matemáticas desea que sus estudiantes comprendan cómo se comporta el azar cuando se repite muchas veces un mismo experimento. Para ello, propone simular 1000 lanzamientos de un dado equilibrado utilizando Python dentro de GeoGebra (Figura 1). La idea es que los estudiantes observen cuántas veces aparece cada cara, comparen esas frecuencias con el valor teórico esperado y reflexionen sobre por qué, aun tratándose de un proceso aleatorio, empiezan a aparecer patrones cuando se trabaja con una cantidad grande de datos.

Figura 1.
Resultados de la simulación de 1000 lanzamientos de un dado en el entorno Python de GeoGebra



Nota. La figura muestra el código utilizado para realizar la simulación, así como la tabla generada con las frecuencias absolutas y relativas de cada cara del dado.

Al revisar los resultados de la simulación, lo primero que llama la atención es que ninguna cara del dado aparece con una frecuencia muy distinta de las demás. Cada una se mueve alrededor del 16 o 17 por ciento, que es justamente lo que uno esperaría si el dado fuera equilibrado y todas las caras tuvieran la misma probabilidad de salir. Hay pequeñas variaciones entre ellas, pero

son normales y esperables en cualquier proceso aleatorio: algunas caras aparecen unas pocas veces más o menos, sin que eso signifique que el dado "prefiera" una en particular.

Lo interesante de este ejercicio es que permite ver cómo el azar empieza a mostrar regularidades cuando se trabaja con un número grande de lanzamientos. Aunque cada tiro es impredecible, el conjunto de los 1000 ensayos deja entrever un patrón estable que se acerca bastante al valor teórico. Esa estabilidad no se aprecia en lanzamientos aislados, pero se hace evidente cuando se dispone de suficientes datos. Justamente ahí está el valor educativo de la simulación: ayuda a que los estudiantes comprendan que la probabilidad no es una idea abstracta, sino algo que se puede observar y analizar cuando se mira el comportamiento de muchos casos juntos.

b) Las interpretaciones contemporáneas de la probabilidad como recursos complementarios

Aunque históricamente la probabilidad se enseñó casi exclusivamente desde la interpretación clásica, la investigación reciente señala la importancia de introducir múltiples enfoques. Biehler (2018) argumenta que la comprensión profunda del azar requiere que los estudiantes conozcan las tres grandes perspectivas:

- clásica, basada en casos equiprobables;
- frecuencial, donde la probabilidad es un límite de frecuencias;
- bayesiana, donde se interpreta como grado de creencia actualizado con información disponible.

En los últimos años, el enfoque bayesiano ha adquirido renovada relevancia en la educación matemática debido a su cercanía con situaciones reales donde la información evoluciona. Investigaciones como las de Benavoli y Zaffalon (2022) explican que este enfoque ayuda a comprender fenómenos como diagnósticos médicos, evaluaciones de riesgo y procesos de toma de decisiones donde las probabilidades dependen del conocimiento previo.

c) La formalización como construcción progresiva de significado

Uno de los avances didácticos más significativos de la última década es la incorporación de la argumentación como un componente central del aprendizaje probabilístico. Zieffler et al. (2018) señalan que la enseñanza de la probabilidad no puede limitarse a repetir procedimientos ni a resolver ejercicios de manera automática. Lo realmente formativo ocurre cuando los estudiantes deben explicar por qué un resultado es más plausible que otro y qué evidencias sostienen esa idea. En ese proceso, se ven obligados a pensar en la relación entre sus intuiciones, los datos que observan y la coherencia de sus

propias explicaciones. El aula se convierte así en un espacio donde las ideas se ponen a prueba, donde las afirmaciones deben justificarse y donde los argumentos se vuelven tan importantes como los números.

Esta forma de trabajar desplaza el enfoque tradicional centrado en algoritmos y abre la puerta a una comprensión más profunda de la probabilidad. Cuando los estudiantes discuten, revisan y confrontan sus razonamientos, comienzan a ver la probabilidad como una herramienta para interpretar el mundo y no solo como un tema más del currículo. La disciplina se articula entonces con la toma de decisiones, el análisis crítico de la información y la valoración de la evidencia, habilidades indispensables en la vida cotidiana y en cualquier campo profesional. En este sentido, argumentar no es un añadido complementario, sino una vía para que la probabilidad cobre sentido y se conecte con la manera en que las personas enfrentan la incertidumbre.

d) La probabilidad como forma de interpretar la realidad

El aprendizaje formal de la probabilidad tiene efectos más amplios que el dominio de técnicas matemáticas. Permite comprender fenómenos complejos del entorno: desde la epidemiología hasta la economía, desde las tendencias sociales hasta el análisis de datos científicos. McGrayne (2014) recuerda que los modelos probabilísticos han transformado la manera en que las sociedades toman decisiones, evalúan riesgos y predicen comportamientos.

En el aula, la probabilidad se convierte así en una herramienta para pensar críticamente la incertidumbre. Cuando los estudiantes comprenden que no todos los eventos pueden predecirse, pero sí pueden analizarse, y que los datos pueden orientar decisiones basadas en evidencia, el azar deja de ser un misterio y se convierte en una ventana para interpretar el mundo contemporáneo.

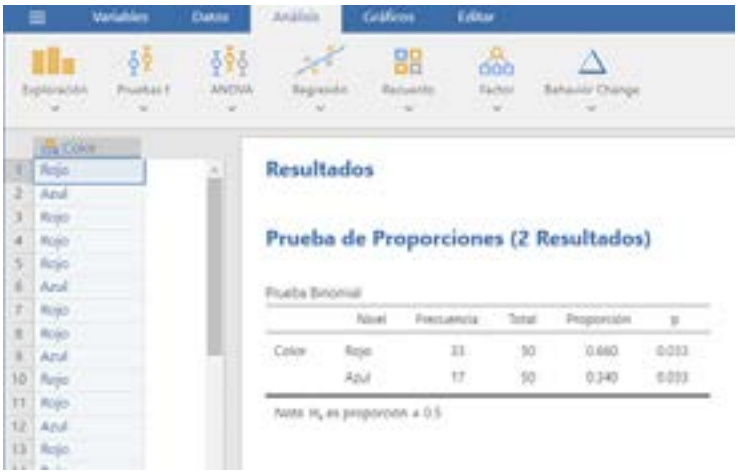
Por ejemplo, en una actividad de aula, el docente plantea que una caja contiene fichas de dos colores: rojas y azules, en una proporción desconocida para los estudiantes. Con el propósito de explorar cómo se comportan los fenómenos aleatorios, cada estudiante debe simular 50 extracciones con reemplazo y registrar el resultado de cada una (Figura 2).

Los resultados muestran que, de las 50 extracciones realizadas, el color rojo apareció 33 veces y el azul 17 veces. Esto significa que, en la muestra obtenida, aproximadamente el 66 % de las fichas fueron rojas y el 34 % azules. Si la probabilidad de obtener rojo y azul fuera la misma (0,5 cada una), esperaríamos frecuencias más equilibradas, cercanas a 25 y 25.

Sin embargo, la prueba binomial arroja un valor p = 0,033, lo que indica que es poco probable obtener una diferencia tan marcada únicamente por azar si la probabilidad real fuera 0,5.

En términos sencillos, los datos sugieren que el color rojo tiene una probabilidad mayor que el azul en este experimento, es decir, en la caja hay, efectivamente, más fichas rojas que azules y eso se refleja en la simulación. Desde la perspectiva del aula, este resultado permite a los estudiantes ver que, aunque cada extracción individual es impredecible, al reunir muchos datos aparecen patrones que ayudan a entender mejor el comportamiento del azar y a tomar decisiones basadas en evidencias y no solo en intuiciones.

Figura 2.
Resultados de la simulación de 50 extracciones aplicada a la variable "Color"



Nota. La figura muestra la distribución observada de 50 extracciones con reemplazo entre los colores rojo y azul en jamovi.

Experimentos aleatorios, sucesos y espacio muestral

Introducir la probabilidad en el aula implica conducir a los estudiantes hacia la comprensión de que el azar no es una fuerza misteriosa ni un capricho de la naturaleza, sino un fenómeno que puede describirse, analizarse y modelarse. En la vida cotidiana todos hemos experimentado situaciones donde los resultados no pueden predecirse con certeza: un balón que rebota de forma irregular, una rifa escolar, la aparición o no de un medicamento en una farmacia, el éxito de una semilla plantada por estudiantes en una actividad de Ciencias Naturales. Estas experiencias espontáneas son, en esencia, experimentos aleatorios, y constituyen un entorno pedagógico privilegiado para iniciar la reflexión probabilística.

El experimento aleatorio como experiencia fundante de la incertidumbre

Cuando introducimos la probabilidad en el aula, no partimos de fórmulas ni de definiciones formales, sino de una experiencia muy humana: la sensación de no saber qué va a pasar. Lanzar una moneda, mezclar fichas en una bolsa, esperar el resultado de un sorteo, sembrar semillas en un experimento escolar o registrar la lectura de un sensor en condiciones ligeramente cambiantes son ejemplos cercanos de situaciones donde el resultado es incierto, aunque el procedimiento sea claro. A estas situaciones las llamamos experimentos aleatorios.

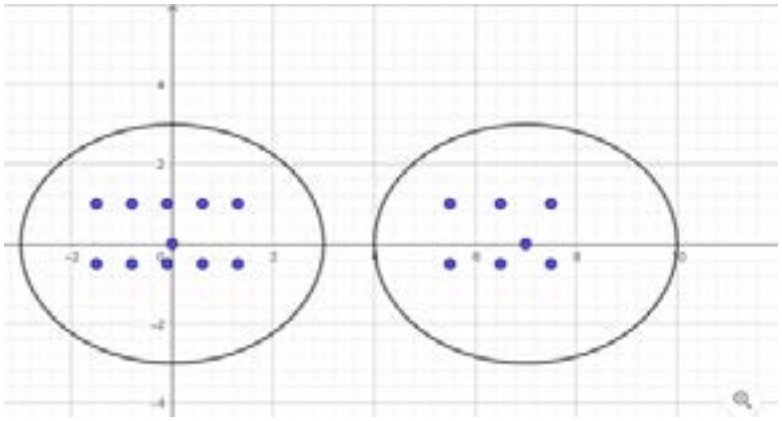
En términos sencillos, un experimento aleatorio es una acción que se realiza bajo ciertas reglas y que puede repetirse muchas veces, pero cuyo resultado individual no puede predecirse con certeza. La didáctica de la probabilidad se apoya precisamente en ese contraste: los estudiantes se sorprenden ante secuencias que les parecen "raras" y, a partir de esa sorpresa, se abre espacio para discutir qué significa realmente hablar de azar y de probabilidad.

Un aspecto clave es que los experimentos aleatorios permiten trabajar no solo con contenidos matemáticos, sino también con habilidades transversales. Aizikovitsh-Udi y Amit (2011) mostraron que una unidad didáctica centrada en la probabilidad en la vida cotidiana, con discusión colectiva y actividades de investigación, favorece el desarrollo del pensamiento crítico y creativo en el alumnado. Esto refuerza la idea de que no se trata únicamente de "calcular probabilidades", sino de aprender a formular conjeturas, justificar decisiones y revisar creencias a partir de evidencias.

En una actividad de ciencias, puede interesarnos el suceso "la planta germina antes del día siete"; en un contexto digital, "la plataforma registra más de diez accesos en una hora". Los sucesos no están "dados" de antemano en la realidad, sino que son selecciones analíticas que dependen de las preguntas que formulamos. Esta idea aparece con fuerza en los trabajos recopilados por Batanero, Chernoff, Engel, Lee y Sánchez (2016), donde se muestra cómo las decisiones sobre qué eventos se estudian condicionan el tipo de razonamiento probabilístico que surge en el aula.

Desde una perspectiva didáctica, es fundamental que los estudiantes aprendan a describir y representar los sucesos de distintas maneras. El uso de diagramas de Venn (Figura 3), tablas, esquemas verbales, materiales manipulativos o simulaciones en computadora permite visualizar relaciones como la unión, la intersección y la complementariedad.

Figura 3.
Representación de sucesos mutuamente excluyentes en un experimento aleatorio



Nota. La gráfica muestra dos conjuntos que representan los sucesos "sacar un caramelo de fresa" y "sacar un caramelo de limón".

La gráfica ayuda a que el razonamiento probabilístico que ocurre en el aula tome forma concreta. Cuando el estudiante ve los dos círculos separados y los puntos dentro de cada uno, entiende sin necesidad de fórmulas que ciertos sucesos no pueden darse al mismo tiempo y que todo se decide por los resultados posibles que allí aparecen. Esa representación sencilla les permite reconocer ideas como la unión, la intersección o la ausencia de ella, y mirar la probabilidad como una relación entre sucesos y no solo como un número que aparece por cálculo. En otras palabras, el diagrama actúa como un apoyo visual que organiza el pensamiento del alumnado y les permite discutir, comparar y justificar con mayor claridad por qué un suceso es compatible con otro o por qué no lo es.

En la obra colectiva Teaching and Learning Stochastics: Advances in Probability Education Research, se muestran múltiples experiencias en las que estas representaciones ayudan a superar malentendidos frecuentes, por ejemplo, cuando el alumnado confunde sucesos mutuamente excluyentes con sucesos independientes, o interpreta la probabilidad solo como "frecuencia aproximada" sin considerar la estructura del experimento.

Todo ello nos lleva a una conclusión importante: los experimentos aleatorios y los sucesos son, al mismo tiempo, herramientas matemáticas y dispositivos didácticos. Al diseñar actividades, el profesorado tiene la oportunidad de elegir qué experiencias de azar ofrecer, qué sucesos proponer al análisis y qué preguntas formular para que los estudiantes no se queden solo con la anécdota ("salió cara muchas veces"),

sino que avancen hacia una reflexión más estructurada ("¿es razonable este resultado?, ¿qué esperábamos?, ¿cómo podemos comprobarlo?").

El espacio muestral: mapear lo posible para comprender lo probable

Si el experimento aleatorio es el escenario y los sucesos son los focos que elegimos encender, el espacio muestral es el mapa completo de lo posible. Lo definimos como el conjunto de todos los resultados que pueden darse en un experimento, sin omitir ninguno y sin repetirlos. Aunque la definición parece sencilla, su construcción suele ser un punto crítico para el aprendizaje.

Las investigaciones recogidas en Research on Teaching and Learning Probability muestran que muchos estudiantes tienen dificultades para enumerar correctamente el espacio muestral, sobre todo en experimentos compuestos. No es raro que, al trabajar con el lanzamiento de dos dados, consideren que (2, 5) y (5, 2) son el mismo resultado, o que en un problema de selección de equipos no distingan entre "orden" y "combinación". Estas dificultades no son simples errores de cálculo: revelan problemas más profundos en la comprensión de la estructura del experimento.

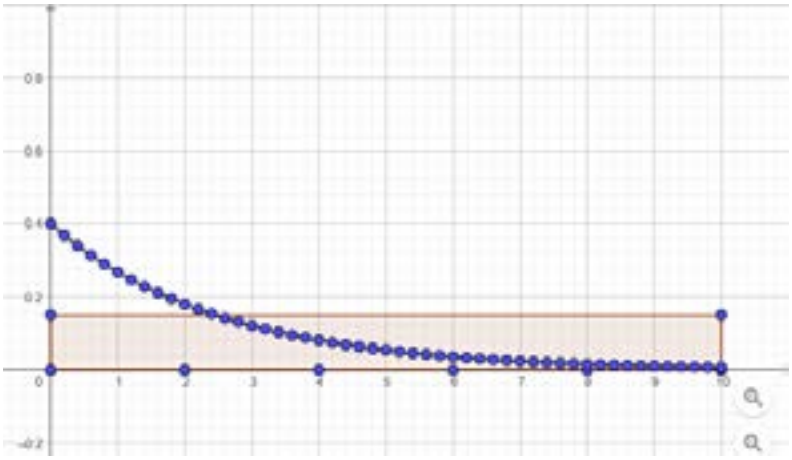
Por esa razón, Batanero y Borovcnik (2016) recomiendan trabajar el espacio muestral mediante representaciones sistemáticas: listas ordenadas, diagramas en árbol, tablas de doble entrada, cuadros combinatorios, entre otros. Estas herramientas ayudan a que el alumnado vea la organización interna del conjunto de resultados y descubra, por ejemplo,

por qué la suma 7 en dos dados tiene más formas de obtenerse que la suma 2 o la suma 12. A partir de ahí, se abre el camino para discutir si todos esos resultados son igualmente probables o no, y qué consecuencias tiene esto para el cálculo posterior.

Ahora bien, el espacio muestral no siempre es uniforme ni discreto. En muchos fenómenos reales el conjunto de resultados posibles es muy amplio o incluso infinito. En estos casos, no se trata de enumerar uno por uno todos los elementos, sino de describir el espacio muestral con modelos más globales (intervalos, distribuciones, funciones).

La gráfica (Figura 4) ilustra cómo, en muchos fenómenos reales, el espacio muestral no puede describirse como una lista finita de resultados, sino solo mediante modelos continuos.

Figura 4.
Representación del espacio muestral continuo



Nota. La figura ilustra cómo, en un fenómeno continuo, el espacio muestral no se describe enumerando resultados individuales, sino a través de modelos globales.

El intervalo horizontal representa todos los valores posibles del tiempo de espera entre 0 y 10 minutos: no es una colección de puntos aislados, sino una infinitud de posibilidades en un rango continuo. Sobre ese mismo eje, la barra uniforme muestra un modelo simplificado donde todos los tiempos serían igualmente probables, mientras que la curva exponencial evidencia un comportamiento más cercano a situaciones reales, donde los valores pequeños son más frecuentes y la probabilidad decae a medida que aumenta el tiempo. Esta superposición permite visualizar que, en contextos de mediciones continuas o datos generados por sensores, no tiene sentido enumerar caso por caso, sino que se requiere una función o distribución que describa cómo se reparte la probabilidad

Los trabajos recientes de Chernoff y Sriraman (2020) insisten en la necesidad de que, incluso en niveles iniciales, el profesorado ayude a los estudiantes a distinguir entre la idea de "lista finita de resultados" y la de "rango de valores posibles", preparando el terreno para una comprensión futura de las distribuciones de probabilidad.

La literatura también muestra que la construcción del espacio muestral está estrechamente vinculada con el desarrollo de la alfabetización probabilística del profesorado y del estudiante. En el volumen editado por Batanero y Chernoff (2016), diversos capítulos analizan cómo las creencias docentes, su formación previa y su familiaridad con las representaciones del azar influyen directamente en las tareas que proponen y en la manera de orientar la discusión en clase. Cuando el profesorado concibe el

espacio muestral solo como un requisito técnico para "aplicar la fórmula", tiende a reducir las actividades a ejercicios mecánicos de listado. En cambio, cuando lo entiende como un modelo que organiza lo posible y que da sentido a la noción de probabilidad, diseña experiencias donde los estudiantes participan activamente en su construcción, contrastan sus propias enumeraciones y discuten distintas maneras de representar lo mismo.

En términos didácticos, esto implica trabajar de manera articulada las tres nociones: experimento aleatorio, sucesos y espacio muestral. El estudiante debe acostumbrarse a seguir una especie de "ruta de modelización":

- Describir claramente el experimento: qué se hace, cómo se repite, qué condiciones se mantienen.
- Identificar todos los resultados posibles y construir el espacio muestral de forma organizada.
- Seleccionar los sucesos que interesan y representarlos como subconjuntos del espacio muestral.
- Analizar, a partir de este marco, si tiene sentido hablar de probabilidades iguales, diferentes, aproximadas o empíricas.

Cuando esta ruta se vuelve habitual, el cálculo de probabilidades deja de ser un procedimiento aislado y se convierte en la consecuencia natural de una forma de pensar. Como sintetiza Batanero et al. (2016) al hablar del "sentido estadístico", el objetivo no es solo que el alumnado responda correctamente a un ejercicio, sino que pueda explicar por qué una probabilidad es razonable en función del contexto, del experimento y de los posibles resultados.

Reglas básicas de la probabilidad y su interpretación

Las reglas básicas de la probabilidad constituyen el núcleo conceptual desde el cual se articulan modelos, simulaciones y razonamientos sobre fenómenos inciertos. En la práctica educativa, su enseñanza requiere superar la visión algorítmica centrada en fórmulas, y avanzar hacia un enfoque interpretativo que ayude a comprender por qué estas reglas tienen sentido y cómo se vinculan con procesos reales. Esto implica transitar de intuiciones primarias a representaciones más formales, reconociendo que la probabilidad, lejos de ser un mero cálculo, es un lenguaje para describir la incertidumbre (Kahneman, 2011).

Reglas esenciales del razonamiento probabilístico: unión, intersección y complemento de sucesos

La regla de la adición permite calcular la probabilidad de que ocurra al menos uno de dos sucesos. El principio central es evitar la doble contabilidad de resultados compartidos: $\mathbf{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$ Esta expresión no solo es una relación algebraica, sino una representación del modo en que se distribuyen los resultados dentro de un espacio muestral. Cuando los sucesos son mutuamente excluyentes, la fórmula se simplifica, ya que no existe intersección:$\mathbf{P(A \cup B) = P(A) + P(B)}$

Por ejemplo, si en una biblioteca hay 40 libros de probabilidad, 30 de estadística y 10 que pertenecen a ambas áreas, la probabilidad de que un libro seleccionado al azar pertenezca a al menos una de esas categorías es:

$$\mathbf{P\left(P \cup E\right) = \frac{40}{100} + \frac{30}{100} - \frac{10}{100} = \frac{60}{100}}$$

Este tipo de situaciones favorece el razonamiento visual mediante diagramas de Venn, los cuales, según Duval (2017), permiten coordinar el registro gráfico con el registro simbólico, promoviendo una comprensión más profunda.

La regla de la multiplicación establece cómo calcular la probabilidad de que dos sucesos ocurran simultáneamente. Su formulación depende de la relación entre los sucesos:

- $\mathbf{P(A \cap B) = P(A) \cdot P(B)}$, si A y B son independientes.
- $\mathbf{P(A \cap B) = P(A) \cdot P(B|A)}$, en el caso general.

Este es uno de los puntos donde más errores conceptuales se observan, especialmente entre estudiantes que asumen que la independencia es una condición natural o "intuitiva" (Garfield & Ben-Zvi, 2008). Para ilustrarlo, imaginemos dos extracciones sucesivas sin reemplazo desde una urna con 5 fichas rojas y 5 azules. La probabilidad de extraer dos fichas rojas consecutivas es:

$$\mathbf{P\left(R_1 \cap R_2\right) = \frac{5}{10} \cdot \frac{4}{9} = \frac{20}{90}}$$

Aquí el segundo suceso depende del primero, porque el espacio muestral cambia tras la primera extracción. Este razonamiento es clave para comprender fenómenos dependientes como pruebas diagnósticas, procesos de calidad o cadenas de eventos.

La regla del complemento establece que la probabilidad de que no ocurra un suceso es:

$$\mathbf{P(A^c) = 1 - P(A)}$$

Aunque simple, esta regla posee un alto valor heurístico. En análisis de riesgo, fiabilidad o epidemiología, suele ser más fácil determinar la probabilidad del suceso contrario y restarlo de 1.

Por ejemplo, si la probabilidad de que un examen clínico produzca un falso positivo es 0,08, entonces la probabilidad de que no lo produzca es: P(no falso positivo)=$1 - 0.08 = 0.92$

Esta regla también permite repensar situaciones acumulativas. Por ejemplo: "¿Cuál es la probabilidad de que ninguna de las tres máquinas falle durante una jornada?". Si la probabilidad de que cada máquina falle es baja e independiente, es más eficiente calcular el complemento del suceso "al menos una falla".

Dimensiones interpretativas de las reglas: clásico, frecuencial y bayesiano

Las reglas básicas de la probabilidad no deberían enseñarse como herramientas aisladas ni como un conjunto de recetas que el estudiantado memoriza sin comprender. Su sentido profundo emerge cuando se interpretan desde diferentes perspectivas epistemológicas que, históricamente, han dado forma al concepto de probabilidad. Reconocer estas interpretaciones ayuda a evitar una visión reducida del azar y favorece un aprendizaje más reflexivo.

En términos generales, la literatura especializada identifica tres grandes marcos interpretativos:

1. **Interpretación clásica:** adecuada para espacios equiprobables y finitos, donde los resultados se consideran igualmente posibles. Es la visión heredada de Laplace, para quien la probabilidad consistía en el cociente entre casos favorables y casos posibles (Hacking, 2006). Ejemplo: al lanzar un dado perfecto, la probabilidad de obtener un número par se determina dividiendo los tres resultados favorables (2, 4, 6) entre los seis posibles.
2. **Interpretación frecuencial:** enfatiza la probabilidad como límite de las frecuencias relativas de un suceso cuando un experimento se repite un gran número de veces. Esta interpretación, asociada a von Mises (1981), fundamenta la probabilidad en la estabilidad estadística que emerge a largo plazo. Ejemplo: si se repite el lanzamiento de una moneda diez mil veces y se observa que "cara" aparece aproximadamente la mitad de las veces, se considera que la probabilidad converge hacia 0.5.
3. **Interpretación subjetiva o bayesiana:** concibe la probabilidad como un grado razonable de creencia basado en la información disponible. Esta visión, impulsada desarrollada actualmente en estadística bayesiana, sostiene que la probabilidad refleja el nivel de confianza que un individuo asigna a la ocurrencia de un evento dado cierto conocimiento previo (Gelman et al., 2021). Ejemplo: un médico puede

asignar una probabilidad preliminar del 20 por ciento a que un paciente tenga una enfermedad, basándose en su experiencia previa y en la prevalencia poblacional, antes incluso de solicitar pruebas diagnósticas.

Las simulaciones, especialmente con herramientas digitales como GeoGebra, R o Jamovi, se han convertido en un recurso esencial para articular estas tres interpretaciones. Al repetir un proceso miles de veces, el estudiantado observa de primera mano cómo la frecuencia relativa se aproxima a la probabilidad teórica, fenómeno que ya von Mises (1981) identificaba como fundamento de la estabilidad del azar. Esta aproximación visual, dinámica y experimental permite que la probabilidad deje de ser un concepto meramente abstracto y se convierta en un patrón observable.

Por ejemplo, al simular diez mil lanzamientos de una moneda en Jamovi, se obtiene una gráfica donde la proporción de caras oscila fuertemente al inicio, pero se estabiliza alrededor de 0.5 conforme aumentan las repeticiones. Esto permite una conversación didáctica rica: ¿por qué se observan fluctuaciones al principio?, ¿qué significa la "estabilidad" del azar?, ¿qué relación tiene este comportamiento con la interpretación frecuencial?

Del mismo modo, en GeoGebra se puede construir un applet que represente el lanzamiento de un dado con un botón que ejecuta simulaciones incrementales: 10, 50, 100, 500, 1000 intentos. A medida que el número de repeticiones aumenta, la frecuencia relativa de cada cara se va alineando con la distribución equiprobable esperada de la interpretación clásica. Esta actividad brinda al estudiantado la oportunidad de confrontar sus propias intuiciones y reconocer que el azar no es caos absoluto, sino variabilidad organizada.

Para Batanero, Contreras y Díaz (2016), estas experiencias constituyen un medio poderoso para desarrollar el razonamiento probabilístico: permiten comparar la intuición inicial con la evidencia empírica, discutir fenómenos como la ley de los grandes números o los sesgos cognitivos y construir gradualmente una comprensión estructurada del azar. Las simulaciones actúan como "escenarios de exploración controlada", donde el docente guía y provoca preguntas significativas.

Técnicas de conteo: permutaciones, variaciones y combinaciones

Comprender la probabilidad en el marco de la interpretación clásica exige disponer de un conjunto de herramientas que permitan determinar el número de resultados posibles asociados a un experimento aleatorio. Estas herramientas, conocidas como

técnicas de conteo o métodos combinatorios, constituyen un puente entre la estructura del espacio muestral y las reglas básicas de la probabilidad.

La combinatoria no surge como una colección aislada de procedimientos, sino como una forma rigurosa de responder preguntas del tipo "¿cuántos arreglos posibles existen?", "¿cuántas formas hay de seleccionar elementos?", o "¿de cuántas maneras puede organizarse una situación?". Estas preguntas aparecen en innumerables contextos: distribución de turnos, asignación de puestos, códigos de seguridad, rutas posibles, juegos de azar, muestreos, entre otros. Desde una mirada didáctica, este carácter transversal permite emplear problemas contextualizados que ayuden a los estudiantes a visualizar la estructura del conteo antes de recurrir a expresiones formales (Lockwood, 2013).

En este subepígrafe se profundiza en los fundamentos de las técnicas de conteo mediante cuatro bloques: (a) el principio aditivo y el principio multiplicativo; (b) el concepto de factorial como mecanismo de ordenación; (c) permutaciones y variaciones; y (d) combinaciones. Se incorporan ejemplos, interpretaciones y dificultades didácticas, así como una articulación explícita con el cálculo probabilístico.

Principios básicos: aditivo y multiplicativo

Los dos principios fundamentales de la combinatoria son sorprendentemente simples, pero se encuentran en el origen de todos los métodos posteriores.

pero se encuentran en el origen de todos los métodos posteriores.

a) Principio aditivo

Si un suceso puede ocurrir de mmm maneras y otro suceso diferente puede ocurrir de nnn maneras, y ambos sucesos no pueden darse simultáneamente, entonces el número total de maneras en que puede ocurrir uno u otro es: m+n. Este principio refleja situaciones de alternativa excluyente. Por ejemplo, si un estudiante puede elegir entre 5 proyectos de matemáticas y 3 de ciencias, y no puede combinar áreas, tiene 8 opciones posibles. Desde la didáctica, este principio suele confundirse con situaciones multiplicativas, de modo que trabajar listas exhaustivas o diagramas de árbol iniciales ayuda a fortalecer la distinción (Brousseau, 1997).

b) Principio multiplicativo

Si un procedimiento se compone de dos etapas independientes, donde la primera puede realizarse de m maneras y la segunda de n maneras, entonces el número total de procedimientos posibles es: m×n.m. Este principio modela situaciones de elección sucesiva. Por ejemplo, si un candado tiene 4 posiciones posibles en la primera rueda y 6 en la segunda, entonces hay 4×6=244 códigos posibles.

Ambos principios constituyen la base para construir espacios muestrales complejos sin necesidad de enumeración. Al enseñarles, es útil alternar diagramas de árbol, tablas y representaciones gráficas, pues facilitan la transición desde lo intuitivo hacia la formalización (Duval, 2017).

La factorial: fundamento del ordenamiento

El símbolo factorial, idenotado como n!, expresa el número de maneras de ordenar $\mathbf{n!}$, elementos distintos: $\mathrm{n!} = \mathrm{n} \times (\mathrm{n}-1) \times (\mathrm{n}-2) \times \ldots \times 2 \times 1$ .Su importancia es crucial: aparece en permutaciones, variaciones, combinaciones, coeficientes binomiales e incluso en modelos probabilísticos como la distribución binomial y la distribución hipergeométrica.

Desde una perspectiva didáctica, se observa que los estudiantes suelen entender intuitivamente que "ordenar elementos" genera muchas opciones, pero no siempre conectan esta idea con la factorial. Proponer tareas que incluyan organizar personas en una fila, ordenar libros o generar claves ayuda a establecer un puente entre la intuición y la expresión formal.

Ejemplo: ¿Cuántas maneras hay de ordenar 5 libros distintos en un estante?

$5! = 120$.

Permutaciones y variaciones: ordenaciones con y sin restricción

a) Permutaciones sin repetición

Una permutación consiste en ordenar todos los elementos disponibles. Si hay n elementos distintos, el número de permutaciones es: $\mathrm{P(n)} = \mathrm{n!}$

Ejemplo: Ordenar 7 banderas diferentes en una ceremonia:

$7! = 5040$

b) Permutaciones con repetición

Si algunos elementos se repiten, la fórmula incorpora divisiones factoriales:

$$\mathrm{P}\left(\mathrm{n_1, n_2, \ldots, n_k}\right) = \frac{\mathrm{n!}}{\mathrm{n_1! n_2! \cdots n_k!}}$$

Ejemplo: La palabra "MATEMÁTICA" tiene 10 letras, con repeticiones de A (3), M (2) y T (2): $\frac{10!}{3!2!2!} = 151200$

Este tipo de tareas suele ser cognitivamente más complejas, pues exige reconocer la causa de la sobrecontabilidad. Se recomienda trabajarlas desde manipulaciones de cartas o bloques antes de introducir la fórmula.

Variaciones (con y sin repetición)

Las variaciones modelan situaciones donde se seleccionan algunos elementos y el orden importa.

Variaciones sin repetición: $\mathrm{V}\left(\mathrm{n, k}\right) = \frac{\mathrm{n!}}{(\mathrm{n}-\mathrm{k})!}$

Ejemplo: Seleccionar y ordenar 3 estudiantes de un grupo de $V(10,3) = 720$

Variaciones con repetición: $VR(n,k) = n^k$

Ejemplo: Generar códigos de 4 dígitos usando los números 0–9: $10^4 = 10000$

Las variaciones son particularmente valiosas para modelar claves, rutas, asignaciones y sistemas de muestreo. En la práctica docente, su mayor dificultad radica en distinguir cuándo el orden importa y cuándo no; tareas comparativas ayudan a clarificar esta idea.

Combinaciones: selección sin orden

Las combinaciones se utilizan cuando se seleccionan elementos sin importar el orden. La fórmula general es:

$$C\left(n,k\right) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Ejemplo: ¿De cuántas maneras puede formarse un comité de 4 personas entre 12 candidatos? C(12,4)=495.

Desde la enseñanza, es frecuente que los estudiantes confundan variaciones y combinaciones, lo que exige diseñar secuencias de actividades que exploren explícitamente el rol del orden mediante representaciones alternativas (diagramas, árboles, listados parciales, simulaciones digitales).

Dificultades didácticas y orientaciones pedagógicas

La investigación en didáctica de la probabilidad y la combinatoria (Jones et al., 2007) ha identificado las siguientes dificultades habituales:

1. Confusión entre orden y no orden:
   Los estudiantes tienden a asumir que ordenar y seleccionar son equivalentes. Estrategia: actividades comparativas con objetos manipulables.
2. Uso mecánico de fórmulas:
   Aplican expresiones sin comprender la estructura combinatoria subyacente.
   Estrategia: construir primero el conteo mediante diagramas, tablas y listados parciales.
3. Sobrecarga cognitiva:
   La combinatoria implica coordinar varios niveles de abstracción.
   Estrategia: integrar tecnología (GeoGebra, Python, Jamovi) para visualizar procesos.
4. Dificultad para identificar el tamaño del espacio muestral:
   Estrategia: trabajar casos pequeños y aumentar progresivamente la complejidad.
   De manera general las técnicas de conteo constituyen un componente esencial para comprender y aplicar la probabilidad desde la interpretación clásica. Su enseñanza debe

equilibrar la formalización matemática con estrategias didácticas que permitan a los estudiantes construir significado, visualizar estructuras de selección y ordenamiento, y conectar estos procedimientos con el cálculo probabilístico. Con apoyo de herramientas digitales, representaciones múltiples y problemas contextualizados, es posible transformar la combinatoria en un campo accesible, potente y formativo dentro del estudio del azar.

Estrategias didácticas y mediaciones tecnológicas para el desarrollo del razonamiento probabilístico

El desarrollo del razonamiento probabilístico constituye uno de los desafíos más relevantes en la educación matemática contemporánea. No se trata únicamente de enseñar a calcular probabilidades, sino de favorecer que el estudiantado comprenda el carácter incierto de numerosos fenómenos y pueda interpretar, modelar y argumentar sobre situaciones donde la variabilidad y el azar juegan un papel esencial. Investigaciones recientes han mostrado que las intuiciones iniciales sobre el azar suelen ser frágiles, guiadas por ideas espontáneas, heurísticos o sesgos cognitivos que dificultan una comprensión formal (Kahneman, 2011). Por esta razón, es indispensable construir ambientes didácticos que permitan vincular la experiencia intuitiva del azar con su conceptualización matemática, articulando estrategias de enseñanza con herramientas tecnológicas que apoyen la visualización, la experimentación y la simulación digital.

La didáctica de la probabilidad reconoce al menos tres ejes fundamentales para formar un pensamiento probabilístico robusto: (a) la exploración de fenómenos aleatorios mediante experimentos reales o virtuales; (b) la explicitación y el cuestionamiento de intuiciones erróneas; y (c) la construcción progresiva de modelos simbólicos y gráficos que permitan interpretar la incertidumbre (Jones et al., 2007). Integrar estos ejes exige diseñar secuencias didácticas donde la manipulación, el análisis exploratorio y la representación visual sean tan importantes como el cálculo formal. Es aquí donde la tecnología adquiere un rol protagónico: las simulaciones digitales permiten repetir procesos aleatorios miles de veces en pocos segundos, observar patrones emergentes, contrastar hipótesis intuitivas con evidencia empírica y comprender la estabilidad estadística.

Estrategias didácticas centradas en la exploración, la indagación y el diálogo

Una estrategia central para el desarrollo del razonamiento probabilístico es promover situaciones de exploración guiada, donde el estudiantado pueda manipular objetos, experimentar

con fenómenos aleatorios sencillos y formular conjeturas. Actividades como lanzar monedas, seleccionar fichas de una urna, explorar juegos de azar o analizar situaciones de incertidumbre cotidiana permiten generar preguntas, comparar resultados y construir gradualmente ideas clave, como la noción de espacio muestral, frecuencia relativa, independencia y variabilidad.

El valor pedagógico de estas experiencias no reside solo en la manipulación física, sino en la forma en que se articulan con la discusión socio matemática: ¿por qué los resultados no siguen un patrón fijo?, ¿por qué a veces parecen acercarse a un valor estable?, ¿cómo distinguir lo aleatorio de lo determinista?

Otra estrategia clave es el uso de representaciones múltiples, tal como sugiere Duval (2017). Tablas, diagramas de árbol, gráficos de barras, simulaciones y listas permiten reorganizar la información de distintas maneras, facilitando la comprensión de relaciones que no siempre son visibles en una sola representación. Aquí es importante no "imponer" la fórmula desde el inicio; más bien, se busca que los estudiantes construyan primero la estructura combinatoria o probabilística del problema y solo después formalicen el procedimiento.

La modelación didáctica también constituye un recurso poderoso: crear situaciones donde el estudiantado deba construir un modelo probabilístico, validarlo con datos simulados o reales, y argumentar su pertinencia, refuerza la conexión entre teoría y práctica. En este tipo de tareas, los estudiantes deben justificar por qué su espacio muestral es adecuado, qué suposiciones realizan (por ejemplo, equiprobabilidad), y cómo sus predicciones se contrastan con los resultados observados.

Tecnología y visualización: simulaciones del azar, experimentación digital y análisis dinámico

Las tecnologías digitales ofrecen un entorno privilegiado para explorar el azar. Herramientas como GeoGebra, Python, Desmos, Jamovi o plataformas de simulación permiten ejecutar cientos o miles de repeticiones de un experimento aleatorio, mostrar la evolución de la frecuencia relativa, comparar resultados, identificar convergencias y visualizar oscilaciones. Este enfoque no solo acelera procesos que serían muy lentos de ejecutar manualmente, sino que también promueve una comprensión profunda de la interpretación frecuencial de la probabilidad y del carácter estable de las leyes estadísticas (von Mises, 1981; Batanero, Contreras & Díaz, 2016).

Estas herramientas también facilitan el trabajo con representaciones dinámicas. Por ejemplo, un gráfico interactivo que actualiza la frecuencia relativa en tiempo real mientras se ejecuta una simulación permite observar las oscilaciones iniciales, la progresiva reducción de la variabilidad y la emergencia de un valor estable.

Integración didáctica: de la intuición a la formalización mediante experiencias tecnológicas

Integrar estrategias didácticas con herramientas tecnológicas implica diseñar secuencias donde la exploración inicial del azar dé paso a la modelación simbólica y al análisis formal. Una posible estructura didáctica podría incluir:

1. Experiencia intuitiva

El estudiantado observa o manipula un fenómeno aleatorio (por ejemplo, extraer fichas de una urna).

2. Discusión de predicciones e hipótesis

Se expresan expectativas informales: "creo que debería salir más o menos la mitad", "creo que a veces se equilibrará".

3. Simulación digital

Se ejecutan muchas repeticiones para contrastar intuiciones con resultados empíricos.

4. Construcción de representaciones

Se analizan gráficos de frecuencias, histogramas o diagramas dinámicos.

5. Formalización matemática

Se introducen expresiones como probabilidad teórica, espacio muestral, independencia o combinaciones, ya conectadas con la experiencia.

6. Reflexión metacognitiva

Se discute cómo la simulación ayuda a comprender la probabilidad y qué sesgos se han superado.

Conclusiones

Este capítulo ha mostrado que entender la probabilidad va mucho más allá de aprender fórmulas. Supone reconocer que, cuando hablamos de azar, nuestras primeras intuiciones suelen estar llenas de ideas parciales, patrones que creemos ver donde no los hay y expectativas que no siempre se sostienen al mirar los datos con calma. Trabajar con experimentos aleatorios sencillos, discutir qué esperamos que ocurra y comparar esas expectativas con lo que realmente sucede permite que el estudiantado tome conciencia de esa distancia entre intuición y realidad, y empiece a construir una mirada más crítica sobre la incertidumbre.

A lo largo del capítulo también se vio que conceptos como experimento aleatorio, espacio muestral, sucesos, independencia o reglas básicas de la probabilidad solo adquieren sentido cuando se vinculan con situaciones concretas. No basta con enunciarlos: hay que explorarlos mediante ejemplos, representaciones múltiples y conversaciones en el aula que ayuden a responder preguntas del tipo "qué puede pasar", "qué consideramos posible" y "cómo contamos los casos". De este modo, la probabilidad deja de ser un recetario de técnicas y se convierte en un lenguaje para describir y analizar fenómenos donde intervenir directamente no es posible, pero sí es posible modelar y anticipar.

Finalmente, el capítulo ha subrayado el papel de la tecnología como aliada para pensar el azar. Las simulaciones digitales, los gráficos dinámicos y la posibilidad de repetir un experimento miles de veces en segundos permiten ver cómo las frecuencias relativas se estabilizan, cómo emergen regularidades y cómo se ponen a prueba los modelos teóricos. Integrar estas herramientas en la enseñanza no solo facilita los cálculos, sino que ayuda a que las y los estudiantes construyan un razonamiento probabilístico más profundo, capaz de sostener decisiones informadas en contextos reales donde la incertidumbre está siempre presente.

Referencias

Aizikovitsh-Udi, E., & Amit, M. (2011). Developing the skills of critical and creative thinking by probability teaching. Procedia – Social and Behavioral Sciences, 15, 1087–1091. https://doi.org/10.1016/j.sbspro.2011.03.243

Batanero, C., & Borovcnik, M. (2016). Statistics and probability in high school. Springer. https://doi.org/10.1007/978-3-319-23986-6

Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). Research on teaching and learning probability. En C. Batanero, E. J. Chernoff, J. Engel, H. S. Lee, & E. Sánchez (Eds.), Research on teaching and learning probability (pp. 1–33). Springer. https://doi.org/10.1007/978-3-319-31625-3_1

Batanero, C., Contreras, J. M., & Díaz, C. (2016). Teaching and learning probability. En D. Ben-Zvi & K. Makar (Eds.), The teaching and learning of statistics (pp. 37–54). Springer. https://doi.org/10.1007/978-3-319-23470-0_4

Benavoli, A., & Zaffalon, M. (2022). The Bayesian approach: A modern view. Entropy, 24(2), 254. https://doi.org/10.3390/e24020254

Biehler, R. (2018). Perspectives on modeling variability. En R. Biehler, B. Frischemeier, & M. Reading (Eds.), Statistics and data science education: New challenges for teaching and learning (pp. 45–67). Springer.

Brousseau, G. (1997). Theory of didactical situations in mathematics. Kluwer Academic Publishers.

Chance, B., Cobb, G., Rossman, A., Calder, J., & Swanson, T. (2016). Simulation-based inference in introductory statistics. Wiley.

Chernoff, E. J., & Sriraman, B. (Eds.). (2016). Teaching and learning stochastics: Advances in probability education research. Springer.

Duval, R. (2017). Understanding the mathematical way of thinking: The registers of semiotic representations. Springer.

Fischbein, E., & Malkiel, E. (2019). Intuitions of randomness revisited. International Journal of Mathematical Education in Science and Technology, 50(1), 132–147.

Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. Springer.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2021). Bayesian data analysis (3rd ed.). CRC Press.

Hacking, I. (2006). The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511817557

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability education. En F. K. Lester (Ed.), Second handbook of research on mathematics teaching and learning (pp. 909–955). Information Age.

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

Lockwood, E. (2013). A model of students' combinatorial thinking. Journal of Mathematical Behavior, 32(2), 251–265. https://doi.org/10.1016/j.jmathb.2013.02.008

McGrayne, S. (2014). The theory that would not die. Yale University Press.

Prodromou, T. (2019). Data modelling and statistical reasoning with dynamic data. Springer. https://doi.org/10.1007/978-3-030-03633-1

Retamal, C., & Alsina, Á. (2022). Las creencias intuitivas del azar en estudiantes chilenos. Bolema, 36(73), 449–468.

Von Mises, R. (1981). Probability, statistics and truth. Dover Publications.

Zieffler, A., Garfield, J., & Fry, E. (2018). Reasoning about uncertainty. Springer.

Capítulo III

Variables aleatorias, distribuciones y modelación estadística

Introducción

Comprender los fenómenos que observamos en la realidad implica reconocer que no todos se comportan de manera fija o predecible. Las calificaciones de un curso, el tiempo que tarda en llegar un bus, el número de mensajes que recibimos en un día o la cantidad de personas que acuden a un servicio son ejemplos cotidianos en los que la variación está siempre presente. Este capítulo parte de esa idea: para describir y explicar la variabilidad necesitamos herramientas que permitan capturarla, representarla y analizarla con sentido. Las variables aleatorias y las distribuciones estadísticas cumplen justamente ese papel, convirtiéndose en un puente entre los datos que observamos y los modelos que elaboramos para comprenderlos.

A lo largo del capítulo, el lector encontrará un recorrido que inicia en el concepto de variable aleatoria como una manera de formalizar la incertidumbre y organizar los posibles valores que puede tomar un fenómeno. Desde allí se avanza hacia las funciones de probabilidad y densidad, que permiten describir patrones, identificar regularidades y anticipar comportamientos. Este enfoque no se limita a la técnica: busca mostrar cómo los modelos estadísticos ofrecen una forma de narrar la variabilidad y de interpretar tanto lo esperado como lo excepcional dentro de un conjunto de datos.

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

Finalmente, el capítulo invita a pensar las distribuciones no como figuras abstractas, sino como representaciones de historias reales. Cada distribución dice algo sobre el fenómeno que la originó: si los resultados son simétricos o están sesgados, si los valores extremos son comunes o raros, si los eventos ocurren de manera uniforme o concentrada. Comprender estas características es fundamental para modelar fenómenos reales, analizar patrones y tomar decisiones informadas. Con este propósito, el capítulo combina teoría, ejemplos contextualizados y reflexiones didácticas que ayudan a dar sentido a una de las dimensiones más potentes y formativas de la estadística.

Variable aleatoria: concepto, sentido y ejemplos contextualizados

La variabilidad forma parte de nuestra vida diaria de maneras tan sutiles que, con frecuencia, pasa desapercibida. La hora exacta en que llega el bus, el número de mensajes que recibimos durante la mañana, el tiempo que tarda en calentarse el agua o la cantidad de estudiantes que faltan un lunes cualquiera son situaciones donde los resultados nunca son totalmente predecibles. Aunque solemos convivir con estas fluctuaciones sin pensarlo demasiado, representan el punto de partida para comprender uno de los conceptos más poderosos de la probabilidad: la variable aleatoria.

Lejos de ser una idea puramente técnica, la variable aleatoria es una forma de dar estructura matemática a la incertidumbre, permitiendo analizar fenómenos que no se comportan de manera fija. Desde esta perspectiva, comprenderla es comprender que el mundo no es estático ni exacto, pero sí presenta patrones y regularidades. Esta idea, central en el pensamiento probabilístico, fundamenta los modelos que articulan este capítulo.

Con este propósito, el epígrafe se desarrolla en tres momentos:

1. Fundamentos conceptuales y cognitivos, donde se reconstruye el significado profundo del concepto;
2. Clasificación y modelación, donde se vincula la variable aleatoria con sus representaciones y usos;
3. Dificultades comunes, donde se identifican obstáculos que interfieren en su comprensión.

A lo largo del desarrollo, se incorporan múltiples ejemplos cotidianos que demuestran que este concepto vive en el corazón de nuestra experiencia diaria.

La variable aleatoria como estructura de la variabilidad: fundamentos conceptuales y cognitivos

Comprender una variable aleatoria implica, ante todo, reconocer la naturaleza constante de la variación en el entorno. Pensemos, por ejemplo, en la hora exacta en que un estudiante se conecta a una clase virtual. Aunque todos saben que la sesión

empieza a las 08h00, cada día se observa una ligera variación: 07h58, 08h01, 08h03... Esa diferencia, pequeña pero persistente, muestra que incluso las actividades rutinarias están atravesadas por fluctuaciones que no responden al azar caótico, sino a la combinación de múltiples factores: tráfico, conexión a internet, temperatura del dispositivo, horarios familiares.

Batanero (2001) afirma que el principal desafío es desmontar la expectativa determinista que la escuela ha cultivado durante años: la idea de que repetir un procedimiento debería producir siempre el mismo resultado. Sin embargo, la vida diaria demuestra lo contrario. Cuando un estudiante realiza cinco veces un mismo experimento de caída libre con sensores digitales, los tiempos nunca coinciden exactamente. Cuando grabamos un audio, la duración de cada fragmento presenta pequeñas variaciones.

La variable aleatoria surge para conceder significado matemático a esa variabilidad natural.

Stewart (2013) ilustra que, aunque cada resultado individual sea incierto, la colección de muchos resultados revela un comportamiento regular. Esto explica por qué fenómenos como el tiempo de carga de una página web, la cantidad de pasos que da una persona en la mañana o el número de usuarios en una cafetería durante una hora tienden a seguir patrones que pueden observarse, representarse e incluso modelarse.

Moore (2010) sugiere que, en lugar de comenzar con definiciones formales, es más apropiado partir de actividades donde el estudiante experimente la variación. Por ejemplo:

- Temperatura del aula a lo largo del día.

Aunque se mida con el mismo termómetro, los valores fluctúan: 22.1°C, 22.3°C, 21.8°C...

- Tiempo que tarda en hervir agua en distintas cocinas.

Incluso usando la misma olla y la misma cantidad de agua, el tiempo cambia ligeramente.

- Número de interrupciones durante una clase virtual.

Algunas sesiones tienen una sola interrupción, otras cuatro o cinco.

Estos fenómenos no necesitan grandes laboratorios para ser observados; se encuentran en cualquiera de nuestras rutinas. Identificarlos y analizarlos permite que el estudiante capte de forma natural el papel de la variable aleatoria.Bakker (2004) y Ben-Zvi (2000) sostienen que este tipo de experiencias facilita que el concepto emerja como un modelo interno que organiza la variabilidad. Es decir, el estudiante aprende que la variable aleatoria no representa un número, sino el abanico de posibilidades de un fenómeno que nunca se repite idénticamente.

Borovcnik (2016) agrega que este proceso no es inmediato, pues exige reconocer tres niveles de estructuración:

- El fenómeno variable (lo que observamos).
- Los valores posibles (el espacio de resultados).
- El significado numérico asignado a cada resultado (la formalización).

La combinación de estos tres niveles da origen a un marco conceptual potente que permite comprender la variación como un fenómeno ordenado.

Apoyo didáctico: Desde el punto de vista didáctico el docente propone la siguiente situación: si se registra durante varios días la hora real de llegada de los estudiantes, teniendo como referencia la hora de entrada oficial (08h00). Observa llegadas como 07h55, 07h58, 08h02, 08h07, etc. El objetivo es mostrar que la puntualidad no es fija, sino una variable aleatoria continua ligada a muchos factores.

El fenómeno que nos interesa analizar es el "grado de puntualidad", que puede definirse a partir de dos formas de medición: la hora exacta de llegada o, de manera más operativa para el trabajo estadístico, los minutos de adelanto o retraso con respecto a las 08h00. Esta segunda opción resulta especialmente útil en herramientas como Jamovi, ya que permite expresar la puntualidad como una variable aleatoria continua, capaz de asumir múltiples valores dentro de un intervalo razonable, por ejemplo, desde –10 hasta +15 minutos. Desde la perspectiva didáctica, trabajar con este tipo de variable ofrece oportunidades claras para el aprendizaje: por un lado, ayuda al estudiante a reconocer la variabilidad como una característica natural de los fenómenos reales, y por otro, permite hacer visible la importancia del comportamiento global de los datos, mostrando que lo fundamental no es un día aislado sino el modo en que el conjunto se comporta en su totalidad.

Para efectos didácticos, puedes trabajar con un conjunto pequeño (Tabla 1), por ejemplo, 20 observaciones de "minutos respecto de las 08h00" (valores negativos = llegan antes; positivos = llegan después).

Como se observa en la Figura 1, la puntualidad del grupo se organiza en torno a una ligera tardanza promedio. La media es de 0,78 minutos y la mediana de 1 minuto, lo que indica que, en general, los estudiantes llegan apenas después de la hora oficial. La desviación típica de 3,07 minutos muestra una variabilidad moderada, con llegadas que oscilan entre 4 minutos antes y 5 minutos después de las 08h00.

Tabla 1.
Registro diario del grado de puntualidad de los estudiantes

Día	Estudiante	Minutos_08h00
1	A	-3
1	B	2
1	C	5
2	A	-1
2	B	1
2	C	4
3	A	-4
3	B	0
3	C	3
...	...	...

Nota. La tabla presenta los valores observados de minutos de adelanto o retraso respecto de las 08h00 para cada estudiante en tres días consecutivos.

Figura 1.
Estadísticos descriptivos del grado de puntualidad respecto de las 08h00

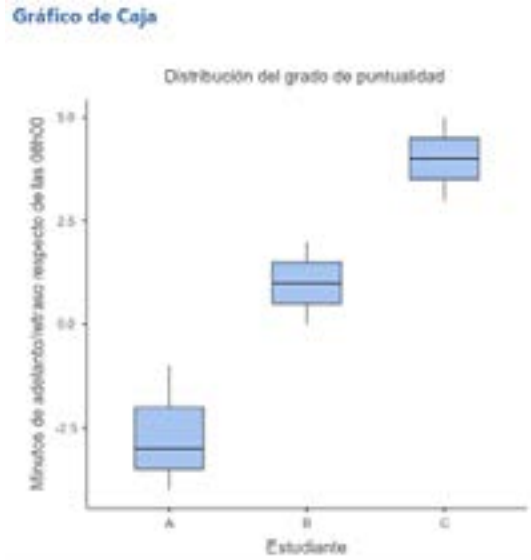
Descriptivas

	Minutos_08h00
N	9
Perdidos	0
Media	0.778
Mediana	1
Desviación típica	3.07
Mínimo	-4
Máximo	5
25percentil	−1.00
50percentil	1.00
75percentil	3.00

Nota. La figura presenta los estadísticos descriptivos correspondientes a los minutos de adelanto o retraso respecto de las 08h00, registrados en una muestra de nueve observaciones sin valores perdidos.

El gráfico de cajas (Figura 2) muestra con claridad las diferencias en el grado de puntualidad de los tres estudiantes, expresado en minutos de adelanto o retraso respecto de las 08h00.

Figura 2.
Distribución del grado de puntualidad de los estudiantes respecto de las 08h00



Nota. La figura representa el grado de puntualidad de tres estudiantes, expresado en minutos de adelanto o retraso con relación a las 08h00.

El estudiante A presenta valores negativos agrupados alrededor de –3 minutos, lo que indica que llega sistemáticamente antes de la hora acordada. El estudiante B muestra un patrón más cercano a la puntualidad, con una mediana ligeramente superior a 1 minuto y una variabilidad moderada. Por su parte, el estudiante C se caracteriza por retrasos consistentes entre 3 y 5 minutos, evidenciando un comportamiento más tardío de forma estable. La dispersión dentro de cada caja es baja, lo que sugiere que las conductas individuales se mantienen relativamente constantes.

Representar, clasificar y modelar la variabilidad: tipos de variables, ejemplos y proyecciones didácticas (versión ampliada con ejemplos)

Una vez que el estudiante reconoce que la variabilidad es natural y que los resultados no son fijos, surge la necesidad de representarla y clasificarla. Aquí es donde la variable aleatoria adopta sus formas más conocidas: discreta y continua.

Variables aleatorias discretas

Son aquellas que toman valores contables. Este tipo de variable aparece continuamente en la vida cotidiana, aunque pocas veces lo notemos. Por ejemplo:

- Número de llamadas telefónicas recibidas en una mañana.
- Cantidad de veces que un ascensor se detiene entre pisos.
- Número de estudiantes que entregan una tarea el mismo día.
- Cuántas frutas defectuosas aparecen en una caja.

DeVeaux, Velleman y Bock (2019) explican que estas variables son ideales para introducir la relación entre frecuencia y probabilidad, porque permiten observar variación sin perder precisión en el conteo.

Variables aleatorias continuas

Modelan fenómenos donde los valores no se cuentan, sino que se miden. Por ejemplo:

- Tiempo que tarda en descargarse un archivo.
- Cantidad de lluvia acumulada en un día.
- Nivel de ruido en decibelios en una avenida céntrica.
- Frecuencia cardíaca en reposo a lo largo de varios minutos.

Hastie et al. (2009) señalan que estos fenómenos requieren funciones de densidad y gráficas que distribuyen probabilidades de manera fluida. Por ello, las variables continuas constituyen la base de muchos modelos en ingeniería, medicina, meteorología y ciencias sociales.

Para reforzar el sentido del concepto de variable aleatoria, es útil recurrir a situaciones cotidianas que permitan a los estudiantes reconocer cómo la incertidumbre se manifiesta en fenómenos reales. Por ejemplo, la temperatura del café de la mañana nunca es exactamente igual, lo que la convierte en un caso típico de variable continua: puede ser 78.4 °C, 79.1 °C o 78.9 °C, entre muchos otros valores posibles.

Algo distinto ocurre con el número de reproducciones de un video en redes sociales, que cambia conforme a la actividad de los usuarios y solo puede tomar valores enteros, por lo que constituye una variable discreta. También el tiempo que tarda en llegar un taxi solicitado por aplicación fluctúa de manera natural, aunque la plataforma anuncie "4 minutos": en la práctica puede ser 3.7, 4.2 o 4.9 minutos, evidenciando nuevamente un comportamiento continuo. Por último, contar cuántas personas pasan por la entrada de una tienda cada media hora implica trabajar con cantidades enteras y, por tanto, con una variable discreta.

Además, comprender cómo una variable aleatoria se vincula con su función de probabilidad permite reconocer que los fenómenos no solo producen resultados diversos, sino que también lo hacen con diferentes grados de frecuencia. Algunos valores aparecen de manera más habitual porque responden a regularidades propias del proceso, mientras que otros son menos probables, aunque posibles.

Este reconocimiento, que muchas veces no surge de manera intuitiva, es fundamental para interpretar adecuadamente cualquier distribución y para desarrollar habilidades críticas al analizar informes, simulaciones o representaciones gráficas. Tal como enfatiza James (2021), la función de probabilidad no es

solo una herramienta matemática, sino una forma de narrar por qué ciertos resultados tienen mayor peso en el comportamiento del sistema.

Desde esta perspectiva, la variabilidad deja de ser un obstáculo y pasa a convertirse en una fuente de información valiosa para explicar el fenómeno observado. En correspondencia con lo planteado por Wild y Pfannkuch (1999), reconocer la estructura probabilística de los datos permite que las decisiones dentro del ciclo se sostengan en criterios sólidos, potenciando un razonamiento estadístico más profundo y fundamentado.

Integrar esta comprensión también fortalece las fases de planificación y análisis, pues la función de probabilidad ayuda a justificar por qué se selecciona una determinada variable aleatoria y cómo deben interpretarse sus patrones de variación.En este sentido Wild y Pfannkuch (1999), plantean que el grupo toma decisiones clave:

- Problema: ¿qué tan predecible es el tiempo de entrega?
- Plan: ¿qué variables registramos?
- Datos: número de paradas y tiempo total.
- Análisis: comparación, visualización, identificación de patrones.
- Conclusión: existe variabilidad, pero no caos; hay un intervalo probable y factores que explican las desviaciones.

Esta toma de decisiones es reconocida como ciclo PPDAC. Un ejemplo que el docente puede concebir puede tener como objetivo en comprender cómo un mismo fenómeno puede involucrar variables aleatorias discretas y continuas, analizar su variabilidad y vincularlo con las etapas del ciclo PPDAC.

Situación: "Imagina que pides comida o un servicio de mensajería por una aplicación móvil. A veces llega rápido, otras veces tarda un poco más, aunque parezca que siempre recorre la misma ruta. ¿Por qué ocurre esto? ¿Qué características del recorrido podrían explicar la variación?"

El grupo de docentes registró información de 15 pedidos reales en diferentes días y horarios (Tabla 2). Se observó:

- Número de paradas del repartidor (semáforos, intersecciones, congestión). Variable aleatoria discreta (valores: 3, 5, 4, 6, 7...).
- Tiempo total de entrega (minutos y segundos desde la confirmación hasta la recepción). Variable aleatoria continua (valores entre 14.3 y 23.8 minutos).

Tabla 2.
Registro de pedidos y características del trayecto en distintos horarios del día

Pedido	Paradas	Tiempo_min
1	3	14.3
2	5	17.8
3	4	19.1
4	6	21.4
5	7	23.8
6	4	18.2
7	5	17.5
8	6	19.9
9	3	15.0
10	7	22.3
11	5	18.7
12	4	16.9
13	6	20.5
14	5	19.0
15	7	21.8

Nota. Los datos corresponden a 15 pedidos reales registrados en distintos días y horarios. La variable Paradas representa un conteo discreto del número de detenciones del repartidor durante el trayecto, mientras que Tiempo_min indica el tiempo total de entrega en minutos, medido como una variable continua.

Los estadísticos descriptivos (Figura 3) revelan que el tiempo típico de entrega se sitúa alrededor de los 19 minutos, dado que la media y la mediana prácticamente coinciden. Esta cercanía sugiere que no existen valores extremos que distorsionen la distribución.

El histograma del tiempo de entrega (Figura 4) evidencia que el proceso no es completamente regular, pero tampoco caótico. La mayoría de los pedidos se concentra en un intervalo central cercano a los 18–20 minutos, lo que sugiere un "tiempo típico" de entrega alrededor de ese rango. Los valores más cortos (cercanos a 14–15 minutos) y los más largos (por encima de 22 minutos) aparecen con menor frecuencia, lo que indica que representan situaciones menos habituales, posiblemente asociadas a condiciones de tráfico poco usuales, cambios en la demanda o particularidades de la ruta.

Figura 3.
Estadísticos descriptivos del tiempo total de entrega (en minutos)

Descriptivas

Descriptivas	
	Tiempo_min
N	15
Media	19.1
Mediana	19.0
Mínimo	14.3
Máximo	23.8

Nota. La tabla presenta los principales estadísticos descriptivos del tiempo total de entrega registrado en 15 pedidos realizados mediante aplicación móvil.

La diferencia entre el tiempo mínimo (14.3 minutos) y el máximo (23.8 minutos) evidencia que, aunque el fenómeno es variable, las entregas se mantienen dentro de un intervalo relativamente acotado. Estos resultados permiten anticipar patrones que luego pueden relacionarse con la función de densidad y con la interpretación probabilística del fenómeno.

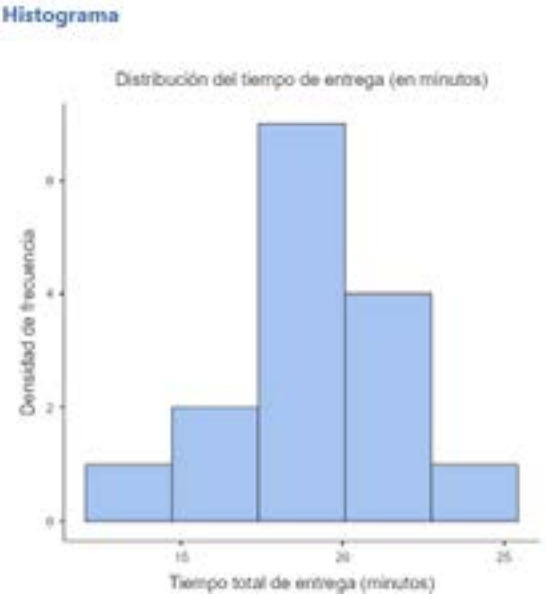
El histograma del tiempo de entrega (Figura 4) evidencia que el proceso no es completamente regular, pero tampoco caótico. La mayoría de los pedidos se concentra en un intervalo central cercano a los 18–20 minutos, lo que sugiere un "tiempo típico" de entrega alrededor de ese rango. Los valores más cortos (cercanos a 14–15 minutos) y los más largos (por encima de 22 minutos) aparecen con menor frecuencia, lo que indica que representan situaciones menos habituales, posiblemente asociadas a condiciones de tráfico poco usuales, cambios en la demanda o particularidades de la ruta.

En conjunto, la gráfica muestra una distribución unimodal con ligera asimetría hacia la derecha, consistente con la idea de que, en algunos casos, el pedido puede demorarse más de lo esperado, pero es poco común que llegue extremadamente rápido.

Este histograma es muy útil para que los estudiantes descubran que la variabilidad del tiempo de entrega no es un "defecto del sistema", sino una característica natural del fenómeno. A partir de la gráfica pueden discutir preguntas como: ¿qué intervalo de tiempos consideraríamos razonable para un pedido?, ¿qué tan raro es obtener un tiempo muy corto o muy largo?,

¿qué factores del contexto podrían mover la distribución hacia la derecha o hacia la izquierda (hora pico, lluvia, saturación de pedidos)? De este modo, el gráfico se convierte en un punto de partida para hablar de variable aleatoria continua, de frecuencia y de probabilidad de manera cercana a su experiencia cotidiana.

Figura 4.
Histograma del tiempo total de entrega de pedidos por aplicación



Nota. El gráfico muestra la distribución del tiempo total de entrega (en minutos) de 15 pedidos realizados mediante una aplicación móvil.

Errores comunes y dificultades conceptuales (versión ampliada con ejemplos cotidianos)

A pesar de que la variabilidad forma parte de la vida cotidiana, muchos estudiantes tienen dificultades para reconocer su papel dentro del análisis estadístico. Estas dificultades surgen de ideas intuitivas pero equivocadas que es necesario problematizar desde la didáctica. Entre las más comunes se encuentran las siguientes:

a) Interpretar la variabilidad como un error

Es frecuente que el estudiante piense que, si los datos cambian, es porque "algo salió mal". Sin embargo, la variación aparece incluso en situaciones totalmente habituales: el tiempo que tarda en hervir el agua cambia, aunque se utilice la misma olla, la duración efectiva de una canción varía por milésimas al reproducirla varias veces y un dispositivo móvil nunca consume exactamente la misma batería en dos días similares. Superar esta creencia permite comprender que la variabilidad no es una falla, sino un rasgo natural del fenómeno que merece ser estudiado.

b) Suponer que todos los resultados son equiprobables

Esta idea suele provenir de ejemplos escolares idealizados, como lanzar una moneda o extraer fichas al azar. En la vida real, la equiprobabilidad es excepcional: algunas rutas de entrega son sistemáticamente más rápidas que otras, ciertas horas concentran mayor tráfico vehicular y los usuarios de redes sociales tienden a ser más activos en momentos específicos del día. Reconocer estas diferencias ayuda a entender que, en la mayoría de los fenómenos, no todos los resultados tienen la misma probabilidad de ocurrir.

c) Confundir la variable aleatoria con sus valores

Muchos estudiantes se quedan únicamente con el dato 3 llamadas, 7 mensajes, 5 minutos, y pierden de vista que la variable aleatoria representa el fenómeno completo. Los valores son solo manifestaciones puntuales de ese fenómeno, no la variable en sí misma. Esta distinción es clave para avanzar hacia una comprensión más conceptual.

d) Pensar que toda variable aleatoria debe ser discreta

La predominancia de ejercicios escolares centrados en conteos refuerza esta idea. No obstante, en la vida diaria abundan mediciones continuas: el nivel de glucosa en sangre, la velocidad de conexión a internet, la temperatura ambiental o la intensidad del sonido. Mostrar estos ejemplos ayuda a ampliar el repertorio mental del estudiante y a reconocer la diversidad de variables presentes en su entorno.

e) Asociar variabilidad con caos absoluto

Como señala Watson (2006), muchos estudiantes interpretan la variabilidad como desorden o falta de estructura. En realidad, una variable aleatoria permite identificar patrones globales incluso cuando los valores fluctúan. La variación no implica caos, sino una organización que puede describirse, modelarse y comprenderse

f) Desconectar la variable del fenómeno real

Otro obstáculo frecuente es mencionar una variable sin justificar su pertinencia o sin relacionarla con el fenómeno que se desea estudiar. Esto ocurre cuando no se reflexiona sobre preguntas esenciales: ¿tiene sentido medir el número de carros por minuto o la velocidad promedio?, ¿qué representa realmente el "tiempo de espera"?, ¿los valores posibles son razonables para el fenómeno?

Como advierten Biehler (2018) y Pfannkuch (2019), esta desconexión empobrece el proceso de modelación y limita la capacidad del estudiante para interpretar adecuadamente los datos.

De manera general la variable aleatoria no es una fórmula ni un tecnicismo: es una manera de comprender el mundo tal como es, con toda su variación natural. Permite dar sentido a fenómenos cotidianos, formar modelos, comparar comportamientos y construir las distribuciones que se estudian en los epígrafes siguientes. Su comprensión profunda depende de unir experiencia, representación y modelación; de ahí que los ejemplos cotidianos sean una herramienta esencial para hacer visible la estructura del azar en la vida real.

Función de probabilidad y función de densidad: interpretación didáctica

La comprensión de una variable aleatoria alcanza una nueva profundidad cuando los estudiantes descubren que los valores observados no solo cambian, sino que lo hacen siguiendo patrones que pueden describirse, representarse y modelarse. En este punto, la función de probabilidad y la función de densidad se convierten en herramientas esenciales para entender cómo se distribuye la variación, pues permiten reconocer que algunos resultados aparecen con mayor frecuencia que otros. Como afirma Batanero (2001), una enseñanza significativa de estos conceptos debe partir de las experiencias de variación vividas por el propio estudiante y no únicamente de definiciones formales. En este sentido, la función de probabilidad y la función de densidad no son objetos matemáticos descontextualizados, sino formas organizadas de expresar la estructura del fenómeno aleatorio. Ambas funcionan como una especie de mapa conceptual que ofrece información sobre dónde se concentra la mayor parte de los resultados, cuáles son poco frecuentes y cómo se distribuye el comportamiento global del sistema.

La función de probabilidad como narrativa de la variación discreta.

Cuando se trabaja con variables discretas, la función de probabilidad asigna a cada valor posible un número que representa su probabilidad de ocurrencia. Sin embargo, tal como señala Moore (2010), este concepto se comprende mejor si se construye a partir de la inspección de frecuencias reales antes que desde símbolos matemáticos. Observar, por ejemplo, el número de paradas que realiza un repartidor en distintos pedidos permite descubrir que ciertos valores, como 5 o 6, tienden a ser más frecuentes que otros.

Esta observación inicial abre la puerta a una interpretación más estructurada del fenómeno: la distribución de probabilidades no nace de una fórmula, sino del análisis de la variación observada. Borovcnik (2016) insiste en que este paso es fundamental, pues ayuda al estudiante a entender que la probabilidad no se reparte de manera uniforme, sino que refleja diferencias reales en el comportamiento del fenómeno.

En un ambiente de aula, un docente puede pedir a los estudiantes que organicen los datos en una tabla de frecuencias y luego construyan un gráfico de barras. A partir de ello, las discusiones emergen de manera natural:

- ¿por qué algunas situaciones son más probables que otras?,
- ¿qué condiciones del contexto podrían explicar las diferencias?,
- ¿por qué es razonable que la probabilidad de observar una sola parada sea menor que la de observar cinco?

La función de probabilidad se convierte, entonces, en una manera de narrar la historia del fenómeno discreto, de distinguir entre resultados comunes y raros, y de reflexionar sobre qué factores podrían modificar la distribución. Desde la perspectiva de Ben-Zvi (2000), estas discusiones no solo fortalecen la comprensión matemática, sino también la capacidad del estudiante para vincular los modelos con su propia experiencia cotidiana.

Por ejemplo, un docente propone trabajar con una situación muy cercana a la realidad cotidiana de los estudiantes: el uso del teléfono celular a lo largo del día. Para ello, se plantea analizar cómo dos aspectos del mismo fenómeno: el tiempo que cada persona pasa usando su dispositivo y la cantidad de notificaciones que recibe, pueden comportarse de manera distinta desde el punto de vista estadístico.

Durante una semana, un grupo de 10 estudiantes registró de forma voluntaria y sistemática dos tipos de datos por día:

1. El tiempo total de uso del celular, medido en horas con decimales. Este registro varía de manera continua a lo largo del día, ya que el uso no ocurre en unidades exactas, sino en intervalos que pueden tomar cualquier valor dentro de un rango.
2. El número de notificaciones recibidas, considerando mensajes, alertas de aplicaciones, redes sociales y recordatorios.Este conteo solo puede tomar valores enteros, lo que lo convierte en un ejemplo típico de variable aleatoria discreta.

El objetivo del docente es que los estudiantes descubran cómo un mismo contexto puede contener variables de naturaleza distinta, y cómo esta diferencia influye en la manera

en que se organizan, se representan y se interpretan los datos.Los datos propuestos por el docente son los siguientes (Tabla 3):

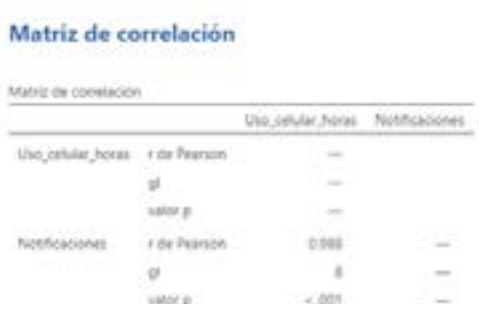
Tabla 3.
Tiempo de uso del celular y número de notificaciones registradas por los estudiantes

Estudiante	Uso_celular_horas	Notificaciones
1	3.2	45
2	4.8	62
3	2.1	30
4	5.4	75
5	3.9	58
6	6.2	81
7	4.1	50
8	5.8	77
9	2.7	36
10	6.5	90

Nota. Los datos corresponden a un registro semanal realizado por 10 estudiantes, quienes anotaron su tiempo total de uso del teléfono celular (en horas) y el número de notificaciones recibidas durante un día típico.

La matriz de correlación (Figura 5) muestra que existe una relación lineal muy fuerte y positiva entre el uso del celular en horas y la cantidad de notificaciones recibidas. El coeficiente de Pearson es r = 0.988, lo cual indica que, a medida que aumenta el tiempo que un estudiante utiliza su celular, también tiende a aumentar el número de notificaciones que recibe.

Figura 5.
Correlación entre las horas de uso del celular y la cantidad de notificaciones recibidas



Nota. La matriz presenta el coeficiente de correlación de Pearson entre el tiempo de uso del celular (en horas) y el número de notificaciones recibidas por cada estudiante.

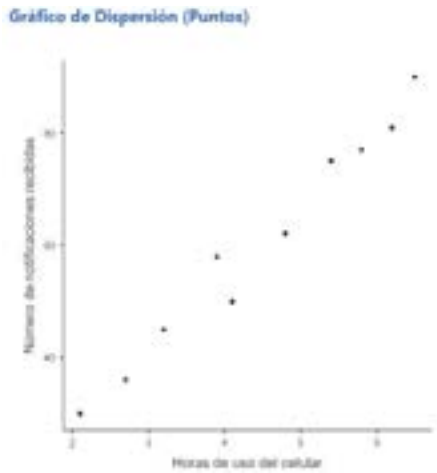
Además, el valor de significación ($p < .001$) confirma que esta relación no es producto del azar, sino que es estadísticamente significativa incluso con un tamaño de muestra reducido ($gl = 8$). En términos prácticos, este patrón sugiere que los estudiantes con mayor exposición al celular están más expuestos a recibir notificaciones, lo cual coincide con comportamientos habituales de uso de redes sociales, mensajería y aplicaciones interactivas.

El gráfico de dispersión (Figura 6) muestra la relación entre el tiempo de uso del celular (en horas) y el número de notificaciones recibidas por cada estudiante. Visualmente, los puntos siguen un patrón ascendente: a medida que aumenta el tiempo de uso, también aumenta la cantidad de notificaciones.

Este comportamiento indica una relación lineal positiva fuerte entre ambas variables. Es decir, los estudiantes que utilizan el celular durante más horas tienden a recibir un mayor número de notificaciones. No se observan valores atípicos ni casos que se alejen del patrón general, lo que sugiere que la relación es consistente en todo el conjunto de datos.

En términos prácticos, esta relación tiene sentido: un mayor tiempo activo en el dispositivo implica mayor interacción con redes sociales, mensajería y aplicaciones que generan notificaciones, por lo que la tendencia ascendente es coherente con el comportamiento real del uso del celular.

Figura 6.
Dispersión entre las horas de uso del dispositivo móvil y la cantidad de notificaciones recibidas



Nota. El gráfico muestra la relación entre las horas de uso del celular y el número de notificaciones recibidas por los estudiantes.

La regresión lineal simple (Figura 7) predice el número de notificaciones recibidas por los estudiantes. El modelo resultó significativo, $R^2 = 0.976$, lo que indica que el 97.6 % de la variabilidad en las notificaciones se explica mediante el tiempo de uso del dispositivo.

El coeficiente de regresión para el predictor fue estadísticamente significativo, $b = 13.15,\ t(8) = 18.12,\ p < 0.001$. Esto implica que, por cada hora adicional de uso del celular, se esperan aproximadamente 13 notificaciones más, en promedio. El intercepto no fue significativo, b = 1.63, p = .646, lo que sugiere que su valor carece de interpretación sustantiva para este contexto.

En conjunto, los resultados muestran una relación lineal fuerte y positiva entre el tiempo de uso del celular y la cantidad de notificaciones recibidas.

Figura 7.
Modelo de regresión lineal entre el uso del celular (horas) y el número de notificaciones recibidas

Regresión Lineal

Medidas de Ajuste del Modelo

Modelo	R	R^2
1	0.988	0.976

Nota. Models estimated using sample size of N=10

Coeficientes del Modelo - Notificaciones

Predictor	Estimador	EE	t	p
Constante	1.63	3.407	0.478	0.646
Uso_celular_horas	13.15	0.726	18.120	< .001

Nota. El modelo de regresión lineal predice la cantidad de notificaciones recibidas a partir de las horas de uso del celular.

Apoyo didáctico: Al finalizar este análisis, es importante que el docente reconozca que trabajar con datos reales no solo fortalece la comprensión estadística, sino que también abre oportunidades para que los estudiantes desarrollen un pensamiento crítico sobre los fenómenos que los rodean.

Desde la mirada de Wild y Pfannkuch (1999), cada situación que se modela en el aula debe invitar a que el estudiante tome decisiones, compare alternativas y se pregunte por las condiciones que generan la variabilidad observada.

De igual manera, la perspectiva de Ben-Zvi (2000) recuerda que la construcción de significado en estadística ocurre cuando los estudiantes logran vincular los modelos con sus propias experiencias, y no cuando memorizar fórmulas es el centro de la actividad. Por ello, se sugiere diseñar tareas donde la exploración de datos, la construcción de gráficos y la interpretación conjunta permitan transformar la incertidumbre en explicaciones razonadas y conscientes. Finalmente, siguiendo a Biehler (2018), orientar la discusión hacia la relación entre contexto y datos favorece que los estudiantes entiendan la estadística como una herramienta para comprender el mundo, y no como una colección de procedimientos aislados. Así, la enseñanza se vuelve más significativa, situada y coherente con los desafíos de la alfabetización estadística contemporánea.

La función de densidad: dar forma gráfica a la variación continua

Cuando el fenómeno que se estudia produce valores que pueden cambiar de manera suave y sin saltos, la variable aleatoria continua se convierte en la herramienta conceptual adecuada para describirlo. En estos casos, la probabilidad ya no se atribuye a valores puntuales, sino a intervalos, y es precisamente aquí donde la función de densidad adquiere sentido. Tal como señalan Hastie et al. (2009), la densidad funciona como una curva que capta la intensidad con la que ciertos valores tienden a aparecer, mostrando la estructura subyacente del fenómeno sin perder la riqueza propia de la variabilidad continua.

Desde un punto de vista didáctico, esta noción suele resultar abstracta al inicio. Por ello, es útil partir de representaciones empíricas, como los histogramas construidos a partir de datos reales. A medida que se acumulan observaciones (los tiempos de descarga de un archivo, los cambios en la temperatura del aula a lo largo de la mañana o la velocidad de conexión durante una videollamada) las barras del histograma permiten ver zonas donde los valores se agrupan con mayor frecuencia. Sin embargo, esta representación, aunque informativa, suele ser rugosa. La función de densidad suavizada, en cambio, transforma esa irregularidad en una curva continua que revela patrones más nítidos. James (2021) enfatiza que esta representación no pretende identificar un valor exacto, sino ayudar a comprender la tendencia general del fenómeno y, sobre todo, a reconocer qué intervalos concentran mayor probabilidad.

Para ilustrarlo, consideremos la fluctuación de la velocidad de conexión a internet durante una videollamada, un fenómeno cotidiano para muchos estudiantes.

Si se registra la velocidad cada poco segundo, aparecen valores como 9.8 Mbps, 11.2 Mbps, 10.7 Mbps, 12.5 Mbps o 8.9 Mbps, que varían sin saltos abruptos. Al graficar estos datos en un histograma, es posible observar que la mayor parte se concentra entre 10 y 12 Mbps, mientras que las velocidades muy bajas o excepcionalmente altas aparecen con menor regularidad. La densidad suavizada construida a partir de estos datos da forma a esa variación: muestra un "pico" alrededor del intervalo donde la conexión suele estabilizarse y una caída progresiva hacia los extremos menos frecuentes, capturando de manera visual lo que los números por sí solos no revelan.

Tabla 4.
Velocidad de conexión a Internet medida en Mbps en 20 casos observados

Caso	Velocidad_Mbps
1	9.8
2	11.2
3	10.7
4	12.5
5	8.9
6	10.3
7	11.0
8	10.9
9	11.5
10	9.7
11	10.8
12	12.0
13	11.7
14	9.9
15	10.5
16	11.3
17	10.1
18	12.2
19	9.5
20	11.1

Nota. Cada registro representa la velocidad promedio de descarga (en Mbps) obtenida en un caso individual mediante una prueba básica de conectividad.

En este ejemplo el docente debe hacer presentar la situación con determinadas reflexiones tales como:

- "Para esta actividad vamos a trabajar con un conjunto de datos que simula la medición real de la velocidad de conexión a internet. Imaginemos que estamos en un laboratorio de informática y queremos analizar qué tan estable es la velocidad de nuestra red durante un minuto. Para ello, realizamos una medición automática cada poco segundo
- Como suele ocurrir en cualquier red doméstica o institucional, la velocidad no se mantiene fija. A veces aumenta ligeramente, otras veces disminuye, y de vez en cuando presenta picos o caídas poco frecuentes. Eso es justamente lo que queremos estudiar: la variación natural en un fenómeno continuo.
- Los 20 valores que vamos a analizar (Tabla 4) representan velocidades registradas en intervalos muy cortos de tiempo. Algunos valores aparecen alrededor de los 10 u 11 Mbps, que es donde la conexión tiende a estabilizarse. Otros valores se alejan un poco más, como 8.9 Mbps o 12.5 Mbps, que corresponden a momentos donde la red está más cargada o más libre. Ninguno de estos cambios es abrupto; simplemente forman parte de la variabilidad propia del sistema.
- Quiero que observen estos datos como lo haría un analista: no se trata solo de mirar números, sino de preguntarse cómo se comporta la velocidad en general. ¿En qué rango se concentra la mayor parte de las mediciones? ¿Qué tan frecuentes son las velocidades extremas? ¿Cómo se vería esta información en un histograma y en una curva de densidad? Este análisis nos permitirá comprender mejor cómo funciona una variable aleatoria continua y cómo la función de densidad nos ayuda a interpretar su comportamiento global."

Una vez ingresados los datos en Jamovi y generados los estadísticos descriptivos (Figura 8) podemos observar que la velocidad de conexión presenta una variación natural pero claramente estructurada. La media de 10.8 Mbps y la mediana de 10.9 Mbps indican que la mayor parte de las mediciones se concentran alrededor de este valor central, lo que sugiere un comportamiento relativamente estable del sistema. La desviación típica, cercana a 0.95 Mbps, confirma que la variabilidad es moderada: las oscilaciones existen, pero no son abruptas.

Figura 8.
Estadísticos descriptivos de la velocidad de conexión medida en intervalos de pocos segundos

Descriptivas

	Velocidad_Mbps
N	20
Media	10.8
Mediana	10.9
Desviación típica	0.954
Mínimo	8.90
Máximo	12.5

Nota. La figura muestra los estadísticos descriptivos correspondientes a 20 mediciones reales de velocidad de conexión (en Mbps) tomadas en intervalos cortos de tiempo.

El histograma (Figura 9) muestra de manera clara cómo varía la velocidad de conexión a internet cuando se realizan mediciones sucesivas en intervalos muy cortos. Lo primero que se observa es que los valores no son fijos ni completamente regulares: se mueven entre aproximadamente 9 y 12.5 Mbps, reflejando la variabilidad natural propia de un fenómeno continuo. Sin embargo, esta variación no es caótica.

La curva de densidad suavizada, superpuesta sobre las barras, ayuda a visualizar el patrón subyacente: existe una concentración evidente alrededor de los 10 y 11 Mbps, donde se registra la mayor frecuencia de observaciones. Esto sugiere que, aunque la velocidad fluctúa, tiende a estabilizarse sobre un rango típico.

Figura 9.
Distribución de la velocidad de conexión (Mbps)

Nota. El histograma muestra 20 mediciones de velocidad de conexión a internet tomadas en intervalos cortos.

El valor pedagógico de esta representación está en que ofrece a los estudiantes un puente entre la experiencia y la abstracción. La curva de densidad permite plantear preguntas que movilizan el razonamiento probabilístico: ¿qué condiciones explican que la mayor densidad esté concentrada en ciertos rangos?, ¿por qué ocurren las caídas de velocidad?, ¿cómo podría modificarse la curva en momentos de alta congestión digital? De acuerdo con Hastie et al. (2009), interpretar la densidad implica reconocer que la variabilidad no es caótica, sino estructurada; y como sugiere James (2021), esa estructura solo se vuelve evidente cuando se observa cómo la probabilidad se distribuye a lo largo de un continuo.

Interpretación conjunta: del dato al modelo y del modelo al fenómeno

Aunque la función de probabilidad y la función de densidad corresponden a naturalezas distintas, ambas comparten un propósito esencial: organizar y explicar la variabilidad del fenómeno aleatorio. Desde una perspectiva conceptual, estas dos funciones no deberían enseñarse como compartimentos aislados, sino como herramientas que permiten observar un mismo proceso desde dos ángulos complementarios. Wild y Pfannkuch (1999),

al presentar el ciclo PPDAC, recuerdan que la fase de Análisis exige interpretar patrones, identificar tendencias y vincular los resultados con el contexto; es decir, pasar del dato al modelo y del modelo de vuelta al fenómeno. Desde esta mirada, tanto la función de probabilidad como la función de densidad ayudan a responder preguntas centrales del razonamiento estadístico: ¿qué valores son habituales?, ¿cuáles resultan inusuales?, ¿qué forma tiene la distribución?, ¿qué revela esa forma sobre la estructura interna del fenómeno?

Para ilustrar esta integración puede utilizarse un ejemplo cercano al entorno escolar: el proceso de puntuación de un videojuego educativo. En muchos juegos, el puntaje por completar determinados retos se expresa mediante números enteros, lo que permite modelar su comportamiento mediante una función de probabilidad. Sin embargo, el tiempo que tarda cada estudiante en completar el nivel, medido en segundos o décimas de segundo, varía de forma continua y requiere una función de densidad para representar su distribución.

Enunciado: En un curso de Matemáticas, el docente incorpora un videojuego educativo sobre fracciones como recurso para fortalecer la comprensión conceptual y el razonamiento de los estudiantes (Tabla 5).

Cada vez que un estudiante completa el nivel 1, el sistema registra dos tipos de información: un puntaje entero asociado a la cantidad de retos superados, que constituye una variable aleatoria discreta, y el tiempo total empleado para completar el nivel, medido en segundos con una cifra decimal, que corresponde a una variable aleatoria continua

El propósito del análisis es examinar cómo se distribuyen los puntajes obtenidos por los estudiantes, cómo se comportan los tiempos registrados y, finalmente, cómo ambos modelos pueden interpretarse en conjunto para describir patrones de desempeño típicos, identificar variabilidad entre estudiantes y comprender mejor el proceso de aprendizaje mediado por el videojuego.

Los resultados descriptivos (Figura10) muestran que el tiempo empleado por los estudiantes para completar el nivel del videojuego es bastante consistente dentro del grupo. La media (48.5 s) y la mediana (47.7 s) son muy similares, lo que indica una distribución equilibrada sin grandes desviaciones. La variabilidad es moderada, con una desviación típica de 4.96 segundos, lo que sugiere que la mayoría de los estudiantes se concentra en un rango de desempeño relativamente estrecho. Los valores mínimos (41.8 s) y máximo (56.4 s) confirman la ausencia de tiempos atípicos, evidenciando que el nivel tiene una dificultad adecuada y que los estudiantes mantienen un ritmo de ejecución similar.

Tabla 5.
Resultados de puntaje y tiempo de resolución del nivel 1 en un videojuego educativo de fracciones

Estudiante	Puntaje	Tiempo_seg
1	30	47.2
2	35	52.8
3	40	44.5
4	30	55.1
5	45	42.7
6	35	49.3
7	40	46.9
8	50	41.8
9	35	53.6
10	45	43.9
11	40	48.2
12	30	56.4

Nota. La tabla presenta los puntajes obtenidos (variable aleatoria discreta) y los tiempos de resolución en segundos con una cifra decimal (variable aleatoria continua) registrados por 12 estudiantes tras completar el nivel 1 del videojuego educativo sobre fracciones.

Figura 10.
Estadísticos descriptivos del tiempo empleado para completar el nivel del videojuego educativo

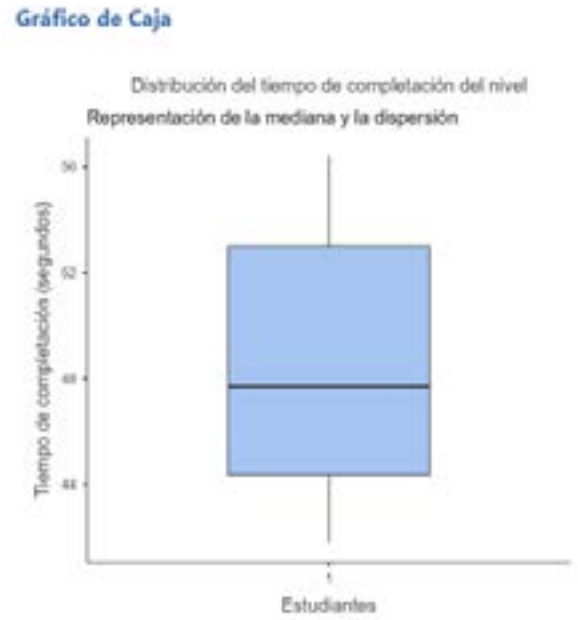
Descriptivas

	Tiempo_seg
N	12
Perdidos	0
Media	48.5
Mediana	47.7
Desviación típica	4.96
Mínimo	41.8
Máximo	56.4

Nota. La figura muestra los valores descriptivos del tiempo total (en segundos) que tardaron los 12 estudiantes en completar el nivel.

Por otra parte, la gráfica (Figura 11) permite observar que, aunque los estudiantes muestran cierta variabilidad en el tiempo que tardan en completar el nivel, la mayoría se concentra en un intervalo relativamente estrecho. La mediana ubicada cerca de los 48 segundos, indica un tiempo típico de ejecución, mientras que la amplitud entre los cuartiles sugiere diferencias moderadas en el ritmo de trabajo de cada estudiante. Los valores más alejados del centro reflejan desempeños más rápidos o más lentos, que pueden deberse a factores como la familiaridad con el videojuego, el nivel de atención o el grado de comprensión de los retos planteados. En conjunto, el diagrama ofrece una visión clara de cómo se distribuyen los tiempos y permite al docente identificar tanto patrones generales como posibles casos que merecen una observación más detallada.

Figura 11.
Distribución de los tiempos empleados por los estudiantes para completar el nivel del videojuego educativo



Nota. El diagrama de caja representa la variabilidad del tiempo (en segundos) que los estudiantes emplearon para completar el nivel.

Cuando ambos modelos se analizan juntos, los estudiantes pueden comprender que el desempeño en el videojuego no es totalmente predecible, pero tampoco arbitrario. Se observan regularidades: ciertos puntajes se repiten porque reflejan niveles de dominio frecuentes, y los tiempos se concentran en un rango porque dependen de la dificultad del nivel, la familiaridad con la mecánica o la concentración del estudiante.

DeVeaux, Velleman y Bock (2019) destacan que esta capacidad para anticipar tendencias sin perder de vista la complejidad del fenómeno es una de las fortalezas del razonamiento probabilístico. Watson (2006) complementa esta idea al señalar que uno de los desafíos centrales en la educación estadística consiste en mostrar que la variabilidad tiene estructura, que el azar no implica caos, sino un patrón que puede interpretarse y explicarse.

En general, la articulación entre las funciones de probabilidad y densidad permite que el estudiante transite de los datos concretos a los modelos abstractos, y de estos, nuevamente al fenómeno que les dio origen. De esta forma, la estadística se convierte en una herramienta para comprender situaciones reales, no solo en una técnica para producir gráficos. Este movimiento entre niveles (dato, modelo y fenómeno) constituye una de las competencias clave en la alfabetización estadística contemporánea y un elemento indispensable para formar ciudadanos capaces de interpretar la incertidumbre de manera crítica y fundamentada.

Distribuciones discretas: Bernoulli, binomial y Poisson

Las distribuciones discretas constituyen un eje fundamental para comprender cómo se organiza la variabilidad cuando los fenómenos se expresan mediante conteos o resultados puntuales. Este tipo de modelos permite que el estudiante avance desde una intuición primaria del azar hacia una interpretación más estructurada del comportamiento probabilístico. Tal como señalan Wild y Pfannkuch (1999), una enseñanza eficaz de la probabilidad requiere que el estudiante aprenda a identificar patrones, explicar tendencias y situar cada resultado dentro del contexto que lo origina. En esta línea, DeVeaux, Velleman y Bock (2019) destacan que las distribuciones discretas son especialmente valiosas porque ofrecen un marco claro y accesible para analizar fenómenos que, aunque variables, conservan una estructura interna reconocible.

A continuación, se desarrollan tres modelos ampliamente utilizados: Bernoulli, binomial y Poisson, cada uno acompañado de un ejemplo contextual y orientaciones didácticas que favorecen el razonamiento probabilístico en el aula.

Distribución de Bernoulli: decisiones binarias y eventos con dos resultados

La distribución de Bernoulli modela fenómenos que solo pueden tener dos resultados mutuamente excluyentes: éxito o fracaso, sí o no, ocurre o no ocurre. Es el modelo discreto más simple: cada ensayo se representa mediante un valor binario (1 o 0). Su

profundidad conceptual reside en que permite introducir la idea de probabilidad como una medida de tendencia de largo plazo más que como una predicción absoluta del próximo resultado.

Cuando un fenómeno se describe mediante una Bernoulli, se asume que:

1. Cada ensayo es independiente: el resultado anterior no influye en el siguiente.
2. La probabilidad de éxito p se mantiene constante.
3. No hay más de dos posibilidades: cualquier matiz o variación se simplifica a "éxito" o "fracaso".

Aunque su estructura es simple, su interpretación didáctica es profunda. Como señala Watson (2006), la comprensión de Bernoulli ayuda a desmontar ideas erróneas comunes, como creer que después de una racha de fracasos "ya es hora" de que llegue un éxito (falacia del apostador). La Bernoulli permite conversar con estudiantes sobre cuándo un fenómeno puede razonablemente considerarse binario y qué implicaciones tiene hacerlo.

Trabajar con Bernoulli posibilita preguntas como:

- ¿Qué factores determinan la probabilidad de éxito?
- ¿Qué significa realmente "éxito" en un fenómeno concreto?
- ¿Qué pasa si la probabilidad cambia a lo largo del tiempo?

Wild y Pfannkuch (1999) sugieren aprovechar estas discusiones para fortalecer la comprensión contextual del fenómeno, ya que la Bernoulli es tanto un modelo matemático como una forma de mirar la realidad. Para estos autores, enseñar una distribución no debería limitarse a presentar su fórmula o a resolver ejercicios mecánicos; por el contrario, es fundamental ayudar a los estudiantes a reconocer que detrás de cada variable dicotómica existe una historia, una situación y un proceso de toma de decisiones. La distribución Bernoulli invita a distinguir entre eventos que ocurren y eventos que no ocurren, pero esta simplicidad aparente encierra una manera poderosa de comprender el mundo: muchas experiencias humanas pueden describirse mediante esta lógica binaria.

Caso de estudio: Reconocimiento de fracciones mediante un videojuego educativo

En una clase de Matemáticas de séptimo año, el docente introduce un videojuego educativo centrado en el reconocimiento de fracciones simples. Cada estudiante debe completar una actividad inicial del juego en la que se presentan diez tarjetas digitales y, en cada una, se evalúa si el estudiante identifica correctamente la fracción representada. El sistema registra para cada estudiante una variable dicotómica llamada Reconoce_frac, que toma el valor de 1 si el estudiante reconoce correctamente la fracción del ejercicio final del nivel, y 0 si no lo logra.

El propósito del docente es analizar el desempeño general del grupo en esta tarea inicial para identificar patrones de logro, determinar la proporción de estudiantes que presentan dificultades y tomar decisiones pedagógicas informadas para los siguientes niveles del videojuego.

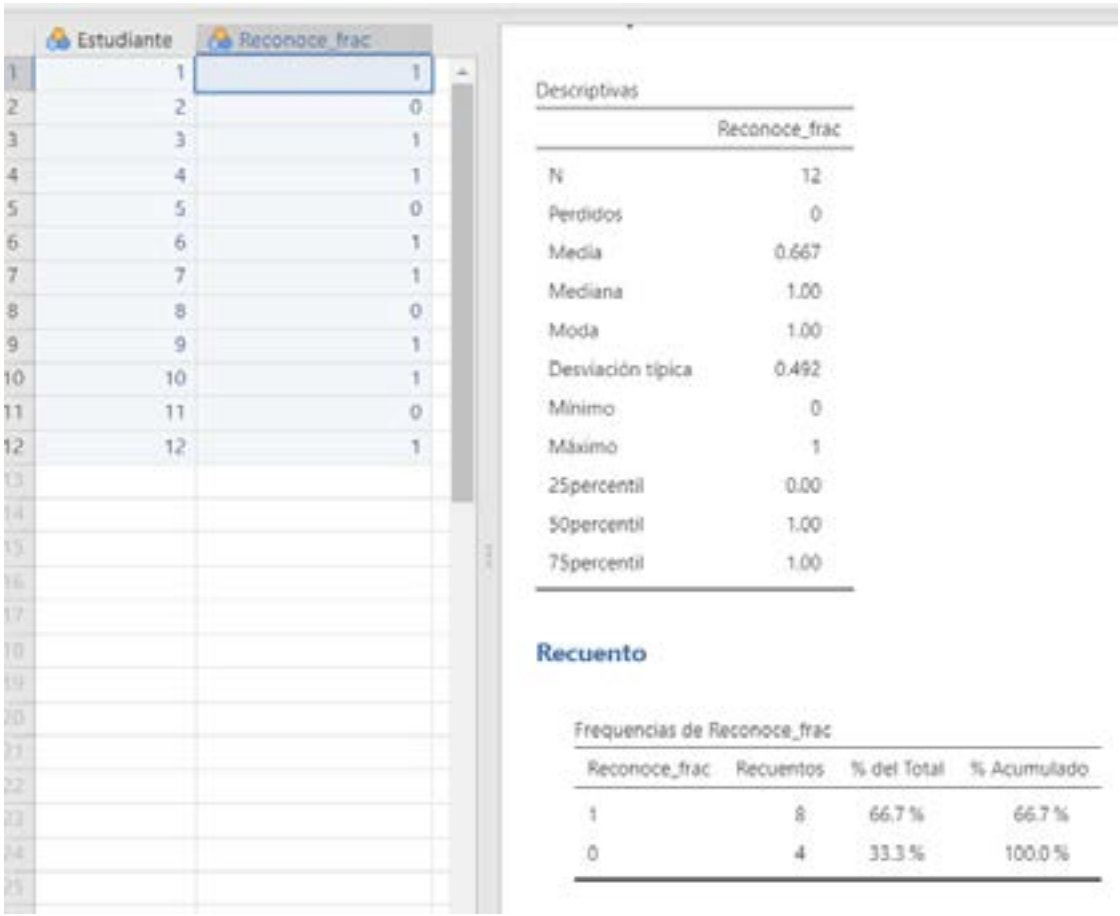
Los resultados descriptivos (Figura 12) muestran que participaron doce estudiantes y no hubo datos perdidos. La media del desempeño es 0.667, lo cual indica que, en promedio, aproximadamente dos tercios del grupo logra reconocer correctamente la fracción presentada en el nivel inicial. La mediana y la moda, ambas iguales a 1, refuerzan esta tendencia: la mayoría de estudiantes tiene un resultado correcto. La desviación típica, de 0.492, sugiere una variabilidad moderada dentro del grupo, coherente con el hecho de que solo existen dos valores posibles (0 y 1) en esta variable.

El análisis de frecuencias aporta una lectura más clara: 8 de 12 estudiantes (66.7%) reconocen correctamente la fracción, mientras que 4 estudiantes (33.3%) no lo logran. Este comportamiento dicotómico usual en variables discretas es fundamental para la toma de decisiones pedagógicas. Por ejemplo, el docente puede deducir que, aunque la mayoría de los estudiantes muestra un dominio adecuado del contenido, existe un grupo significativo (un tercio del curso) que requiere apoyo adicional antes de avanzar a actividades más complejas dentro del videojuego.

La Figura 13 permite visualizar de manera sencilla cómo se distribuyen las respuestas de los estudiantes en la tarea inicial de reconocimiento de fracciones. El gráfico de barras muestra una clara diferencia entre quienes logran identificar correctamente la fracción presentada y quienes no lo consiguen. De los 12 estudiantes evaluados, ocho responden correctamente, mientras que cuatro presentan dificultades, lo que equivale a un 66.7 % y un 33.3 %, respectivamente.

Esta asimetría evidencia que, aunque la mayoría del grupo domina el concepto evaluado, existe un porcentaje importante que requiere apoyo pedagógico adicional antes de avanzar a niveles más complejos del videojuego educativo. La interpretación conjunta del gráfico y de las estadísticas descriptivas sugiere que los estudiantes con errores podrían beneficiarse de experiencias de refuerzo, explicaciones visuales suplementarias o retroalimentación inmediata dentro del propio entorno digital.

Figura 12.
Distribución de estudiantes que reconocen correctamente la fracción en el nivel inicial del videojuego educativo
.



	Estudiante	Reconoce_frac
1	1	1
2	2	0
3	3	1
4	4	1
5	5	0
6	6	1
7	7	1
8	8	0
9	9	1
10	10	1
11	11	0
12	12	1

Descriptivas

	Reconoce_frac
N	12
Perdidos	0
Media	0.667
Mediana	1.00
Moda	1.00
Desviación típica	0.492
Mínimo	0
Máximo	1
25percentil	0.00
50percentil	1.00
75percentil	1.00

Recuento

Frequencias de Reconoce_frac

Reconoce_frac	Recuentos	% del Total	% Acumulado
1	8	66.7 %	66.7 %
0	4	33.3 %	100.0 %

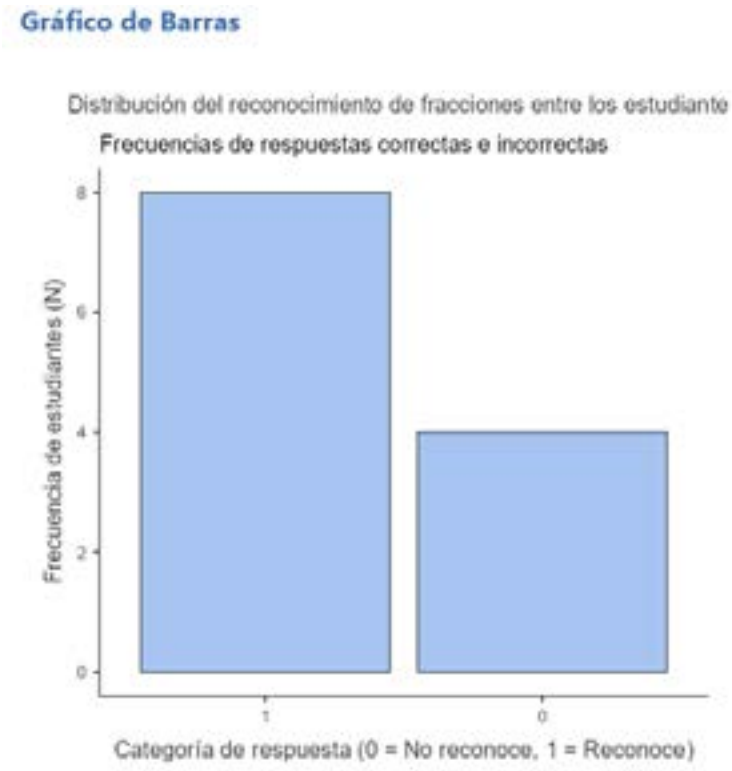
Nota. La figura presenta los resultados de 12 estudiantes en la variable dicotómica Reconoce_frac, donde 1 indica reconocimiento correcto de la fracción y 0 indica error.

Desde el punto de vista didáctico, este caso permite a los estudiantes comprender cómo una variable aleatoria discreta puede analizarse mediante herramientas estadísticas simples, y cómo la interpretación de frecuencias y medidas de tendencia central se utiliza para caracterizar comportamientos de aprendizaje. Además, contextualiza la estadística dentro de un entorno cercano y motivador, como el uso de videojuegos educativos, reforzando el sentido práctico del análisis de datos en situaciones reales del aula.

Wild y Pfannkuch destacan que cuando el docente vincula el modelo con el contexto, se genera una comprensión más rica porque los estudiantes dejan de ver la probabilidad como un cálculo aislado y comienzan a interpretarla como una herramienta para explicar fenómenos reales.

De este modo, la distribución Bernoulli se convierte en un puente entre el razonamiento matemático y la lectura crítica del entorno, permitiendo a los estudiantes identificar patrones, formular hipótesis y analizar situaciones cotidianas desde una perspectiva estadística más profunda.

Figura 13.
Frecuencia de respuestas correctas e incorrectas en el reconocimiento de fracciones



Nota. El gráfico muestra la distribución de respuestas de 12 estudiantes en la tarea de reconocimiento de fracciones. La categoría 1 corresponde a estudiantes que identificaron correctamente la fracción, mientras que la categoría 0 indica respuestas incorrectas.

Distribución binomial: la acumulación de éxitos en múltiples intentos

La binomial extiende la lógica de Bernoulli a un conjunto de n ensayos independientes, todos con la misma probabilidad de éxito. Su esencia radica en describir cuántos éxitos se obtienen en un número fijo de intentos. Este modelo introduce un nuevo nivel de complejidad porque ya no se trata de analizar un solo evento binario, sino la variación acumulada de muchos.

La distribución binomial requiere cuatro elementos:

1. Número fijo de ensayos n: no se detiene hasta completar los intentos.
2. Independencia entre intentos.
3. Probabilidad constante de éxito p.
4. Interés en el conteo de éxitos, no en el orden de aparición.

De acuerdo con DeVeaux, Velleman y Bock (2019), la binomial es una herramienta poderosa para investigar patrones: revela que ciertos resultados son más probables que otros (por ejemplo, obtener 5 aciertos en 10 intentos cuando p=0.5), y que los extremos son menos frecuentes. Esto permite introducir gráficas que muestran cómo se concentra la probabilidad alrededor de valores centrales, favoreciendo la comprensión del comportamiento típico del fenómeno.

La enseñanza de la binomial permite trabajar preguntas como:

- ¿Cuáles son los resultados "esperables"?
- ¿Cómo cambia la distribución si aumenta el número de ensayos?
- ¿Cómo influye un cambio en p sobre la forma de la distribución?
- ¿Por qué los resultados muy altos o muy bajos son menos probables?

Caso de estudio: "La calidad en una planta de producción de botellas"

Una empresa dedicada a la producción de botellas plásticas utiliza un sistema automático de inspección que detecta si cada botella fabricada cumple con los estándares de calidad. Cada botella puede clasificarse como defectuosa (0) o no defectuosa (1). A partir de registros históricos, los ingenieros determinaron que la probabilidad de que una botella salga correctamente fabricada es aproximadamente p = 0,92, ya que en promedio el 8 % presenta algún defecto. Además:

1. En cada lote se inspeccionan exactamente 20 botellas, por lo que el número de ensayos es fijo.
2. La inspección es independiente para cada botella.
3. La probabilidad de que una botella salga en buen estado se mantiene relativamente estable durante un mismo ciclo de producción.
4. El interés está en contar cuántas botellas no defectuosas hay en un lote, no en el orden de inspección.

Estas condiciones permiten modelar el número de botellas correctas en un lote mediante una distribución binomial con parámetros n=20 y p=0,92.

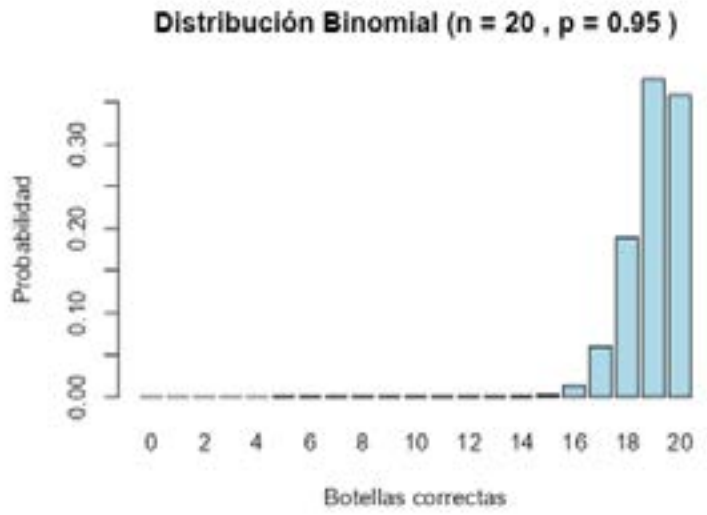
Sea la variable aleatoria: X=número de botellas en buen estado dentro del lote $\mathbf{X}$ ~Entonces: X=Binomial (n=20, p=0,92)

Preguntas que la empresa desea responder

- ¿Cuál es la probabilidad de obtener exactamente 19 botellas correctas en un lote de 20?
- ¿Cuál es la probabilidad de que el lote tenga al menos 18 botellas correctas?
- ¿Qué tan probable es obtener 16 o menos, lo cual implicaría activar una alarma de control de calidad?
- ¿Cuántos defectos son "esperables" en un lote típico?
- ¿Qué valores serían considerados "atípicos" o señal de un problema en la línea de producción?

La distribución mostrada en la Figura 14 permite comprender de manera directa cómo se comporta un proceso productivo cuando la probabilidad de obtener un resultado "correcto" es muy alta. En este caso, el modelo binomial describe el número de botellas correctamente envasadas dentro de un lote de 20 unidades, asumiendo que cada botella tiene una probabilidad de éxito del 95 %.

Figura 14 .
Distribución binomial de botellas correctas en un lote de 20 unidades con probabilidad de éxito p = 0.95.



Nota. La figura presenta la distribución teórica de probabilidades del número de botellas correctas bajo un modelo binomial con n = 20 y p = 0.95.

Al observar la gráfica, se vuelve evidente que los valores más probables se concentran en la parte derecha del eje horizontal: obtener 18, 19 o incluso las 20 botellas correctas no solo es posible, sino que representa la situación más habitual. Estos resultados reflejan un proceso estable, con una línea de producción que comete muy pocos errores y que, por tanto, genera lotes de calidad consistente.

Este tipo de comportamiento es especialmente útil para que los estudiantes visualicen la idea de "resultado esperable", un concepto que suele ser abstracto cuando se explica únicamente desde ecuaciones o fórmulas. Al trabajar con la distribución binomial en un contexto concreto como el control de calidad, se puede mostrar que las probabilidades no son simples números, sino herramientas para anticipar patrones y reconocer cuándo un proceso funciona como debería.

Desde un enfoque pedagógico, la gráfica ofrece una oportunidad para promover discusiones valiosas en el aula. Los estudiantes pueden reflexionar sobre cómo cambia la forma de la distribución si la probabilidad de éxito disminuye, o qué implicaciones tiene aumentar el tamaño del lote. También permite introducir la distinción entre lo "posible" y lo "probable", y comprender por qué ciertos valores casi nunca se observan, aunque no sean imposibles. En resumen, este tipo de representación visual, sostenida por un caso cercano al mundo laboral, ayuda a que la estadística deje de verse como un conjunto de procedimientos mecánicos y se convierta en una herramienta para interpretar fenómenos reales con mayor claridad y sentido.

James (2021) resalta que este tipo de análisis ayuda al estudiante a visualizar la estructura del fenómeno, no solo a memorizar fórmulas. Esta observación es crucial porque, en la enseñanza habitual, la distribución binomial suele reducirse a un conjunto de procedimientos que parecen abstractos y desligados de la experiencia. Sin embargo, cuando se trabaja desde situaciones concretas como acertar preguntas en un cuestionario, registrar éxitos en un experimento o evaluar intentos en un videojuego educativo; la binomial deja de ser una expresión algebraica y se convierte en una herramienta para comprender cómo se acumulan los resultados en escenarios donde el azar interviene repetidamente.

Distribución de Poisson: conteo de eventos en intervalos

La distribución de Poisson se utiliza para modelar eventos que ocurren de forma independiente dentro de un intervalo, ya sea de tiempo, espacio o cualquier unidad continua. A diferencia de Bernoulli y binomial, no existe un número fijo de ensayos: la Poisson describe la frecuencia de aparición de sucesos relativamente raros, regidos por una tasa promedio estable.

Un fenómeno sigue una Poisson cuando:

1. Los eventos son infrecuentes respecto al tamaño del intervalo.
2. Los eventos ocurren uno a la vez, sin simultaneidad.
3. La tasa promedio – es constante.

4. La ocurrencia en un intervalo no afecta otro intervalo (independencia).

Hastie et al. (2009) explican que Poisson es especialmente útil para describir procesos donde no es posible identificar claramente el número de ensayos, pero sí observar un patrón de ocurrencias. Por ello, la Poisson permite captar regularidades invisibles a simple vista: aunque los valores fluctúen día a día, la tasa promedio mantiene una estabilidad sorprendente.

A nivel educativo, trabajar con la Poisson abre la puerta a reflexiones como:

- ¿Qué significa que un fenómeno tenga una "tasa estable"?
- ¿Qué diferencia a la Poisson de la binomial?
- ¿Por qué algunos días ocurren muchos eventos y otros casi ninguno?
- ¿Qué factores del contexto pueden modificar la tasa promedio?

James (2021) enfatiza que este tipo de modelos permite comprender fenómenos aparentemente caóticos desde una perspectiva analítica que revela patrones, no azar puro. Para este autor, la clave no reside solamente en calcular probabilidades, sino en aprender a "leer" el comportamiento de los eventos a través del lente adecuado.

Caso de estudio: La llegada de llamadas a un centro de soporte técnico

En una institución educativa que ofrece soporte tecnológico a docentes y estudiantes, existe un pequeño centro de atención encargado de resolver problemas relacionados con plataformas virtuales, contraseñas, conexión a redes, videoconferencias y accesos a la intranet institucional. Con el crecimiento de la matrícula y la migración hacia entornos digitales, la demanda de soporte ha aumentado en los últimos ciclos académicos. Para mejorar la gestión del servicio, la coordinación decide analizar cuántas llamadas recibe el centro por hora durante los momentos de mayor actividad.

Después de revisar los registros de varias semanas, el equipo observa que, durante las horas pico de la mañana, el número de llamadas que ingresan cada hora fluctúa entre 6 y 14, con un promedio estable cercano a 10 llamadas por hora. Además, las llamadas parecen llegar de manera independiente unas de otras, sin patrones definidos más allá de la intensidad propia de ese horario.

Debido a estas características se propone modelar el número de llamadas por hora mediante una Distribución de Poisson con parámetro $\lambda = 10$ llamadas por hora.

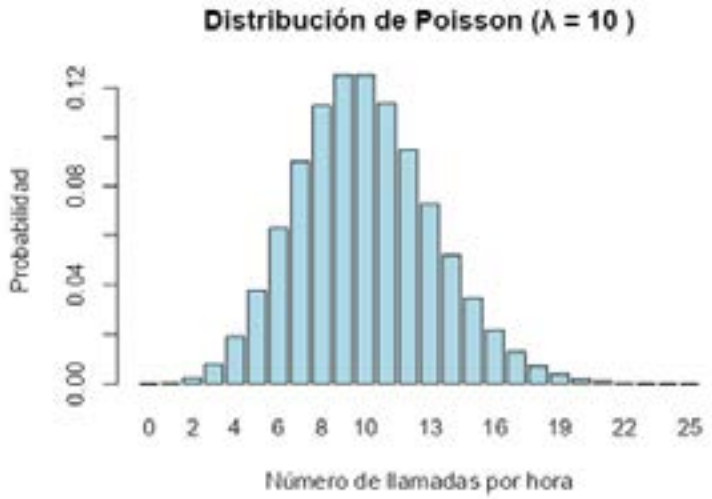
Preguntas que el área de soporte desea responder

1. ¿Cuál es la probabilidad de recibir exactamente 8 llamadas en una hora?
2. ¿Qué tan probable es recibir 15 o más llamadas, lo que podría saturar la capacidad del personal?
3. ¿Qué valor de llamadas por hora puede considerarse típico y cuál sería un valor atípico que debería activar una alerta?
4. ¿Cuál es el número esperado de llamadas por hora y qué tan grande puede ser su variabilidad?
5. ¿Cómo podría cambiar esta distribución si se incrementa el número de usuarios o si una plataforma falla de manera inesperada?

Voy a asumir el modelo que definimos antes: $\mathrm{X} \sim \mathrm{Poisson}(\lambda = 10)$ donde X es el número de llamadas por hora.

La distribución presentada en la figura 15 permite observar cómo se comporta el número de llamadas que llegan por hora a un centro de soporte cuando el proceso sigue un patrón de ocurrencia aleatoria con una tasa promedio estable. El valor de – = 10 indica que, en condiciones normales, se esperan alrededor de diez llamadas por hora. Esto se evidencia en la forma del gráfico: la barra correspondiente a este valor es una de las más altas y constituye el centro de la distribución.

Figura 15 .
Distribución de Poisson del número de llamadas por hora en un centro de soporte ($\lambda = 10$)



Nota. La figura muestra la distribución teórica de probabilidades para el número de llamadas que llegan en una hora, modelada mediante una Poisson con parámetro $\lambda = 10$

De igual manera, se aprecia que los valores inmediatamente cercanos a la media como 8, 9, 11 o 12 llamadas, mantienen probabilidades relativamente altas. Esto sugiere que, aunque el promedio es diez, existe una fluctuación natural que hace completamente razonable observar ligeras variaciones de una hora a otra. Hacia ambos extremos, en cambio, las barras disminuyen de manera pronunciada. Los resultados muy bajos (por ejemplo, menos de 4 llamadas) o muy altos (16 o más) aparecen con probabilidades considerablemente pequeñas, lo que indica que representan situaciones poco habituales.

La forma ligeramente sesgada hacia la derecha, propia de la distribución de Poisson, muestra que, aunque es más probable registrar un volumen de llamadas similar al promedio, siempre existe una pequeña posibilidad de que la demanda del servicio aumente más de lo esperado. Esta "cola" hacia la derecha es especialmente importante desde una perspectiva operativa, ya que permite anticipar episodios de saturación y planificar recursos de manera preventiva. En síntesis, la distribución ofrece una mirada detallada y realista sobre la variabilidad del sistema, permitiendo identificar con claridad qué situaciones deben considerarse normales y cuáles podrían requerir atención o intervención del equipo de soporte.

Apoyo didáctico: El docente puede aprovechar y reflexionar con los estudiantes en relación a los procedimientos matemáticos que fundamentan la interpretación:

1. Probabilidad de recibir exactamente 8 llamadas

Si $\mathbf{X} \sim \mathrm{Poisson}(\lambda)$

Entonces: . $$\mathbf{P}\left(\mathbf{X}=\mathbf{8}\right)=\frac{\mathbf{e}^{-\mathbf{10}}\mathbf{10}^{\mathbf{8}}}{\mathbf{8!}}$$

Para k=8 y $\lambda = 10$

$$\mathbf{P}\left(\mathbf{X}=\mathbf{8}\right)=\frac{\mathbf{e}^{-\mathbf{10}}\mathbf{10}^{\mathbf{8}}}{\mathbf{8!}}$$

2. Probabilidad de recibir 15 o más llamadas

Se busca: $P(X \geq 15)$.
Para calcularlo, se usa la probabilidad complementaria:

$$\mathbf{P(X \geq 15) = 1 - P(X \leq 14)}$$

Dónde: $$\mathbf{P\left(X \leq 14\right) = \sum_{k=0}^{14} \frac{e^{-10}10^{k}}{k!}}$$

Por eso: $$\mathbf{P\left(X \geq 15\right) = 1 - \sum_{k=0}^{14} \frac{e^{-10}10^{k}}{k!}}$$

3. Valores típicos y atípicos
Con el criterio usual (probabilidad < 5 %), usamos:
Para el límite inferior (valores muy bajos):

$$\mathbf{P\left(X \leq 14\right) = \sum_{k=0}^{14} \frac{e^{-10}10^{k}}{k!}}$$

Para el límite superior (valores muy altos):

$$\mathbf{P(X \geq 16) = 1 - P(X \leq 15)}$$

Dónde:

$$\mathbf{P\left(X \leq 15\right) = \sum_{k=0}^{15} \frac{e^{-10}10^{k}}{k!}}$$

Valores con probabilidad menor a 0.05 se consideran atípicos.

4. Valor esperado y variabilidad de la Poisson
Media: $E(X) = \lambda$

Varianza: $Var(X) = \lambda$

Desviación estándar: $\sigma = \sqrt{\lambda}$

Para $\lambda = 10$: $E(X) = 10$ y $\sigma = \sqrt{10} \approx 3.16$

5.Cómo cambia la distribución si aumenta la demanda

La forma de la distribución depende solo de λ:

Si aumenta la demanda (más usuarios entonces más llamadas):

$$\lambda_{\text{nuevo}} > 10$$

- La media aumenta.
- La desviación estándar aumenta.
- La distribución se desplaza hacia la derecha.
- Se vuelve más probable observar valores altos.

No hay fórmula adicional: solo cambia el parámetro λ en las fórmulas anteriores.

En general, modelar un fenómeno mediante distribuciones discretas como Bernoulli, Binomial o Poisson ofrece una vía privilegiada para dotar de sentido a comportamientos que, a primera vista, pueden parecer erráticos o difíciles de anticipar. Estas herramientas permiten revelar patrones estables, variaciones significativas y señales tempranas de cambio que no se perciben de forma intuitiva. Lo que inicialmente se presenta como una colección dispersa de datos adquiere estructura y se convierte en una representación cuantificable de los procesos que operan en el fondo.

El uso de distribuciones discretas como Bernoulli, Binomial o Poisson se convierte en una herramienta clave para que docentes y estudiantes comprendan el comportamiento del azar más allá de la intuición. Estas distribuciones permiten reconocer patrones, estimar variabilidad y anticipar situaciones que, sin un marco probabilístico, parecerían simplemente caprichosas. Como afirma James (2017), el valor pedagógico de la probabilidad radica en mostrar que la incertidumbre puede ser organizada y analizada mediante modelos conceptuales que revelan la estructura del fenómeno. Bajo esta mirada, la enseñanza de la probabilidad deja de centrarse únicamente en cálculos para transformarse en una forma de leer la realidad, interpretarla con mayor nitidez y tomar decisiones didácticas y profesionales más informadas.

Distribuciones continuas: uniforme, normal, exponencial

Las distribuciones continuas constituyen uno de los pilares centrales para el análisis del azar en contextos donde la variable aleatoria puede asumir infinitos valores dentro de un intervalo real. Este tipo de distribuciones resulta fundamental en el modelamiento moderno de datos, especialmente cuando se estudian fenómenos que varían en el tiempo, procesos de medición, puntajes de desempeño, duraciones, rendimientos o fluctuaciones de condiciones físicas. A diferencia de las distribuciones discretas como la Bernoulli, binomial o Poisson; las

distribuciones continuas no cuentan sucesos aislados, sino que describen la densidad con la que la probabilidad se reparte sobre un conjunto ininterrumpido de posibilidades.

De Veaux, Velleman y Bock (2019) destacan que la adopción de distribuciones continuas es esencial porque muchos fenómenos reales no pueden describirse adecuadamente mediante conteos, sino mediante mediciones: un tiempo, una distancia, una temperatura, una calificación continua, un nivel de esfuerzo. En estas situaciones, la naturaleza del fenómeno exige considerar la variación no como saltos discretos, sino como una progresión suave en la que cada punto del intervalo es posible, aunque con distinta probabilidad.

Comprender las distribuciones continuas implica reconocer cómo se concentra la probabilidad en ciertas zonas del intervalo, cómo cambian las formas de una curva y cómo estos modelos permiten identificar regularidades ocultas. Montgomery y Runger (2018) subrayan que el análisis de datos reales requiere, en primera instancia, distinguir si la variable responde a un modelo continuo o discreto, pues de ello dependen los métodos de estimación, inferencia y predicción que pueden aplicarse.

Tres distribuciones continuas ocupan un lugar central en este marco conceptual: la uniforme, la normal y la exponencial. Cada una responde a una lógica distinta de variación y modela una familia de fenómenos característicos. Su estudio permite desarrollar en los estudiantes una comprensión profunda del azar y una mirada crítica hacia los datos observados.

Distribución Uniforme: igualdad de posibilidades en un intervalo

La distribución uniforme continua en el intervalo [a,b] se caracteriza por asignar la misma probabilidad a cualquier subintervalo de igual longitud dentro del rango. Esto significa que la densidad de probabilidad es constante, una idea que Rice (2007) considera fundamental para iniciar el recorrido conceptual hacia modelos continuos más elaborados.

Su función de densidad es: $f\left(x\right) = \frac{1}{b-a},\ a \leq x \leq b$

y su probabilidad acumulada crece de manera lineal:

$$F\left(x\right) = \frac{x-a}{b-a}S$$

A nivel conceptual, la distribución uniforme es clave porque introduce la idea de "densidad constante", un principio que aparece más adelante en modelos de simulación, procesos aleatorios y algoritmos de generación de números pseudoaleatorios.

Caso de estudio: Selección aleatoria de intervalos para pruebas de carga en un servidor educativo

En una universidad que trabaja con una plataforma virtual para educación a distancia, el equipo de ingeniería necesita realizar pruebas de carga para evaluar la resistencia del servidor en diferentes momentos del día. Para evitar introducir sesgos como realizar pruebas solo en horas de baja actividad; se decide que el instante exacto en que se ejecutará cada prueba debe ser completamente aleatorio dentro de un intervalo de diez horas (8:00 a 18:00).

El equipo define que cualquier minuto del intervalo es igualmente probable, sin preferir momentos de mayor uso o descanso del servidor. Bajo este supuesto, el tiempo de inicio de cada prueba puede modelarse mediante una distribución uniforme continua en el intervalo [0,600] minutos.

De acuerdo con Rice (2007), la distribución uniforme es adecuada cuando se parte de un principio de equidad o de ignorancia total sobre qué valor será más probable, y proporciona una base conceptual sólida para introducir la idea de densidad constante en variables continuas. Además, Richard De Veaux, Velleman y Bock (2019) recuerdan que este tipo de modelos ayuda a evitar patrones involuntarios que distorsionan un análisis experimental.

Preguntas que el equipo desea responder

- ¿Cuál es la probabilidad de que una prueba ocurra entre las 11:00 y las 11:30?
- ¿Con qué frecuencia las pruebas podrían coincidir con el horario de clase virtual sin ser programadas así?
- ¿Qué tan dispersos se distribuyen los tiempos de inicio cuando la asignación es realmente uniforme?

La simulación (Figura 16) confirma la idea de que los tiempos de inicio de las pruebas se comportan como una distribución uniforme en [0, 600] minutos (de 8:00 a 18:00). El histograma es prácticamente plano: en todos los tramos del eje horizontal hay frecuencias similares, lo que indica que ningún momento del intervalo es claramente más frecuente que otro. Las descriptivas numéricas apuntan en la misma dirección: la media es ≈ 298.5 minutos y la mediana ≈ 296.7, ambas muy cercanas al centro teórico del intervalo (300 minutos). La desviación típica ≈ 172 coincide con el valor esperado para una uniforme [0, 600], y los valores mínimo y máximo se aproximan bastante a 0 y 600, como cabría esperar en una simulación grande.

Los dos valores de probabilidad que aparecen (≈ 0.0485 y ≈ 0.1976) responden a las preguntas del equipo: la primera es la probabilidad estimada de que una prueba caiga entre las 11:00

y las 11:30 (unos 30 minutos dentro de las 10 horas totales), y la segunda es la probabilidad de que coincida con la franja de clase virtual considerada (aprox. 20 % del intervalo total). En resumen, estos resultados muestran que el procedimiento de asignación aleatoria logra lo que se buscaba: las pruebas se distribuyen de forma equitativa a lo largo del día, sin concentrarse en horas de mayor o menor actividad, y las probabilidades de coincidir con intervalos específicos dependen únicamente de la longitud de esos intervalos, no de un sesgo del sistema.

Figura 16.
Distribución simulada de los tiempos de inicio de las pruebas de carga en un servidor educativo

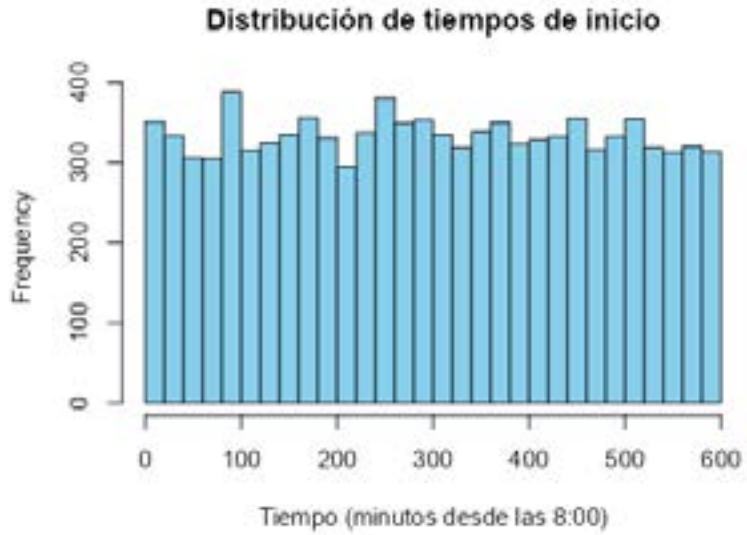
Rj Editor

```
 [1] 172.55 472.98 245.39 529.81 564.28  27.33 316.86 535.45 330.86 273.97

[1] 0.0485

[1] 0.1976

  Media Mediana Desv_Tipica Minimo Maximo
1 298.5   296.7         172 0.0392    600
```



Nota. El histograma muestra la distribución de 10 000 tiempos de inicio generados bajo una distribución uniforme continua en el intervalo de 0 a 600 minutos, equivalente al periodo comprendido entre las 8:00 y las 18:00.

Este ejercicio ofrece una oportunidad valiosa para que los estudiantes comprendan cómo una distribución uniforme permite modelar situaciones donde todas las posibilidades tienen la misma probabilidad. Al simular miles de tiempos de inicio y observar que el histograma se mantiene prácticamente plano, se hace evidente que no existe un "momento privilegiado" dentro del intervalo, lo que ayuda a romper ideas intuitivas pero erróneas sobre la aleatoriedad. Además, al calcular probabilidades asociadas a tramos específicos del día, los estudiantes pueden apreciar que estas dependen únicamente de la longitud del intervalo considerado, y no de supuestos subjetivos sobre el comportamiento del sistema. En resumen, la actividad muestra cómo la simulación en R y jamovi se convierte en una herramienta didáctica poderosa: permite visualizar conceptos abstractos, contrastar expectativas con resultados empíricos y fortalecer la comprensión de la probabilidad como modelo para describir fenómenos reales

La distribución normal: variación alrededor de un centro

La distribución normal es quizá la distribución continua más influyente en estadística y en la enseñanza del análisis de datos. Su carácter central proviene de su capacidad para modelar fenómenos donde los valores tienden a agruparse alrededor de una media, y donde las desviaciones extremas son cada vez menos frecuentes.

Definida por su media μ y desviación estándarσ , su densidad es:

$$\mathbf{f}\left(\mathbf{x}\right) = \frac{\mathbf{1}}{\sqrt{\mathbf{2}\square\square^{\mathbf{2}}}}\mathbf{e}^{-\frac{(\mathbf{x}-\square)^{\mathbf{2}}}{\mathbf{2}\square^{\mathbf{2}}}}$$

Esta fórmula representa una curva simétrica con forma de campana. La mayor parte de la probabilidad se concentra en el intervalo $\mu \pm \sigma$ /, y casi toda en$\mu \pm 3\sigma$. Esto explica la conocida "regla empírica" del 68–95–99.7 %.

Según De Veaux et al. (2019), la normal no solo describe numerosos fenómenos naturales como estaturas, errores de medición o variaciones fisiológicas; sino que también sirve como base para procedimientos inferenciales clásicos.

Sin embargo, Montgomery y Runger (2018) destacan que su poder explicativo no depende tanto de su forma, sino del Teorema Central del Límite (TCL). Este teorema establece que el promedio de muchas variables independientes tiende a distribuirse normalmente, incluso cuando las variables originales no lo son.

Por ello:

- el puntaje promedio de una clase,
- el tiempo promedio de ejecución de un algoritmo,
- la producción diaria promedio de una máquina,

tienden a ajustarse a una normal.

Caso de estudio: Variabilidad de los puntajes en una prueba estandarizada de razonamiento matemático

En un colegio con programas de evaluación continua, los docentes aplican una prueba estandarizada que mide razonamiento matemático en estudiantes de décimo año. Los resultados, al ser analizados, muestran una clara concentración alrededor de 70 puntos sobre 100, con simetría ligera y pocas notas extremadamente bajas o altas.

El equipo académico decide entonces modelar la distribución de los puntajes mediante una distribución normal con media aproximada$\mu=70$ y desviación estándar$\sigma=8$.

Según De Veaux, Velleman y Bock (2019), este tipo de patrón es típico en muchos fenómenos educativos, porque el rendimiento surge de múltiples factores que actúan de manera acumulativa y cuyo promedio tiende naturalmente a formar la conocida "curva de campana". Montgomery y Runger (2018) destacan que esta regularidad está amparada por el Teorema Central del Límite, que explica por qué la normal aparece incluso cuando las variables individuales no siguen una distribución normal.

Preguntas que los docentes desean responder

- ¿Qué porcentaje de estudiantes podría esperarse dentro del rango "esperado" entre 62 y 78 puntos?
- ¿Cuántos estudiantes podrían considerarse con desempeño inusualmente alto (percentil 95 o superior)?
- ¿Qué tan extremo es un puntaje de 50? ¿Debe interpretarse como un caso aislado o como señal de dificultades sistemáticas?

La Figura 17 muestra la distribución simulada de los puntajes de 120 estudiantes en una prueba estandarizada de razonamiento matemático. La forma general del histograma, acompañada de la curva normal teórica N(70, 8), permite aproximarnos a las tres inquietudes que el equipo docente desea resolver. La concentración de barras alrededor de 70, con disminución progresiva hacia ambos extremos, sugiere que la mayor parte del grupo se ubica en un rango de desempeño típico, mientras que solo unos pocos estudiantes alcanzan resultados particularmente altos o muy bajos.

Figura 17.
Distribución simulada de puntajes en una prueba estandarizada de razonamiento matemático (n = 120).

Nota. La figura muestra una simulación basada en un modelo de distribución normal con media 70 y desviación estándar 8, que representa el comportamiento esperado de los puntajes en una prueba estandarizada.

1. ¿Qué porcentaje de estudiantes podría esperarse dentro del rango entre 62 y 78 puntos?

Visualmente, la mayor densidad de barras se concentra justamente entre 62 y 78 puntos, lo que coincide con el tramo central de la curva de distribución. Este intervalo abarca el sector donde la campana alcanza su forma más ancha, lo que indica una probabilidad elevada de que los puntajes se ubiquen en ese rango. La simulación confirma esta lectura: aproximadamente dos tercios del grupo se sitúan en ese intervalo, dato que armoniza con la regla empírica de la normal, según la cual alrededor del 68 % de los valores se encuentran dentro de una desviación estándar de la media.

2. ¿Cuántos estudiantes presentan un desempeño inusualmente alto (percentil 95 o superior)?

En el extremo derecho de la distribución se observa una presencia escasa de barras, lo que señala que solo una proporción muy pequeña de estudiantes alcanza puntajes superiores a 83, valor que corresponde al percentil 95 del modelo teórico. La simulación sugiere que, en un grupo de 120 estudiantes, es esperable encontrar entre 5 y 6 casos con un rendimiento excepcionalmente alto. La poca densidad en ese sector del histograma confirma que se trata de resultados estadísticamente raros y pedagógicamente significativos, útiles para identificar logros sobresalientes.

3. ¿Qué tan extremo es un puntaje de 50? ¿Corresponde a un caso aislado o advierte problemas más profundos?

Al observar el extremo izquierdo de la figura, se nota que los puntajes muy bajos (cercanos a 50) prácticamente no aparecen. Esa escasez se corresponde con la probabilidad teórica: menos del 1 % de los estudiantes debería obtener un puntaje igual o inferior a 50. En consecuencia, un puntaje de esa magnitud se interpreta como un valor atípico, es decir, una observación aislada que se aleja del patrón general del grupo. Por sí solo, un resultado tan extremo no implica necesariamente un problema sistémico, pero sí invita a analizar de manera individual las condiciones del estudiante, la pertinencia de las tareas evaluadas o la presencia de factores que hayan afectado su desempeño.

En resumen, la simulación permite a los docentes visualizar el patrón esperado de rendimiento y, al mismo tiempo, situar cada pregunta en un contexto estadístico claro. El rango 62–78 funciona como referencia de normalidad, los valores superiores al percentil 95 identifican desempeños particularmente elevados, y un puntaje de 50 constituye un caso inusual que merece atención individualizada. Estas interpretaciones respaldan decisiones pedagógicas informadas y ayudan a comprender mejor la variabilidad natural del aprendizaje en contextos reales.

Los resultados obtenidos en la simulación permiten abrir una discusión pedagógica sobre cómo comprender la variabilidad del rendimiento estudiantil y cómo traducir estas evidencias en decisiones más justas y pertinentes al interior del aula. La forma de la distribución, dominada por una concentración amplia alrededor del promedio y una presencia muy reducida de casos extremos, nos recuerda que el rendimiento académico rara vez puede explicarse desde un único factor. Más bien, emerge de una combinación compleja de condiciones individuales, experiencias previas, formas de enseñanza, intereses, y apoyos disponibles. Este reconocimiento no solo es estadístico sino profundamente pedagógico, porque invita a mirar al grupo como un conjunto diverso, en el que las diferencias no deben ser leídas como fallas, sino como puntos de partida distintos.

El hecho de que la mayoría de los estudiantes se concentre entre 62 y 78 puntos sugiere que, para ellos, los contenidos y exigencias de la prueba se encuentran dentro de un rango de desafío razonable. Esta franja central, que reúne aproximadamente a dos tercios del grupo, sirve como una referencia valiosa para ajustar la enseñanza: indica qué habilidades están siendo alcanzadas de manera generalizada y cuáles podrían requerir ampliación, profundización o un trabajo más contextualizado. En este sentido, los puntajes "esperados" no deben asumirse como

una meta cerrada, sino como un punto de equilibrio desde donde impulsar procesos de mejora sin desatender la heterogeneidad de trayectorias.

Por otro lado, los estudiantes ubicados en el percentil 95 o superior representan un segmento cuya presencia es estadísticamente pequeña, pero pedagógicamente relevante. Su rendimiento sobresaliente no debe verse solo como un indicador de excelencia individual, sino como una señal para la institución: la necesidad de ofrecer retos adicionales, fortalecer itinerarios diferenciados, promover proyectos de profundización y, sobre todo, evitar que estos estudiantes queden desatendidos bajo la idea errónea de que "ya dominan" todo lo necesario. La educación inclusiva también se expresa en la capacidad de ampliar los límites para quienes avanzan a ritmos más acelerados, sin que ello implique desatender al resto del grupo.

Finalmente, la rareza de un puntaje de 50 abre una reflexión importante sobre las dificultades extremas. La estadística muestra que estos casos son poco frecuentes, pero su existencia demanda atención cuidadosa. Un puntaje muy bajo puede responder a múltiples causas: una comprensión insuficiente de los contenidos, ansiedad frente a la evaluación, experiencias previas de frustración, o incluso barreras externas vinculadas al contexto familiar o social. Por ello, más que etiquetar o atribuir déficit, resulta necesario adoptar una mirada diagnóstica que considere dimensiones emocionales, cognitivas y pedagógicas. Cada caso de este tipo constituye una invitación a revisar no solo la trayectoria del estudiante, sino también la claridad de las instrucciones, la adecuación del instrumento evaluativo y la disponibilidad de apoyos adicionales.

La distribución exponencial: tiempos de espera y falta de memoria

La distribución exponencial pertenece a la familia de las distribuciones de tiempos de espera. Modela la duración entre eventos sucesivos que ocurren de forma aleatoria e independiente, característica propia de sistemas que no poseen "historial".
Su parámetro es la tasa λ

Su densidad es: $\mathbf{f\left(x\right) = \square e^{-\square x},\ x \geq 0}$

La función de supervivencia es: $\mathbf{P\left(X > x\right) = e^{-\square x}}$

Wasserman (2010) subraya que la principal característica de la exponencial es su propiedad de falta de memoria:

$$\mathbf{P(X > t + s \mid X > t) = P(X > s)}$$

Esto significa que la probabilidad de esperar un tiempo adicional no depende del tiempo ya transcurrido.

Ejemplos aplicados

1. Centros de soporte y sistemas de colas

Los tiempos entre llamadas en una mesa de ayuda suelen modelarse mediante distribuciones exponenciales, especialmente cuando el flujo es irregular e impredecible.

2. Ingeniería electrónica

El tiempo entre fallas de componentes electrónicos o sensores puede aproximarse por una exponencial, especialmente cuando se supone una tasa de falla constante.

3. Educación digital

En plataformas educativas masivas, los intervalos entre accesos de estudiantes pueden ajustarse a distribuciones de tipo exponencial, lo que permite planificar capacidad del servidor o predecir picos de demanda.

Caso de estudio: Tiempos de espera entre llamadas en un centro de soporte universitario

Un centro de soporte tecnológico de una universidad recibe diariamente decenas de solicitudes de ayuda sobre plataformas académicas, contraseñas, fallos de conectividad y uso de software. Al analizar los tiempos entre llamadas consecutivas, el equipo observa que los intervalos son variables, pero tienden a ser cortos cuando hay mayor demanda.

Tras un análisis exploratorio, detectan que los tiempos entre llamadas se ajustan razonablemente a una distribución exponencial con parámetro –=0.2, lo que indica un promedio de cinco minutos entre solicitudes.

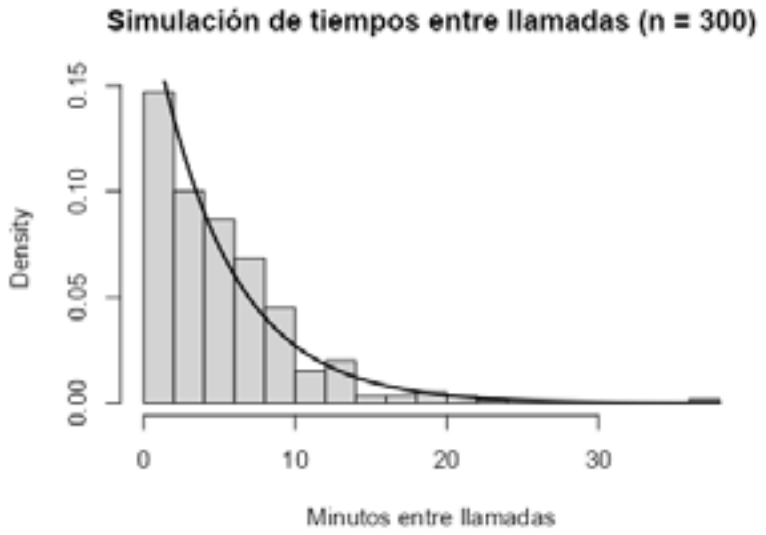
Wasserman (2010) destaca que la distribución exponencial es la herramienta natural para modelar tiempos entre eventos independientes en sistemas sin memoria, donde el tiempo ya transcurrido no afecta la probabilidad del siguiente evento.

Preguntas que el área de soporte desea responder

- ¿Cuál es la probabilidad de esperar más de 10 minutos entre una llamada y otra?
- ¿Qué tan probable es recibir dos solicitudes con diferencia de menos de un minuto cuando hay alta demanda?
- ¿Cuál es la distribución esperada de tiempos en horas pico y qué implicaciones tiene para el número de operadores necesarios?

La Figura 18 representa la simulación de 300 intervalos entre llamadas en un centro de soporte universitario, modelados mediante una distribución exponencial con parámetro λ = 0.2. La forma característica del histograma con barras altas cerca del cero y una disminución progresiva hacia la derecha, evidencia que los tiempos de espera cortos son mucho más frecuentes que los intervalos largos. La curva teórica superpuesta refuerza esta lectura: a medida que transcurren más minutos desde la última llamada, la probabilidad de que llegue una nueva disminuye rápidamente.

Figura 18.
Simulación de tiempos entre llamadas en un centro de soporte universitario (n = 300).



Nota. El histograma muestra 300 intervalos simulados entre llamadas según una distribución exponencial con parámetro λ = 0.2, equivalente a un tiempo medio de cinco minutos entre solicitudes.

Al responder las preguntas se puede concluir:

1. **Probabilidad de esperar más de 10 minutos entre una llamada y otra**

La caída pronunciada de la curva y la escasez de barras más allá de los 10 minutos indican que los intervalos largos son poco comunes. El cálculo teórico arroja una probabilidad aproximada del 13 %, lo que coincide visualmente con la figura: existe la posibilidad de que se presenten tiempos de espera prolongados, pero no constituyen el comportamiento usual. Para el equipo de soporte, esto significa que depender de estos lapsos tranquilos para organizar el trabajo podría resultar arriesgado, pues ocurren de manera esporádica.

2. Probabilidad de recibir dos solicitudes con diferencia de menos de un minuto

En el extremo izquierdo de la figura se observa una concentración notable de barras, todas agrupadas muy cerca del cero. Este patrón refleja que los tiempos muy cortos (intervalos menores de un minuto) son relativamente frecuentes. El cálculo de la simulación confirma esta tendencia: alrededor del 18 % de los intervalos entre llamadas ocurre en menos de un minuto. Desde una perspectiva operativa, esto implica que el sistema puede experimentar "picos súbitos" de solicitudes que exigen una reacción inmediata. Para un centro de soporte, esta realidad hace evidente la necesidad de contar con suficiente. personal en momentos de alta demanda o, en su defecto, mecanismos alternativos que ayuden a absorber las solicitudes acumuladas.

3. Distribución esperada en horas pico y su relación con el número de operadores

La simulación permite proyectar cuántas llamadas podrían recibirse en períodos de mayor actividad. Con una tasa $\lambda = 0.2$ por minuto, el modelo predice un promedio de 12 llamadas por hora, aunque con variaciones. La gráfica, al mostrar que la probabilidad de intervalos cortos es alta, sugiere que en franjas específicas podrían acumularse varios requerimientos en poco tiempo. Además, la probabilidad de superar las 15 llamadas por hora es cercana al 13 %, lo cual no es despreciable.

En términos prácticos, estos resultados implican que un solo operador tendría dificultades para mantener un flujo eficiente durante las horas pico. La combinación de muchos intervalos cortos y la posibilidad de recibir múltiples llamadas en rápida sucesión justifica la presencia de al menos dos operadores para garantizar tiempos de respuesta adecuados y evitar congestiones.

En su resumen, la figura y los resultados analíticos revelan un sistema donde predominan los intervalos breves, ocasionalmente interrumpidos por lapsos más largos. Este comportamiento, típico de la distribución exponencial, ayuda a comprender la naturaleza irregular de la demanda: aunque los promedios son útiles, la gestión operativa requiere anticipar escenarios de presión inmediata. La simulación se convierte así en una herramienta valiosa para dimensionar el equipo, planificar turnos y diseñar estrategias de respuesta que mantengan la calidad del servicio incluso en momentos de alta demanda.

Apoyo didáctico: El recorrido por las distribuciones continua uniforme, normal y exponencial permite comprender que cada una de ellas ofrece una forma distinta de interpretar fenómenos reales donde los valores posibles no son discretos, sino parte de un intervalo continuo. Desde una perspectiva didáctica, reconocer

las diferencias conceptuales entre estos modelos ayuda a que los estudiantes desarrollen un pensamiento probabilístico más flexible y ajustado a la variedad de situaciones que pueden encontrarse en contextos académicos, científicos y profesionales.

La distribución uniforme introduce la idea de equidad: todos los resultados dentro del intervalo tienen la misma probabilidad. Su sencillez favorece que los estudiantes transiten de situaciones discretas a continuas sin perder de vista la noción fundamental de probabilidad. Es un punto de partida ideal para explorar cómo se comportan los valores cuando no existe preferencia por ningún resultado específico y cómo se distribuyen de manera homogénea en un rango determinado.

La distribución normal, por su parte, incorpora la noción de concentración alrededor de un valor central. Su forma de "campana" no solo aparece en los libros de estadística, sino que se manifiesta en una gran cantidad de fenómenos cotidianos: aprendizajes, mediciones biológicas, desempeño académico, variaciones naturales, entre otros. Desde la enseñanza, la normal permite discutir ideas clave como el comportamiento de la variabilidad, el papel del promedio y la importancia de los desvíos respecto a ese centro. Más aún, invita a los estudiantes a leer datos desde una mirada integral, entendiendo que la mayoría de valores se agrupan alrededor del centro y que los casos extremos, aunque posibles, ocurren con menor frecuencia.

Finalmente, la distribución exponencial aporta un modelo adecuado para fenómenos en los que se espera que los eventos ocurran de manera repentina y sin memoria. Su carácter asimétrico ayuda a los estudiantes a diferenciar situaciones donde el tiempo entre eventos tiene un comportamiento decreciente: muchos intervalos cortos y muy pocos intervalos largos. Este enfoque es especialmente útil para interpretar procesos dinámicos, como tiempos de espera, flujos de llamadas o llegadas a un sistema, en los que la incertidumbre se expresa de forma distinta a la simetría de la normal.

En conjunto, el estudio de estas tres distribuciones no solo amplía el repertorio de modelos disponibles, sino que permite comparar estructuras, formas, niveles de concentración y significados probabilísticos. A nivel didáctico, este contraste favorece una comprensión más profunda de la probabilidad continua, pues permite que los estudiantes reconozcan que cada modelo responde a una lógica particular y se ajusta mejor a ciertos tipos de fenómenos. Más aún, invita a pensar la estadística como un lenguaje que describe comportamientos diversos, y no como un conjunto rígido de fórmulas.

Síntesis conceptual y didáctica: hacia una comprensión profunda de las distribuciones y la modelación estadística

El estudio de las variables aleatorias, sus distribuciones y los modelos que permiten describir fenómenos reales constituye un eje fundamental dentro de la educación estadística contemporánea. A lo largo de este capítulo se han presentado distintos tipos de distribuciones cada una con sus propiedades particulares, pero todas relacionadas por la necesidad de comprender la variabilidad inherente a los datos y la manera en que los modelos permiten interpretarla y anticiparla. Este epígrafe final tiene como propósito integrar los conceptos estudiados, ofrecer una lectura pedagógica articulada y mostrar cómo la modelación sirve como puente entre la teoría y los problemas reales que docentes y estudiantes enfrentan en el aula.

Tal como señalan Bakker (2004) y Batanero (2001), aprender estadística no consiste simplemente en manipular fórmulas o memorizar definiciones, sino en desarrollar un pensamiento que permita comprender patrones, incertidumbres y relaciones dentro de los datos. Desde esta perspectiva, las distribuciones continuas no se presentan como objetos estáticos, sino como herramientas que ayudan a interpretar fenómenos que se expresan en términos de densidad, probabilidad acumulada y comportamiento global. Comprender su forma, sus parámetros y su utilidad en contextos prácticos permite al estudiante construir un sentido estadístico que se fortalece mediante la experiencia, la discusión y la simulación.

Uno de los aportes más relevantes en educación estadística proviene del trabajo de Wild y Pfannkuch (1999), quienes sostienen que el pensamiento estadístico implica reconocer la necesidad de los datos, transitar entre modelos y realidad, y comprender que toda inferencia lleva implícita una dosis de incertidumbre. En ese marco, la integración de las distribuciones continúa siendo un pilar formativo, pues obliga a los estudiantes a observar la forma de los datos, analizar su dispersión, describir su comportamiento típico y explicar los casos excepcionales. La enseñanza de las distribuciones es, en este sentido, una enseñanza sobre cómo pensar la variabilidad, cómo razonarla y cómo comunicarla.

Asimismo, autores como Borovcnik (2016) han señalado que el desarrollo de la alfabetización probabilística requiere no solo comprender modelos teóricos, sino también interpretar qué significan en contextos donde la información incompleta y la incertidumbre son parte del problema.

Este enfoque permite que la estadística deje de percibirse como una disciplina rígida para convertirse en un terreno de exploración conceptual donde modelos como la distribución normal o la exponencial se interpretan como aproximaciones útiles, pero no exactas, de fenómenos que observamos en el mundo real.

La incorporación de simulaciones en este capítulo responde precisamente a esta visión contemporánea de la enseñanza estadística: utilizar herramientas tecnológicas no solo para calcular, sino para experimentar y visualizar comportamientos. Ben-Zvi (2000) destaca que el uso de software, gráficos dinámicos y herramientas de simulación amplía la capacidad del estudiante para comprender fenómenos que, de otro modo, resultan abstractos o difíciles de representar mentalmente.

En consonancia con este enfoque, el objetivo de este epígrafe es reunir los elementos fundamentales que permiten comprender las distribuciones continuas desde una perspectiva conceptual y pedagógica. Esto incluye analizar el papel de los parámetros y los momentos, reconocer la importancia de la variabilidad en la modelación de situaciones reales, y reflexionar sobre cómo estas ideas contribuyen al desarrollo del pensamiento estadístico y la alfabetización necesaria para enfrentar problemas del mundo contemporáneo, tal como subrayan Moore (2010), Watson (2006) y Pfannkuch (2019).

Parámetros, momentos y significado pedagógico de la forma de una distribución

Comprender una distribución de probabilidad implica, ante todo, leer su forma. Esa forma simétrica, sesgada, aplanada, concentrada o dispersa, no es solo una característica visual, sino una expresión de cómo se comportan los datos, qué valores tienden a aparecer con mayor frecuencia y cuáles son menos probables. En educación estadística, enseñar a interpretar la forma de una distribución representa uno de los desafíos más relevantes, tal como destacan Batanero (2001), Borovcnik (2016) y Watson (2006), porque exige que el estudiante transite de observar datos puntuales a comprender patrones agregados que resumen un fenómeno.

Desde un punto de vista formal, la descripción de una distribución se basa en sus parámetros y momentos. Los parámetros,como la media, la varianza, la desviación estándar o los percentiles; permiten identificar el comportamiento central y la medida de dispersión. Los momentos aportan información adicional sobre la forma: si la distribución es simétrica o está inclinada hacia un lado, si presenta colas ligeras o pesadas, o si concentra la mayoría de valores alrededor de la media.

Sin embargo, reducir la enseñanza a la simple definición de estos conceptos sería insuficiente. Como señalan Bakker (2004) y Wild y Pfannkuch (1999), el desafío pedagógico radica en ayudar a los estudiantes a comprender lo que estas medidas significan en los datos, y no solo cómo se calculan. Por ejemplo, entender que una desviación estándar pequeña implica que la mayoría de los valores están cerca del promedio, mientras que una desviación grande revela heterogeneidad, diversidad o variabilidad significativa.

Esta comprensión se vuelve especialmente importante cuando se comparan distribuciones. La comparación no consiste únicamente en determinar cuál media es mayor, sino en analizar cómo cambia la forma global y qué implicaciones tiene para interpretar el fenómeno. Tal como explican Moore (2010) y Stewart (2013), aprender estadística implica moverse constantemente entre representaciones: de los datos individuales al histograma, del histograma a los parámetros, y de estos a una interpretación contextualizada.

a) La media y la mediana como indicadores del comportamiento central

El primer momento de una distribución es su media. Desde una perspectiva estadística, representa el punto de equilibrio del conjunto de datos, aquello que resume de forma sintética el comportamiento típico. Sin embargo, desde la educación estadística y siguiendo a Watson (2006), es crucial enseñar que la media no es necesariamente el valor más frecuente, ni siempre el más representativo. Los estudiantes suelen confundir "promedio" con "valor típico", lo que exige actividades donde la media se compare con la moda, con la mediana o con el rango intercuartílico.

La mediana, por su parte, aporta un elemento de interpretación muy valioso cuando la distribución es asimétrica. En la distribución normal media y mediana coinciden, pero en la distribución exponencial la mediana es siempre menor que la media debido a la presencia de colas largas hacia la derecha. Esta diferencia es una oportunidad didáctica para discutir con los estudiantes la importancia de la forma en la interpretación de los parámetros.

Los trabajos de Biehler (2018) subrayan que enseñar variabilidad implica enseñar a interpretar el "típico" no como un valor único, sino como un rango razonable de resultados. De ahí que la media cobre verdadero sentido cuando se acompaña de una medida de dispersión, como la desviación estándar.

b) La varianza y la desviación estándar como medidas de variabilidad

El segundo momento de una distribución describe su dispersión. La varianza y la desviación estándar permiten comprender qué tan concentrados o dispersos están los valores con respecto a la media. Ahora bien, desde la didáctica, explicar la varianza puede ser una tarea compleja, pues el estudiante tiene dificultades para interpretar el uso del cuadrado en su cálculo y para visualizar qué significa en términos de datos reales.

En cambio, la desviación estándar al estar en las mismas unidades que la variable original se vuelve una herramienta mucho más intuitiva. Moore (2010) insiste en que su enseñanza debe apoyarse en ejemplos visuales, gráficos y simulaciones que permitan observar cómo aumenta o disminuye la dispersión en el histograma cuando la desviación cambia. Una actividad didáctica habitual consiste en tomar una muestra y generar nuevas simulaciones con diferentes grados de variabilidad, de modo que los estudiantes puedan ver cómo cambia la forma de la distribución en función de la dispersión.

En el caso de la distribución normal, la relación entre la desviación estándar y las áreas bajo la curva constituye un recurso pedagógico muy potente. Permite comprender por qué valores como 62 o 78 puntos en el caso de los resultados de la prueba simulada representan desempeños comunes, mientras que puntajes como 83 o 50 son menos frecuentes o incluso excepcionales. Esto conecta directamente la medida de dispersión con la interpretación contextual: la estadística deja de ser un número y se convierte en un argumento.

argumento.

c) Asimetría, colas y significado contextual

La asimetría es quizá uno de los elementos menos abordados en los cursos iniciales, a pesar de su importancia para la interpretación de muchas distribuciones reales. Tal como recuerda Borovcnik (2016), la mayor parte de los fenómenos aleatorios no presentan simetría perfecta, sino que muestran sesgos más o menos pronunciados. En la práctica, fenómenos como tiempos de espera, rendimiento en tareas complejas o ingresos económicos tienden a presentar colas largas hacia la derecha, como ocurre con la distribución exponencial que se utilizó en la simulación de llamadas.

d) Curtosis y concentración

La curtosis analiza si una distribución presenta colas más pesadas o más ligeras que la normal. Aunque puede parecer un

concepto avanzado, resulta útil en situaciones donde se necesita identificar fenómenos con alta presencia de valores extremos. En contextos educativos, este parámetro permite discutir sobre dispersiones irregulares, situaciones de riesgo o variabilidad muy alta, elementos que Watson (2006) considera fundamentales para que el estudiante comprenda la incertidumbre inherente a los datos reales.

Caso de estudio: Interpretación de parámetros, momentos y forma en los resultados de una prueba diagnóstica

Contexto general

Una institución educativa ha iniciado un programa de refuerzo en Matemática para estudiantes de primer curso de bachillerato. Como punto de partida, se aplicó una prueba diagnóstica de 40 ítems (cada uno vale 1 punto) a 150 estudiantes. El puntaje total de cada estudiante puede variar entre 0 y 40 puntos y se almacena en una base de datos junto con un identificador de estudiante.

Los docentes no quieren limitarse a obtener un promedio general. Su objetivo es:

- Describir cómo se distribuyen los puntajes en el grupo (tendencia central y dispersión).
- Analizar qué tan homogénea es la cohorte (variabilidad de los resultados).
- Identificar si la distribución presenta asimetría (más estudiantes con puntajes bajos o altos).
- Explorar niveles de desempeño (bajo, medio, alto) para orientar mejor el plan de refuerzo.
- Extraer implicaciones didácticas: ¿necesitan trabajar habilidades básicas, resolución de problemas, comprensión de funciones, etc.?

La distribución presentada en la Figura 19 permite observar con bastante claridad cómo se comportarían los puntajes de un grupo de estudiantes si enfrentaran una prueba diagnóstica de 40 preguntas bajo condiciones promedio.

El histograma muestra una forma aproximadamente simétrica, con una mayor concentración de estudiantes alrededor de los 22 puntos, valor que coincide con la media calculada mediante la simulación.

Este comportamiento sugiere que la mayoría del grupo presenta un nivel de dominio intermedio sobre los contenidos evaluados. Es decir, una buena parte de los estudiantes logra responder correctamente un poco más de la mitad de los ítems, mientras que solo unos pocos alcanzan valores muy altos o muy bajos. La presencia de una ligera dispersión hacia la derecha indica que existen estudiantes que obtienen puntuaciones más elevadas, aunque representan un porcentaje reducido.

Figura 19.
Distribución de puntajes simulados en la prueba diagnóstica de Matemática

Nota. La figura muestra la distribución de los puntajes simulados de 150 estudiantes en una prueba diagnóstica de 40 ítems. Los datos se generaron mediante una distribución binomial con parámetros n=40 y p=0,55.

El análisis también muestra que los puntajes extremos tanto muy bajos como muy altos, son relativamente poco frecuentes. Esto es un indicio de que, en general, el grupo no se encuentra polarizado, sino que comparte un nivel de desempeño relativamente homogéneo.

Desde una perspectiva pedagógica, este tipo de distribución resulta útil porque permite anticipar el nivel de apoyo que será necesario brindar al inicio del curso: un promedio moderado, variabilidad controlada y pocos valores atípicos suelen asociarse con grupos que requieren refuerzo puntual, pero no necesariamente intervenciones intensivas.

Modelación de fenómenos reales: variabilidad, patrones y predicción

La estadística se vuelve verdaderamente significativa cuando permite comprender fenómenos reales. En el aula, esto implica ayudar a los estudiantes a interpretar variabilidad, identificar patrones y utilizar modelos para anticipar comportamientos posibles.

En coherencia con este enfoque, Biehler (2018) sostiene que la modelación estadística representa un pilar para el desarrollo del pensamiento estadístico porque permite que los estudiantes pasen de describir datos a razonar con ellos y a tomar decisiones fundamentadas en la incertidumbre.

Este subepígrafe busca mostrar cómo las distribuciones continuas funcionan cómo los parámetros permiten realizar inferencias y predicciones razonables. A través de los casos simulados trabajados previamente, se demuestra que modelar no consiste en replicar la realidad de manera exacta, sino en construir una representación que ilumine relaciones relevantes y ayude a responder preguntas pertinentes.

a) De los datos al modelo: interpretar la variabilidad como rasgo del fenómeno

En educación estadística, uno de los desafíos más complejos es lograr que los estudiantes no interpreten la variabilidad como "error" o "ruido", sino como una característica natural del fenómeno que se estudia. Watson (2006) y Pfannkuch (2019) enfatizan que desarrollar una mirada estadística implica reconocer que los datos reales rara vez siguen un patrón perfecto y que esta imperfección es justamente lo que hace necesaria la modelación.

El caso de los puntajes en la prueba estandarizada es un ejemplo claro. La variabilidad entre estudiantes no es un problema a corregir: es un rasgo esperado, producto de múltiples factores cognitivos, emocionales, socioculturales y pedagógicos. Tal como explica DeVeaux, Velleman y Bock (2019), los rendimientos suelen agruparse alrededor de un nivel promedio, generando una forma aproximadamente simétrica que se asemeja a la curva normal. Esto no significa que la población esté "perfectamente normalizada", sino que la normal es un modelo adecuado para representar tendencias globales cuando los factores que influyen son variados e independientes.

Al construir la simulación en R y en Jamovi y observar la curva superpuesta, los estudiantes pueden comprender visualmente cómo el modelo refleja patrones reales: concentración alrededor del promedio, dispersión moderada, presencia ocasional de puntajes altos y baja probabilidad de valores extremos. En términos pedagógicos, esta vinculación entre contexto, datos y modelo es lo que permite que la estadística deje de ser un conjunto de cálculos y se convierta en una forma de interpretar fenómenos.

b) Distribuciones para fenómenos distintos: normalidad, uniformidad y tiempos exponenciales

Cada distribución continua describe una estructura interna distinta. Enseñar esta diferencia tiene implicaciones directas tanto en la comprensión conceptual como en la capacidad de aplicar el modelo adecuado a cada caso. Cada distribución continua describe una estructura interna distinta. Enseñar esta diferencia tiene implicaciones directas tanto en la comprensión conceptual como en la capacidad de aplicar el modelo adecuado a cada caso.

La distribución normal: concentración, simetría y comportamiento típico

La distribución normal aparece en procesos donde intervienen múltiples factores pequeños que actúan de manera acumulativa. Montgomery y Runger (2018) explican que esta regularidad está respaldada por el Teorema Central del Límite, lo que convierte a la normal en un modelo robusto para describir fenómenos como calificaciones, mediciones biológicas, errores instrumentales y desempeños humanos. La simulación de los 120 puntajes muestra precisamente esta estructura: valores concentrados entre 62 y 78, simetría ligera y casos extremos poco frecuentes.

La enseñanza debe poner énfasis en que la normal no es un "molde perfecto", sino una herramienta para describir tendencias. Como sugiere Moore (2010), los estudiantes deben aprender a identificar cuándo la forma observada se acerca a una normal y cuándo no, para decidir si es un modelo pertinente.

La distribución uniforme: equidad y ausencia de concentración

La distribución uniforme permite modelar fenómenos donde todos los valores dentro de un intervalo son igualmente probables. Aunque menos frecuente en aplicaciones reales, es fundamental desde una perspectiva didáctica porque ayuda a introducir la idea de densidad constante y a contrastarla con distribuciones más complejas. Stewart (2013) sostiene que su valor pedagógico radica en mostrar una forma ideal que raramente se observa en la práctica, pero que permite comprender principios fundamentales sobre intervalos, continuidad y probabilidad.

Casos reales como la selección de números pseudoaleatorios, la simulación de ubicaciones geográficas o la asignación de horarios pueden ilustrar este comportamiento. Cuando se simulan datos uniformes en R o Jamovi, la ausencia de picos o concentraciones facilita la discusión sobre qué significa "igual probabilidad" en contextos continuos.

La distribución exponencial: tiempos de espera y eventos sin memoria

El caso de los intervalos entre llamadas en un centro de soporte universitario es un ejemplo excelente de cómo la distribución exponencial describe fenómenos donde predominan los intervalos cortos, pero existen probabilidades no despreciables de esperas más largas. Wasserman (2010) resalta que la exponencial se utiliza para modelar tiempos entre eventos independientes en sistemas sin memoria, fenómeno que no puede ser descrito adecuadamente con una distribución normal.

La simulación realizada en R revela una caída pronunciada en los primeros minutos y una cola larga hacia la derecha, lo que permite responder preguntas como:

- ¿Cuál es la probabilidad de esperar más de 10 minutos entre llamadas?
- ¿Qué tan probable es que dos llamadas lleguen con menos de 1 minuto de diferencia?
- ¿Cómo impacta esta variabilidad en el número de operadores necesarios?

Este tipo de análisis es clave para mostrar que la estadística no solo describe, sino que **ayuda a tomar decisiones en escenarios reales.**

c) Predicción razonada: del modelo a la toma de decisiones

La estadística no predice valores exactos; predice comportamientos probables. James (2017) insiste en que la enseñanza de la incertidumbre debe incluir la idea de "predicción razonada": el modelo no garantiza lo que sucederá, pero orienta lo que puede esperarse.

En el caso de la distribución normal de puntajes, el modelo permite estimar la proporción de estudiantes que probablemente se ubiquen en el rango esperado (62–78) y detectar casos atípicos que requieren atención pedagógica. Esto es crucial para la toma de decisiones educativas: identificar brechas, planificar refuerzos, valorar desempeños atípicos o reconocer logros excepcionales.

En el caso de la distribución exponencial, la predicción permite anticipar picos de demanda y ajustar recursos: número de operadores, distribución de turnos, tiempos de respuesta. Efron y Tibshirani (1993) han contribuido significativamente a este enfoque a través del bootstrapping y la simulación como métodos para generar intervalos de confianza en situaciones donde la teoría tradicional resulta insuficiente. Integrar estas técnicas ayuda a los estudiantes a comprender que los modelos estadísticos no solo describen, sino que también permiten anticipar lo que podría ocurrir bajo distintas condiciones.

Autores como Hastie et al. (2009) y James et al. (2021) destacan que la predicción estadística se fortalece cuando se combina la teoría con herramientas computacionales. En el aula, esto implica promover experiencias donde los estudiantes simulen, modifiquen parámetros y observen cómo los modelos responden, permitiendo un aprendizaje más profundo y significativo.

Caso de estudio: Predicción de la demanda de energía eléctrica en un campus universitario

Contexto general

Un campus universitario de tamaño medio ha comenzado a experimentar aumentos inesperados en su consumo diario de energía eléctrica. Estos cambios dificultan la gestión operativa

el área administrativa debe anticipar gastos, programar mantenimientos, prevenir sobrecargas en los edificios y garantizar que la infraestructura responda adecuadamente a la demanda del estudiantado y del personal docente.

Para comprender mejor el comportamiento del consumo, la institución implementa un sistema de monitoreo que registra mediciones horarias durante un trimestre completo. Con esta información, el equipo técnico busca identificar patrones, reconocer fuentes de variabilidad y construir modelos predictivos que permitan tomar decisiones más eficientes y anticipar situaciones críticas.

El propósito central del estudio es modelar el comportamiento del consumo energético, explicar sus fluctuaciones y generar pronósticos confiables para su gestión institucional.

Datos disponibles

Se recopilaron mediciones horarias durante 90 días consecutivos. Las variables registradas fueron:

- Hora del día (0–23)
- Consumo eléctrico (kWh) por edificio
- Condiciones climáticas (temperatura y nubosidad)
- Tipo de jornada (laboral, fin de semana o feriado)
- Eventos especiales (seminarios, congresos, actividades masivas)

Preguntas que el equipo desea responder

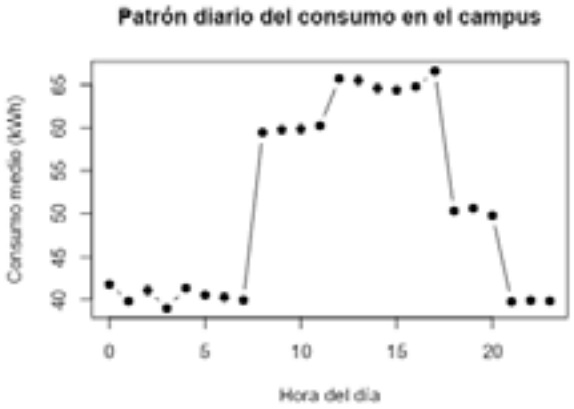
El análisis se orienta a resolver interrogantes clave que permitan comprender y anticipar el comportamiento energético del campus. Entre las principales preguntas se encuentran:

1. ¿Cómo varía el consumo de energía a lo largo del día y qué tan estable es este comportamiento?
2. ¿Existen diferencias significativas entre el consumo registrado en días laborales y en fines de semana?
3. ¿En qué medida la temperatura contribuye al incremento o disminución del consumo diario?
4. ¿Cuáles son los patrones horarios o semanales más característicos y cuáles representan los picos de mayor demanda?
5. ¿Qué escenarios de consumo pueden anticiparse ante cambios en la temperatura o ante la realización de eventos institucionales?
6. ¿Qué decisiones operativas o presupuestarias pueden tomarse a partir de las predicciones obtenidas?

La figura 20 evidencia que el consumo energético del campus universitario no se mantiene constante a lo largo del día, sino que sigue un patrón claramente diferenciado según el nivel de actividad institucional. Durante la madrugada y las primeras horas de la mañana, el consumo promedio se mantiene en valores cercanos a 40–43 kWh, lo que coincide con periodos de baja

ocupación y uso limitado de instalaciones. A partir de las 8:00, se observa un ascenso pronunciado que alcanza niveles entre 58 y 60 kWh, representando el inicio de la jornada académica y el encendido de equipos, oficinas y laboratorios.

Figura 20.
Patrón diario del consumo energético promedio en el campus universitario



Nota. La figura muestra la variación del consumo medio de energía eléctrica (kWh) a lo largo de las 24 horas del día, calculado a partir del registro horario de un trimestre académico.

El consumo llega a su punto más alto entre las 11:00 y las 15:00, donde se registran valores cercanos a 65 kWh, momento que coincide con el mayor flujo de estudiantes y actividades operativas. Hacia las 17:00 el consumo desciende gradualmente hasta situarse nuevamente en torno a los 40 kWh por la noche.

La figura 21 muestra de manera clara que el consumo energético diario presenta comportamientos distintos entre días laborales y fines de semana. En los días laborales, los valores oscilan aproximadamente entre 30 y 100 kWh, con una mediana cercana a los 55–60 kWh, lo que indica una demanda sostenidamente alta durante la actividad académica regular. La amplitud del rango y la presencia de consumos elevados sugiere una mayor variabilidad vinculada al uso intensivo de aulas, oficinas, equipos eléctricos y circulación constante de personas.

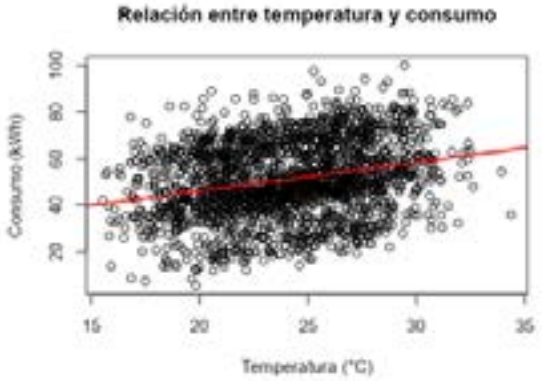
En los fines de semana, en cambio, el consumo se reduce de forma notable: los valores se concentran entre 15 y 70 kWh, con una mediana alrededor de los 35–40 kWh, reflejando una actividad institucional más baja y homogénea. La aparición de algunos puntos atípicos sobre los 70–80 kWh sugiere la realización ocasional de eventos o actividades extraordinarias que elevan temporalmente la demanda

Figura 21.
Consumo energético en días laborales y fines de semana

Nota. La figura compara la distribución del consumo diario de energía (kWh) entre días laborales y fines de semana.

La figura 21 evidencia que, aunque el consumo energético horario presenta una amplia dispersión, existe una tendencia general al incremento conforme la temperatura ambiente aumenta. Para temperaturas entre 15 y 20 °C, los consumos observados suelen situarse entre 20 y 60 kWh, con varios puntos que descienden incluso por debajo de los 20 kWh. A medida que la temperatura alcanza valores intermedios, entre 22 y 28 °C, la nube de puntos se vuelve más densa y los consumos se concentran principalmente entre 40 y 80 kWh, lo que sugiere una mayor demanda de equipos de ventilación, climatización o incremento de la actividad en el campus.

Figura 22.
Relación entre la temperatura y el consumo energético en el campus

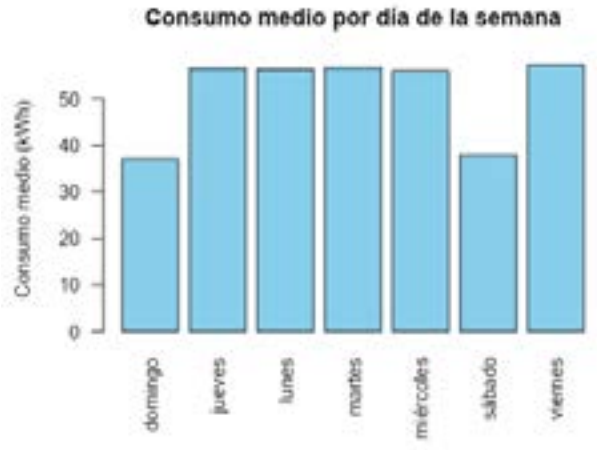


Nota. La figura muestra la relación entre la temperatura ambiente (°C) y el consumo energético horario (kWh).

En el rango más alto, por encima de 30 °C, se observan consumos que superan con facilidad los 80 kWh, llegando en algunos casos a valores cercanos a los 100 kWh. La línea de tendencia ascendente confirma que la relación entre temperatura y consumo, aunque moderada, es positiva: temperaturas más elevadas tienden a vincularse con mayores niveles de gasto energético. Este patrón cuantitativo permite inferir que el clima puede convertirse en un factor relevante para planificar la gestión de la demanda, especialmente en jornadas de calor intenso.

La figura 22 revela que el consumo energético promedio varía de manera notable a lo largo de la semana, lo que sugiere la presencia de patrones asociados al ritmo de actividad del campus. Los valores más bajos se registran los domingos y los sábados, con promedios cercanos a 38 kWh, lo que coincide con la reducción natural de actividades académicas y administrativas durante el fin de semana.

Figura 23.
Consumo medio de energía por día de la semana en el campus universitario

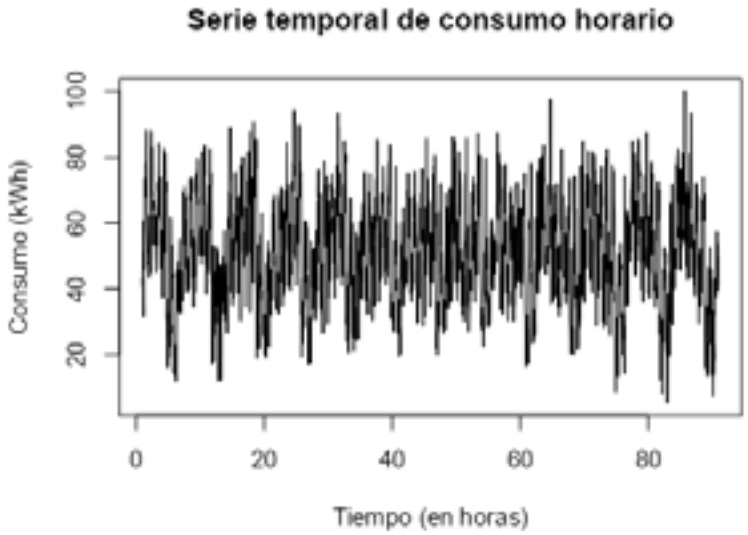


Nota. La figura muestra el consumo promedio de energía (kWh) registrado para cada día de la semana durante el período de observación.

En contraste, los días laborables muestran consumos significativamente superiores. Entre ellos, lunes, martes, miércoles y jueves presentan valores muy similares, rondando los 55 kWh, lo que evidencia un comportamiento estable y elevado de la demanda en la primera parte de la semana. El viernes, aunque mantiene un promedio elevado (aproximadamente 58 kWh), se ubica como uno de los días con mayor consumo, posiblemente debido al cierre de actividades semanales, el uso intensivo de laboratorios o la concentración de eventos institucionales.

La figura 23 muestra un comportamiento horario del consumo energético caracterizado por una marcada variabilidad, con oscilaciones frecuentes entre valores bajos cercanos a los 20–30 kWh y picos que superan los 90 kWh en distintos momentos del periodo observado.

Figura 24.
Serie temporal del consumo horario de energía en el campus universitario



Nota. La figura presenta la evolución del consumo de energía (kWh) registrada hora a hora durante el periodo analizado.

Este patrón refleja la dinámica propia de un campus universitario, en el que la demanda eléctrica se incrementa en franjas asociadas a actividades académicas, uso de laboratorios, encendido de sistemas de climatización o eventos puntuales, mientras disminuye en horarios de menor ocupación. Aunque no se aprecia una tendencia clara al alza o a la baja durante el periodo analizado, sí se observan ciclos recurrentes de aumento y disminución que sugieren comportamientos diarios relativamente estables, influenciados por rutinas institucionales.

Desarrollo del pensamiento estadístico y lectura crítica de la incertidumbre

La enseñanza de la estadística no puede reducirse a la transmisión de fórmulas, métodos o procedimientos. Implica, ante todo, desarrollar en los estudiantes una forma particular de pensar: un pensamiento estadístico capaz de interpretar la incertidumbre, cuestionar patrones aparentes, reconocer la variabilidad como un elemento inherente a los fenómenos reales y utilizar

modelos probabilísticos de manera crítica y contextualizada. Wild y Pfannkuch (1999) fueron pioneros en describir este enfoque, afirmando que el pensamiento estadístico se basa en cuatro componentes esenciales: la necesidad de los datos, el entendimiento de la variabilidad, la construcción de modelos y la integración de estos modelos en el razonamiento empírico.

estos modelos en el razonamiento empírico.

En coherencia con esta visión, este subepígrafe se centra en cómo el estudio de las distribuciones continuas contribuye al desarrollo de una alfabetización estadística profunda, crítica y orientada a la toma de decisiones. La meta no es solo que los estudiantes sean capaces de identificar la forma de una distribución, sino que puedan interpretar su significado, evaluar su pertinencia y emplearla para explicar y comprender fenómenos reales.

a) La importancia de la incertidumbre como objeto de enseñanza

En la mayoría de los contextos educativos, los estudiantes se familiarizan con la noción de error, pero no con la noción de variabilidad. Mientras que el error se percibe como algo que debe evitarse o corregirse, la variabilidad debe ser comprendida como una propiedad inherente de los fenómenos y un insumo clave para el análisis estadístico. Autores como Borovcnik (2016) y Pfannkuch (2019) han destacado que la incertidumbre no debe abordarse desde una perspectiva puramente matemática, sino también desde una perspectiva epistemológica y didáctica: ¿qué implica que un fenómeno sea incierto?, ¿por qué no podemos predecir un valor exacto?, ¿qué significa hablar de probabilidades en lugar de certezas?

Cuando los estudiantes trabajan con simulaciones como las realizadas en este capítulo comienzan a observar que, incluso bajo un mismo modelo, los resultados pueden variar notablemente entre repeticiones. Esta experiencia es crucial para romper la idea de que la estadística proporciona respuestas deterministas. Simular 300 tiempos entre llamadas con una distribución exponencial produce siempre un patrón reconocible, pero nunca idéntico. Esta irregularidad dentro de una regularidad general es una puerta de entrada privilegiada al pensamiento estadístico.

Bakker (2004) sostiene que este tipo de experiencias prácticas permite que los estudiantes desarrollen una sensibilidad hacia la estructura del fenómeno: comprenden qué aspectos son más estables (como la forma general de la curva) y cuáles son más fluctuantes (como los valores puntuales de cada simulación).

Esta capacidad de distinguir tendencias de ruido es una de las competencias centrales de la alfabetización estadística.

b) La simulación como herramienta para visualizar patrones y consolidar conceptos

El uso de simulaciones ocupa un lugar cada vez más central en la educación estadística contemporánea. Ben-Zvi (2000) ha mostrado que las herramientas tecnológicas permiten a los estudiantes visualizar patrones que de otro modo permanecen ocultos, manipular parámetros en tiempo real y observar cómo se altera la forma de la distribución cuando cambia la media, la desviación estándar o el parámetro λ. Esta interactividad transforma la relación con la estadística: de una actividad abstracta basada en procedimientos, a una actividad exploratoria que promueve la indagación y el descubrimiento.

En este capítulo, las simulaciones generadas en Jamovi y R permitieron recrear:

- Una distribución normal de puntajes, donde los estudiantes pudieron identificar un comportamiento concentrado alrededor del promedio, observar la simetría y reconocer la baja probabilidad de casos extremos;
- una distribución exponencial de tiempos de espera, donde se evidenció la presencia de muchos intervalos cortos y pocos intervalos largos, reforzando el concepto de "falta de memoria" del modelo;
- una distribución uniforme, útil para explicar estructuras sin concentración, como la selección aleatoria en un intervalo continuo.

Estas experiencias no solo fortalecen la comprensión conceptual, sino que promueven en los estudiantes una actitud interrogativa: ¿por qué la curva tiene esa forma?, ¿qué sucede si cambio el parámetro?, ¿cómo interpretar las colas largas?, ¿qué implican los casos raros para la toma de decisiones? Esta actitud inquisitiva es un componente clave del pensamiento estadístico, como señalan James (2017) y Watson (2006).

c) Competencias de lectura crítica: interpretar modelos sin idolatrarlos

Un desafío importante en la enseñanza consiste en evitar que los estudiantes atribuyan a los modelos un estatus absoluto. La estadística trabaja con aproximaciones, no con certezas. Efron y Tibshirani (1993) recuerdan que todo modelo es una representación parcial del fenómeno, no su réplica exacta.

Por ello es esencial que los estudiantes desarrollen una lectura crítica que les permita preguntarse si un modelo es adecuado, si los supuestos se cumplen y si los resultados se interpretan correctamente.

El caso de los puntajes simulados con una distribución normal ilustra esta necesidad. Aunque la distribución normal fue un modelo adecuado, no implica que todos los fenómenos educativos presenten simetría o concentración alrededor de un promedio. Del mismo modo, la distribución exponencial permitió modelar razonablemente los tiempos entre llamadas, pero sería inapropiado utilizarla para describir fenómenos donde existe dependencia temporal o donde los eventos no son aleatorios.

La lectura crítica también implica comprender las consecuencias prácticas del modelo. En el caso exponencial, por ejemplo, la probabilidad de intervalos menores de un minuto tiene implicaciones directas para la gestión de personal: es necesario considerar momentos de alta carga y evitar decisiones basadas únicamente en promedios. Tal como advierten James et al. (2021), la estadística aplicada requiere interpretar los modelos en su contexto, reconociendo sus límites y alcances.

d) La alfabetización estadística como competencia para el siglo XXI

La alfabetización estadística es considerada hoy una competencia esencial. Watson (2006) y Pfannkuch (2019) destacan que esta competencia no se reduce al dominio técnico, sino que incluye elementos éticos, comunicativos y epistemológicos: saber qué afirmaciones son legítimas, qué incertidumbres deben explicitarse, cómo comunicar resultados sin inducir a error y cómo tomar decisiones públicas o institucionales fundamentadas en evidencia.

Este capítulo contribuye a esa alfabetización mostrando que:

- los modelos estadísticos permiten comprender fenómenos complejos de manera más clara y manejable;
- la variabilidad no es un problema, sino una fuente de información;
- los parámetros y la forma de una distribución cuentan historias sobre el fenómeno;
- las simulaciones son herramientas pedagógicas poderosas para explorar, cuestionar y consolidar ideas;
- la lectura crítica es indispensable para evitar conclusiones simplistas o modelos inapropiados.

El recorrido realizado a lo largo de este epígrafe permitió articular de manera integrada tres dimensiones fundamentales de la educación estadística contemporánea: (a) la comprensión de los parámetros y momentos como expresiones del comportamiento

global de una distribución; (b) la modelación de fenómenos reales a través de distribuciones continuas como la uniforme, la normal y la exponencial; y (c) el desarrollo de un pensamiento estadístico orientado a la interpretación crítica de la incertidumbre y la variabilidad.

Esta integración es indispensable para formar ciudadanos capaces de comprender datos, evaluar riesgos, interpretar tendencias y tomar decisiones fundamentadas, tal como señalan Watson (2006) y Pfannkuch (2019). Una enseñanza que se limite a presentar fórmulas pierde de vista la riqueza conceptual que caracteriza a la estadística como disciplina y como práctica social. Por ello, autores como Bakker (2004), Batanero (2001) y Moore (2010) insisten en que la variabilidad debe convertirse en el eje central del currículo: no como un elemento accesorio, sino como la clave para entender por qué necesitamos modelos probabilísticos y cómo estos nos permiten pensar fenómenos reales.

Conclusiones

El Capítulo 3 mostró que comprender variables aleatorias y distribuciones continuas es esencial para interpretar la variabilidad inherente a los fenómenos reales, donde se enfatiza en que la estadística no debe enseñarse como un conjunto de fórmulas aisladas, sino como una forma de pensar que permite identificar patrones, reconocer incertidumbres y analizar datos desde una perspectiva crítica.

Las distribuciones uniforme, normal y exponencial no se presentan como objetos abstractos, sino como modelos que ayudan a describir distintos comportamientos: equidad en la uniforme, concentración y simetría en la normal, y tiempos de espera asimétricos en la exponencial.

A través de simulaciones desarrolladas en R y Jamovi, el capítulo mostró el valor educativo de visualizar patrones y contrastar lo esperado con lo observado. Este enfoque, ayuda al estudiantado a comprender que la estadística opera en escenarios donde la certeza es imposible, pero donde sí es posible razonar de manera informada. Las simulaciones permiten experimentar, ajustar parámetros y explorar cómo cambia la forma de una distribución, fortaleciendo la comprensión conceptual y el pensamiento estadístico.

Referencias

Bakker, A. (2004). Design research in statistics education [Tesis doctoral, Utrecht University].

Batanero, C. (2001). Didáctica de la probabilidad y la estadística. Editorial Síntesis.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. Mathematical Thinking and Learning, 2(1), 127–155. https://doi.org/10.1207/S15327833MTL0202_3

Biehler, R. (2018). Perspectives on modeling variability. En R. Biehler, B. Frischemeier, & M. Reading (Eds.), Statistics and data science education: New challenges for teaching and learning (pp. 45–67). Springer.

De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2019). Intro Stats (5th ed.). Pearson.

Efron, B., & Tibshirani, R. (1993). An introduction to the bootstrap. Chapman & Hall.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

James, G. (2017). Probability and the art of teaching uncertainty. Oxford University Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An introduction to statistical learning: With applications in R (2nd ed.). Springer. https://doi.org/10.1007/978-1-0716-1418-1

Montgomery, D. C., & Runger, G. C. (2018). Applied statistics and probability for engineers (7th ed.). Wiley.

Moore, D. S. (2010). The basic practice of statistics (6th ed.). W. H. Freeman.

Pfannkuch, M. (2019). Reimagining statistical literacy for the 21st century. En S. Prodromou (Ed.), Advances in statistics education (pp. 75–92). Springer.

Rice, J. A. (2007). Mathematical statistics and data analysis (3rd ed.). Cengage Learning.

Stewart, I. (2013). Concepts of modern mathematics. Dover.

Wasserman, L. (2010). All of statistics: A concise course in statistical inference. Springer.

Watson, J. (2006). Statistical literacy at school: Growth and goals. Lawrence Erlbaum Associates.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223–265. https://doi.org/10.1111/j.1751-5823.1999.tb00442.x

Capítulo IV

Pensamiento inferencial: del dato a la argumentación

Introducción

Cuando trabajamos con datos en el aula o en la investigación social, tarde o temprano aparece una pregunta que no puede responderse solo mirando una tabla o un gráfico: ¿podemos extender lo que vemos en esta muestra a un grupo más amplio? Este capítulo parte justamente de ese punto. La estadística inferencial nos permite dar el salto entre lo conocido y lo posible, entre la información que ya tenemos y las conclusiones que buscamos construir. No se trata de "adivinar" ni de confiar ciegamente en los números, sino de aprender a razonar con incertidumbre y a tomar decisiones informadas a partir de evidencias que nunca son perfectas, pero sí pueden ser suficientemente sólidas.

A lo largo de estas páginas exploraremos cómo conceptos como población, muestra, sesgo, error estándar o pruebas de hipótesis se enlazan para formar una estructura de pensamiento que nos ayuda a interpretar situaciones reales. Veremos que detrás de cada estimación hay elecciones que deben hacerse

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

con cuidado: a quién incluir, cómo medir, qué instrumento usar, cuánta variabilidad aceptar y qué tipo de pregunta se desea responder. Cuando estas decisiones se vuelven conscientes, el análisis deja de ser un procedimiento mecánico y se transforma en una forma de lectura crítica de la realidad.

Este capítulo invita a mirar los datos con una actitud reflexiva. Estimar un parámetro, construir un intervalo de confianza o contrastar una hipótesis no tiene sentido si se realiza sin interpretar el contexto que les da origen. Por eso, además de presentar las herramientas inferenciales más comunes, se propone una lectura que conecte los resultados con escenarios educativos, sociales y cotidianos. La intención es que el lector descubra que la inferencia estadística no consiste en aplicar pruebas, sino en argumentar con evidencias, reconocer límites, entender la variabilidad y, sobre todo, aprender a justificar con claridad aquello que los datos realmente permiten afirmar.

Población, muestra y sesgo: decisiones de muestreo en la práctica

Comprender cómo se define una población y cómo se selecciona una muestra es un paso decisivo para desarrollar un pensamiento inferencial sólido. La población nunca es solo un número o una lista: es una construcción conceptual que orienta la mirada sobre el fenómeno y define el alcance de la indagación. Como explica Batanero (2001), los estudiantes suelen tener dificultades para distinguir entre "lo que se quiere estudiar" y "los datos que efectivamente se tienen", por lo que enseñar a delimitar una población implica enseñar a pensar el fenómeno desde su origen.

Construyendo significado: hacia una comprensión crítica de la muestra

Desde esta perspectiva, la muestra adquiere un sentido más amplio que el de un simple subconjunto, una muestra representa una ventana hacia algo más grande; no pretende replicar la población de manera exacta, sino proporcionar una base razonable para realizar inferencias. Esto significa que seleccionar una muestra requiere preguntarse quiénes están incluidos, quiénes quedan fuera y por qué. Estos interrogantes, que pueden parecer menores, son imprescindibles para desarrollar una actitud crítica frente a los datos.

Ejemplo 1: La encuesta sobre hábitos de lectura

Supongamos que una institución quiere conocer los hábitos de lectura de los estudiantes de bachillerato.

- Población: todos los estudiantes de bachillerato de la institución.
- Muestra no probabilística: únicamente quienes asisten a la biblioteca en la hora del recreo.

Aunque puede parecer práctico, este tipo de muestra introduce un sesgo de selección, porque quienes frecuentan la biblioteca suelen tener mayor afinidad por la lectura. Este tipo de decisiones afecta de manera directa la validez del estudio, ya que se genera una imagen distorsionada del fenómeno.

En la práctica educativa también resulta útil distinguir entre métodos probabilísticos y no probabilísticos. Los primeros buscan garantizar que todos los individuos de la población tengan alguna posibilidad conocida de ser seleccionados, mientras que los segundos suelen basarse en la accesibilidad y la disponibilidad.

Ejemplo 2: Muestreo aleatorio simple en una clase

Un docente desea conocer el nivel de satisfacción de sus estudiantes con el trabajo en proyectos. En lugar de encuestar solo a quienes llegaron temprano o levantaron la mano (muestra sesgada), decide:

1. Numerar a todos los estudiantes del curso.
2. Usar una aplicación para seleccionar 10 números al azar.

Este procedimiento reduce la influencia de preferencias personales, tiempos de llegada o motivaciones individuales. Horton (2015) señala que mostrar estos pasos a los estudiantes complementa la enseñanza del muestreo con una dimensión ética y metodológica.

Imaginemos ahora que se quiere estudiar la percepción sobre el uso de tecnología en el aula en un colegio con una proporción desigual entre hombres y mujeres. Si se toma una muestra aleatoria simple, pueden quedar sobre o subrepresentados ciertos grupos, lo que afectaría la posibilidad de comparar experiencias o actitudes entre géneros. El muestreo estratificado permite organizar la población por género y seleccionar al azar dentro de cada estrato, de manera que las voces de todos los grupos aparezcan en la proporción que les corresponde.

El concepto de sesgo también juega un papel central en este epígrafe. No se trata de un error accidental, sino de una deformación estructural que compromete la inferencia y puede conducir a conclusiones equivocadas sobre el comportamiento de una población. Para Biehler (2018), visibilizar el sesgo fortalece la transparencia metodológica, favorece posturas críticas y ayuda a comprender que la producción de datos no es neutral. En investigaciones educativas, reconocer posibles fuentes de sesgo no solo mejora la validez del estudio, sino que forma a los estudiantes en una mirada más ética y reflexiva sobre el análisis de datos.

El sesgo de no respuesta aparece con frecuencia en estudios educativos que utilizan formularios digitales o encuestas en línea. Imaginemos a un profesor que desea conocer la calidad de acceso a internet en los hogares de sus estudiantes. Aunque el formulario se envía a toda la clase, tienden a responder quienes tienen buena conectividad, mientras que aquellos con acceso limitado no logran completarlo o directamente no pueden abrirlo. Este tipo de sesgo es especialmente problemático porque excluye precisamente a quienes enfrentan las mayores dificultades, generando un panorama distorsionado sobre la brecha digital. Así, lo que aparenta ser un problema de "baja participación" se transforma en una amenaza real para la validez de las conclusiones.

Durante los últimos años, Gelman (2021) ha subrayado que la selección de muestras debe entenderse dentro de un ecosistema más amplio de inferencia. No basta con pensar en cuántos estudiantes responden; es necesario comprender cómo cada decisión metodológica influye en la calidad del análisis estadístico. Trabajar con múltiples muestras, compararlas y examinar sus diferencias no es un error metodológico, sino una oportunidad pedagógica para mostrar que la variación entre muestras es inherente al proceso de recolección de datos. Cuando los estudiantes analizan estos contrastes, desarrollan una comprensión más profunda de por qué las conclusiones no siempre coinciden incluso cuando se estudia la misma población.

La comparación de dos muestras distintas permite ilustrar con claridad cómo el sesgo afecta las inferencias. Supongamos que dos grupos de trabajo desean estudiar el tiempo que los estudiantes dedican a estudiar matemáticas. El primero selecciona una muestra completamente al azar, mientras que la segunda encuesta únicamente a quienes suelen entregar tareas puntualmente. Aunque ambos grupos investigan a la misma población, sus conclusiones inevitablemente divergen: el segundo grupo encontrará tiempos de estudio más altos debido a la selección sesgada.

Caso de estudio: Acceso digital, tiempo de estudio y sesgos de muestreo

El análisis conjunto del acceso digital, el tiempo de estudio y los distintos sesgos de muestreo ofrece un escenario ideal para que los estudiantes comprendan cómo la calidad de los datos influye en la validez de las conclusiones estadísticas. En este caso de estudio partimos de una situación cercana a la realidad educativa: no todos los estudiantes tienen las mismas condiciones de conectividad ni las mismas oportunidades de participar en una encuesta, y estas diferencias pueden sesgar los resultados sin que el investigador lo advierta.

A través de la simulación y el contraste entre distintos tipos de muestra, el epígrafe permite observar cómo se modifican las distribuciones, las tendencias y las relaciones entre variables cuando el muestreo no se diseña cuidadosamente. Esta mirada integradora ayuda a reconocer que la estadística no solo describe datos sino que evalúa los procesos que los originan, invitando a desarrollar una actitud crítica y reflexiva frente a la información que usamos para argumentar en contextos educativos.

Una universidad quiere analizar cómo se relacionan el acceso a internet en el hogar, el tiempo semanal dedicado a estudiar matemáticas y la percepción sobre el uso de tecnología en clase. La cohorte está formada por 300 estudiantes de primer año: 60 por ciento mujeres y 40 por ciento hombres. Los responsables del estudio deciden:

1. Definir la población y estratos

- Población: los 300 estudiantes matriculados en Matemáticas
- Estratos: género (mujer, hombre).
- Variables principales:
  - genero (Mujer/Hombre)
  - acceso (bueno/limitado)
  - tiempo_estudio (horas de estudio de matemáticas por semana)
  - uso_tecnologia (escala 1–5 de acuerdo con el uso de recursos digitales en clase).

2. Plan de muestreo estratificado por género

- Se selecciona una muestra de 120 estudiantes manteniendo la proporción de la población: 72 mujeres y 48 hombres.
- Dentro de cada estrato, la selección es aleatoria simple.
- Objetivo: estimar el promedio de horas de estudio y la media de uso_tecnologia para cada género, comparando sus intervalos de confianza.

3. Introducción del sesgo de no respuesta

- El cuestionario se aplica en línea. Responden con mayor probabilidad quienes tienen buen acceso y con menor probabilidad quienes tienen acceso limitado.
- En la simulación se puede modelar, por ejemplo, que:

a. Estudiantes con acceso bueno responden con proba bilidad 0.85.
b. Estudiantes con acceso limitado responden con probabilidad 0.40.

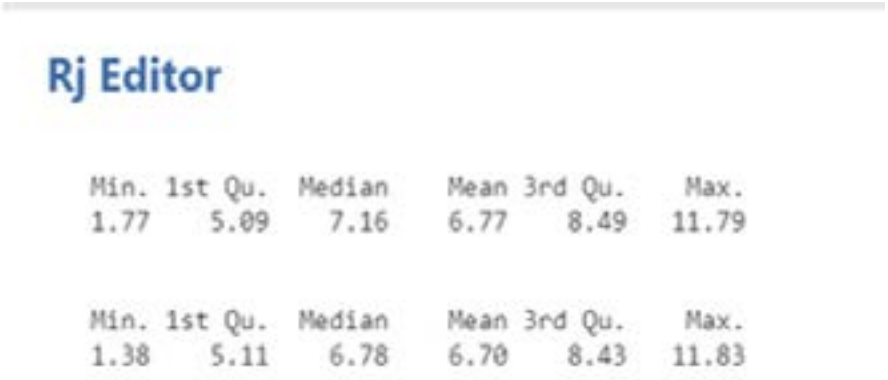
- Esto generará una muestra "realmente observada" donde los estudiantes con dificultades de conectividad quedan subrepresentados, lo que afectará la estimación de tiempo_estudio y uso tecnología.

4. Comparación de dos muestras distintas sobre la misma población

- Grupo A (muestra cuidadosamente diseñada): usa el plan estratificado anterior y controla el sesgo de no respuesta (por ejemplo, llamando por teléfono a quienes no contestan).
- Grupo B (muestra sesgada): difunde el enlace del cuestionario solo por el aula virtual y toma como muestra "a quienes respondan", sin control adicional.
- Ambos grupos calculan medias, desviaciones estándar e intervalos de confianza de tiempo_estudio y uso_tecnologia. Posteriormente comparan resultados y discuten cómo el diseño muestral y el sesgo de no respuesta han influido en las conclusiones.

La Figura 1 muestra que, en términos generales, los dos tipos de muestra producen distribuciones de tiempo de estudio muy parecidas, pero con matices que vale la pena mirar con lupa. En la primera fila, correspondiente a la muestra estratificada observada, la media (6.77 horas) y la mediana (7.16 horas) indican que la mayoría de estudiantes se sitúa alrededor de las 7 horas semanales de estudio, con un 50 por ciento de los valores entre aproximadamente 5.09 y 8.49 horas.

Figura 1.
Estadísticos descriptivos del tiempo de estudio en dos tipos de muestra.



Nota. La primera fila corresponde a la muestra estratificada observada, afectada por el sesgo de no respuesta según las probabilidades definidas en la simulación. La segunda fila representa la muestra espontánea, recolectada únicamente entre quienes respondieron voluntariamente.

En la segunda fila, que representa la muestra espontánea recolectada solo entre quienes respondieron voluntariamente, la media (6.70) y la mediana (6.78) son ligeramente menores y el rango intercuartílico va de 5.11 a 8.43 horas. Es decir, los valores centralesson cercanos, pero la muestra estratificada tiende a

concentrar un poco más el tiempo de estudio en torno a valores algo superiores, mientras que la muestra espontánea se "desplaza" levemente hacia abajo. Estas pequeñas diferencias, aunque numéricamente modestas, son importantes desde el punto de vista metodológico. La muestra estratificada intenta respetar la estructura de la población y controlar el sesgo de no respuesta, por lo que ofrece una imagen más equilibrada del tiempo real que el conjunto del estudiantado dedica a matemáticas.

La muestra espontánea, en cambio, depende de quién decide contestar el cuestionario, de modo que ciertos perfiles, por ejemplo: estudiantes con menos tiempo o con dificultades de acceso, pueden quedar infrarrepresentados o sobrerrepresentados. El resultado es que dos estudios que "miden lo mismo" y analizan la misma población terminan con resúmenes descriptivos similares, pero no idénticos, recordándonos que la forma en que se selecciona a los participantes condiciona las conclusiones que podemos sacar de los datos.

El gráfico (Figura 2) revela que ambas muestras presentan una distribución similar en torno al tiempo de estudio, aunque con matices que reflejan el impacto del diseño muestral.

Figura 2.
Distribución del tiempo de estudio según el tipo de muestra

Nota. La figura compara la variación del tiempo semanal dedicado al estudio entre dos estrategias de muestreo.

En la muestra estratificada se aprecia una mayor estabilidad alrededor de los valores centrales, lo que sugiere que esta estrategia logra capturar mejor la diversidad real de la población. La muestra espontánea, en cambio, tiende a mostrar mayor dispersión y una ligera caída en los valores centrales, lo que se relaciona con la subrepresentación de estudiantes con menor acceso tecnológico y tiempos de estudio más bajos.

Los histogramas (Figura 3) permiten apreciar diferencias claras en la forma de las distribuciones entre las dos muestras. En el panel correspondiente a la muestra estratificada se observa una distribución más equilibrada, con varios picos moderados que reflejan la diversidad real de hábitos de estudio dentro de la población original.

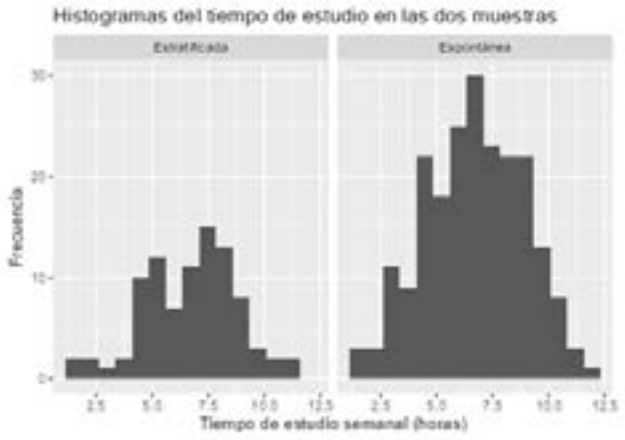
En contraste, la muestra espontánea muestra una mayor acumulación de estudiantes en los valores centrales, lo que indica una tendencia hacia la homogeneización causada por el sesgo de respuesta: quienes estudian más horas o tienen mejor conectividad tienden a participar en mayor proporción.

Este contraste evidencia cómo dos muestras que provienen de la misma población pueden ofrecer retratos distintos del fenómeno cuando la selección de participantes no garantiza igualdad de oportunidades para responder.

Figura 3.
Histogramas del tiempo de estudio en las dos muestras

Rj Editor

```
'data.frame':   303 obs. of  6 variables:
 $ id             : int  109 273 254 217 284 123 266 153 39 41 ...
 $ genero         : Factor w/ 2 levels "Hombre","Mujer": 2 2 2 2 2 2 2 2 2 2
 $ acceso         : Factor w/ 2 levels "Bueno","Limitado": 1 1 2 1 1 1 1 1
 $ tiempo_estudio: num  5.92 5.44 2.69 6.74 8 ...
 $ uso_tecnologia: num  5 3 3 4 4 5 2 4 4 5 ...
 $ tipo_muestra   : Factor w/ 2 levels "Estratificada",..: 1 1 1 1 1 1 1 1
```

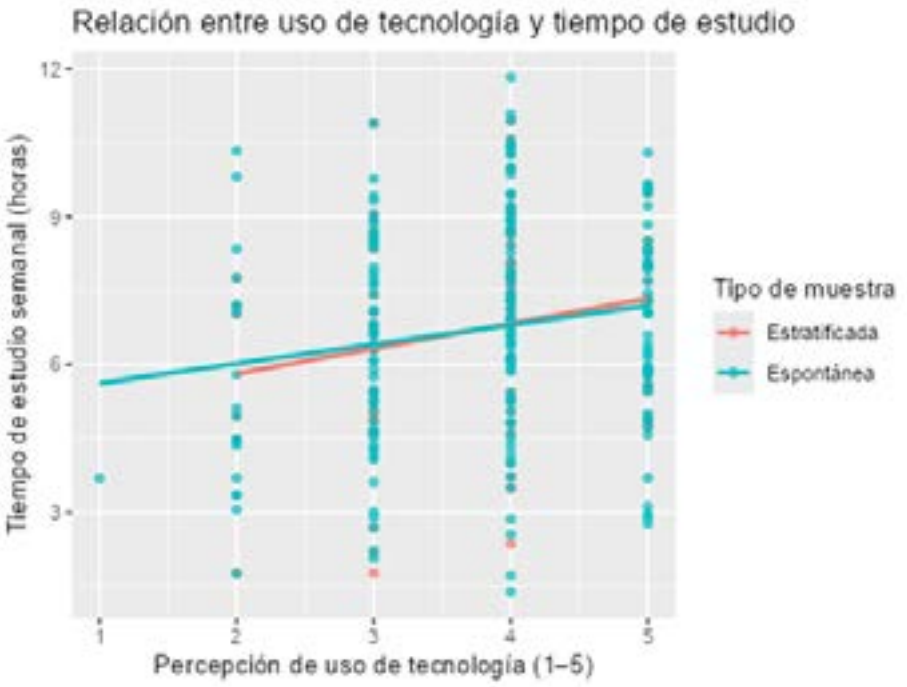


Nota. La figura presenta la distribución del tiempo semanal dedicado al estudio en dos estrategias de muestreo: una muestra estratificada observada y una muestra espontánea recolectada sin control de no respuesta.

La figura 4 evidencia una relación positiva, aunque moderada, entre la percepción del uso de tecnología en el aula y el tiempo dedicado al estudio: a mayor valoración del uso tecnológico, los

estudiantes tienden a reportar más horas de trabajo académico semanal. Sin embargo, la pendiente ligeramente más pronunciada en la muestra espontánea sugiere que la ausencia de control sobre la no respuesta puede magnificar la relación observada, posiblemente porque los estudiantes más motivados o con mejor conectividad participan en mayor proporción en ese tipo de muestreo. En la muestra estratificada, la línea de tendencia es más suave, reflejando un patrón más equilibrado y representativo de la población original.

Figura 4.
Relación entre uso de tecnología y tiempo de estudio



Nota. La figura muestra la relación entre la percepción del uso de tecnología en clase (escala de 1 a 5) y el tiempo semanal dedicado al estudio, diferenciada por tipo de muestra.

La experiencia con estas simulaciones ofrece al docente un punto de partida poderoso para que los estudiantes comprendan que el análisis de datos no es un ejercicio mecánico, sino un proceso que exige decisiones razonadas sobre la forma en que se recolecta la información.

Al comparar muestras estratificadas, espontáneas y afectadas por distintos tipos de sesgo, el profesorado puede mostrar que los resultados numéricos dependen de cómo se selecciona a los participantes. Esta constatación ayuda a que los estudiantes entiendan que la estadística no solo describe datos, sino que

también evalúa la calidad de esos datos, y que instrumentos como los diagramas de caja, histogramas y gráficos de dispersión son herramientas para interpretar críticamente el proceso de muestreo.

Apoyo didáctico: Desde una perspectiva didáctica, trabajar estos contenidos en el aula permite fomentar la reflexión sobre la validez de las conclusiones y el carácter ético de la producción de información. El docente puede guiar actividades en las que los estudiantes diseñen sus propias encuestas, identifiquen posibles fuentes de sesgo y analicen cómo estos influyen en los resultados. Esto no solo desarrolla habilidades técnicas en el uso de software como Jamovi, sino que fortalece el pensamiento inferencial: la capacidad de cuestionar, argumentar y justificar decisiones metodológicas. En definitiva, el abordaje de estos contenidos contribuye a formar estudiantes más críticos, conscientes de la importancia de la representatividad y capaces de tomar decisiones informadas en contextos reales de análisis de datos.

Estimación puntual y por intervalos: comprensión del error y la precisión

En el análisis de datos educativos, es habitual que docentes e investigadores trabajen con muestras y no con la totalidad de la población. Esto implica tomar decisiones a partir de información parcial, reconociendo que cada estadístico que calculamos es una aproximación y no un valor definitivo. Tal como explica Batanero y Ben-Zvi (2013), razonar estadísticamente exige comprender que el dato muestral no es una copia del dato poblacional, sino una expresión condicionada por la variabilidad inherente al muestreo. Garfield y delMas (2008) señalan que muchos errores conceptuales surgen cuando los estudiantes interpretan la media o la proporción muestral como si fueran verdades exactas, sin considerar cuánto podrían variar esos valores si se repitiera el estudio.

Por ello, este epígrafe desarrolla la estimación puntual, el error estándar y los intervalos de confianza como herramientas complementarias para razonar con incertidumbre. A partir de ejemplos concretos y del aporte de investigadores en educación estadística. El propósito es que el docente pueda explicar por qué un intervalo de confianza no es un "decorado", sino un instrumento esencial para comunicar precisión, reconocer límites y tomar decisiones fundamentadas.

Estimación puntual de medias y proporciones: la primera aproximación al parámetro

La estimación puntual consiste en utilizar un valor único, por lo general la media o la proporción muestral, para aproximar un

parámetro poblacional. En la práctica educativa, este procedimiento suele presentarse como un primer acercamiento al razonamiento inferencial, porque permite visualizar cómo un dato resumen puede representar, de manera tentativa, a un conjunto mucho más amplio. Sin embargo, esta misma simplicidad es la que suele provocar malentendidos iniciales: muchas y muchos estudiantes interpretan ese valor puntual como si fuera la "verdadera" realidad de la población, sin considerar la inevitable variabilidad asociada al muestreo.

Según Batanero y Ben-Zvi (2013), es justamente en esta etapa donde emergen las primeras confusiones en el aula. El estudiantado, especialmente cuando no ha trabajado con muestras múltiples, tiende a asumir que la media o proporción obtenida es un reflejo exacto de la población, como si el proceso de muestreo no introdujera ninguna variación. Esta percepción es comprensible, pues el pensamiento intuitivo suele preferir certezas, y la idea de que diferentes muestras puedan producir resultados distintos se percibe, al inicio, como una especie de "error" o inconsistencia del procedimiento estadístico.

Garfield y delMas (2008) enfatizan que antes de avanzar hacia los intervalos de confianza o hacia técnicas inferenciales más complejas, es indispensable que el estudiantado experimente la variabilidad muestral de manera concreta. Observar cómo dos o más muestras extraídas de la misma población generan estimaciones puntuales diferentes ayuda a comprender que estas no son verdades absolutas, sino aproximaciones sujetas al azar. Este paso pedagógico, aunque sencillo, es crucial: constituye la base para entender por qué una estimación siempre debe ir acompañada de una medida de incertidumbre y por qué la inferencia estadística, más que ofrecer certezas, nos permite razonar con niveles de precisión y confianza.

Ejemplo 3. Estimación puntual de una media y una proporción

Una docente encuesta a 12 estudiantes sobre su tiempo semanal de estudio y si poseen acceso estable a internet.

- Horas de estudio: 4, 5, 6, 5, 7, 8, 6, 5, 6, 7, 4, 6
- Acceso estable (1 = sí, 0 = no): 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1

Media muestral: $\bar{x} = \frac{69}{12} \approx 5.75 \text{horas}$

Proporción muestral: $\hat{p} = \frac{9}{12} = 0.75$

Estos valores representan estimaciones puntuales del tiempo medio de estudio y de la proporción de estudiantes con acceso estable en la población. Aunque son útiles como una primera

aproximación, no deben interpretarse como descripciones definitivas de la realidad poblacional. En esencia, funcionan como una fotografía tomada desde un ángulo particular: ofrecen información, pero no capturan la totalidad del fenómeno.

Sin embargo, como advierte Bakker (2004), una segunda muestra, seleccionada bajo los mismos criterios y con el mismo procedimiento, podría arrojar valores distintos. Esta variabilidad no implica un error en el cálculo ni una falla metodológica, sino una característica natural del trabajo con datos muestrales. Cada muestra recoge una porción diferente de la población y, por tanto, refleja matices propios de su composición.

Ese diferencial recibe el nombre de error muestral, una consecuencia inevitable de todo proceso de selección y un recordatorio de que nuestras inferencias siempre llevan consigo un margen de incertidumbre. Comprenderlo es fundamental en el aula: ayuda a que el estudiantado no se quede únicamente con el valor puntual, sino que reconozca que toda estimación es, en realidad, una aproximación sujeta al azar y que solo adquiere sentido cuando se analiza junto con la variabilidad que la acompaña.

Ejemplo 4. Comparando dos cursos para comprender la estimación puntual

Una institución educativa quiere explorar si existen diferencias iniciales entre dos cursos paralelos de primer año respecto a su dedicación académica y a las condiciones tecnológicas con las que estudian. Para ello, una docente decide tomar una muestra piloto de 10 estudiantes por curso, con el fin de realizar una primera aproximación a los parámetros poblacionales. La intención no es llegar a conclusiones definitivas, sino construir una mirada preliminar que oriente futuras decisiones pedagógicas.

Datos recolectados

Curso A

Horas semanales de estudio: 6, 5, 7, 4, 5, 6, 8, 7, 5, 6

Acceso estable a internet (1 = sí, 0 = no): 1, 1, 1, 0, 1, 1, 1, 1, 0, 1

Curso B

Horas semanales de estudio: 3, 4, 5, 4, 3, 5, 4, 3, 4, 5

Acceso estable a internet (1 = sí, 0 = no): 0, 1, 0, 0, 1, 0, 1, 0, 0, 1

1. Cálculo de la media muestral en cada curso

Para cada curso se calcula la media de horas de estudio:

$$\overline{\mathbf{x}}_{\mathbf{A}} = \frac{\mathbf{6+5+7+4+5+6+8+7+5+6}}{\mathbf{10}} = \frac{\mathbf{59}}{\mathbf{10}} = \mathbf{5.9 horas}$$

$$\overline{\mathbf{x}}_{\mathbf{B}} = \frac{\mathbf{3+4+5+4+3+5+4+3+4+5}}{\mathbf{10}} = \frac{\mathbf{40}}{\mathbf{10}} = \mathbf{4.0 horas}$$

Estas medias sugieren que, en promedio, el Curso A dedica más tiempo al estudio que el Curso B. Sin embargo, siguiendo las advertencias de Batanero y Ben-Zvi (2013), estas cifras no deben asumirse como descripciones definitivas de la población. Una nueva muestra podría mostrar valores diferentes o incluso invertir el patrón. Se trata de un primer vistazo a la posible realidad de ambos grupos.

2. Cálculo de la proporción muestral de acceso estable

Como la variable de acceso toma valores 0 y 1, la proporción de estudiantes con acceso estable en cada curso se calcula dividiendo el número de estudiantes con valor 1 para esa variable para el total de estudiantes del curso.

En el Curso A, de los 10 estudiantes, 8 declaran tener acceso estable:

$$\widehat{\mathbf{p}}_{\mathbf{A}} = \frac{\mathbf{8}}{\mathbf{10}} = \mathbf{0.80}$$

En el Curso B, solo 4 de los 10 estudiantes tienen acceso estable:

$$\widehat{\mathbf{p}}_{\mathbf{B}} = \frac{\mathbf{4}}{\mathbf{10}} = \mathbf{0.40}$$

La interpretación inicial es que la proporción de estudiantes con acceso estable a internet en el Curso A es aproximadamente el doble que en el Curso B. De nuevo, se trata de estimaciones puntuales que dependen de la muestra seleccionada y que podrían variar si se repitiera el muestreo con otro grupo de estudiantes.

La resolución del ejercicio en R (Figura 5) confirma de manera precisa los valores obtenidos mediante el cálculo manual. Al agrupar los datos se obtiene que la media de horas de estudio es de 5.9 horas en el Curso A y 4.0 horas en el Curso B, reproduciendo exactamente las fracciones 59/1059/1059/10 y 40/1040/1040/10 derivadas previamente. Del mismo modo, el tratamiento de la variable acceso como binaria (0 = no, 1 = sí) permite calcular la proporción de acceso estable a internet como la media de dicha variable, dando como resultado 0.80 para el Curso A y 0.40 para el Curso B.

Estas salidas muestran cómo R facilita el procesamiento de datos y valida los procedimientos estadísticos fundamentales: la estimación puntual de una media y de una proporción. Además, el uso de dplyr permite generar una tabla resumen más completa, desde la cual se aprecia que la estructura de los datos coincide con la teoría presentada y que las estimaciones puntuales pueden reproducirse y verificarse con herramientas computacionales que favorecen la comprensión del proceso inferencial desde una perspectiva más aplicada.

Figura 5.
Relación entre uso de tecnología y tiempo de estudio

```
Medias por curso con aggregate:

  curso horas
1     A   5.9
2     B   4.0


Proporciones de acceso (media de 0/1) con aggregate:

  curso acceso
1     A    0.8
2     B    0.4



Tabla resumen por curso:

# A tibble: 2 × 6
  curso     n suma_horas media_horas suma_acceso_estable proporcion_acceso
  <fct> <int>      <dbl>       <dbl>                <dbl>             <dbl>
1 A        10         59         5.9                    8               0.8
2 B        10         40         4                      4               0.4
```

Nota. La figura muestra las estimaciones puntuales obtenidas a partir de una muestra de 10 estudiantes por curso.

Error estándar e intervalos de confianza: comunicar la incertidumbre

Una vez obtenidos los valores puntuales, la pregunta relevante es: ¿qué tan estables son estas estimaciones? En otras palabras, ¿podemos confiar en que la media o la proporción obtenida reflejan, al menos de manera aproximada, lo que ocurre en la población? El error estándar ofrece una primera pista para responder a esa inquietud, porque indica cuánto podrían fluctuar las estimaciones si repitiéramos el muestreo una y otra vez bajo las mismas condiciones.

A partir de esta idea surge una herramienta clave para el razonamiento inferencial: los intervalos de confianza. Más que presentar un único valor como si fuera definitivo, estos intervalos ofrecen un rango plausible donde podría encontrarse el parámetro poblacional. Para el estudiantado, comprender esta transición del valor puntual al intervalo, es crucial para desarrollar pensamiento estadístico maduro. Ya no se trata solo de "calcular una media", sino de reconocer que toda estimación es parte de un proceso sujeto al azar y que la incertidumbre no es un defecto del método, sino una característica inherente de la realidad cuando trabajamos con datos muestrales.

Ejemplo 5: una muestra de 25 estudiantes registra un promedio de 6,2 horas de estudio semanal, con una desviación estándar de 1,5 horas, lo que refleja una variabilidad moderada entre los

tiempos reportados. Con este tamaño muestral, la estimación de la media se vuelve razonablemente estable, permitiendo construir un intervalo de confianza que no solo ofrece un valor central, sino también un rango plausible donde podría encontrarse la media poblacional real. Este tipo de análisis resulta especialmente útil en contextos educativos, porque ayuda a interpretar la información más allá de un único número y a comprender hasta qué punto el promedio observado puede generalizarse a un grupo más amplio.

- Error estándar: $\mathbf{SE} = \frac{\mathbf{1.5}}{\sqrt{\mathbf{25}}} = \mathbf{0.30}$
- Margen de error: $\mathbf{ME} = \mathbf{2.064} \times \mathbf{0.30} \approx \mathbf{0.62}$
- Intervalo de confianza **(95 %): [5,58; 6,82]**

Siguiendo la explicación de Gelman y Hill (2014), vale la pena insistir en que un intervalo de confianza no debe tomarse como si asegurara que "la media verdadera está ahí dentro". Lo fundamental no es ese intervalo en sí, sino el proceso con el que se construye. Si tuviéramos la posibilidad de repetir el estudio una y otra vez, recogiendo nuevas muestras y calculando un nuevo intervalo cada vez, descubriríamos que aproximadamente el 95 % de ellos sí terminaría conteniendo el valor real del parámetro. Ese es el sentido de hablar de "confianza": no es que este intervalo particular tenga una probabilidad del 95 % de ser correcto, sino que el método funciona bien en el largo plazo. Entenderlo así nos ayuda a evitar lecturas exageradas y a ver el intervalo como lo que realmente es: una expresión de la estabilidad del procedimiento frente a la variabilidad natural de los datos.

Ejemplo 6. Una docente desea comprender cuán estable es la estimación del tiempo que sus estudiantes dedican al estudio semanal. Para ello, toma una muestra de 30 estudiantes y obtiene una media de 6,8 horas con una desviación estándar de 1,9 horas. A partir de estos valores, calcula el error estándar, dividiendo la desviación estándar entre la raíz cuadrada del tamaño muestral:

$$\mathbf{EE} = \frac{\mathbf{1.9}}{\sqrt{\mathbf{30}}} \approx \mathbf{0.35}$$

Este valor indica que, si el docente repitiera el muestreo bajo las mismas condiciones, la media muestral podría oscilar aproximadamente ±0,35 horas alrededor del valor observado. Con esta información, construye un intervalo de confianza del 95 %, utilizando la fórmula habitual $\bar{x} \pm 1.96 \times EE$. El resultado es:

$6{,}8 \pm 1.96(0{,}35) = [6{,}1 \text{ horas}, 7{,}5 \text{ horas}]$.

Estos resultados se pueden comprobar en R tal como muestra la gráfica (Figura 5) que ilustra la idea central del intervalo de confianza desde una perspectiva geométrica. El punto ubicado en el centro corresponde a la media muestral de 6,8 horas, obtenida a partir de la muestra. A su izquierda y derecha se muestran los límites inferiores (6,12) y superior (7,48), conectados mediante un segmento horizontal que representa el intervalo de confianza del 95 %.

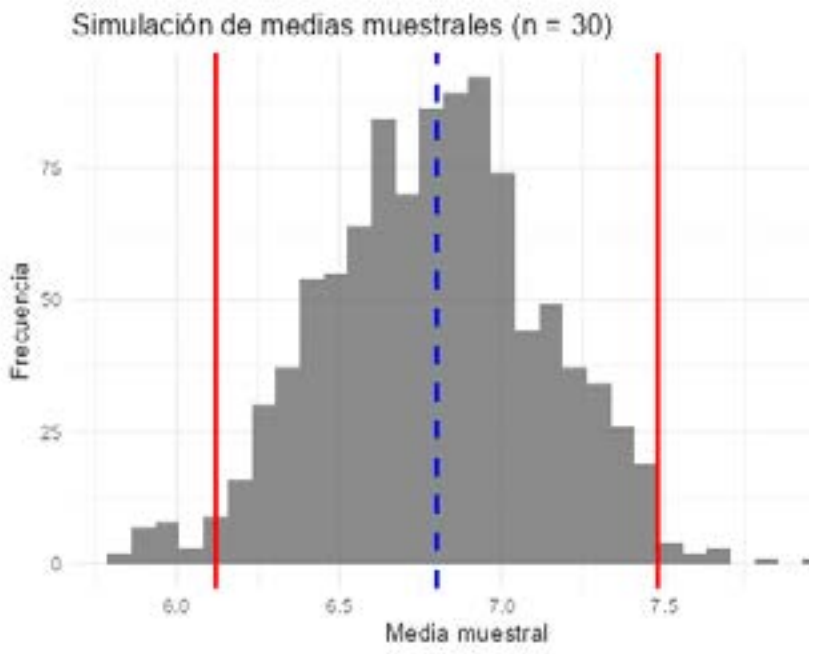
Figura 6.
Representación geométrica de la media muestral y su intervalo de confianza del 95 %

Nota. La figura muestra la media muestral de 6,8 horas ubicada en el centro de la recta, acompañada del intervalo de confianza del 95 %, cuyos límites inferiores y superiores son aproximadamente 6,12 y 7,48 horas, respectivamente.

Esta visualización permite comprender que la media muestral no debe interpretarse como un valor exacto o definitivo, sino como el punto más representativo dentro de un rango más amplio de posibles valores. El intervalo funciona, así como una "zona plausible" donde, considerando la variabilidad propia del muestreo, es razonable suponer que se encuentra la media poblacional. Cuanto más corto sea este segmento, mayor será la precisión de la estimación; en cambio, un intervalo más extenso indicaría mayor incertidumbre. Esta representación gráfica facilita que estudiantes y docentes visualicen la incertidumbre estadística como una distancia sobre la recta numérica, haciendo más accesible el concepto de error de estimación.

La gráfica (Figura 6) representa la distribución de 1000 medias muestrales obtenidas mediante simulación, lo que permite visualizar cómo varían las medias cuando se repite el proceso de muestreo muchas veces bajo las mismas condiciones. El histograma muestra que la mayoría de las medias simuladas se concentran alrededor del valor central de 6,8 (indicado por la línea azul punteada), lo cual refleja el principio de que el promedio de las medias muestrales tiende a aproximarse a la media poblacional verdadera. Las líneas rojas delimitan el intervalo de confianza del 95 % construido a partir de una sola muestra, y es posible observar que gran parte de las medias simuladas cae dentro de ese rango.

Figura 7.
Distribución simulada de medias muestrales y ubicación del intervalo de confianza del 95 %



Nota. La figura muestra la distribución de 1000 medias muestrales obtenidas mediante simulación bajo una población teórica con media 6,8 y desviación estándar 1,9, utilizando un tamaño de muestra de 30.

Esta correspondencia visual evidencia que, si bien cada media puede fluctuar debido al azar del muestreo, el intervalo de confianza suele capturar con alta probabilidad el valor real del parámetro.

En conjunto, la figura ilustra que el intervalo no es un dato aislado, sino una representación geométrica de la incertidumbre, basada en cómo se comportarían muchas posibles muestras tomadas de la misma población.

En síntesis, el estudio del error estándar y de los intervalos de confianza permite que el estudiantado comprenda que toda estimación basada en una muestra está inevitablemente sujeta a variación. Más que ofrecer resultados exactos, estas herramientas ayudan a valorar la estabilidad de los cálculos y a reconocer que el promedio observado no es un punto fijo, sino una aproximación influida por el azar del muestreo. Visualizar esta incertidumbre ya sea a través de rectas numéricas o de simulaciones con múltiples muestras, facilita que las y los estudiantes desarrollen un sentido más profundo de la inferencia estadística. De este modo, la enseñanza no se limita al cálculo mecánico, sino que invita a interpretar los resultados con cautela, a estimar rangos plausibles y a fundamentar afirmaciones basadas en evidencia más sólida y reflexiva.

Variabilidad, tamaño muestral y precisión: comprender qué significa "estimar bien"

La amplitud de un intervalo de confianza depende directamente de la variabilidad de los datos y del tamaño muestral, dos elementos que determinan cuánta información aportan realmente la muestra sobre la población. Horton (2015) recuerda que una estimación más precisa no implica necesariamente que sea correcta, sino que refleja una menor dispersión alrededor del valor observado. En otras palabras, la precisión habla de estabilidad, no de veracidad absoluta.

Esta distinción es clave en la enseñanza, porque permite al estudiantado comprender que incluso un intervalo muy estrecho sigue siendo una aproximación influida por el azar. En la misma línea, Biehler (2018) recomiendan trabajar con representaciones visuales de intervalos superpuestos, de modo que el estudiantado pueda observar de manera intuitiva cómo los intervalos se vuelven más cortos a medida que aumenta el tamaño de la muestra. Este enfoque visual ayuda a entender por qué una muestra grande no garantiza una estimación perfecta, pero sí una mayor consistencia entre muestras sucesivas, fortaleciendo así la interpretación crítica de la incertidumbre estadística.

Ejemplo 7: Un ejemplo sencillo permite apreciar esta idea. Supongamos que una docente mide el tiempo de lectura diaria de dos grupos diferentes. En el primer caso, toma una muestra pequeña de 10 estudiantes y obtiene una media de 25 minutos con un intervalo de confianza del 95 % entre 18 y 32 minutos.

En el segundo caso, repite el estudio con una muestra mucho mayor, de 80 estudiantes, y obtiene una media muy similar, de 24 minutos, pero con un intervalo más estrecho: de 22 a 26 minutos.

Aunque ambas muestras proporcionan valores cercanos, los intervalos muestran historias distintas. El primero es amplio y refleja gran incertidumbre: la media poblacional podría estar bastante por encima o por debajo del valor observado. El segundo, en cambio, ofrece un rango mucho más preciso, mostrando que con más datos la estimación se vuelve más estable y confiable. Para el estudiantado, visualizar y comparar estos dos intervalos facilita comprender por qué el tamaño muestral influye tanto en la precisión y por qué no basta con un solo valor puntual para interpretar un fenómeno con rigor estadístico.

La figura 7 permite ver cómo cambia la precisión de una estimación cuando se trabaja con muestras de distinto tamaño. En el primer caso, con solo diez estudiantes, el error estándar es bastante alto y el intervalo de confianza termina siendo muy ancho, lo que deja claro que la media obtenida puede alejarse bastante del valor real en la población. En cambio, cuando se amplía la muestra a ochenta estudiantes, la situación mejora de forma notable: el error estándar disminuye y el intervalo se estrecha, lo que indica que la media calculada es mucho más estable y confiable.

Figura 8.
Comparación del error estándar, desviación estándar e intervalos de confianza en muestras pequeñas y grandes.

Nota. La figura muestra los valores de error estándar, desviación estándar e intervalos de confianza reconstruidos para dos escenarios hipotéticos: un primer caso con una muestra pequeña de 10 estudiantes y un segundo caso con una muestra grande de 80 estudiantes.

Esta comparación ayuda a entender un punto esencial en estadística: no basta con obtener una media, sino que es importante reconocer cuánta incertidumbre la acompaña y cómo

esa incertidumbre se reduce a medida que se cuenta con más información. El desarrollo de este epígrafe permite comprender que la estimación puntual y la estimación por intervalos son dos caras complementarias de un mismo proceso: aproximarnos al conocimiento de una población asumiendo con claridad la incertidumbre que implica trabajar con muestras. La media, la proporción u otros estadísticos muestrales constituyen un primer acercamiento al parámetro, pero solo adquieren sentido pleno cuando se analizan junto con el error estándar y los intervalos de confianza, que cuantifican la variabilidad inherente al muestreo y nos ofrecen un rango plausible para interpretar los resultados. Los ejemplos desarrollados muestran de forma concreta cómo la precisión aumenta al incrementar el tamaño muestral, pero también evidencian que la exactitud de una estimación depende tanto de la variabilidad como de la calidad del diseño de muestreo. De este modo, "estimar bien" no consiste únicamente en aplicar fórmulas, sino en valorar críticamente los supuestos, las condiciones y los límites del procedimiento utilizado.

Desde una perspectiva didáctica, el epígrafe subraya la importancia de enseñar la inferencia estadística como un proceso argumentativo y no como un conjunto de cálculos mecánicos. Hacer visible la incertidumbre, comparar muestras, construir intervalos y discutir su interpretación ayuda al estudiantado a reconocer que la estadística es una herramienta para razonar con evidencia, no para producir verdades absolutas. Finalmente, comprender el papel de la variabilidad, la precisión y la incertidumbre fortalece una alfabetización estadística crítica, capaz de guiar decisiones educativas informadas y coherentes con los retos contemporáneos de análisis de datos.

Pruebas de hipótesis: sentido, pasos y lectura crítica

En el ámbito educativo es frecuente encontrarse con preguntas que requieren más que una simple descripción de datos. El docente quiere saber si una intervención realmente mejoró el rendimiento de su grupo, si dos metodologías producen efectos distintos sobre el aprendizaje, o si un programa de formación genera cambios significativos en las actitudes del estudiantado. En cada uno de estos casos, observar diferencias en los datos no es suficiente: se necesita un modo de evaluar si dichas diferencias pueden atribuirse a la intervención y no al azar propio del muestreo. Este es el papel que cumplen las pruebas de hipótesis, un tipo de razonamiento inferencial que nos permite valorar la evidencia y tomar decisiones fundamentadas.

Las pruebas de hipótesis son uno de los contenidos más difíciles para el estudiantado, no porque las fórmulas sean complejas, sino porque exigen comprender ideas abstractas como azar,

variabilidad, evidencia y razonamiento condicional. Por ello, el reto del docente es presentar este tema de manera cercana, conceptual y gradual, mostrando que el contraste de hipótesis no es un algoritmo, sino una forma de pensar con datos.

El propósito inferencial: por qué contrastamos hipótesis en educación

La finalidad de una prueba de hipótesis no es "probar que algo es verdad", sino evaluar si la evidencia empírica es suficientemente fuerte como para cuestionar una afirmación inicial. Esa afirmación inicial se denomina hipótesis nula (H_0) y representa un punto de referencia técnico: igualdad de medias, ausencia de cambio o estabilidad de proporciones. No es una postura personal ni una creencia, sino un marco que permite ordenar el análisis y comparar lo que observamos con lo que esperaríamos si nada hubiera cambiado.

Como explica Ben-Zvi y Garfield (2004), el valor pedagógico de enseñar pruebas de hipótesis radica en desarrollar un pensamiento crítico que permita leer resultados de manera contextualizada. En educación, contrastar hipótesis es útil para determinar si una metodología es realmente mejor que otra, si un programa tuvo un impacto apreciable, o si los grupos estudiados muestran diferencias que merecen atención. Trabajar con hipótesis ayuda además a que el estudiantado comprenda que las conclusiones estadísticas no son absolutas: dependen del tamaño de la muestra, de la variabilidad de los datos y del nivel de evidencia que estemos dispuestos a aceptar. Este enfoque evita interpretaciones simplistas y fomenta una mirada más cuidadosa, donde cada resultado se analiza en función del contexto, las limitaciones y las decisiones que se pretenden fundamentar.

Ejemplo 8: un docente implementa un programa de tutorías para mejorar los hábitos de estudio de su grupo. Para evaluar si el programa tiene efecto, registra las horas de estudio semanal de 10 estudiantes antes y después de la intervención.

Los datos (en horas) se muestran en la siguiente tabla 1:

Pregunta clave: ¿Es razonable atribuir este aumento de 0,7 horas al programa de tutorías, o podría deberse simplemente al azar?

Comparación Pretest vs. Postest en horas de estudio

La prueba t para muestras pareadas mostró una diferencia estadísticamente significativa entre las horas de estudio antes y después del programa de tutorías.

Los resultados de la prueba t para muestras pareadas (Tabla 2) indican que existe una diferencia estadísticamente significativa entre las horas de estudio antes y después de la intervención

docente, t(9) = −4.58, p = .001. La diferencia media observada fue de 0.70 horas, lo que revela un incremento en el tiempo que los estudiantes dedicaron al estudio semanal tras participar en el programa de tutorías. El intervalo de confianza del 95 por ciento [−1.05, −0.354] confirma que este aumento es consistente y no atribuible al azar. En conjunto, estos resultados sugieren que la intervención tuvo un efecto positivo y significativo en los hábitos de estudio del grupo evaluado.

Tabla 1.
Horas de estudio registradas antes y después de la intervención en el programa de tutorías

Estudiante	Pretest (antes)	Postest (después)
1	4	5
2	5	6
3	5	6
4	6	7
5	6	7
6	7	7
7	5	6
8	7	7
9	6	7
10	7	7

Nota. Los datos corresponden a las horas de estudio semanal reportadas por los estudiantes en dos momentos: antes del programa de tutorías (pretest) y después de su aplicación (postest).

Los resultados descriptivos muestran un patrón bastante claro (Tabla 3): después de la intervención, las y los estudiantes estudian más y de forma más homogénea. La media de horas de estudio pasa de 5.80 en el pretest a 6.50 en el postest, y la mediana aumenta de 6 a 7 horas. Es decir, no solo sube el promedio, sino que también el valor "típico" se desplaza hacia un mayor número de horas, lo que sugiere que el cambio no se debe a uno o dos casos extremos, sino a un ajuste general del grupo.

Además, la desviación estándar disminuye de 1.033 a 0.707 y el error estándar también se reduce, lo que indica que, tras la intervención, las respuestas se concentran más alrededor de la media. En términos sencillos, antes del programa había más dispersión en los hábitos de estudio; después, el grupo no solo estudia un poco más, sino que sus comportamientos son más parecidos entre sí. Todo esto respalda la idea de que la intervención de tutorías contribuyó a mejorar y estabilizar las horas de estudio semanal del grupo.

Tabla 2.
Resultados de la prueba t para muestras pareadas en horas de estudio antes y después de la intervención

Estadístico	Valor
t	-4.58
gl	9
p	0.001
Diferencia media (post - pre)	-0.700
Error estándar (EE)	0.153
IC 95%	[-1.05 ; -0.354]

Nota. La prueba t pareada compara las puntuaciones del pretest y postest de las horas de estudio.

Tabla 3.
Estadísticos descriptivos de las horas de estudio antes y después de la intervención

Variable	N	Media	Mediana	DE	EE
pre	10	5.80	6.00	1.033	0.327
post	10	6.50	7.00	0.707	0.224

Nota. a tabla presenta los estadísticos descriptivos correspondientes a las horas de estudio registradas en el pretest y el postest.

El análisis de situaciones donde una misma muestra es evaluada en dos momentos o condiciones distintas como ocurre en los estudios pre-post, antes-después o en comparaciones repetidas en el tiempo; constituye una oportunidad privilegiada para desarrollar razonamiento estadístico profundo en el aula. La literatura ha mostrado que este tipo de tareas permite trabajar simultáneamente la variabilidad, el significado del cambio y la capacidad de argumentar con datos, tres dimensiones que son esenciales en la alfabetización estadística moderna (Batanero, 2001; Garfield & Ben-Zvi, 2008).

Además, analizar diferencias dentro del mismo grupo facilita que los estudiantes comprendan cómo evoluciona cada individuo y no solo los promedios, reforzando la idea de que la estadística es una herramienta para interpretar fenómenos reales más que un conjunto de algoritmos aislados (Bakker, 2004). Desde un enfoque más amplio del pensamiento estadístico, diversos autores destacan que estas actividades fortalecen la capacidad de conectar información, interpretar evidencia y justificar conclusiones de manera crítica.

Se presenta a continuación un procedimiento general, aplicable a cualquier situación donde se comparan dos mediciones en una misma muestra, estructurado para favorecer una comprensión conceptual, interpretativa y contextualizada.

Procedimiento didáctico para resolver problemas con comparaciones de medidas pareadas

1. Comprender el propósito de la comparación

El proceso comienza con la lectura cuidadosa del contexto: qué se midió, por qué se midió dos veces y qué se espera responder con los datos. Esta fase inicial es clave porque evita que el análisis se limite a ejecutar una prueba, permitiendo que el estudiantado construya una pregunta de investigación con sentido . En este punto se aclara qué representa cada medición y cómo encaja dentro del fenómeno estudiado.

2. Identificar que los datos son pareados

El siguiente paso consiste en reconocer que las dos mediciones provienen de las mismas personas, grupos o unidades de análisis. Esto diferencia claramente esta situación de los problemas con muestras independientes y justifica el uso de procedimientos que trabajan con diferencias individuales.

Explorar descriptivamente las dos mediciones

Antes de realizar cualquier contraste formal, es necesario revisar medias, medianas, dispersión y patrones de cambio. Esta exploración permite que los estudiantes generen expectativas razonables acerca del resultado inferencial como un componente esencial del aprendizaje estadístico significativo. La intención es comprender la "historia" de los datos antes de pasar al modelo (Batanero, 2001)

3. Calcular y analizar las diferencias individuales

En este paso se construye una nueva variable: la diferencia entre la segunda y la primera medición. Este tipo de reformulación del problema ayuda a centrar el análisis en la unidad fundamental del cambio y favorece el desarrollo del pensamiento estadístico. Se revisa cuántas diferencias son positivas, negativas o nulas, y qué magnitud tienen. Esto convierte un problema potencialmente abstracto en una lectura simple: ¿hay evidencia de que la mayoría de las personas cambió?

4. Formular hipótesis estadísticas comprensibles

Aquí se presenta la estructura formal de la inferencia: la hipótesis nula establece que la diferencia media en la población es cero, mientras que la alternativa plantea la existencia de un cambio.

5. Aplicar la prueba estadística adecuada

Una vez construida la variable diferencia, se selecciona la prueba apropiada: t pareada si se asume normalidad, o Wilcoxon si el patrón de diferencias no es normal. Se recomiendan usar

software como Jamovi o R para que el estudiantado dedique más tiempo a interpretar y menos a ejecutar cálculos mecánicos (Garfield & delMas ,2008).

6. Interpretar los resultados en el contexto del problema

La interpretación combina varios elementos: el valor p, el intervalo de confianza, la dirección y magnitud del cambio y el tamaño del efecto.

7. Elaborar conclusiones fundamentadas y reflexivas

Finalmente, se redacta una conclusión que responda directamente la pregunta inicial, resuma el cambio observado y considere limitaciones y posibles implicaciones. Esta etapa es donde se evidencia el pensamiento estadístico maduro: no se trata de "ganar" una prueba, sino de interpretar datos para comprender y mejorar una situación concreta.

Hipótesis nula y alternativa: del lenguaje cotidiano al lenguaje académico

Antes de aplicar cualquier contraste, es necesario formular con claridad la hipótesis nula (H_0) y la alternativa (H_0). Garfield y delMas (2008) señalan que esta formulación constituye una de las principales barreras conceptuales para el estudiantado: con frecuencia confunden las hipótesis con sus expectativas, creencias personales o con el resultado que desean obtener. Desde esta perspectiva, la dificultad no reside únicamente en la notación estadística, sino en comprender que las hipótesis funcionan como "puntos de partida" para evaluar evidencia y no como afirmaciones que deban defenderse o demostrar de manera absoluta.

Desde la educación estadística con enfoque constructivista, autores como Batanero y Díaz (2011) destacan que el aprendizaje de las pruebas de hipótesis requiere que el estudiantado reconstruya el significado de conceptos como variabilidad, incertidumbre y evidencia. Sin este andamiaje conceptual, las hipótesis tienden a interpretarse como etiquetas rígidas en lugar de ser entendidas como herramientas para razonar con datos.

Por ejemplo:

Lenguaje cotidiano: *"Creo que el grupo que usa la plataforma estudia más que el grupo que no la usa."*

Formulación estadística:

$$\mathbf{H}_0 : \mu_1 = \mu_2 \qquad \mathbf{H}_1 : \mu_1 > \mu_2$$

Esta traducción facilita que el estudiantado establezca conexiones entre sus ideas previas y los conceptos formales.

Ejemplo 9: En un curso de estadística, el docente decide probar una plataforma digital de estudio. La mitad del grupo trabaja con la plataforma durante un mes; la otra mitad sigue estudiando solo con el material impreso habitual. Al final del periodo, el docente les pide que registren cuántas horas a la semana dedicaron al estudio de la asignatura.

Los resultados (resumidos) son los siguientes:

Grupo 1: con plataforma $n_1 = 28$

- Media de horas de estudio: $\overline{X}_1 = 6.3$
- Desviación estándar: $s_1 = 1.2$

Grupo 2: sin plataforma $n_2 = 30$

- Media de horas de estudio: $\overline{X}_2 = 5.5$
- Desviación estándar: $s_2 = 1.3$

La pregunta del docente es muy simple en lenguaje cotidiano: "¿Realmente la plataforma ayuda a que el estudiantado estudie más, o esta diferencia se podría explicar solo por el azar?"

A partir de aquí entramos al terreno formal.

Hipótesis en lenguaje cotidiano y en lenguaje estadístico

Primero recogemos la intuición:

Lenguaje cotidiano: *"Creo que el grupo que usa la plataforma estudia más horas que el grupo que no la usa."*

Eso lo traducimos a notación estadística:

- μ_1: media poblacional de horas de estudio de quienes usan la plataforma.
- μ_2: media poblacional de horas de estudio de quienes no la usan.

Entonces:

- **Hipótesis nula (H_0)**: la plataforma no cambia el tiempo medio de estudio.

$$H_0 : \square_1 = \square_2$$

- **Hipótesis alternativa (H_1):** la plataforma aumenta el tiempo medio de estudio.

$$H_1 : \square_1 > \square_2$$

Tomamos un nivel de significación habitual: α=0,05, prueba unilateral (solo nos interesa si aumenta).

1. Tipo de prueba

Tenemos dos grupos distintos de estudiantes, sin emparejamiento, con medias y desviaciones estándar conocidas. Lo más natural es usar una prueba t para muestras independientes (versión de Welch, que no supone varianzas iguales).

En términos más pedagógicos: vamos a comparar "media con plataforma" frente a "media sin plataforma", teniendo en cuenta la variabilidad de cada grupo y el tamaño muestral.

2. Diferencia observada entre las medias

Calculamos primero la diferencia simple: $\overline{X}_1 - \overline{X}_2 = 6.3 - 5.5 = 0.8$ "horas"

Es decir, en promedio, el grupo con plataforma declara estudiar 0,8 horas más por semana. La pregunta es:

¿0,8 horas es una diferencia suficientemente grande como para no atribuirla al azar?

3. Cálculo del error estándar de la diferencia

La diferencia entre medias no se interpreta en el vacío; hay que considerar cuánta dispersión hay en cada grupo y cuántos estudiantes participaron. Para eso usamos el error estándar de la diferencia:

Elevamos al cuadrado las desviaciones estándar:

$$\mathbf{s_1^2 = (1.2)^2 = 1\ .44}$$

$$\mathbf{s_2^2 = (1.3)^2 = 1\ .69}$$

Dividimos cada varianza entre su tamaño muestral:

$$\mathbf{\frac{s_1^2}{n_1} = \frac{1.44}{28} \approx 0.0514}$$

$$\mathbf{\frac{s_2^2}{n_2} = \frac{1.69}{30} \approx 0.0563}$$

Sumamos esos dos valores:

$$\mathbf{0.0514 + 0.0563 \approx 0.1077}$$

Hacemos la raíz cuadrada:

$$\mathbf{EE_{dif} = \sqrt{0.1077} \approx 0.33}$$

Ese 0,33 es el error estándar de la diferencia de medias: una medida de cuánto esperaríamos que se mueva la diferencia entre grupos solo por el azar del muestreo.

4. Estadístico t

Ahora comparamos la diferencia observada con el tamaño de ese error estándar:

$$\mathbf{t = \frac{\overline{X}_1 - \overline{X}_2}{EE_{dif}} = \frac{0.8}{0.33} \approx 2.44}$$

Mientras más grande es t en valor positivo, más lejos está nuestra diferencia de "cero diferencias" en unidades de error estándar.

5. Grados de libertad y p-valor (idea general)

Con la versión de Welch, los grados de libertad son aproximadamente 56. No hace falta que el estudiantado memorice la fórmula exacta; lo importante es saber que se usan para buscar el valor crítico o el p-valor en la distribución t.

Para unos 56 grados de libertad y una prueba unilateral: un valor t cercano a 2,44 da un p-valor alrededor de 0,01.

Eso significa que, si en realidad no hubiera diferencia entre los grupos (si $\mathbf{H_0}$ fuera cierta), la probabilidad de obtener una diferencia tan grande como 0,8 horas, o incluso mayor, solo por azar, sería de alrededor del 1 %. Es decir, bastante baja.

8. Decisión estadística

Como el p-valor es ≈ 0,01 y nuestro nivel de significación era α=0,05:

- $p < \alpha$
- Rechazamos H_0.

Los datos que hemos observado no encajan bien con la idea de que la plataforma no produce ninguna diferencia. Hay evidencia suficiente para pensar que sí tiene efecto.

9. Interpretación en lenguaje humano

Una forma de comunicar el resultado, sin tecnicismos innecesarios, podría ser:

Con la información recogida, la diferencia de 0,8 horas semanales de estudio entre el grupo que usó la plataforma y el que no la usó es demasiado grande como para atribuirla solo al azar. Los análisis estadísticos sugieren que la plataforma digital sí está asociada con un mayor tiempo de estudio, al menos en este contexto y con este grupo de estudiantes.

- Desde un enfoque frecuentista, hablas de probabilidad de observar esos datos si la hipótesis nula fuera cierta.
- Desde una mirada didáctica y constructivista, enfatizas el proceso: partir de una conjetura en lenguaje cotidiano, traducirla a hipótesis formales, analizar la variabilidad y, finalmente, volver a un lenguaje comprensible para tomar decisiones educativas.

La figura 8 permite comparar de forma precisa el comportamiento de ambos grupos a partir de sus valores estimados.

El grupo que trabajó con la plataforma obtuvo una media de 6,3 horas de estudio por semana, con un intervalo de confianza del 95 % entre 5,84 y 6,77 horas. En contraste, el grupo sin plataforma registró una media menor, 5,5 horas, cuyo intervalo de confianza se ubica entre 5,02 y 5,99 horas.

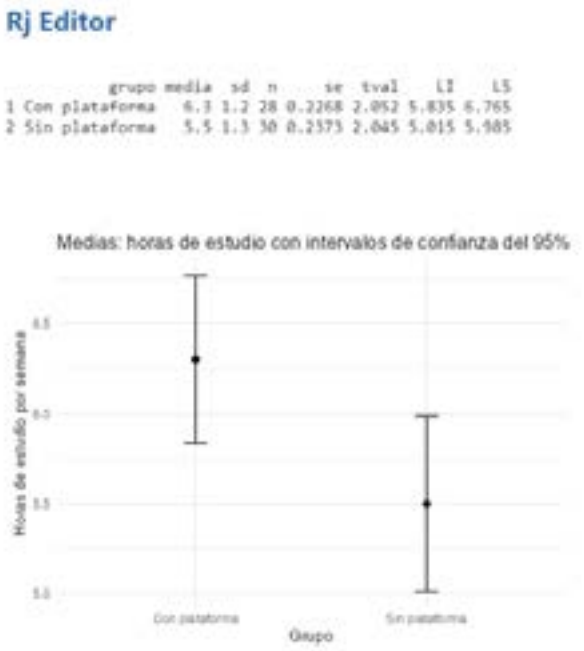
Cuando se observan juntos, estos intervalos revelan un patrón interesante: aunque existe una ligera superposición entre ambos rangos, la mayor parte del intervalo del grupo con plataforma queda por encima del intervalo correspondiente al grupo sin plataforma. Esto, junto con la diferencia en las medias, sugiere una ventaja consistente a favor del uso de la plataforma digital. Además, el error estándar es menor en ambos casos (0,2268 y 0,2373), lo cual indica que las estimaciones son razonablemente estables para tamaños muestrales de 28 y 30 estudiantes.

Comprender la diferencia entre la hipótesis nula y la hipótesis alternativa es un paso esencial para que el estudiantado pueda interpretar con sentido cualquier contraste estadístico. Cuando los estudiantes logran traducir sus intuiciones y preguntas cotidianas al lenguaje formal de la estadística, dejan de ver las hipótesis como frases abstractas y empiezan a reconocerlas

como herramientas que permiten organizar el razonamiento y dar estructura a la investigación.

Esta transición del lenguaje común al académico no solo clarifica qué se quiere evaluar, sino que también ayuda a evitar malentendidos frecuentes, como creer que la hipótesis nula expresa una opinión personal o un deseo de resultado. En realidad, la hipótesis nula funciona como un punto de referencia que permite valorar la evidencia, mientras que la alternativa expresa la dirección o el tipo de cambio que interesa analizar. Cuando el docente acompaña este proceso con ejemplos cercanos y explicaciones sencillas, la formulación de hipótesis deja de ser un trámite técnico y se convierte en un ejercicio de reflexión que fortalece la capacidad de pensar con rigor y claridad.

Figura 9.
Comparación de las medias de horas de estudio entre los grupos con y sin plataforma, con intervalos de confianza del 95 %



Nota. La figura muestra las medias semanales de estudio para los dos grupos: uno que trabajó con la plataforma digital y otro que utilizó únicamente material impreso.

Pruebas Z y t de Student: criterios para decidir

La elección entre una prueba Z y una prueba t de Student no es un asunto meramente técnico; constituye una decisión clave para garantizar que las inferencias realizadas reflejen adecuadamente la incertidumbre presente en los datos. En términos generales, esta decisión depende de dos elementos fundamentales:

el tamaño de la nuestra y la disponibilidad de la desviación estándar poblacional. Sin embargo, diversos autores han mostrado que este criterio básico debe matizarse para evitar interpretaciones erróneas o conclusiones precipitadas.

Gelman y Hill (2014) advierten que aplicar la prueba Z sin considerar el origen de la desviación estándar suele conducir a un exceso de confianza en las conclusiones, porque ignora la variabilidad adicional asociada a la estimación muestral. Este planteamiento coincide con la perspectiva clásica de Student (1908), quien desarrolló su distribución precisamente para corregir el sesgo que se introducía al trabajar con muestras pequeñas.

Desde un enfoque más didáctico, Batanero y Díaz (2011) señalan que muchos estudiantes creen que ambas pruebas son equivalentes o intercambiables, lo que genera errores conceptuales frecuentes. Por ello, recomiendan enfatizar en la enseñanza que la prueba Z se justifica únicamente cuando σ es conocida o cuando los tamaños muestrales son lo suficientemente grandes como para que la estimación de la desviación estándar sea estable.

En la literatura reciente, autores como Cumming (2014) destaca que la prueba t incorpora explícitamente la incertidumbre del error estándar, lo que la vuelve más adecuada para investigaciones en ciencias sociales y educativas, donde los tamaños de muestra tienden a ser moderados o reducidos.

En resumen, aunque la regla general indica que la prueba Z se utiliza cuando σ es conocida o $\mathbf{n > 30}$, y la prueba t cuando σ es desconocida o la muestra es pequeña, los autores coinciden en que la prueba t ofrece una representación más fiel de la incertidumbre en la mayoría de contextos educativos. Su uso no solo ajusta de manera más adecuada el error estándar, sino que promueve una comprensión más crítica del proceso inferencial.

Ejemplo 10. Situación cercana a una prueba Z: σ conocida y muestra grande

Imagina que el Ministerio aplica cada año una prueba estandarizada de matemáticas a todo el país. A partir de muchos años de aplicación, se sabe que los puntajes se distribuyen aproximadamente con media poblacional $\mu_0 = 500$ puntos y desviación estándar poblacional $\sigma = 100$ puntos.

Este año, una provincia que ha implementado un nuevo programa de apoyo reporta una muestra de 100 estudiantes con una media de $\overline{X} = 530$ puntos. La pregunta es:

¿Se puede concluir que el programa está asociado con un aumento real en los puntajes, o esa diferencia de 30 puntos podría deberse solo al azar?

Los resultados del análisis (Tabla 4) indican que la media observada de los puntajes (530 puntos) se encuentra tres errores estándar por encima de la media poblacional histórica de 500

puntos. El estadístico Z = 3 y su p-valor bilateral (p = .0027) muestran que esta diferencia de 30 puntos es muy poco probable que se deba únicamente al azar. El intervalo de confianza del 95 % (510.4 a 549.6) tampoco incluye el valor de la media poblacional, reforzando esta conclusión.

Tabla 4,
Resultados de la prueba Z para evaluar el efecto del programa en los puntajes

Estadístico	**Valor**
Media muestral	530
Media poblacional (H_0)	500
Error estándar (SE)	10
Estadístico Z	3.00
p (bilateral)	0.0027
IC 95 %	[510.4 ; 549.6]

Nota. La tabla resume los resultados de una prueba Z aplicada para comparar la media observada de una muestra de 100 estudiantes con la media poblacional histórica de 500 puntos.

En términos prácticos, hay evidencia estadística sólida para afirmar que el programa implementado está asociado con una mejora real en el rendimiento académico de los estudiantes en la prueba estandarizada.

Ejemplo 11. Situación típica de prueba t: σ desconocida y muestra pequeña

Ahora pensemos en un escenario mucho más habitual en educación: un docente en un solo curso quiere saber si su estrategia de evaluación formativa ha incrementado el tiempo de estudio semanal de su grupo. Para ello recoge datos de 12 estudiantes y obtiene:

- Media muestral: $\overline{\mathbf{X}} = \mathbf{6}$
- Desviación estándar muestral: $\mathbf{s} = \mathbf{1.2}$
- Hipótesis de referencia: antes de la intervención, el grupo estudiaba en torno a 6 horas.

Los resultados de la prueba t de una muestra (Figura 9) indican que el grupo estudia, en promedio, 6,18 horas semanales, muy cerca de las 6 horas que se tomaron como valor de referencia. La desviación estándar es 0,753, y con 12 estudiantes esto se traduce en un error estándar de 0,217.

Los resultados de la prueba t de una muestra (Figura 9) indican que el grupo estudia, en promedio, 6,18 horas semanales, muy cerca de las 6 horas que se tomaron como valor de referencia. La desviación estándar es 0,753, y con 12 estudiantes esto se traduce en un error estándar de 0,217.

Figura 10.
Resultados de la prueba t de una muestra para el tiempo de estudio semanal

Prueba t de una Muestra

Prueba t de una Muestra

							Intervalo de Confianza al 95%	
		Estadístico	gl	p		Tamaño del Efecto	Inferior	Superior
horas	T de Student	0.843	11.0	0.417	La d de Cohen	0.243	−0.336	0.813

Nota. H_a $\mu \neq 6$

Descriptivas

	N	Media	Mediana	DE	EE
horas	12	6.18	6.20	0.753	0.217

Nota. La figura presenta los resultados de una prueba t de una muestra aplicada a 12 estudiantes, con el fin de comparar la media observada de horas de estudio semanal (6,18 horas) con el valor de referencia de 6 horas.

Al comparar la media observada con el valor de 6 horas, Jamovi calcula un estadístico t(11) = 0,843 con un p = 0,417. Este p-valor es mucho mayor que 0,05, de modo que no hay evidencia estadística para rechazar la hipótesis de que la media real sea 6 horas; la diferencia de 0,18 horas puede explicarse perfectamente por la variabilidad normal entre estudiantes.

El tamaño del efecto de Cohen (d = 0,243) también refuerza esta lectura: se trata de un efecto pequeño, lo que sugiere que, aun si existiera algún cambio real, su magnitud sería modesta. En resumen, con estos datos no podemos afirmar que la estrategia de evaluación haya modificado de forma clara el tiempo de estudio del grupo; más bien, los resultados son coherentes con la idea de que el alumnado sigue estudiando alrededor de 6 horas por semana.

Apoyo didáctico: Trabajar con pruebas de hipótesis en el aula no debería convertirse en un ejercicio mecánico de "aplicar fórmulas", sino en una oportunidad para que el estudiantado comprenda qué significa realmente comparar evidencia con una afirmación inicial. En el ejemplo analizado, los resultados muestran que la media observada no es estadísticamente distinta del valor de referencia, lo que invita a pensar que los datos, por sí solos,

no sostienen la idea de un cambio relevante en el tiempo de estudio. Desde una perspectiva formativa, esto permite conversar con los estudiantes sobre la importancia de mirar más allá de la diferencia numérica y atender a la variabilidad, al tamaño de la muestra y al contexto en el que se recogen los datos.

Este tipo de análisis también ayuda a desmontar la expectativa frecuente de que toda intervención debe producir un aumento inmediato y visible. Entender que la ausencia de diferencias significativas no es un "fracaso", sino un resultado posible y valioso, fomenta en el aula una postura más crítica y honesta frente a la evidencia. A la vez, abre la puerta para que los estudiantes reflexionen sobre qué otros factores podrían estar influyendo, qué ajustes metodológicos serían necesarios o qué tipo de evidencias complementarias podrían recabarse.

Errores tipo I y tipo II en el razonamiento estadístico

En el corazón de toda prueba de hipótesis existe una tensión entre dos riesgos inevitables: equivocarse por exceso de confianza o equivocarse por exceso de cautela. Estos riesgos se formalizan en la estadística inferencial como error tipo I y error tipo II, conceptos centrales para comprender que toda decisión basada en datos está sujeta a incertidumbre y que nunca trabajamos con conclusiones absolutas, sino con grados de evidencia.

El error tipo I (α) ocurre cuando se rechaza la hipótesis nula siendo esta verdadera. Es, en términos prácticos, declarar que "hay un efecto" donde realmente no lo hay. En educación, este error puede llevar a pensar que una intervención didáctica produjo mejoras cuando, en realidad, los resultados observados podrían explicarse por el azar o la variabilidad natural de los grupos. Su importancia radica en que, si no se controla, se corre el riesgo de implementar estrategias o programas ineficaces, generando expectativas infundadas, uso inadecuado de recursos o conclusiones pedagógicas equivocadas.

Por otro lado, el error tipo II (β) aparece cuando no se rechaza la hipótesis nula siendo esta falsa. En este caso, el peligro consiste en pasar por alto un efecto real: creer que una intervención "no funciona" cuando en verdad sí produce un impacto. Este tipo de error tiene consecuencias relevantes en entornos educativos, porque puede llevar a descartar prácticas valiosas, no dar continuidad a estrategias prometedoras o desestimar cambios que requieren más tiempo o condiciones más estables para hacerse visibles. Su probabilidad está asociada al tamaño de la muestra, la variabilidad de los datos y la magnitud del efecto: cuanto más pequeños sean los grupos o más dispersa sea la información, mayor es el riesgo de no detectar diferencias que sí existen.

Comprender ambos errores es fundamental para la toma de decisiones informada. En la práctica, no se trata simplemente de memorizar definiciones, sino de entender que toda inferencia estadística implica un equilibrio: minimizar el error tipo I reduce la probabilidad de falsas alarmas, pero aumenta el riesgo de cometer un error tipo II, es decir, de dejar pasar efectos reales. De la misma manera, aumentar la sensibilidad para detectar cambios puede elevar el riesgo de concluir que un resultado es significativo cuando en realidad no lo es. Este juego de compensaciones obliga a reflexionar sobre el contexto de investigación, las implicaciones pedagógicas y los costos de equivocarse en uno u otro sentido.

Ejemplos que el docente puede desarrollar:

Ejemplo 12. Evaluación de un nuevo método de enseñanza

Una institución introduce un método innovador de enseñanza de matemáticas y quiere comprobar si esta mejora el rendimiento respecto al enfoque tradicional. Se comparan las calificaciones de ambos grupos mediante una prueba estadística.

Error tipo I (falso positivo)

El análisis indica que el nuevo método mejora el rendimiento, por lo que la escuela decide implementarlo. Sin embargo, en realidad no existe una mejora real; la diferencia observada se debe al azar o a factores externos (motivación inicial, afinidad del docente, clima del aula).

La consecuencia es que la institución adopta una estrategia que solo parece eficaz, invirtiendo tiempo, recursos y expectativas en algo que no produce los resultados prometidos.

Error tipo II (falso negativo)

El análisis no encuentra diferencias significativas y se concluye que el método no aporta beneficios. Sin embargo, el método sí mejora el rendimiento, pero el estudio no fue capaz de detectarlo (por ejemplo, por usar una muestra demasiado pequeña).

El resultado es que la escuela descarta una herramienta realmente útil que podría haber apoyado el aprendizaje del estudiantado.

Ejemplo 13. Programa de tutorías para mejorar hábitos de estudio

Una docente aplica un programa de tutorías para fortalecer la organización del tiempo en un grupo de estudiantes y evalúa si aumentan sus horas de estudio semanales.

Error tipo I

Se concluye que las tutorías aumentaron las horas de estudio, pero en realidad los cambios se deben a un examen cercano que motivó al grupo a estudiar más.

El riesgo es atribuirle al programa un efecto que no proviene del programa, generando una falsa sensación de éxito.

Error tipo II

Los análisis no detectan un aumento significativo, por lo que se piensa que las tutorías no funcionan. Pero el programa sí generó cambios, solo que estos fueron pequeños o requieren más tiempo para consolidarse.

El riesgo es abandonar una intervención que sí aportaba mejoras, aunque de forma progresiva.

Ejemplo 14. Detección temprana de dificultades de aprendizaje

Se utiliza una prueba diagnóstica para identificar a estudiantes con riesgo de dificultades lectoras.

Error tipo I

La prueba señala que un niño tiene riesgo cuando en realidad no lo tiene.

En la práctica, esto puede llevar a intervenciones innecesarias, ansiedad para la familia o desvío de recursos educativos.

Error tipo II

La prueba no detecta a una niña que realmente sí necesita apoyo.

Como consecuencia, la estudiante no recibe la ayuda oportuna y sus dificultades podrían profundizarse con el tiempo.

Desde un enfoque formativo, enseñar los errores tipo I y tipo II permite desarrollar en el estudiantado una mirada crítica sobre las afirmaciones basadas en datos. Los ayuda a reconocer que la estadística no busca certezas absolutas, sino decisiones razonables apoyadas en evidencia. También favorece el desarrollo de un pensamiento más matizado: no todo resultado "significativo" implica un cambio real, ni toda ausencia de significación indica falta de impacto. Al integrar estos conceptos en el análisis educativo, se promueve una comprensión más profunda del carácter provisional y contextual de las conclusiones estadísticas, fortaleciendo la capacidad de interpretar datos de manera reflexiva y responsable.

Contrastes más comunes en contextos educativos (proporciones, medias, diferencias)

Comprender los contrastes más habituales en investigación educativa implica mucho más que aplicar procedimientos estadísticos; supone, como argumentan Batanero y Díaz (2011), reconocer la incertidumbre inherente a los datos y leerla de forma crítica para tomar decisiones informadas. Garfield y delMas (2008) añaden que la fortaleza de estos contrastes radica en su capacidad para articular preguntas educativas reales con herramientas cuantitativas transparentes.

Desde esta mirada, los contrastes se presentan en dos grandes grupos: los que se aplican a proporciones y variables categóricas, y los que se aplican a medias y diferencias entre grupos. La diferenciación no es meramente técnica.

Contrastes para proporciones y variables categóricas

En educación, una gran parte de la información que se recoge no es numérica continua sino categórica: aprobar o no aprobar, asistir o no asistir, alto o bajo desempeño, riesgo o no riesgo, logro o no logro. Estas variables permiten describir situaciones relevantes, pero exigen métodos específicos para poder contrastar hipótesis o estudiar asociaciones entre grupos. Como plantean Batanero y Díaz (2011), comprender cómo se analizan las proporciones es fundamental para desarrollar un pensamiento estadístico auténtico, pues la interpretación de estos contrastes se vincula directamente con decisiones educativas reales: asignación de recursos, intervenciones focalizadas o evaluación de programas.

Los contrastes para proporciones permiten responder preguntas como:

- ¿La proporción de estudiantes con logro mejora luego de una intervención?
- ¿Dos grupos muestran patrones similares de aprobación?
- ¿Existe relación entre la participación y el nivel de desempeño?

Garfield y delMas (2008) subrayan que estos contrastes son especialmente pedagógicos porque ayudan al estudiantado a pasar de la observación de frecuencias "a simple vista" a una interpretación rigurosa basada en evidencia y variabilidad.

A continuación, se desarrollan los contrastes más habituales.

a) Comparación de proporciones: prueba Z para una o dos proporciones

La prueba Z es uno de los contrastes más utilizados para comparar proporciones cuando el tamaño de la muestra es lo suficientemente grande como para justificar la aproximación normal. Su fundamento radica en que, al incrementarse el número de observaciones, la distribución muestral de la proporción tiende a adoptar una forma aproximadamente normal, lo que permite calcular errores estándar y realizar inferencias con mayor precisión.

Este comportamiento se explica por el principio general de las aproximaciones asintóticas, ampliamente discutido por Gelman y Hill (2014), quienes señalan que muchos procedimientos inferenciales ganan estabilidad conforme aumenta la información disponible en los datos. En ese sentido, la prueba Z no solo se apoya en un criterio técnico, sino en una lógica estadística que busca garantizar conclusiones más fiables a partir de la variabilidad

muestral y del comportamiento esperado de la distribución bajo condiciones de gran tamaño de muestra.

Se utiliza cuando:

1. La variable es dicotómica (sí/no).
2. El tamaño muestral garantiza la aproximación normal:
3. $\mathrm{np} \geq 5 \quad \mathrm{and} \quad \mathrm{n}(1-\mathrm{p}) \geq 5.$
4. Se comparan proporciones con un valor de referencia (una proporción) o entre grupos (dos proporciones).
5. El objetivo es evaluar si la diferencia observada refleja un cambio real o solo fluctuaciones aleatorias.

Como advierte Gelman y Hill (2014), la clave pedagógica no es solo calcular la prueba, sino interpretar el resultado en términos de incertidumbre y tamaño del efecto.

Ejemplo 15. Prueba Z para una proporción

Una institución desea saber si al menos el 70 % de estudiantes alcanza el nivel esperado en comprensión lectora.

Datos:

45 de 60 estudiantes lograron el nivel $\longrightarrow \widehat{\mathrm{p}} = 0.75$

Hipótesis:

- H_0: p = 0.70
- H_1: p > 0.70

Interpretación:

La diferencia entre 0.70 y 0.75 puede parecer relevante, pero la prueba Z ayuda a evaluar si este incremento podría aparecer por azar. Un valor p pequeño sugiere que la proporción real en la población podría estar por encima de 0.70.

Cumming (2014) recomienda acompañar siempre esta prueba con un intervalo de confianza del 95 %, porque permite comunicar con claridad la incertidumbre alrededor de la proporción observada.

Ejemplo 15. Prueba Z para dos proporciones

Situación:

Una docente implementa un módulo digital y desea saber si aumentó la proporción de estudiantes aprobados.

- Antes: 18/30 aprueban entonces 0.60
- Después: 24/30 aprueban entonces 0.80

Se usa Z para dos proporciones porque:

- Se comparan dos momentos distintos.
- Ambas proporciones provienen de muestras independientes.
- Las frecuencias cumplen los supuestos de normalidad aproximada.

Interpretación: La diferencia de 0.20 es notable, pero la pregunta clave es:

¿es suficientemente grande como para atribuirla al programa y no al azar?

El contraste Z y su intervalo de confianza aportan evidencia más robusta que una simple comparación descriptiva.

b) Prueba χ^2 (chi cuadrado) de independencia

La prueba χ^2 es uno de los contrastes más utilizados para determinar si dos variables categóricas están asociadas o si sus distribuciones son independientes. En el ámbito educativo, su utilidad es particularmente relevante porque permite identificar patrones que no siempre se distinguen a simple vista en las tablas de frecuencias. Agresti (2018) explica que este contraste resulta esencial cuando se investigan relaciones entre atributos cualitativos, pues permite evaluar si las diferencias entre valores observados y esperados son atribuibles al azar o responden a un efecto real. Es adecuada cuando:

- Ambas variables son categóricas.
- Se desea saber si las categorías de una variable están asociadas a las categorías de otra.
- Las frecuencias esperadas son ≥ 5 (Howell, 2017).
- El interés no es casual, sino relacional.

Ejemplo 16. Se quiere saber si el nivel lector se relaciona con asistir o no a un programa de apoyo. **(χ^2 de independencia en lectura y asistencia)**

Tabla 5.
Asistencia a tutorías según el nivel lector del estudiantado

Nivel lector	Asiste	No asiste
Adecuado	12	8
Bajo	16	4

Nota. La tabla presenta la distribución conjunta del nivel lector y la asistencia a tutorías.

Desde una perspectiva didáctica, Batanero y Díaz (2011) destacan que entender la lógica de este contraste ayuda al alumnado a interpretar adecuadamente la variabilidad y desarrollar un razonamiento estadístico más crítico y fundamentado. En conjunto, estos aportes consolidan a la prueba $_0$ como una herramienta indispensable para analizar perfiles de riesgo, estrategias de estudio y diversas dinámicas educativas basadas en datos categóricos.

c) Pruebas exactas y alternativas no paramétricas

En el análisis de datos educativos, la elección de una prueba estadística apropiada depende tanto del tipo de variable como de los supuestos que estas cumplen, por lo que comprender la lógica detrás de los métodos disponibles es fundamental para realizar inferencias sólidas. Agresti (2018) enfatiza que, antes de aplicar cualquier contraste, es necesario evaluar si las condiciones teóricas se sostienen, ya que de ello depende la validez de los resultados.

Del mismo modo, Batanero y Díaz (2011) destacan que muchos errores en la interpretación estadística provienen de una comprensión insuficiente sobre la naturaleza de las distribuciones y la variabilidad, lo que subraya la importancia de enseñar no solo técnicas, sino también sus fundamentos conceptuales. Por su parte, Batanero y Ben-Zvi (2013) muestran que el razonamiento sobre variabilidad es clave para interpretar correctamente las diferencias entre grupos y tomar decisiones basadas en evidencia.

Cuando los supuestos de χ^2 no se cumplen, por ejemplo, cuando alguna frecuencia es menor a 5; por tanto, es necesario recurrir a alternativas más robustas, como:

- Prueba exacta de Fisher
- Prueba de McNemar (cuando las mediciones son pareadas)

Mayo (2018) subraya que estas pruebas "severas" refuerzan la validez del análisis, pues evitan conclusiones erróneas basadas en frecuencias insuficientes. Las contribuciones de Bakker (2004) se suman a esta perspectiva al señalar que el uso de herramientas y métodos adecuados permite representar de manera más precisa los patrones que emergen de los datos, algo especialmente relevante en entornos educativos donde las muestras no siempre cumplen los supuestos ideales.

Ejemplo 17. Fisher exacta con grupos pequeños

Un grupo pequeño de 10 estudiantes participa en un micro taller. Se desea saber si el taller se relaciona con completar una tarea.

Tabla 6.
Relación entre asistencia a tutorías y cumplimiento de tareas

Completa tarea	Sí	No
Asiste	5	1
No asiste	2	2

Nota. La tabla muestra la distribución de estudiantes según su asistencia a las tutorías y el cumplimiento de las tareas asignadas.

Las frecuencias son bajas en algunas celdas.

Por ello, χ^2 no es recomendable y se utiliza Fisher exacta, que calcula la probabilidad exacta de obtener una distribución igual o más extrema bajo H_0.

Interpretación:

Si $p < 0.05$, se concluye que el patrón entre asistencia y completar la tarea no es aleatorio. Dado el tamaño reducido, esta prueba evita sobre interpretar diferencias que podrían ser producto del azar. Los contrastes para proporciones y variables categóricas son herramientas especialmente útiles para interpretar situaciones educativas que, a simple vista, pueden parecer triviales o evidentes.

A través de estos procedimientos es posible reconocer cambios, asociaciones y tendencias que no siempre se muestran de forma clara en los datos brutos.

Lo verdaderamente valioso, desde el punto de vista formativo, es que permiten al estudiantado mirar más allá de la intuición inicial y preguntarse qué tan probable es que una diferencia observada se deba realmente al fenómeno que interesa o, por el contrario, sea producto del azar.

El análisis de proporciones invita a pensar en la variabilidad, en la forma en que se distribuyen los resultados y en la importancia de considerar el tamaño de la muestra antes de sacar conclusiones apresuradas. Esta reflexión es clave en contextos educativos, donde las decisiones deben basarse en evidencia confiable.

Contrastes para medias y diferencias entre grupos

Comparar medias es una de las tareas más frecuentes en investigación educativa: analizar si un grupo estudia más que otro, si una intervención mejora puntajes, si dos modalidades de enseñanza producen diferencias en el rendimiento o si los estudiantes cambian su desempeño antes y después de un programa. Estas comparaciones son habituales porque permiten evaluar el impacto de las prácticas pedagógicas, contrastar enfoques metodológicos y comprender mejor cómo evolucionan los aprendizajes en diferentes contextos.

Como señalan Kline (2013) y Howell (2017), elegir correctamente un contraste supone atender tres aspectos esenciales: la naturaleza de las variables, el cumplimiento de los supuestos estadísticos y la forma en que se organiza el diseño del estudio. Cuando estas decisiones se toman con criterios sólidos, el análisis de medias deja de ser un mero ejercicio aritmético y se convierte en una herramienta de interpretación que ayuda a explicar por qué ciertos grupos se comportan de manera distinta, qué cambios son atribuibles a una intervención y qué variabilidad responde más al azar que a efectos reales.

Como señalan Kline (2013) y Howell (2017), elegir correctamente un contraste requiere considerar tres aspectos esenciales:

1. el tamaño de la muestra,
2. la forma en que se distribuyen los datos y
3. si se conoce o no la desviación estándar poblacional.

De la combinación de estos criterios surgen diferentes pruebas para comparar medias. A continuación, se desarrollan las más utilizadas en educación, junto con sus fundamentos conceptuales y ejemplos explicados paso a paso.

1. Prueba Z para una media: σ conocida y tamaño grande

La prueba Z se utiliza en situaciones donde se conoce la desviación estándar poblacional (σ)o cuando el tamaño muestral es

lo suficientemente grande para que la distribución de la media se aproxime a la normal.

Aunque en educación rara vez se conoce σ, este contraste sigue siendo útil para fines didácticos, sobre todo para introducir el razonamiento inferencial.

Ejemplo 18: Una institución afirma que, en promedio, sus estudiantes dedican 6 horas semanales al estudio. Una muestra grande de n = 80 estudiantes reportan:

Media muestral = 6.4 horas

σ poblacional (conocida por estudios previos) = 1.5 horas

Hipótesis:

H_0: $\mu = 6$

H_1: $\mu \neq 6$

Interpretación:

Si el valor Z resulta significativo, indica que el tiempo promedio en esta muestra difiere del valor institucional, lo que sugiere que podrían haberse producido cambios reales en los hábitos de estudio. Sin embargo, el punto clave como explica Cumming (2014), no es el estadístico Z en sí, sino evaluar si la diferencia observada es coherente con la variabilidad que razonablemente podría esperarse en este tipo de datos. Esta perspectiva enfatiza que la interpretación no debe centrarse únicamente en la significación, sino en comprender cuán plausible es el cambio a la luz de la variabilidad muestral y del intervalo de confianza que la acompaña.

2. Prueba t para muestras independientes: comparar dos grupos distintos

Se usa cuando se desea comparar las medias de dos grupos independientes, como:

- estudiantes de dos cursos,
- dos metodologías diferentes,
- grupos con y sin intervención,
- condiciones de aprendizaje presenciales y virtuales.

Howell (2017) subraya que las condiciones clave para esta prueba son:

- Independencia entre los grupos
- Aproximada normalidad en cada conjunto
- Homogeneidad de varianzas (cuando se usa la versión clásica)

Cuando no se cumple la homogeneidad, se utiliza la corrección de Welch, más robusta.

Ejemplo 20

Una docente quiere saber si un módulo digital mejora el desempeño. Compara:

Tabla 7.
Estadísticos descriptivos de los puntajes según participación en la plataforma virtual

Grupo	n	Media	DE
Con plataforma	30	6.3	1.2
Sin plataforma	30	5.5	1.3

Nota. La tabla presenta los estadísticos descriptivos de los puntajes obtenidos por los estudiantes que trabajaron con la plataforma virtual y aquellos que no la utilizaron.

Hipótesis:

- $\mathbf{H_0 : \mu_1 = \mu_2}$
- $\mathbf{H_1 : \mu_1 > \mu_2}$
- Interpretación:

Si el valor *p* es menor que 0.05, los datos sugieren que quienes utilizaron la plataforma estudiaron más. No obstante, este resultado solo indica que la diferencia difícilmente se deba al azar; no dice nada sobre cuán grande o importante es. Por eso, el tamaño del efecto —como el *d* de Cohen— es indispensable para valorar si la diferencia observada es realmente significativa en términos educativos. Mientras el p refleja la evidencia estadística, el tamaño del efecto permite comprender la magnitud del cambio y si este tiene un impacto que pueda considerarse relevante para la práctica docente o para la toma de decisiones institucionales.

3. Prueba t para muestras pareadas: antes y después

La prueba t para muestras pareadas resulta especialmente útil cuando se quiere evaluar un cambio real en un mismo grupo de personas tras una intervención, actividad o experiencia educativa. A diferencia de los análisis que comparan grupos distintos, aquí cada participante se convierte en su propio punto de referencia: se observa cómo estaba antes y cómo se encuentra después.

Este enfoque permite aislar de mejor manera el efecto de la intervención, porque elimina la variabilidad que existe entre individuos. En contextos educativos, esta prueba ayuda a responder preguntas muy habituales, como si un programa de tutorías mejora el rendimiento, si una estrategia didáctica incrementa la motivación, o si una herramienta tecnológica facilita el aprendizaje.

Se usa cuando las mediciones proceden del mismo grupo en dos momentos:

- pretest y postest,
- antes y después de una intervención,
- dos tareas realizadas por los mismos estudiantes.

Ejemplo 21

Una docente aplica tutorías durante cuatro semanas. Antes y después, registra las horas de estudio:

- Pretest: media = 5.8
- Postest: media = 6.5
- Diferencia media = 0.7 horas

Hipótesis:

- $\mathbf{H_0 : \mu_{dif} = 0}$
- $\mathbf{H_1 : \mu_{dif} > 0}$

Interpretación:

Si la diferencia promedio resulta estadísticamente significativa, puede interpretarse que el aumento en las horas de estudio está asociado al programa. Sin embargo, como destaca Cumming (2014), la significación por sí sola no basta para comprender el alcance real del cambio. El intervalo de confianza de la diferencia proporciona información esencial sobre cuánta mejora es razonable esperar y cuánta incertidumbre existe alrededor de esa estimación. De este modo, no solo se identifica que el programa produce un efecto, sino también la magnitud y la precisión con que puede describirse dicho efecto, lo que permite tomar decisiones más informadas y realistas en el ámbito educativo.

4. ANOVA de un factor: tres o más medias

Howell (2017) explica que el ANOVA contrasta la variabilidad entre grupos con la variabilidad dentro de ellos. Cuando la variación entre las medias grupales es mucho mayor que la dispersión que existe dentro de cada grupo, resulta poco probable que esa diferencia sea producto de fluctuaciones aleatorias. En ese caso se concluye que existe evidencia de diferencias significativas entre los grupos. Esta lógica convierte al ANOVA en una herramienta especialmente valiosa para estudiar fenómenos educativos en los que se desea comparar varios métodos, intervenciones o niveles de desempeño sin perder rigor estadístico.

Ejemplo 22. Tres modalidades de estudio (A, B y C) producen los siguientes resultados:

Interpretación:

Un ANOVA significativo indica que al menos una de las medias difiere del resto, pero no especifica entre qué grupos se encuentran esas diferencias. Esta información es crucial en la interpretación, ya que el análisis global solo señala la existencia de una variación sistemática, sin precisar su origen. Por ello, una vez obtenida una F significativa, es necesario complementar el análisis con comparaciones post hoc, que permiten identificar con precisión qué pares de medias presentan diferencias estadísticamente relevantes.

Estas pruebas controlan el error asociado a realizar múltiples comparaciones y ofrecen una visión más detallada del efecto, facilitando conclusiones válidas sobre las modalidades que realmente se distinguen entre sí.

En síntesis, los contrastes para medias constituyen una herramienta esencial para comprender cómo varían los aprendizajes y las prácticas educativas en distintos grupos y contextos. Más allá de las fórmulas, su verdadero valor radica en ayudar a interpretar si las diferencias observadas reflejan cambios reales o simplemente variabilidad propia del muestreo. Elegir entre una prueba t, una prueba Z o un ANOVA implica considerar con cuidado el diseño del estudio, el tamaño de la muestra y el tipo de datos disponibles; pero, sobre todo, exige una lectura crítica que no se limite al valor p, sino que incorpore los intervalos de confianza y el tamaño del efecto como parte del argumento pedagógico. Cuando se enseñan y aplican desde esta perspectiva, estos contrastes se transforman en un medio para fortalecer la toma de decisiones informada, fomentar la reflexión sobre la evidencia y promover una comprensión más profunda de los procesos educativos que se busca analizar.

Tabla 8.
Estadísticos descriptivos de los puntajes según modalidad de estudio

Grupo	Media	DE	n
A	6.1	1.1	25
B	5.6	1.3	25
C	6.8	1.0	25

Nota. Los datos presentan las medias, desviaciones estándar y tamaños muestrales de tres modalidades de estudio (A, B y C).

Conclusiones

El capítulo 4 mostró que la inferencia estadística es, ante todo, una forma de pensar con datos y no solo un conjunto de fórmulas. Trabajar con población, muestra y sesgo permitió comprender que toda conclusión se apoya en decisiones de muestreo que pueden acercarnos o alejarnos de la realidad que queremos estudiar. La idea de que "los datos no hablan solos" atraviesa todo el capítulo: es necesario preguntarse quiénes participaron, cómo se recogió la información y qué tipo de preguntas se pretende responder para que el análisis tenga sentido.

Al abordar la estimación puntual y por intervalos, así como el tamaño muestral y la precisión, el capítulo insistió en que toda

ISBN 978-9942-596-47-5 | 2025
https://editorial.risei.org

estimación va acompañada de incertidumbre. Lejos de ser un error, esa incertidumbre es una característica propia del trabajo con muestras y debe aprender a interpretarse con herramientas como el error estándar, los intervalos de confianza y el análisis de la variabilidad. En esa misma lógica, las pruebas de hipótesis se presentaron como un puente entre el lenguaje cotidiano y el lenguaje académico: formular $\mathbf{H_0}$ y $\mathbf{H_1}$, elegir la prueba adecuada y analizar el valor p y el tamaño del efecto solo tiene sentido si se vincula con una pregunta real sobre cambios, diferencias o relaciones entre grupos.

Finalmente, el capítulo resaltó la importancia de formar una actitud crítica frente a los resultados estadísticos. La inferencia no entrega verdades absolutas, sino evidencias que deben leerse con prudencia, reconociendo sus límites y el contexto en que se producen. Cuando el estudiantado aprende a justificar sus conclusiones, a discutir la calidad de los datos y a reconocer qué puede y qué no puede afirmarse a partir de un análisis, la estadística deja de ser una lista de procedimientos para convertirse en una herramienta de argumentación.

Referencias

Agresti, A. (2018). Statistical methods for the social sciences (5.ª ed.). Pearson.

Bakker, A. (2004). Design research in statistics education: On symbolizing and computer tools [Tesis doctoral, Utrecht University]. CD-$_0$ Press.

Batanero, C. (2001). Didáctica de la estadística. Universidad de Granada.

Batanero, C., & Ben-Zvi, D. (2013). Reasoning about variability in statistics education. Springer.

Batanero, C., & Díaz, C. (2011). Estadística y probabilidad en educación matemática. Editorial Síntesis.

Ben-Zvi, D., & Garfield, J. (2004). The challenge of developing statistical literacy, reasoning and thinking. Springer.

Biehler, R. (2018). Perspectives on modeling variability and statistical reasoning in education. Springer.

Cumming, G. (2014). The new statistics: Why and how. Psychological Science, 25(1), 7–29. https://doi.org/10.1177/0956797613504966

Garfield, J., & delMas, R. (2008). Research on reasoning about variability. En D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning and thinking (pp. 201–226). Springer.

Garfield, J., & delMas, R. (2008). Helping students learn to reason about statistical inference: A review of research. Statistics Education Research Journal, 7(2), 39–54.

Gelman, A. (2021). Regression and other stories. Cambridge University Press.

Gelman, A., & Hill, J. (2014). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Horton, N. J. (2015). Helping students think with data: Tools for improving statistical reasoning. Journal of Statistics Education, 23(2), 1–12. https://doi.org/10.1080/10691898.2015.11889730

Howell, D. C. (2017). Statistical methods for psychology (8th ed.). Cengage Learning.

Kline, R. B. (2013). Beyond significance testing: Statistics reform in the behavioral sciences (2nd ed.). American Psychological Association.

Mayo, D. G. (2018). Statistical inference as severe testing: How to get beyond the statistics wars. Cambridge University Press.

Sorto, M. A. (2006). Identifying content knowledge for teaching statistics in middle grades [Tesis doctoral, Texas State University–San Marcos].

Student. (1908). The probable error of a mean. Biometrika, 6(1), 1–25. https://doi.org/10.1093/biomet/6.1.1

Tukey, J. W. (1977). Exploratory data analysis. Addison-Wesley.