



Un acercamiento al potencial de los chatbots para generar “diálogos inteligentes” en educación estocástica

An approach to the potential of chatbots for generating
“Intelligent dialogues” in stochastic education

 **Santiago Inzunza Cazares**  | sinzunza@uas.edu.mx
Universidad Pedagógica del Estado de Sinaloa

Recepción: 11 de febrero de 2026 | Aceptación: 30 de junio de 2026

Resumen

En el artículo reflexionamos sobre el potencial de herramientas de inteligencia artificial —como los chatbots— para generar diálogos en la clase de estocástica. A partir de un contexto de muestreo, se inicia un diálogo ficticio entre un estudiante de bachillerato que cursa Temas Selectos de Estadística y ChatGPT, en el que el estudiante desea comprender algunas propiedades del muestreo aleatorio y la construcción de una distribución de muestreo mediante simulación. En el diálogo se plantean dudas sobre aspectos inferenciales que le permitan al estudiante comprender cómo se construye una inferencia estadística. Se analizan las respuestas del chat y se observan aspectos conceptuales bien respondidos, que incluso superan las preguntas planteadas por el usuario. Se concluye que los chatbots pueden constituir una herramienta de apoyo tanto para el profesor como para el estudiante, dentro y fuera del aula de clases.

Palabras clave

Inteligencia artificial, distribuciones muestrales, chatbots, educación estocástica.

 Autor de correspondencia | Santiago Inzunza Cazares

Inzunza Cazares, S. (2026). Un acercamiento al potencial de los chatbots para generar “diálogos inteligentes” en educación estocástica. *Revista Enseñanza de las Matemáticas y Experiencias Docentes*, 2(2), 117–132.
<https://doi.org/10.24844/REMEDI/0202.07>

Abstract

In this article, we reflect on the potential of artificial intelligence tools—such as chatbots—to generate dialogues in the stochastics classroom. Based on a sampling context, a fictional dialogue is initiated between a high school student taking Selected Topics in Statistics and ChatGPT, in which the student seeks to understand some properties of random sampling and the construction of a sampling distribution through simulation. The dialogue raises questions about inferential aspects that may help the student understand how statistical inference is constructed. The chatbot's responses are analyzed, and well-developed conceptual aspects are observed, some of which even go beyond the questions posed by the user. It is concluded that chatbots can serve as a support tool for both teachers and students, inside and outside the classroom.

Keywords

Artificial intelligence, sampling distributions, chatbots, stochastic education.

Introducción

Con el surgimiento reciente de los chatbots basados en inteligencia artificial (IA), se ha generado mucha expectativa sobre el impacto que pueden tener estas herramientas en las aulas de clase, pero aún se desconoce el rol que pueden desempeñar los docentes en su implementación, así como las ventajas y posibilidades que ofrecen a los estudiantes. Existen al respecto distintas opiniones en la literatura, desde las optimistas que afirman que la IA transformará la educación (Escotet, 2024; Porcelli, 2020), hasta las que consideran que el uso de la IA traerá como consecuencia la pérdida de habilidades de escritura, razonamiento matemático y pensamiento crítico, así como el plagio de tareas y exámenes, entre otras (Akçapınar & Sidan, 2024; Rudolph et al., 2023).

En particular, en el caso de México, las autoridades de la Secretaría de Educación Pública (SEP) han señalado sobre la pertinencia de utilizar la IA en el sistema educativo nacional; el titular de la SEP en la conferencia “IA y la transformación educativa en México”, impartida en la UNESCO el 24 de enero de 2025 señaló que “esta nueva herramienta ofrece la posibilidad de optimizar el tiempo de las maestras y los maestros, para el diseño de mejores estrategias pedagógicas y enfocarse en el factor humano y socioemocional, [...] la IA permite que las y los estudiantes accedan a plataformas de aprendizaje personalizado, autoformación y tutorías virtuales con experiencias interactivas y materiales educativos acorde a sus necesidades” (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO], 2025). En un primer análisis, las características de estas herramientas muestran diferencias notables en su diseño respecto a otras herramientas de software desarrolladas hasta ahora;

entre las principales, figura que el usuario se comunica, por vía oral, escrita o diagramática, con el dispositivo para ordenarle el desarrollo de un proceso matemático, de algún análisis estadístico o la resolución de un problema. En el software clásico o convencional, el usuario genera figuras geométricas, introduce datos para su análisis o una ecuación para su solución, y selecciona los pasos y procesos necesarios, dispuestos en la configuración del software; es decir, solo puede realizar una serie de acciones predeterminadas en su diseño. Estamos ante una herramienta con características distintas en lo que concierne a la comunicación e interacción humano-máquina, a las representaciones que genera y a los análisis que realiza para resolver las situaciones planteadas. Por sus características, estas herramientas pueden aplicarse en otros ámbitos del proceso educativo que no se contemplan en el software convencional (e.g., la planeación de clases y las evaluaciones).

Ante ello, cabe preguntarnos: ¿pueden los chatbots inteligentes desarrollar procesos matemáticos precisos y explícitos en términos de representaciones similares a las que emplea el software educativo tradicional? ¿Pueden realizar análisis de datos y simulaciones para ayudar a la comprensión de conceptos, al razonamiento y al pensamiento estadístico de los estudiantes? En tanto constituyen una herramienta de comunicación y diálogo entre usuario y máquina, ¿en qué medida estos diálogos pueden considerarse inteligentes?

Antecedentes

Una revisión que realizamos en Google Académico introduciendo las palabras “chatbots IA mathematics education” proporciona 18,200 artículos a partir de 2022 que han sido publicados con la temática indicada. Con apoyo de ChatGPT los hemos resumido en seis categorías: predominan los estudios centrados en aplicaciones educativas (35%). En segundo lugar, se encuentran los trabajos relacionados con el diseño y el desarrollo de chatbots educativos (23%). Un 18% analiza el impacto en el aprendizaje de las matemáticas, evaluando el rendimiento académico, la comprensión conceptual y la motivación de los estudiantes. Por su parte, el 17% de los artículos trata investigaciones más amplias sobre la IA, sin centrarse exclusivamente en los chatbots. Las áreas menos estudiadas incluyen estudios sobre percepción, ética y aceptación de la IA (7%) así como revisiones sistemáticas y trabajos teóricos (5%).

Labadze et al. (2023) en una revisión de 67 estudios relevantes sobre el uso de chatbots en educación, encontraron numerosos beneficios de su integración en las aulas, tanto para los docentes como para los estudiantes, entre los que se enumeran: tareas en casa y ayuda de estudio, aprendizaje personalizado y desarrollo de múltiples habilidades.

Moral et al. (2023) realizaron un análisis de la satisfacción en el aprendizaje de las matemáticas en estudiantes españoles de primaria sobre temas de geometría, en el que mostraron un alto grado de satisfacción con sus creaciones de IA y señalaron que este tipo de experiencias puede trasladarse a otras asignaturas y contextos educativos.

Wahba et al. (2024), por su parte, realizaron un estudio para evaluar el impacto de un entorno de aprendizaje basado en ChatGPT en el fomento del razonamiento estadístico y de las actitudes hacia la estadística entre estudiantes universitarios árabes. Los resultados muestran efectividad del uso de ChatGPT en el desarrollo del razonamiento estadístico y en el fomento de actitudes positivas hacia la estadística y de motivación por parte de los estudiantes para aprenderla.

Mientras tanto, Pepin et al. (2025) realizaron una amplia revisión sobre el potencial, desafíos e implicaciones en la educación matemática a nivel universitario, e identificaron siete áreas de investigación en las que se están realizando estudios:

1. Comprensión de conceptos matemáticos básicos.
2. Asistencia a los profesores en la planeación de lecciones.
3. Evaluaciones automatizadas y personalizadas.
4. Promoción del aprendizaje personalizado y de estrategias autorreguladas.
5. Desafíos para fomentar el pensamiento crítico.
6. Facilitación del aprendizaje colaborativo y del trabajo en grupos.
7. Actitudes positivas de profesores y estudiantes, limitadas por consideraciones éticas.

Marco conceptual

La inteligencia artificial y los chatbots

Sin pérdida de generalidad, el concepto de IA puede entenderse como la capacidad de las computadoras, mediante algoritmos, de crear máquinas capaces de realizar actividades que requieren inteligencia humana (Porcelli, 2020). El lenguaje es el medio mediante el cual se establece la comunicación entre humanos y máquinas, lo que constituye un punto crucial y desafiante para su funcionamiento; esto implica que las máquinas deben comprender el lenguaje humano. En este contexto se inscriben los chatbots.

Los chatbots corresponden a una categoría de tecnologías digitales para conversar con la computadora, el celular o una tableta electrónica, y responden a cada pregunta que se les plantea de modo similar a la comunicación entre seres humanos, y, más específicamente, como lo hace un profesor cuando un estudiante le plantea una pregunta. Un caso muy

conocido es ChatGPT, al que hacemos referencia en este artículo, con el que documentamos ejemplos del área de la estocástica.

La interacción con los chatbots se realiza mediante una instrucción del usuario, por escrito o de voz, de modo que el dispositivo da una respuesta redactada de forma “coherente”, al menos es la expectativa de su diseño, y queda al usuario verificarla. Esta instrucción es denominada *prompt*, y juega un papel central en las respuestas que el chat proporciona al usuario; en la medida en que el *prompt* es más preciso, en esa medida la respuesta también lo es. El *prompt* puede indicar que un proceso matemático se realice paso a paso para que el estudiante observe los detalles, o bien que utilice distintas estrategias para resolver un problema, pudiendo ser algunas más sencillas que otras, y, con ello, estar en condiciones de promover el pensamiento matemático de los usuarios. “El objetivo del chatbot es mantener la comunicación con el usuario simulando una comunicación humana inteligente, de tal forma que el interlocutor tenga una experiencia tan similar como sea posible a la conversación con otra persona utilizando lenguaje natural” (Cheng & Jiang, 2022).

Diálogos inteligentes

En el proceso educativo, el diálogo entre docentes y estudiantes es una premisa central. También es importante el diálogo entre los estudiantes mismos cuando participan en una tarea en clase o fuera de ella. Cuando el proceso educativo es mediado por una herramienta tecnológica, se establece un diálogo entre el usuario y la herramienta: el usuario proporciona las instrucciones y la herramienta responde a ellas. Es, por tanto, el proceso educativo un proceso interaccional y comunicativo entre quien enseña y quien aprende, pudiendo ser que quien enseña sea incluso una máquina. A este último tipo de diálogos —al del usuario con la herramienta— es al que nos referimos en este artículo.

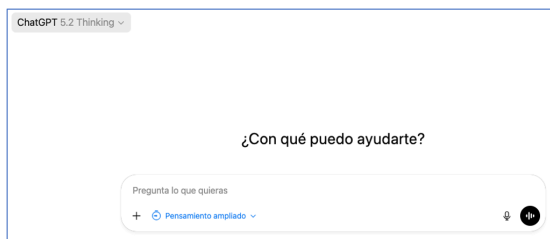
En el caso de los chatbots, la interfaz de comunicación consiste en una ventana de diálogo mediante la cual se puede introducir texto, voz, figuras y diagramas (Figura 1). Se trata de herramientas de conversación con las que se pueden establecer diálogos como los que se generan en el aula de clase entre estudiantes y docentes, ya que, ante una indicación del usuario, el chatbot emite una respuesta.

El chatbot dispone de elementos que muestran cierta inteligencia desde el diseño mismo; por ejemplo, ante una pregunta mal redactada con errores de ortografía, puede hacer correcciones antes de responder, puede emitir una respuesta y hacer sugerencias sobre alguna ampliación o modificación de la pregunta; incluso, puede anticipar preguntas posteriores que el usuario tiene en mente y que están estrechamente relacionadas con las preguntas

previas. Por parte del usuario, que es quien hace las preguntas, debe tener certeza sobre lo que quiere resolver y cómo solicitarlo al chatbot; al menos una idea de por dónde iniciar.

Figura 1

Interfaz de ChatGPT (ventana de diálogo)



En la medida en que obtenga respuestas, puede plantear nuevas preguntas, agregar elementos adicionales a las ya planteadas y, en ese sentido, establecer un diálogo con la herramienta hasta que le satisfaga con su respuesta. El usuario está ante un experto, a quien puede preguntar prácticamente cualquier tema; por lo tanto, es de suma importancia saber qué preguntar y cómo hacerlo. Ante lo anterior, cabe señalar el rol del docente, quien debe iniciar al estudiante en el uso efectivo de estas herramientas.

En suma, los diálogos con la IA suelen denominarse inteligentes porque se dan conversaciones entre el chat y el usuario, en las que el chat se adapta a las preguntas y ofrece explicaciones y retroalimentación a medida que el usuario va preguntando. En vez de solo dar información, cuando están bien diseñados, los chats pueden funcionar como un andamiaje que pregunta, detecta errores, propone ejemplos y guía el razonamiento.

Muestreo e inferencia

La selección de una muestra para inferir características de su población es una acción recurrente hoy en día en diversas aplicaciones de interés para el ser humano; esta acción da lugar a un complejo entramado de conceptos que se dan en el ámbito muestral y poblacional. Con el surgimiento y desarrollo de la tecnología educativa, se ha hecho posible el estudio de ellos por otra vía más accesible en términos didácticos para un conglomerado más amplio de personas que no cuentan con los antecedentes matemáticos que requiere el enfoque formal de la inferencia estadística, además de que da paso al tratamiento de estos conceptos a edades más tempranas, como son la secundaria y el bachillerato. Por ello, Burrill y Biehler (2011) contemplan a la inferencia como una de las ideas centrales que deben formar parte de cualquier currículo.

Dierdorp et al. (2016, p. 39) identifican como parte de la red conceptual que subyace al muestreo las ideas de: tamaño de muestra, proceso aleatorio, distribución, intervalo intuitivo de variabilidad y relación entre muestra y población. Por tanto, los estudiantes requieren comprender que:

- El incremento en el tamaño de la muestra conduce a estimaciones más precisas de la probabilidad y de características poblacionales.
- Las mediciones repetidas de un mismo fenómeno son un proceso aleatorio que conduce a resultados distintos y que las inferencias están influenciadas por la muestra seleccionada.
- La extracción de muchas muestras da lugar a una colección o distribución que puede describirse mediante medidas de centro y variabilidad.
- Una manera razonable de estimar la variabilidad alrededor de un valor esperado en un proceso aleatorio es un intervalo de confianza.

Aspectos metodológicos

Se implementó un diálogo ficticio entre el usuario, que es un estudiante de bachillerato y ChatGPT. (versión 5.2), con una intención didáctica: analizar los aspectos conceptuales de las respuestas que proporciona el chat y evaluar si las representaciones eran apropiadas y podían contribuir a la comprensión de las ideas del muestreo con una mirada hacia la inferencia estadística. En particular, se implementó el siguiente guion, basado en las ideas de Dierdorp et al. (2016):

- ¿Qué es el muestreo aleatorio?
- ¿Qué es la variabilidad muestral?
- ¿Puede ser confiable una muestra para conocer sobre una población?
- Distribución de muestreo
- Relación entre tamaño de muestra y variabilidad muestral

Las preguntas planteadas están alineadas con una ruta conceptual para acceder intuitivamente a conceptos de muestreo, sus propiedades y su papel en la realización de inferencias estadísticas. El rol del usuario es el de un estudiante de bachillerato que desconoce el bagaje conceptual, pero tiene una idea basada en las clases que le ha impartido su profesor; además, le interesa adentrarse en el tema.

1. Análisis de un diálogo con ChatGPT sobre el tema de muestreo aleatorio orientado a un proceso de inferencia estadística.

Usuario: *Me interesa el tema del muestreo aleatorio en la clase de estadística, ¿podríamos iniciar un diálogo sobre ello? Recuerda que soy un principiante de nivel bachillerato.*

ChatGPT: *¡Claro! Vamos a platicarlo desde cero, sin tecnicismos raros.*

1) ¿Qué es “muestreo aleatorio”?

Es una forma de elegir una parte de un grupo grande (una muestra) de manera que todos tengan la misma oportunidad de ser elegidos.

Población: el grupo completo que te interesa (ej. “todos los alumnos de Cobaes”).

Muestra: una parte de esa población (ej. “120 alumnos encuestados”).

Muestreo aleatorio: el método para elegir esa muestra “al azar” (como una rifa).

La idea es que, si eliges bien tu muestra, lo que observes en ella te sirve para estimar cosas de toda la población.

2) ¿Por qué es importante que sea aleatorio?

Porque ayuda a evitar el sesgo, que es cuando los datos se “inclinan” hacia un lado. Ejemplo de sesgo: Si quieres saber “cuántas horas estudian los alumnos” pero sólo encuestas a los del cuadro de honor, te va a salir un promedio más alto de lo real. Con muestreo aleatorio, reduces la probabilidad de elegir “puro cierto tipo de gente”.

3) Un ejemplo simple

Quieres estimar la estatura promedio de todos los alumnos de una escuela.

Población: 2,000 alumnos

Muestra: 100 alumnos

Aleatorio: asignas un número a cada alumno (1 a 2000) y “sacas” 100 números al azar (con una app o tabla de números aleatorios).

En este diálogo, el ChatGPT explica intuitivamente qué es el muestreo aleatorio, proporciona un ejemplo y señala que el muestreo es la base para realizar estimaciones sobre una población. Señala la importancia de que el muestreo sea aleatorio y menciona el sesgo como un aspecto que se puede evitar con este tipo de muestreo; incluso, intenta aclarar qué es el sesgo en una muestra. Termina proponiendo un ejemplo simple en el contexto de una escuela y de sus estudiantes, e incluso ejemplifica cómo realizar la selección de una muestra.

Usuario: Ya me explicaste qué es el muestreo aleatorio, por qué es importante y me pusiste un ejemplo simple que entiendo. Ahora quiero saber un poco sobre la variabilidad en el muestreo porque leí algo al respecto.

ChatGPT: La variabilidad del muestreo (o variabilidad muestral) es la idea de que si tomas varias muestras aleatorias del mismo tamaño, no te van a dar exactamente el mismo resultado, aunque todas estén “bien hechas”.

1) ¿Por qué pasa?

Porque el azar hace que cada muestra salga con personas diferentes. Y como las personas son distintas, tus números cambian un poquito. Piensa en esto:

- *La población es fija (tu escuela no cambia de un minuto a otro).*
- *Pero la muestra cambia cada vez que la sacas al azar.*
- *Por eso cambia tu estimación (porcentaje, promedio, etc.).*

2) Ejemplo rápido (porcentaje)

Quieres estimar el porcentaje de estudiantes que desayunan. Tomas tres muestras aleatorias de 50:

- *Muestra 1: 28 desayunan → 56%*
- *Muestra 2: 24 desayunan → 48%*
- *Muestra 3: 26 desayunan → 52%*

Todas pueden ser aleatorias, pero no coinciden. Ese “sube y baja” es la variabilidad del muestreo.

3) ¿De qué depende que varíe más o menos?

Principalmente de dos cosas:

a) Tamaño de la muestra (n).

- *Muestra pequeña → más variabilidad (resultados más “brincadores”).*
- *Muestra grande → menos variabilidad (resultados más estables).*

b) Qué tan “mezclada” es la población.

Si en la población hay mucha diversidad (opiniones muy divididas, por ejemplo), también tiende a haber más variación entre muestras.

4. ¿Para qué sirve entenderla?

Porque te explica por qué en estadística se habla de:

- *margen de error (“podría variar tantito”)*
- *intervalos de confianza (“este rango es razonable”)*

O sea: no es que una muestra esté “mal”; es que toda estimación tiene un error natural por el azar.

Las respuestas que ha dado el ChatGPT al usuario lo orientan a comprender que la variabilidad se debe a que no siempre en las muestras participan los mismos elementos; por ende, las estimaciones suelen variar alrededor de un valor. Eso conduce al concepto de margen de error e intervalo de confianza. Relaciona la variabilidad con el tamaño de la muestra y la uniformidad de los elementos de la población. Todo ello intuitivamente, para que el usuario tenga una idea de la variabilidad muestral y de los factores que inciden en ella.

Usuario: No me resulta fácil entender que si los resultados de cada muestra son siempre o casi siempre distintos, se utilicen para conocer algo de la población de la que se extrajo la muestra, ¿puede ser confiable hacerlo?

ChatGPT: Sí puede ser confiable, aunque suene raro al principio. La clave es esta idea: una muestra aleatoria no te da “la verdad exacta”, pero sí una buena aproximación, y además puedes medir qué tan lejos podría estar del valor real.

1. ¿Por qué confiar si cambian los resultados?

Porque aunque cada muestra sea distinta, no cambian de cualquier manera: si el muestreo es aleatorio, los resultados tienden a “bailar” alrededor del valor real de la población. Imagina un blanco de tiro:

- *El centro es el valor real de la población.*
- *Cada muestra es un “disparo”.*
- *Los tiros no caen en el mismo punto, pero si el método es bueno, caen cerca del centro y se reparten alrededor.*

2. ¿Qué garantiza la confiabilidad?

Dos cosas muy importantes:

a) Aleatoriedad (quita el sesgo)

- *Hace que la muestra no “favorezca” a un grupo.*
- *Eso permite que el error sea más “por azar” que por “trampa” involuntaria.*

b) Tamaño de muestra (reduce la variabilidad)

- *Con muestras pequeñas, el resultado puede variar bastante.*
- *Con muestras más grandes, el resultado se vuelve más estable.*

Ejemplo sencillo:

- *Si encuestas a 5 alumnos, es fácil que te toque un grupito raro.*
- *Si encuestas a 200, es más difícil que salga algo muy diferente a la realidad.*

3. “Ok, pero ¿y si justo me tocó una muestra rara?”

Esa posibilidad existe. Por eso en estadística no solo se da un número, sino una idea de cuánta incertidumbre hay:

- *Margen de error: “mi estimación puede estar más o menos cerca”*
- *Nivel de confianza: “qué tan seguro estoy de ese margen”*

No es magia: es una forma honesta de decir: “Con esta muestra, lo más razonable es que el valor real esté por aquí.”

4. La frase que lo resume todo:

La estadística no promete exactitud perfecta; promete estimaciones buenas + una medida del error.

Ejemplo rápido para que se sienta “real”

Supón que en toda tu escuela el porcentaje real que desayuna es 50%.

- *Con muestras de 30, podrías obtener 43%, 52%, 57%...*
- *Con muestras de 300, saldría más parecido: 48%, 50%, 51%...*

Los resultados no serán idénticos, pero se concentran cerca del 50% cuando la muestra es grande.

De nueva cuenta, el ChatGPT explica la confiabilidad que se puede tener en los resultados de una muestra, señalando que aun cuando estos cambian de una muestra a otra, siempre giran alrededor del valor real poblacional. Relaciona la confiabilidad con el tamaño de la muestra y la aleatoriedad, utilizando ejemplos al nivel de un estudiante de bachillerato. Concluye señalando que en estadística, las estimaciones suelen ser buenas y se proporciona una medida del error.

Usuario: *Me gusta el ejemplo que propones del porcentaje de estudiantes que desayuna es del 50%. ¿Puedes simular la selección de 10 muestras aleatorias y calcular el porcentaje de estudiantes que desayunan?*

ChatGPT. *Simulé 10 muestras aleatorias suponiendo que en la población el porcentaje real que desayuna es $P = 0.50$ (50%) y que cada muestra tiene $n = 50$ estudiantes. Resumen de estas 10 muestras:*

- *Promedio de los porcentajes: 52.4%*

- *Mínimo:* 40%
- *Máximo:* 64%

Si quieres, lo repetimos con otro tamaño de muestra (por ejemplo $n = 20$ y luego $n = 200$) para que veas cómo cambia la variabilidad.

Tabla 1

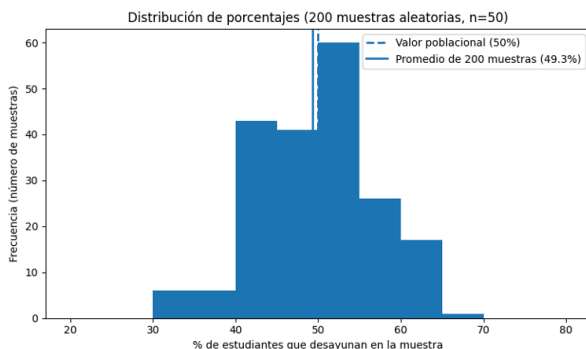
Simulación de 10 muestras aleatorias ($P = 0.50$, $n = 50$)

Muestra	Tamaño (n)	Desayunan
1	50	28(56%)
2	50	24(48%)
3	50	29(58%)
4	50	27 (54%)
5	50	20 (40%)
6	50	32(64%)
7	50	28(56%)
8	50	28(56%)
9	50	21(42%)
10	50	25(50%)

Usuario: *Ahora construye una gráfica con la distribución de las muestras que me ayuden a visualizar los resultados para 200 muestras.*

Figura 2

Distribución de porcentajes muestrales (200 muestras aleatorias, $n=50$)



En esta simulación:

- Promedio de los 200 porcentajes: 49.3%
- Desviación estándar: 6.87 puntos porcentuales
- Mínimo: 30%
- Máximo: 66%

La línea punteada es el 50% real de la población y la línea continua es el promedio de las 200 muestras. Esto muestra que, con muchas muestras, se ve claramente cómo los resultados se concentran alrededor del valor real.

Lo que ha hecho ChatGPT es simular 200 muestras de tamaño 50 de una población con $P = 0.50$ y construir una gráfica con la distribución del muestreo, poniendo especial énfasis en el centro de la distribución (línea continua) y el valor poblacional (línea punteada) para que el usuario observe su relación y cercanía, que es una propiedad de suma importancia en inferencia estadística, y muestra el potencial del muestreo aleatorio para la estimación en poblaciones.

Usuario: *Observo que el 50% real de la población es muy cercano al promedio de las muestras*

ChatGPT: *¡Exacto! Y eso pasa por una idea central en estadística: “cuando el muestreo es aleatorio, el promedio de muchas muestras tiende a acercarse al valor real de la población”.*

¿Qué significa para la confiabilidad?

Una sola muestra puede salir algo lejos del 50% (eso es variabilidad).

Pero si repitieras el proceso muchas veces, el promedio de los resultados se pega al 50%.

Eso nos dice que el método es confiable (no está sesgado).

Si quieres, hacemos una comparación visual rápida: $n=20$ vs $n=200$ (manteniendo 200 muestras), para que veas cómo con muestras más grandes la distribución se “aprieta” alrededor del 50%.

Ya están las dos gráficas con 200 muestras aleatorias comparando:

$n = 20$ (muestras pequeñas)

Promedio: 50.2%

Variabilidad (desv. estándar): ≈ 10.15 puntos %

Rango observado: 20% a 75%

$n = 200$ (muestras grandes)

Promedio: 50.1%

Variabilidad (desv. estándar): ≈ 3.78 puntos %

Rango observado: 39% a 60.5%

Lo que debes notar en las gráficas (ver figura 3):

Con $n = 20$, el histograma está mucho más “extendido”: los porcentajes cambian bastante entre muestras.

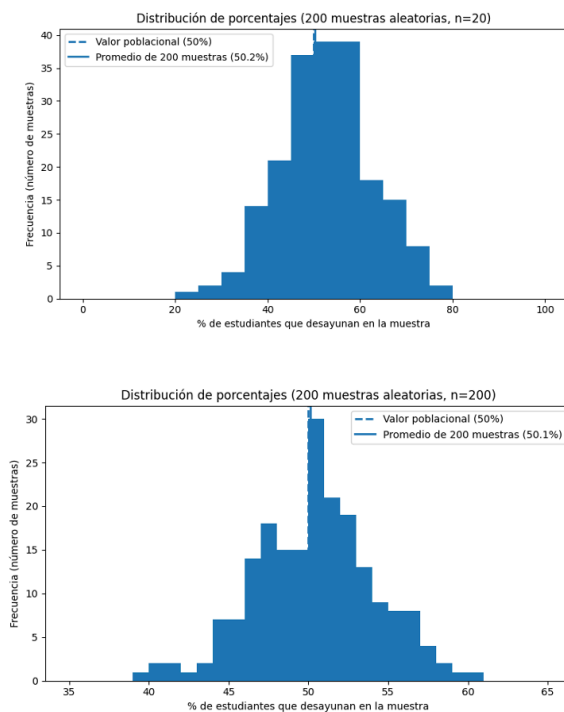
Con $n = 200$, el histograma está más “apretado” alrededor del 50%: los resultados son más estables.

Lo que ha hecho ChatGPT es simular tamaños de muestra de 20 y de 200 de la misma población y comparar ambas distribuciones. Es de crucial importancia que el usuario identifique propiedades como el efecto del tamaño de muestra en la variabilidad del muestreo y que, a pesar de las diferencias entre tamaños de muestra, el centro permanezca prácticamente

invariable; ambas son nociones intuitivas que se derivan del teorema del límite central.

Figura 3

Distribuciones de muestreo para $n=20$ y $n=200$ (200 muestras)



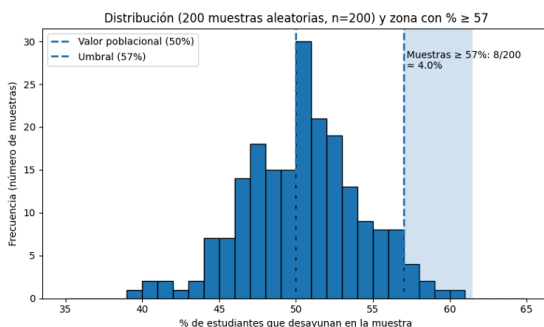
Usuario: Para este último caso (distribución con $n=200$) ¿qué porcentaje de muestras podrían estar reportando que 57% de los estudiantes o más desayunan? Si me lo puedes señalar en la gráfica mucho mejor

ChatGPT: En este último caso (200 muestras, $n = 200$, $P = 50\%$), el porcentaje de muestras que reportan 57% o más es 8 de 200 muestras, es decir, $\approx 4.0\%$. En la gráfica te lo señalé así:

Línea punteada en 57% (el umbral).

Zona sombreada a la derecha = todas las muestras con $\% \geq 57\%$.

Esta última pregunta que el usuario ha planteado a ChatGPT tiene la intención de ver si realiza el cálculo correcto del porcentaje de muestras que están en una sección de la distribución, que en su equivalente teórico significa calcular el valor de la probabilidad de que una muestra contenga el 57% o más de estudiantes que desayunan. Esta cola derecha de la distribución se conoce como p -valor en pruebas de hipótesis. El chat lo ha hecho bien e incluso ha sombreado la porción para mayor visibilidad.

Figura 4*Distribución de muestreo con área delimitada*

Conclusiones e implicaciones para la enseñanza

La definición del *prompt* es sumamente importante para lograr “diálogos inteligentes”; su redacción implica pensar de antemano en los conceptos y procesos involucrados en un problema o situación matemática. Ante ello, se requiere una planeación didáctica de las actividades y de los procedimientos que se solicitarán al chatbot. Otra implicación tiene que ver con la viabilidad de utilizar chatbots inteligentes en la clase de estocástica, pues en los diálogos analizados, las respuestas de ChatGPT son conceptualmente correctas a nivel de bachillerato. Las simulaciones solicitadas se ejecutaron correctamente y las representaciones gráficas utilizadas para mostrar las distribuciones de muestreo cumplen todos los requisitos de escala e incluyen los elementos necesarios para su interpretación.

Los diálogos entre el usuario y ChatGPT sobre el tema de muestreo e inferencia estadística muestran que la IA puede constituirse en una herramienta de apoyo al docente para planear e impartir sus clases y asignar tareas a los estudiantes; por parte de estos, el trabajo en clase o fuera de ella se potencia al contar con un asesor virtual a su alcance. Consideramos que los alumnos pueden, con una planeación didáctica pertinente, acceder a conceptos y razonamientos correctos sobre las ideas del muestreo y la inferencia mediante la simulación.

Un aspecto negativo de la IA que preocupa a muchos es que puede presentar sesgos y errores en la información que proporciona; no ha sido el caso de las respuestas que proporcionó a las actividades que se le plantearon. Considero que las ventajas de cálculo y los diálogos que se generan entre el usuario y la herramienta superan los posibles errores que pueda cometer la IA. Los estudiantes y docentes deben ser conscientes de estos posibles errores y sesgos para utilizar la herramienta racionalmente en el aula de clases. Entre las limitaciones del estudio se encuentra el hecho de que los diálogos entre estudiante y el chat fueron ficticios; aunque se trata de una

temática del bachillerato, la trayectoria de conceptos y razonamientos puede ser desafiante para muchos estudiantes. Los aspectos éticos y posibles errores de los chatbots deben ser inculcados a los estudiantes para un mejor aprovechamiento de la herramienta, como lo señalan algunos autores citados en los antecedentes.

Referencias

- Akçapınar, G., & Sidan, E. (2024). AI chatbots in programming education: Guiding success or encouraging plagiarism. *Discover Artificial Intelligence*, 4(87). <https://doi.org/10.1007/s44163-024-00203-7>
- Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 57–69). Springer. https://doi.org/10.1007/978-94-007-1131-0_10
- Cheng, Y., & Jiang, H. (2022). Customer–brand relationship in the era of artificial intelligence: Understanding the role of chatbot marketing efforts. *Journal of Product & Brand Management*, 31(2), 252–264. <https://doi.org/10.1108/JPBM-05-2020-2907>
- Dierdorp, A., Bakker, A., van Maanen, J., & Eijkelhof, H. (2016). Supporting students to develop concepts underlying sampling and to shuttle between contextual and statistical spheres. En D. Ben-Zvi & K. Makar (Eds.), *The teaching and learning of statistics* (pp. 37–48). Springer. https://doi.org/10.1007/978-3-319-23470-0_4
- Escotet, M. A. (2024). The optimistic future of Artificial Intelligence in higher education. *Prospects*, 54(3–4), 531–540. <https://doi.org/10.1007/s11125-023-09642-z>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(56), 1–17. <https://doi.org/10.1186/s41239-023-00426-1>
- Moral, S., Ruiz, F., & Cebrian, M. (2023). Analysis of artificial intelligence chatbots and satisfaction for learning in mathematics education. *International Journal of Educational Research and Innovation*, 20, 1–14. <https://doi.org/10.46661/ijeri.8196>
- Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*, 11, 9–41. <https://doi.org/10.1007/s40751-025-00172-1>
- Porcelli, A. M. (2020). La inteligencia artificial y la robótica: Sus dilemas sociales, éticos y jurídicos. *Derecho Global: Estudios sobre Derecho y Justicia*, 6(16), 49–105. <https://doi.org/10.32870/dgedj.v6i16.286>

- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2025, 24 de enero). *La inteligencia artificial estará al servicio de la educación y de las personas en México*. <https://www.unesco.org/ethics-ai/es/articles/la-inteligencia-artificial-estara-al-servicio-de-la-educacion-y-de-las-personas-en-mexico>
- Wahba, F., Ajlouni, A., & Abumosa, M. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *EURASIA Journal of Mathematics, Science and Technology Education*, 20(7), 1–14. <https://doi.org/10.29333/ejmste/14726>