



Prácticas innovadoras en los procesos de formación de formadores

DOI Explorando la variabilidad de respuestas en unidades didácticas generadas con inteligencia artificial: un estudio comparativo sobre la enseñanza de la probabilidad



RECIE
REVISTA CARIBEÑA DE INVESTIGACIÓN EDUCATIVA

ISSN (versión digital): 2960-771X
ISSN (versión impresa): 2960-7701

Este trabajo tiene licencia CC BY 4.0.

5^o CONGRESO CARIBEÑO DE
**INVESTIGACIÓN
EDUCATIVA**

Explorando la variabilidad de respuestas en unidades didácticas generadas con inteligencia artificial: un estudio comparativo sobre la enseñanza de la probabilidad

Exploring Response Variability in AI-Generated Didactic Units: A Comparative Study on Probability Teaching

Luz Andrea Romero-Fajardo¹

Daniel Enrique Niño-Porras²

William Alfredo Jiménez³

Resumen

Este experimento analiza los cambios en los resultados generados mediante el uso de inteligencias artificiales en la creación de tareas sobre probabilidad clásica. Mediante el uso explícito de un *prompt*, se generaron cuatro tareas para comparar en diferentes plataformas de inteligencia artificial los resultados, explorando por qué un mismo *prompt* produce respuestas distintas. Para evaluar la calidad de las actividades, se utilizaron algunos parámetros de calidad basados en algunos referentes académicos que miden la claridad conceptual, la orientación pedagógica, la claridad del contenido, la calidad de actividades y el uso de materiales didácticos. Los resultados revelan diferencias significativas en las orientaciones de clase, los tipos de actividades propuestas, la integración de contextos reales y el avance en el desarrollo de contenidos. Este trabajo contribuye al análisis crítico del uso de estas herramientas en la educación.

Palabras clave: inteligencia artificial, probabilidad, prompts, tecnología educativa, unidades didácticas.

Abstract

This experiment analyzes the changes in the results generated by the use of artificial intelligence (AI) in the creation of tasks on classical probability. Through the explicit use of a *prompt*, four tasks were generated to compare the results on different AI platforms, exploring why the same *prompt* produces different responses. To evaluate the quality of the activities, it was analyzed with some quality parameters based on some academic references that measure conceptual clarity, pedagogical orientation, clarity of content, quality of activities and the use of teaching materials. The results reveal significant differences in class orientations, the types of activities proposed, the integration of real contexts and the progress in content development. This work contributes to the critical analysis of the use of these tools in education.

Keywords: artificial intelligence, probability, prompt, educational technology, didactic units.

¹ Universidad Pedagógica Nacional. Colombia, laromero@upn.edu.co, <https://orcid.org/0009-0001-7474-0699>

² Universidad Pedagógica Nacional. Colombia, deninop@upn.edu.co, Universidad Pedagógica Nacional, <https://orcid.org/0009-0007-9241-0021>

³ Universidad Pedagógica Nacional. Colombia wjimenez@pedagogica.edu.co, Universidad Pedagógica Nacional, <https://orcid.org/0000-0003-0795-6059>

1. Introducción

La enseñanza de la probabilidad constituye un desafío significativo en la educación matemática debido a su naturaleza abstracta y a las dificultades que implica conectar conceptos teóricos con aplicaciones prácticas en contextos reales. Según Batanero et al. (2013), es importante proporcionar una cultura que permita al ciudadano participar en la sociedad de la información a partir de la enseñanza de la estadística y la probabilidad. Este reto es particularmente evidente en los niveles básicos y medios de la educación, donde se espera que los estudiantes comprendan conceptos fundamentales como eventos aleatorios, espacios muestrales y el cálculo de probabilidades. En este sentido, siguiendo a Sandoval (2024), las herramientas tecnológicas, y en particular las plataformas de inteligencia artificial (IA), ofrecen nuevas oportunidades para apoyar el diseño de actividades que optimicen el tiempo de planeación para la enseñanza de estos contenidos.

La inteligencia artificial permite a las computadoras hacer tareas siguiendo instrucciones precisas. Usando algoritmos, la IA puede dar resultados esperados y está cambiando muchos campos, entre ellos la educación. Sin embargo, a pesar de su creciente uso, existe poca investigación que explore la consistencia y calidad de los materiales educativos generados mediante IA, especialmente en áreas específicas como las Matemáticas y en un experimento específico para enseñar la probabilidad clásica, que es uno de los temas iniciales reportados en grado 7 para Matemáticas en los Derechos Básicos de Aprendizaje (MEN, 2016).

En este contexto, durante el segundo semestre de 2024, se llevó a cabo en la Universidad Pedagógica Nacional de Colombia el curso titulado Tecnología en Ciencias y Matemáticas, como parte de la Maestría en Docencia de las Matemáticas. Este espacio reunió a 12 profesores de Matemáticas en ejercicio y 3 estudiantes en formación, invitados, de la Licenciatura en Matemáticas, con el propósito de explorar el potencial de las herramientas tecnológicas para mejorar las prácticas pedagógicas. Entre los objetivos del curso, destacó el diseño de actividades interactivas que integraran recursos tecnológicos y pudieran implementarse en entornos digitales, como laboratorios de informática o plataformas en línea.

Una de las tareas de este curso fue generar una unidad didáctica utilizando únicamente herramientas de IA, con la restricción de no modificar los productos generados. Este ejercicio reveló una característica intrigante: ¿Cómo un mismo *prompt* puede producir resultados diferentes en múltiples ejecuciones dentro de una misma plataforma o al ser utilizado en distintas plataformas de IA, realizando la ejecución de manera sincrónica bajo iguales condiciones de conexión a internet y equipos? Esta variabilidad, atribuida al modelo probabilístico que surge en las IA generativas, nos hizo plantar preguntas fundamentales sobre la confiabilidad, la aplicabilidad y el contenido de valor que arrojan los resultados de estas herramientas en la enseñanza, sin realizar modificaciones a sus resultados.

Con las dudas respecto a la primera exploración al crear unidades didácticas haciendo uso de inteligencia artificial, tomamos la decisión de replicar la actividad delimitando el *prompt* para el diseño de una tarea explícita en la introducción de un contenido de estadística.

2. Metodología

El enfoque metodológico de este estudio combina un diseño experimental con un análisis cualitativo de los resultados. Para explorar la variabilidad en las unidades didácticas generadas por IA, se seleccionaron algunas plataformas de IA populares como ChatGPT, Claude.ai, Gemini y Copilot. Para la realización del experimento, se seleccionó el cálculo de probabilidad simple como temática central del estudio.

En un primer momento, se elaboró cuidadosamente un *prompt*, garantizando que fuera lo suficientemente claro y específico, de modo que dichas plataformas generaran productos educativos coherentes y aplicables en un aula de Matemáticas. Esto concuerda con las recomendaciones dadas por Sigel (2023) para elaborar *prompts* objetivos. Este *prompt*, en el que se le indicó a la IA adoptar un rol de profesor de Matemáticas de Bachillerato, de Colombia, con vastos conocimientos de estadística y probabilidad, se ejecutó dentro de cada plataforma para observar la variabilidad en los resultados generados. Se tomó registro de video de las actividades propuestas, así como de las explicaciones entregadas por las IA.

Figura 1
Prompt diseñado para el experimento realizado

Asume el rol de un/a profesor/a colombiano/a de Matemáticas de grado séptimo de Bachillerato con profundos conocimientos de estadística y probabilidad. Tienes preferencia por introducir las temáticas a partir de situaciones contextualizadas y tu enfoque metodológico principal es la Formulación y Resolución de Problemas de Pólya. Tus estudiantes están inmersos en los deportes de fútbol, baloncesto, tenis de mesa, danzas y videojuegos. Crea una tarea que haga uso de material tangible para introducir a la probabilidad simple, que tenga un ejemplo contextualizado y un ejemplo numérico. Además, es necesario que esta tarea tenga, por lo menos, cinco ejercicios que impliquen desarrollar problemas en contextos relacionados con la cotidianidad de los estudiantes.

Luego de obtener los resultados, se realizó un análisis comparativo entre los productos generados por las cuatro plataformas. Este análisis se llevó a cabo cualitativamente, identificando patrones comunes y discrepancias significativas en los enfoques pedagógicos y conceptuales. Además, se tuvo en cuenta a Barragán & Barrera (2021), quienes proponen en su trabajo, de acuerdo con el modelo exhaustivo, unas categorías de análisis de tareas recopiladas de distintas editoriales educativas, de las cuales se resaltan *objetivos, contenidos, conexiones, actividades, metodología, lenguaje, tecnología, enfatización y recursos generales*. Además, se consideraron cuatro categorías de análisis basado en el *prompt*: calidad del contenido matemático, situaciones contextualizadas, enfoque de Pólya, uso de material tangible.

3. Resultados

Teniendo en cuenta que se utilizó el mismo *prompt*, los resultados del experimento evidenciaron una notable variabilidad en las instrucciones de las tareas generadas por las plataformas de inteligencia artificial. El análisis permitió identificar diferencias en términos de precisión conceptual, diseño pedagógico y contextualización en las propuestas para introducir la probabilidad simple en el séptimo grado.

En términos conceptuales, todas las IA lograron acercarse de forma correcta a los conceptos básicos de probabilidad simple, como eventos simples y cálculo de probabilidades. Por un lado, los resultados generados por ChatGPT y Claude.ai incluyeron explicaciones más detalladas y completas, con pasos claros y justificaciones matemáticas sólidas. Por otro lado, Gemini y Copilot presentaron definiciones más breves y directas, adecuadas para estudiantes de niveles básicos, pero con menos interés en el razonamiento que implica este contenido.

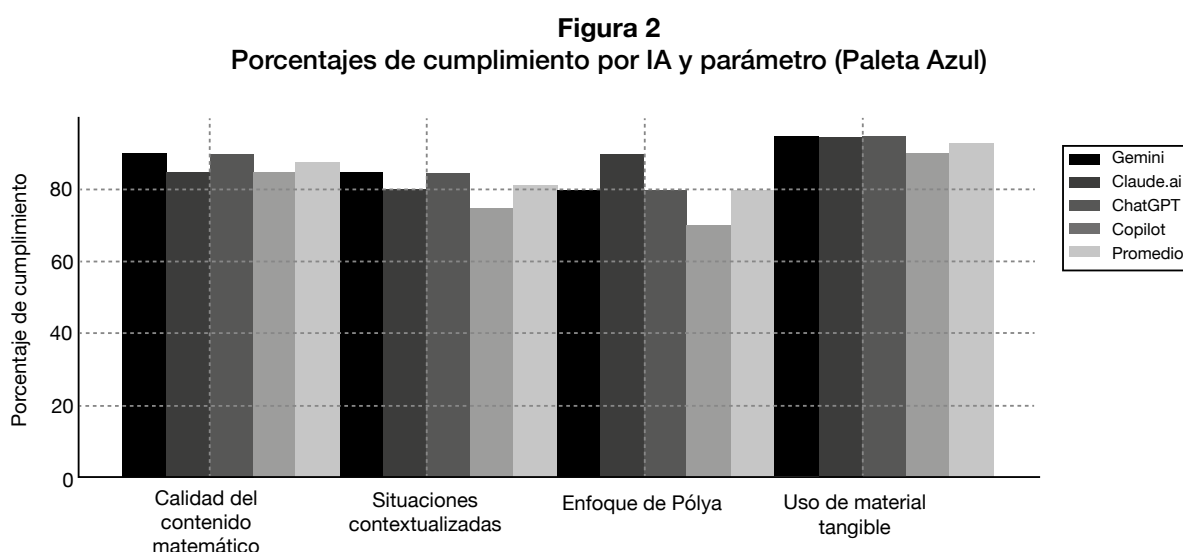
En cuanto al diseño pedagógico (parámetro analizado en la Tabla 1), las propuestas cambiaron notablemente. Claude.ai fue la más consistente al aplicar la metodología de Pólya, estructurando las actividades en fases claras (comprensión, diseño, ejecución y verificación), lo que facilitó un enfoque reflexivo para los estudiantes. ChatGPT, aunque no incluyó explícitamente las etapas de Pólya en todas las actividades, diseñó tareas bien estructuradas que promovían el aprendizaje autónomo y la exploración. Por otro lado, Gemini y Copilot se centraron en ejercicios más cerrados y repetitivos, con menos oportunidades para que los estudiantes desarrollaran estrategias propias o reflexionaran sobre los resultados.

Tabla 1
Ejemplo de parámetro analizado a partir
de la percepción de los autores (interpretación subjetiva)

| Parámetro evaluación | IA evaluación | % Cumplimiento | Observaciones | Puntos a mejorar |
|----------------------|---------------|----------------|--|--|
| Enfoque de Pólya | Gemini | 80 % | El método de Pólya está explícitamente integrado y bien aplicado. | Incluir ejemplos explícitos de reflexión y mejora en la retrospección para fortalecer el último paso. |
| Enfoque de Pólya | Claude.ai | 90 % | Se menciona claramente el enfoque de Pólya y se guía a los estudiantes a través de sus etapas. | Mejorar el componente de retrospección, guiando a los estudiantes a reflexionar sobre errores comunes. |
| Enfoque de Pólya | ChatGPT | 80 % | Se menciona el enfoque de Pólya y sus etapas están implícitas en los ejercicios y ejemplos. | Mejorar la etapa de retrospección con preguntas guiadas para reflexionar sobre estrategias y errores. |
| Enfoque de Pólya | Copilot | 70 % | Se menciona explícitamente el enfoque de Pólya, pero no se desarrolla suficientemente en los ejercicios. | Fortalecer la etapa de retrospección y guiar más claramente a los estudiantes en cada fase del método. |

En relación con la relevancia contextual, ChatGPT y Gemini sobresalieron por mostrar ejemplos directamente relacionados con los intereses propuestos para los estudiantes, como deportes (fútbol, baloncesto) y videojuegos, evidenciando una conexión entre los conceptos matemáticos y la vida cotidiana. Sin embargo, los ejemplos de Claude.ai y Copilot tendieron a ser generales y clásicos, con menos vínculos explícitos con el contexto colombiano o los intereses específicos mencionados en el *prompt*.

En la Figura 2 se presenta, a modo de comparación, el porcentaje evaluado para cada parámetro de la tarea generada por cada plataforma de inteligencia artificial.



Este análisis refleja cómo las características estadísticas y probabilísticas de fondo generadas por las IA influyen en la variabilidad y consistencia de los materiales educativos generados. Además, hay que tener en cuenta la importancia de un diseño cuidadoso de *prompts* y una revisión crítica por parte de los docentes para garantizar la calidad de las actividades o los recursos creados por estas herramientas.

4. Discusión y conclusiones

Es fundamental verificar los contenidos que arrojan las diferentes inteligencias artificiales y los resultados, ya que muchas veces pueden cometer errores en los resultados o proporcionar información no necesaria e incluso al margen de lo solicitado en el *prompt*. Además, ninguna de las tareas muestra de manera explícita el cómo. Para el desarrollo de la clase o la secuencia que debe dirigir el profesor de manera paralela durante el desarrollo de la actividad, siempre es necesario el acompañamiento para clarificar o situar específicamente las recomendaciones de lo expuesto en los resultados de la IA.

Aunque algunas aplicaciones como ChatGPT y Claude.ai resaltan por su precisión conceptual y orientación pedagógica, ninguna logra ofrecer por completo una orientación didáctica clara para la clase ni abordar con profundidad el contexto cultural del aula colombiana ((no se les solicitó a las IA abordar contextos de manera muy específica, pero son aspectos que siempre están en la mente de un profesor que tiene presente a la población y su entorno directo e indirecto). En conclusión, las IA son herramientas valiosas para complementar la planificación docente, pero no reemplazan el juicio profesional necesario para adaptar, supervisar y guiar los procesos de aprendizaje. Es crucial continuar investigando cómo maximizar su impacto educativo.

5. Referencias bibliográficas

- Batanero, C., Díaz, C., Contreras, J. M., & Roa, R. (2013). El sentido estadístico y su desarrollo. *Números. Revista de Didáctica de las Matemáticas*. <https://r.issu.edu.do/pu>
- Barragán, M., & Barrera, J. (2021). *El concepto estadístico de dispersión: una caracterización a partir de textos escolares colombianos*. Universidad Pedagógica Nacional. <https://r.issu.edu.do/e4A>
- Garzón, A., & García, M. (2009). Diseño de una secuencia de actividades para la enseñanza de la probabilidad simple en estudiantes de sexto grado: aplicación y validación. *10.º Encuentro Colombiano de Matemática Educativa*.
- Gómez, P., Mora, M., & Velasco, C. (2018). Capítulo 5. Análisis de instrucción. En P. Gómez, M. Mora, & C. Velasco, *Formación de profesores de Matemáticas y Práctica de Aula* (pp. 197-268). <https://r.issu.edu.do/5UU>
- Ministerio de Educación Nacional (MEN). (1998). *Lineamientos curriculares en Matemáticas*. MEN. <https://r.issu.edu.do/va>
- Ministerio de Educación Nacional (MEN). (2006). *Estándares básicos de competencias en Matemáticas*. MEN.
- Ministerio de Educación Nacional (MEN). (2016). *Derechos básicos de aprendizaje en Matemáticas V2*. MEN.
- Sandoval Hernández, M. A., Morales Alarcón, G. J., Vázquez Leal, H., Huerta Chua, J., & Filobello Niño, U. A. (2024). El uso del prompt de ChatGPT como asistente en la educación. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 14(28). <https://doi.org/10.23913/ride.v14i28.1872>
- Sigel, J. (2023). El arte del prompt: Cómo sacar el máximo provecho de la IA generativa. *Microsoft*. <https://r.issu.edu.do/NdB>