



El futuro de la estadística en la investigación científica: una revisión narrativa de las tendencias emergentes

The future of statistics in scientific research: a narrative review of emerging trends

Artículo de revisión documental | Documentary review article

RESUMEN

INTRODUCCIÓN: La rápida expansión del big data, la inteligencia artificial y la ciencia de datos ha modificado sustancialmente el rol de la estadística en la investigación científica. Tradicionalmente considerada el pilar metodológico del método científico, la estadística ha proporcionado las herramientas fundamentales para el diseño experimental, el análisis de datos y la inferencia de incertidumbre. Sin embargo, el desarrollo de técnicas computacionales avanzadas ha reconfigurado el panorama analítico, generando nuevas oportunidades y desafíos. **EL OBJETIVO** de este estudio es analizar las tendencias emergentes que redefinen el papel de la estadística en la investigación científica contemporánea. **METODOLOGÍA:** se adoptó un enfoque cualitativo basado en una revisión narrativa de la literatura científica publicada entre 2020 y 2025. Tras aplicar criterios de inclusión y exclusión en diversas bases de datos académicas, se seleccionaron 20 estudios relevantes para su análisis conceptual. **LOS RESULTADOS** identifican cuatro tendencias principales: 1) la transición hacia marcos metodológicos híbridos que integran la inferencia estadística con el aprendizaje automático; 2) la urgencia de desarrollar modelos predictivos interpretables; 3) el fortalecimiento de las prácticas de ciencia abierta para combatir la crisis de reproducibilidad; y 4) la transformación de la formación estadística hacia competencias computacionales. **EN CONCLUSIÓN**, la revisión sugiere que la estadística no está siendo desplazada, sino que evoluciona de manera interdisciplinar para continuar desempeñando un papel central en la validación de modelos, el manejo de la incertidumbre y la generación de conocimiento científico fiable.

Autor

Vindell González, Juan José^{1,2}

¹ Unidad de Investigación, Facultad de Medicina, Universidad de El Salvador.

² ORCID: <https://orcid.org/0000-0002-4640-4452>

Correspondencia

juan.vindell@ues.edu.sv

Presentado

20 marzo 2026

Aceptado

11 mayo 2026

Handle: <https://hdl.handle.net/20.500.14492/33588>

<https://doi.org/10.66778/SI.vo4no2.07>

Como citar este artículo:

Vindell González, JJ. El futuro de la estadística en la investigación científica: una revisión narrativa de las tendencias emergentes. Rev. Sal. Int. 2026; 4 (2):77-84 <https://doi.org/10.66778/SI.vo4no2.07>

Este contenido está protegido bajo la licencia CC BY
<https://creativecommons.org/licenses/by/4.0/>



Plataforma digital de la revista: <https://revistas.ues.edu.sv/index.php/si/>

PALABRAS CLAVE:

Investigación biomédica, inteligencia artificial, macrodatos, aprendizaje automático.

ABSTRACT:

BACKGROUND: The rapid expansion of big data, artificial intelligence and data science has substantially altered the role of statistics in scientific research. Traditionally considered the methodological pillar of the scientific method, statistics has provided the fundamental tools for experimental design, data analysis, and inference under uncertainty. However, the development of advanced computational techniques has reshaped the analytical landscape, creating new opportunities and challenges. **OBJECTIVE:** This study aims to analyze the emerging trends that are redefining the role of statistics in contemporary scientific research. **METHODOLOGY:** To this end, a qualitative approach was adopted based on a narrative review of scientific literature published between 2020 and 2025. After applying inclusion and exclusion criteria across various academic databases, 20 relevant studies were selected for conceptual analysis. **RESULTS:** The findings reveal four major trends: 1) the transition toward hybrid methodological frameworks that integrate statistical inference with machine learning; 2) the urgent need to develop interpretable predictive models; 3) the strengthening of open science practices to combat the reproducibility crisis; and 4) the transformation of statistical education toward computational competencies. **IN CONCLUSION,** the review suggests that statistics is not being displaced but is rather evolving in an interdisciplinary manner to continue playing a central role in model validation, uncertainty management, and the generation of reliable scientific knowledge.

KEYWORDS:

biomedical research; artificial intelligence; big data; machine learning

INTRODUCCIÓN

La estadística ha sido históricamente uno de los pilares epistemológicos del método científico al proporcionar las herramientas fundamentales para el diseño de estudios, la inferencia sobre poblaciones y la cuantificación de la incertidumbre. Desde su consolidación como disciplina, ha permitido transformar datos empíricos en evidencia sistemática, lo que ha facilitado la toma de decisiones en

múltiples áreas del conocimiento. Sin embargo, durante el último lustro (2020-2025), la estadística ha experimentado una transformación acelerada. Este cambio ha sido fuertemente impulsado por la expansión del *big data*, la consolidación de la ciencia de datos y el rápido desarrollo de la inteligencia artificial (IA) y el aprendizaje automático (1,2). En este nuevo ecosistema científico, caracterizado por la disponibilidad masiva de información y el aumento exponencial de la capacidad computacional, la disciplina ha comenzado a redefinir sus enfoques metodológicos tradicionales para adaptarse a problemas analíticos más complejos. Diversos autores señalan que el futuro de la estadística radica en la creación de marcos analíticos híbridos que integren la modelización clásica con herramientas computacionales y de inferencia causal (3,4).

Simultáneamente, la evolución del ecosistema científico contemporáneo ha puesto de manifiesto complejos desafíos metodológicos. El uso inapropiado de técnicas estadísticas, la interpretación incorrecta de los resultados inferenciales y la dependencia excesiva del valor *p* amenazan la validez de los estudios (5). En consecuencia, la actual crisis de reproducibilidad científica ha desencadenado un debate global sobre la urgencia de adoptar estándares más rigurosos, fomentar la transparencia y promover el acceso abierto a los datos y códigos (6). El auge de la ciencia de datos ha generado nuevas oportunidades, pero también ha generado cuestionamientos sobre el papel de la estadística en los procesos actuales. Investigaciones recientes destacan que, lejos de ser desplazada por métodos algorítmicos puramente predictivos, la estadística mantiene un papel indispensable para garantizar la interpretación rigurosa de los resultados, la gestión de la incertidumbre y la validación científica de los modelos (7,8). Esta realidad también ha incidido en la formación académica de los investigadores. Hoy en día, la enseñanza de la estadística exige incorporar competencias en programación y en aprendizaje automático, fomentando un ecosistema de investigación altamente interdisciplinario, en el que estadísticos, científicos de datos y especialistas de diversas áreas colaboran estrechamente (9). En este contexto de convergencia tecnológica y revisión crítica de los fundamentos científicos, surge una pregunta central para la academia: ¿Cómo evolucionará la estadística para mantener su rigor metodológico y su relevancia en la era de la IA y el análisis masivo de datos? A pesar del creciente volumen de estudios sobre estas nuevas tecnologías, aún existe una síntesis conceptual limitada sobre cómo se está redefiniendo el papel de la estadística. Por ello, este artículo tiene como objetivo analizar y sintetizar las transformaciones, desafíos

y tendencias emergentes que configuran el futuro de la disciplina, a partir de una revisión narrativa de la literatura científica publicada entre 2020 y 2025.

MATERIALES Y MÉTODOS

Diseño del estudio

Esta investigación adoptó un enfoque cualitativo sustentado en una revisión narrativa de la literatura científica, con el propósito de examinar las principales transformaciones metodológicas que atraviesa la estadística en el marco del avance contemporáneo de la inteligencia artificial, la ciencia de datos y el análisis computacional de alto rendimiento

Las revisiones narrativas constituyen una estrategia metodológica ampliamente utilizada para sintetizar y analizar críticamente el conocimiento disponible en campos científicos caracterizados por un desarrollo conceptual y tecnológico acelerado. Este tipo de revisión permite integrar aportes teóricos y metodológicos provenientes de múltiples disciplinas, facilitando la identificación de tendencias emergentes, debates científicos y desafíos futuros en áreas de investigación en constante evolución. En este sentido, el presente estudio busca ofrecer una síntesis conceptual que permita comprender la evolución de la estadística en el ecosistema actual de análisis de datos, caracterizado por la disponibilidad masiva de información, el desarrollo de algoritmos de aprendizaje automático y el creciente uso de metodologías computacionales en la investigación científica.

Estrategia de búsqueda de literatura

La búsqueda bibliográfica se realizó en diversas bases de datos académicas y repositorios científicos, como PubMed, Google Scholar, DOAJ y arXiv, así como en revistas científicas disponibles en plataformas de difusión académica seleccionadas por su amplia cobertura en la literatura científica en áreas relacionadas con la estadística, la ciencia de datos e inteligencia artificial. Para la selección de los estudios se emplearon combinaciones de palabras clave en inglés, utilizando operadores booleanos para ampliar y refinar los resultados de la búsqueda. Entre los principales términos utilizados se incluyen:

- statistics and artificial intelligence
- data science and statistics
- interpretable machine learning
- statistical learning
- future of statistics

Dichas combinaciones facilitaron la identificación de literatura científica vinculada a la transformación metodológica de la estadística en el contexto del análisis de datos actual, así como de estudios que exploran la relación entre la estadística, la inteligencia artificial y la ciencia de datos.

Criterios de inclusión

Para la selección de los estudios se establecieron los siguientes criterios de inclusión:

Artículos publicados entre 2020 y 2025, con el propósito de circunscribir el análisis a la producción científica reciente sobre las transformaciones contemporáneas de la estadística.

Publicaciones en revistas científicas, repositorios académicos o documentos de investigación con relevancia conceptual o metodológica para el tema de estudio.

Estudios que aborden alguno de los siguientes campos temáticos: estadística aplicada, inteligencia artificial, ciencia de datos, aprendizaje automático, inferencia estadística en contextos de big data.

Criterios de exclusión

Se excluyeron del análisis aquellos documentos que cumplieran con alguna de las siguientes condiciones:

- publicaciones anteriores al año 2020,
- trabajos cuyo contenido no guardaba relación directa con la estadística, la ciencia de datos o la inteligencia artificial aplicada al análisis de datos,
- documentos que carecían de aportes conceptuales o metodológicos pertinentes para los fines del estudio.

Proceso de selección de estudios

El proceso de selección de la literatura se realizó en varias etapas. En primer lugar, se efectuó una revisión inicial de los títulos y resúmenes de los estudios identificados en las bases de datos consultadas, con el propósito de evaluar su pertinencia respecto del objetivo de estudio. Posteriormente, los artículos potencialmente relevantes fueron sometidos a una lectura completa, lo que permitió verificar su contribución conceptual o metodológica al análisis del papel de la estadística en el contexto de la inteligencia artificial y de la ciencia de datos.

Tras aplicar los criterios de inclusión y exclusión previamente definidos, se seleccionaron 20 artículos científicos, que constituyeron la base del análisis teórico desarrollado en el estudio.

Método de análisis

El análisis de la literatura se realizó mediante un proceso de síntesis temática, orientado a identificar patrones conceptuales recurrentes en los estudios revisados. Este enfoque permitió agrupar los hallazgos en distintas categorías analíticas relacionadas con la evolución de la estadística en la investigación científica contemporánea. Durante el proceso de análisis, se examinaron los artículos seleccionados con el propósito de identificar las principales líneas de discusión presentes en la literatura científica, entre ellas:

- la integración entre estadística y ciencia de datos,
- el desarrollo de modelos interpretables de aprendizaje automático,
- los desafíos asociados con la reproducibilidad científica,
- las implicaciones éticas y metodológicas del análisis de grandes volúmenes de datos,
- y la transformación de la formación estadística en el contexto de la investigación basada en datos.

A partir de esta síntesis conceptual, se elaboró una interpretación integradora que permitió identificar las principales tendencias que configuran el futuro de la estadística en la investigación científica. Debido a la naturaleza narrativa del estudio, el proceso de búsqueda y selección de la literatura no siguió un protocolo sistemático como el de PRISMA; sin embargo, se procuró incluir estudios relevantes y representativos del debate científico contemporáneo sobre la evolución de la estadística en el contexto de la inteligencia artificial y la ciencia de datos.

RESULTADOS

La revisión de la literatura permitió identificar un conjunto de investigaciones recientes que abordan la evolución de la estadística en el contexto de la inteligencia artificial, la ciencia de datos y el análisis computacional. En la Tabla 1 se presenta una síntesis de los principales estudios considerados en el análisis, destacando sus aportes centrales.

Tabla 1. Síntesis de los artículos analizados presentada en la revisión de la transformación metodológica de la estadística en el contexto del análisis actual de datos y de su relación con la estadística, la inteligencia artificial y la ciencia de datos.

AUTOR	ENFOQUE PRINCIPAL	CONTRIBUCIÓN CLAVE
Donoho (2024)	Ciencia de datos	Convergencia entre estadística y data science.
He & Lin (2020)	Estadística y ciencia de datos	Identificación de áreas emergentes.
Wu (2025)	Evolución de la estadística	Adaptación al big data.
Kirch et al. (2025)	Desafíos metodológicos	Rol en datos masivos.
Min et al. (2024)	Estadística e IA	Integración con inteligencia artificial.
Yang et al. (2025)	Evidencia científica	Síntesis entre IA y estadística.
Allen et al. (2024)	Machine learning interpretable	Retos de interpretabilidad.
Rudin et al. (2021)	Interpretabilidad	Principios de modelos explicables.
Linardatos et al. (2020)	AI explicable	Métodos de interpretación.
Tursunalieva et al. (2024)	Interpretación de modelos	Técnicas de deep learning.
Barber et al. (2021)	Inferencia predictiva	Validación de modelos
Gelman et al. (2020)	Reproducibilidad	Enfoque bayesiano reproducible
Hernández & Colom (2025)	Ciencia abierta	Políticas de reproducibilidad
Horton & Hardin (2021)	Educación estadística	Integración de computación
Torres-Gordillo (2025)	Educación estadística	Impacto de la IA
Felipe et al. (2025)	Aprendizaje estadístico	Mitigación de sesgos y mejora algorítmica
Montero (2022)	Uso de estadística	Problemas en investigación

Benjamini et al. (2021)	Significancia estadística	Declaración sobre el uso adecuado del valor p y su impacto en la replicabilidad
Van Doorn et al. (2020)	Inferencia bayesiana	Introducción a inferencia bayesiana para la prueba de significancia de hipótesis nula
Quatto et al. (2022)	Evaluación del valor p	Revisión crítica de nuevas propuestas bayesianas para evaluar el valor p

En síntesis, el análisis conjunto de la literatura seleccionada revela que la estadística atraviesa un periodo de profunda adaptación y sinergia ante los avances tecnológicos recientes. Los estudios evidencian una convergencia progresiva entre la estadística, la inteligencia artificial y la ciencia de datos, lo que da lugar a un nuevo paradigma analítico de carácter híbrido (2,3,10). En este ecosistema interdisciplinario, las herramientas computacionales avanzadas amplían la capacidad predictiva y de gestión de datos a gran escala, mientras que la inferencia estadística clásica mantiene su rol indispensable como garante del rigor metodológico (7,8). Para sostener esta evolución, la comunidad científica subraya la urgencia de desarrollar modelos interpretables (11–14) y adoptar prácticas de ciencia abierta que mitiguen la crisis de reproducibilidad (6,15). En el marco de esta búsqueda de mayor solidez metodológica y reproducibilidad, surge igualmente la necesidad de cuestionar la dependencia exclusiva de la significancia estadística tradicional y de adoptar y promover enfoques alternativos más sólidos — como los métodos bayesianos— que permitan trascender las restricciones inherentes al valor p (18–20). Simultáneamente a estos cambios, resulta fundamental renovar la formación académica mediante la integración de competencias en programación. Así, lejos de quedar relegada, la estadística se afirma como el fundamento conceptual esencial para convertir el análisis de grandes volúmenes de datos en conocimiento científico sólido y comprobable.

DISCUSIÓN

Los hallazgos de la revisión de la literatura evidencian que la estadística se encuentra en un proceso de profunda transformación impulsado por el crecimiento del big data, el desarrollo de la inteligencia artificial y la expansión de la ciencia de datos. Lejos de sustituir los enfoques tradicionales, estas transformaciones reflejan una evolución metodológica caracterizada por la integración de modelos estadísticos clásicos con algoritmos de aprendizaje automático (1,2).

Diversos estudios coinciden en que la estadística continúa desempeñando un papel central en la investigación científica contemporánea, especialmente en procesos de inferencia, cuantificación de la incertidumbre y validación de modelos analíticos (7,10). En este sentido, la evidencia revisada sugiere que el futuro de la disciplina se orienta hacia una convergencia interdisciplinaria en la que la estadística aporta el marco teórico necesario para garantizar la robustez metodológica. Esta interacción ha favorecido el desarrollo de enfoques híbridos que combinan la inferencia causal con capacidades predictivas. Según Yang et al. (4) y Min et al. (8), esta integración permite abordar problemas analíticos de mayor complejidad, especialmente en contextos de alta dimensionalidad.

No obstante, esta evolución conlleva retos considerables. La incorporación irreflexiva de modelos algorítmicos de “caja negra” conlleva el riesgo de privilegiar la precisión predictiva por encima de la comprensión del fenómeno (11). Ante este escenario, el desarrollo de la inteligencia artificial explicable (XAI, Explainable Artificial Intelligence) emerge como una de las áreas más urgentes de la estadística aplicada (11,13) para restaurar la transparencia en los procesos de inferencia. Investigaciones recientes destacan que comprender la lógica interna de las predicciones es vital para garantizar la fiabilidad de las decisiones (14). Como señalan Rudin et al. (12), la interpretabilidad no solo asegura la transparencia de los modelos, sino que también permite auditarlos empíricamente y convertirlos en verdaderos motores de descubrimiento científico.

Otro hallazgo relevante se relaciona con el debate sobre la reproducibilidad. La denominada crisis de replicabilidad ha evidenciado limitaciones en el uso de métodos estadísticos, como la dependencia excesiva del valor p y la interpretación inferencial errónea (5). En respuesta a ello, investigadores como Gelman et al. (14) cuestionan el entorno académico que privilegia la originalidad en detrimento del rigor metodológico y proponen procesos de trabajo más transparentes, entre ellos los de orientación bayesiana. Del mismo modo, la

literatura especializada subraya la necesidad urgente de implementar políticas de ciencia abierta, códigos verificables y documentación detallada que salvaguarden la validez de los hallazgos (6).

Una de las discusiones metodológicas más relevantes en la estadística contemporánea gira en torno a las limitaciones del enfoque clásico de las pruebas de hipótesis basadas exclusivamente en el valor p . Aunque este indicador continúa siendo ampliamente utilizado en la investigación científica, múltiples estudios han cuestionado su interpretación dicotómica basada en umbrales arbitrarios de significancia estadística, particularmente en contextos asociados a la crisis de reproducibilidad científica (14,18). El grupo de trabajo de la American Statistical Association (ASA) (18) ha señalado que el valor p no debe interpretarse como una medida absoluta de evidencia científica ni utilizarse de forma aislada para la toma de decisiones inferenciales.

En este contexto, la inferencia bayesiana ha resurgido como una alternativa metodológica de creciente relevancia en la ciencia de datos y la inteligencia artificial. A diferencia del paradigma frecuentista clásico, los métodos bayesianos incorporan conocimiento previo, actualizan las probabilidades a partir de nueva evidencia y expresan explícitamente la incertidumbre inherente a los modelos analíticos. Estas ventajas han impulsado su adopción en dominios de alta complejidad, como el aprendizaje automático, la modelización jerárquica y la síntesis de la evidencia científica (19). Asimismo, diversos autores sostienen que la estadística bayesiana ofrece ventajas importantes para fortalecer la reproducibilidad y la transparencia metodológica, especialmente cuando se complementa con prácticas de ciencia abierta y flujos de trabajo reproducibles (20). No obstante, su adopción aún enfrenta desafíos relacionados con la especificación de las distribuciones previas, la complejidad computacional y la escasa formación especializada en métodos probabilísticos avanzados (19). A pesar de ello, la convergencia entre la inferencia bayesiana, el aprendizaje automático y la computación científica constituye una de las tendencias más prometedoras para el futuro de la estadística aplicada.

En esta misma línea, el análisis destaca la importancia crítica de la calidad de los datos y el control de sesgos empíricos. En entornos de big data, un volumen masivo de información puede, paradójicamente, amplificar los sesgos si no existe una validación estadística rigurosa. En este contexto, enfoques como el propuesto por Felice et al. (17), basado en el aprendizaje estadísticamente mejorado (statistically enhanced learning), ofrecen una alternativa metodológica

al integrar principios estadísticos formales en algoritmos de aprendizaje automático, fortaleciendo su capacidad de generalización y reduciendo el riesgo de inferencias espurias.

De este modo, se refuerza la necesidad de aplicar marcos analíticos robustos que combinen ingeniería de variables, validación estadística y aprendizaje automático (17), reafirmando que el verdadero desafío de la época actual no es la potencia computacional, sino el control de la incertidumbre (7). Horton (15) aborda la necesidad de cambios significativos en la educación en estadística para incorporar mejores habilidades de programación y pensamiento computacional, destacando que estas habilidades son fundamentales para la formación moderna en ciencia de datos y estadística. Torres-Gordillo (9) señala que esta transformación requiere integrar la tecnología mediante metodologías activas, preparando a los investigadores para trascender los paradigmas impulsados por los datos (data-driven).

En conjunto, se identifica una tendencia clara hacia la consolidación de un paradigma analítico híbrido. Esta convergencia redefine las herramientas de investigación e impone la necesidad de fortalecer la formación metodológica y las prácticas de ciencia abierta. Por último, conviene reconocer que, al tratarse de una revisión narrativa, el presente estudio presenta ciertas limitaciones, ya que la selección bibliográfica no sigue un protocolo riguroso y sistemático. No obstante, los 20 trabajos examinados constituyen aportaciones fundamentales del debate académico actual en torno al porvenir de la disciplina.

CONCLUSIONES

El presente estudio analizó las tendencias emergentes que redefinen el papel de la estadística en la investigación científica contemporánea. A partir de la revisión de la literatura reciente, se concluye que la estadística no está siendo desplazada por la inteligencia artificial y la ciencia de datos; por el contrario, se encuentra en una etapa de consolidación como pilar fundamental que garantiza el rigor, la interpretabilidad y el manejo de la incertidumbre en ecosistemas analíticos cada vez más complejos.

Las transformaciones identificadas indican un claro movimiento hacia marcos metodológicos híbridos, donde la inferencia clásica converge con el potencial computacional del aprendizaje automático. Para que esta transición sea exitosa en la investigación científica, es indispensable superar retos críticos como la mejora de las prácticas de ciencia

abierta y reproducibilidad, así como la reestructuración de la formación académica para incluir competencias en programación e ingeniería de datos. En definitiva, el futuro de la estadística radica en su capacidad para evolucionar interdisciplinariamente, asegurando que el avance tecnológico se traduzca en conocimiento científico sólido, ético y fiable.

CONSIDERACIONES ÉTICAS

Este estudio corresponde a una revisión de literatura basada en fuentes secundarias de información científica previamente publicadas. Por lo tanto, no requirió aprobación de un comité de ética ni consentimiento informado.

CONFLICTO DE INTERÉS

El autor declara que no existen conflictos de intereses relacionados con la realización de este estudio ni con la publicación de sus resultados.

REFERENCIAS BIBLIOGRÁFICAS

1. Donoho D. Data science at the singularity. *Harv Data Sci Rev.* 2024;6(1). doi: <https://doi.org/10.1162/99608f92.b91339ef>
2. Wu H. Statistics evolution and revolution to meet data science challenges. *Stat Biosci.* 2025;17(3):813-831. doi: <https://doi.org/10.1007/s12561-024-09454-5>
3. He X, Lin X. Challenges and opportunities in statistics and data science: ten research areas. *Harv Data Sci Rev.* 2020;2(3). doi: <https://doi.org/10.1162/99608f92.95388fcb>
4. Yang S, Gamalo M, Fu H. Integrating RCTs, RWD, AI/ML and statistics: next-generation evidence synthesis [Internet]. 2025 [cited 2026 Mar 3]. Available from: <http://arxiv.org/abs/2511.19735> doi: <https://doi.org/10.48550/arXiv.2511.19735>
5. Montero LML. Uso de la estadística en trabajos de investigación en la Universidad San Pedro [Internet]. 2022. Disponible en: <https://dspace.ups.edu.ec/handle/123456789/11202>
6. Hernandez JA, Colom M. Reproducible research policies and software/data management in scientific computing journals: a survey, discussion, and perspectives. *Front Comput Sci.* 2025;6:1491823. doi: <https://doi.org/10.3389/fcomp.2024.1491823>
7. Kirch, C., Lahiri, S., Binder, H., Brannath, W., Cribben, I., Dette, H., Doebler, P., Feng, O., Gandy, A., Greven, S., Hammer, B., Harmeling, S., Hotz, T., Kauer- mann, G., Krause, J., Krempel, G., Nieto-Reyes, A., Okhrin, O., Ombao, H., ... Lederer, J. (2025). Challenges and Opportunities for Statistics in the Era of Data Science. *Harvard Data Science Review*, 7(2). <https://doi.org/10.1162/99608f92.abf14c9d>
8. Min J, Song X, Zheng S, King CB, Deng X, Hong Y. Applied statistics in the era of artificial intelligence: a review and vision [Internet]. 2024 [cited 2026 Apr 10]. Available from: <https://arxiv.org/abs/2412.10331> doi: <https://doi.org/10.48550/arXiv.2412.10331>
9. Torres-Gordillo VA. Inteligencia artificial y aprendizaje automático en la educación estadística. *Rev. Educ Estad.* 2025;4(1):1-21. doi: <https://doi.org/10.29035/redes.4.1.7>
10. Allen GI, Gan L, Zheng L. Interpretable machine learning for discovery: statistical challenges and opportunities. *Annu Rev Stat Its Appl.* 2024;11(1):97-121. doi: <https://doi.org/10.1146/annurev-statistics-040120-030919>
11. Rudin C, Chen C, Chen Z, Huang H, Semenova L, Zhong C. Interpretable machine learning: fundamental principles and 10 grand challenges [Internet]. 2021 [cited 2026 Mar 6]. Available from: <http://arxiv.org/abs/2103.11251> doi: <https://doi.org/10.48550/arXiv.2103.11251>

12. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. *Entropy*. 2020;23(1), 18. doi: <https://doi.org/10.3390/e23010018>
13. Tursunaliyeva A, Alexander DLJ, Dunne R, Li J, Riera L, Zhao Y. Making sense of machine learning: a review of interpretation techniques and their applications. *Appl Sci*. 2024;14(2), 496. doi: <https://doi.org/10.3390/app14020496>
14. Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, et al. Bayesian workflow [Internet]. 2020 [cited 2026 Mar 5]. Available from: <http://arxiv.org/abs/2011.01808> doi: <https://doi.org/10.48550/arXiv.2011.01808>
15. Horton NJ, Hardin JS. Integrating computing in the statistics and data science curriculum: creative structures, novel skills and habits, and ways to teach computational thinking. *J Stat Data Sci Educ*. 2021;29(sup1):S1-S3. doi: <https://doi.org/10.1080/10691898.2020.1870416>
16. Barber RF, Candès EJ, Ramdas A, Tibshirani RJ. Predictive inference with the jackknife+. *Ann Stat*. 2021;49(1): 486-507. doi: <https://doi.org/10.1214/20-AOS1965>
17. Felice F, Ley C, Bordas SPA, Groll A. Boosting any learning algorithm with statistically enhanced learning. *Sci Rep*. 2025;15(1):1605. doi: <https://doi.org/10.1038/s41598-024-84702-8>
18. Benjamini Y, De Veaux R, Efron B, Evans S, Glickman M, Graubard B, et al. ASA President's Task Force statement on statistical significance and replicability. *Harv Data Sci Rev*. 2021;3(3). doi: [DOI:10.1162/99608f92.f0ad0287](https://doi.org/10.1162/99608f92.f0ad0287)
19. Van Doorn J, van den Bergh D, Böhm U, et al. Bayesian alternatives to null hypothesis significance testing in biomedical research: a non-technical introduction to Bayesian inference with JASP. *BMC Med Res Methodol*. 2020; 20:142. doi: <https://doi.org/10.1156/s12874-020-00980-6>
20. Quatto P, Ripamonti E, Marasini D. Beyond $p < .05$: a critical review of new Bayesian proposals for assessing the p-value. *J Biopharm Stat*. 2022;32(2):308-329. doi: <https://doi.org/10.1080/10543406.2021.2009497>