



Foto: Alfredo Pérez Bautista. Hospital Mixto Jesús María Nayarit

PERSPECTIVAS Y METODOLOGÍAS

De la normalidad a la normalización: aportes matemáticos y estadísticos en la investigación médica con pertinencia territorial

Méndez-Domínguez, N.I.^a; Azcorra-Pérez, H.^{b*}

RESUMEN

Objetivo: Este trabajo analiza la evolución de los principales métodos cuantitativos aplicados a la medicina y reflexiona sobre su utilidad para fortalecer la investigación y los sistemas de salud contemporáneos. **Metodología:** Se realizó un análisis crítico mediante una revisión narrativa con enfoque histórico-epistemológico. Se integraron textos clásicos de probabilidad y estadística, desarrollos metodológicos en epidemiología, literatura sobre complejidad e inteligencia artificial, así como contribuciones latinoamericanas.

^a Hospital Regional de Alta Especialidad de la Península de Yucatán (HRAEPY). Servicios de Salud del Instituto Mexicano del Seguro Social para el Bienestar (IMSS-BIENESTAR), Yucatán, México.

^b Centro de Investigaciones Silvio Zavala, Universidad Modelo, Mérida, Yucatán.

* nmendez.hraepy@imssbienestar.gob.mx

Resultados: El análisis se estructuró en torno a la evolución de los métodos estadísticos, las métricas e indicadores, así como los marcos conceptuales para su aplicación contextual en atención a la salud. **Conclusiones:** La evolución de los métodos cuantitativos demuestra que la interpretación de los fenómenos sanitarios requiere trascender de la aplicación acrítica de modelos universales a la incorporación de enfoques contextualizados considerando la complejidad humana, social y territorial. En sistemas públicos como IMSS Bienestar, caracterizados por una alta diversidad poblacional atendida, el uso contextualizado de modelos estadísticos, indicadores epidemiológicos mixtos y enfoques de complejidad puede mejorar la evaluación de intervenciones, la implementación y la toma de decisiones.

Palabras clave: Estadística; Epidemiología; Medicina basada en la evidencia; Inteligencia Artificial; Toma de decisiones.

INTRODUCCIÓN

La medicina contemporánea se sostiene en una estructura matemática profunda, aunque frecuentemente invisible. Desde la modelación de procesos fisiológicos hasta la estimación de riesgos mediante la bioestadística, la matemática constituye un lenguaje fundamental en la práctica clínica y epidemiológica, cuyo alcance no siempre se dimensiona durante la formación médica.

La normalización se ha entendido, por un lado, como la adopción de la distribución normal y referente universal, por otro, como la incorporación acrítica de modelos desarrollados en contextos euro-estadounidenses. Ambas nociones configuran la manera en que se interpretan los fenómenos relacionados con la salud. Este doble proceso, tanto matemático como epistemológico, invisibiliza otras formas de comprender y modelar la diversidad tanto biológica como social.

Este artículo presenta una revisión narrativa mediante un enfoque histórico-epistemológico, cuyo objetivo es analizar las contribuciones de la matemática y la estadística en la medicina, diferenciando entre métodos analíticos, métricas e indicadores, así como discutir su impacto en la formación médica. Asimismo, se integran desarrollos provenientes de modelos de complejidad y aportes situados desde América Latina.

Se parte del papel del Teorema del Límite Central (TLC) como articulador conceptual de la estadística moderna; a la emergencia de los modelos lineales, logísticos y de Poisson; al nacimiento del análisis de supervivencia; y al desarrollo de métodos situados, en particular los generados desde México y América Latina, donde la lectura contextualizada de fenómenos como la estacionalidad ha permitido desafiar paradigmas normalizados.

METODOLOGÍA

Se presenta una revisión narrativa de la literatura con enfoque histórico y epistemológico. La selección de fuentes incluyó textos clásicos de estadística, artículos fundacionales en epidemiología, revisiones metodológicas contemporáneas y literatura sobre epistemología de la ciencia.

Se incluyeron: a) obras fundamentales de probabilidad y estadística aplicadas a la salud; b) desarrollos metodológicos relevantes en epidemiología; c) contribuciones estadísticas generadas fuera del ámbito euro-estadounidense; y d) literatura sobre complejidad y decolonización del conocimiento.

El análisis se organizó en tres niveles incluyendo la descripción de los métodos estadísticos; las métricas e indicadores con operacionalización de fenómenos. Por último, los marcos conceptuales y su interpretación en salud.

RESULTADOS

Desarrollo histórico: de la variabilidad a la medición científica

El componente matemático de la medicina comienza formalmente en el siglo XIX, cuando Carl Friedrich Gauss describió la distribución normal como un modelo de errores y fenómenos naturales (1). Esta curva, originalmente concebida para resolver problemas astronómicos, se volvió rápidamente una metáfora potente para describir la variabilidad fisiológica humana: aquello que se concentra alrededor de la media aritmética se considera normal y lo que se aleja en exceso se interpreta como anormal. Este traslado conceptual tuvo efectos profundos pues convirtió la matemática en un instrumento para categorizar la salud o la ausencia de ésta.

En 1812, Pierre-Simon Laplace elevó este razonamiento a un nivel más general mediante la formulación del TLC (2), que establece que la suma de variables aleatorias independientes tiende a distribuirse de forma normal. El TLC permitió explicar por qué la normalidad emergía en diversos contextos biológicos aun cuando las distribuciones individuales no fueran normales, y con ello sentó las bases para

justificar el uso de pruebas paramétricas, intervalos de confianza y estimaciones basadas en promedios. Así, la matemática no sólo describía fenómenos, sino que los “normalizaba”.

En el siglo XX, William Feller refinó los fundamentos probabilísticos modernos (3), mientras Francis Galton introdujo la correlación (4) y Karl Pearson la regresión y la prueba χ^2 (5,6). Estos avances permitieron describir la dependencia cuantificable y comparar las proporciones esperadas con las observadas, un elemento esencial para la epidemiología y la salud pública.

La prueba *t de Student* de William Gosset en 1908 (7) posibilitó analizar en términos comparativos muestras pequeñas, situación habitual en estudios clínicos. Por su parte, Ronald A. Fisher revolucionó la inferencia estadística con la máxima verosimilitud, el análisis de varianza (ANOVA) y el diseño experimental (8). Su visión estadística estructuró gran parte de los ensayos clínicos contemporáneos, aunque también contribuyó a normalizar un enfoque frecuentista rígido, el cual dominó durante décadas la práctica biomédica.

La estructura matemática de la inferencia: del Teorema del Límite Central a la normalización del conocimiento

El TLC no sólo permitió la adopción masiva de los métodos paramétricos, también moldeó la manera en que la medicina entiende la incertidumbre. Hoy cuando un clínico interpreta una media de glucosa, un puntaje *z* en un estudio antropométrico o la variación de la presión arterial respecto a rangos de referencia, está aplicando el razonamiento que Laplace anticipó hace más de dos siglos.

Sin embargo, la normalización estadística puede ocultar fenómenos importantes en el contexto médico. No todas las distribuciones clínicas son normales, no todas las poblaciones presentan variabilidad homogénea y no todos los entornos sanitarios son comparables. En ocasiones, la confianza excesiva en la normalidad ha llevado a interpretar como un error a las variaciones que tienen significado clínico o epidemiológico.

Modelos de regresión: un lenguaje para las relaciones clínicas

En la medicina moderna, los estudios observacionales analíticos se apoyan ampliamente en tres técnicas comunes: la regresión lineal que puede asociar desenlaces descritos como variables numéricas, la regresión logística empleada para eventos binarios y la regresión de Poisson que explica los factores asociados a tasas y conteos. Estos modelos suelen presentarse después de la estadística descriptiva en los reportes de investigación.

Aunque diferentes en forma, los tres modelos provienen del mismo marco conceptual, el de los Modelos Lineales Generalizados, cuya base matemática está íntimamente conectada con el TLC y con funcio-

nes de enlace específicas. Los modelos de regresión permiten ir más allá de asociaciones simples al incorporar covariables y delimitar efectos de confusión, pues cuantifican la magnitud del efecto y su dirección, estiman predicciones y permiten inferencias causales cuando el diseño del estudio es adecuado. Además, permiten explorar la interacción de variables predictoras y sus posibles contribuciones a la explicación de fenómenos complejos, lo cual ha sido útil, pero también exige cautela: *un modelo no es una descripción de la realidad, sino una aproximación estructurada*, en la **Figura 1** se resumen las diferencias esenciales entre estos tres métodos, su tipo de desenlace y su interpretación clínica.

Figura 1. Comparación entre regresión lineal, logística y de Poisson. Tres modelos diferentes para diferentes tipos de desenlaces en investigación médica

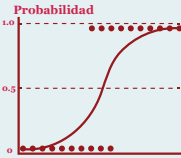
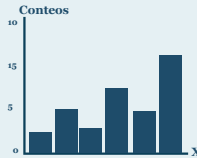
	1. REGRESIÓN LINEAL Desenlace continuo	2. REGRESIÓN LOGÍSTICA Desenlace binario (sí/no)	3. REGRESIÓN DE POISSON Desenlace: conteos o tasas
			
	Ejemplo clínico: Nivel de glucosa (mg/dL) en función de la edad.	Ejemplo clínico: Presencia de enfermedad (sí/no) según presión arterial.	Ejemplo clínico: Número de infecciones respiratorias por paciente-año según edad.
 Distribución asumida	Normal	Bernoulli	Poisson
 Función de enlace	Identidad ($g(\mu) = \mu$)	Logit ($g(p) = \log\left(\frac{p}{1-p}\right)$)	Log ($g(\mu) = \log(\mu)$)
 Parámetro clave	Cambio promedio (β) Unidades del desenlace por unidad del predictor	Odds ratio (OR) Cambio en las probabilidades del evento	Razón de tasas (IRR) Cambio en la tasa de eventos por unidad del predictor
 Interpretación clínica	Incremento promedio esperado del desenlace continuo	Cambio multiplicativo en las odds (probabilidades relativas) del evento	Cambio multiplicativo en la frecuencia del evento (tasas o conteos)
 Supuestos críticos	- Linealidad - Homocedasticidad - Normalidad de los errores - Independencia	- Independencia de observaciones - Especificación correcta del modelo - Ausencia de multicolinealidad severa	- Independencia de observaciones - Igualdad entre media y varianza (puede ajustarse por sobredispersión)
 Ventajas principales	Intuitiva, fácil de interpretar; ampliamente utilizada	Ideal para desenlaces binarios; resultados interpretables (OR)	Adecuada para eventos raros o frecuentes; modela tasas e incidencias
 Limitaciones	No apropiada para desenlaces no continuos; sensible a outliers	OR puede malinterpretarse como riesgo relativo	Sensible a sobredispersión; asume tiempos-persona adecuados
 Usos clínicos típicos	Biomarcadores, antropometría, presión arterial, niveles bioquímicos	Mortalidad (sí/no), respuesta terapéutica, presencia de enfermedad, adherencia	Infecciones, recaídas, hospitalizaciones, eventos adversos, consultas



Foto: Karla Bautista Celestino. Hospital General Cd. Madero

Análisis del tiempo: supervivencia, riesgo y tratamiento

El análisis de supervivencia emergió como un campo propio a mediados del siglo XX. *Kaplan y Meier* introdujeron en 1958 un método no paramétrico para estimar supervivencia con censura (9), mientras Mantel desarrolló la prueba de rango logarítmico (log-rank) para comparar curvas entre grupos (10). El modelo de Cox, presentado en 1972, permitió estimar riesgos instantáneos ajustados sin necesidad de especificar la forma de la distribución basal (11).

El método de *Kaplan–Meier*, hoy indispensable en la investigación clínica, no surgió sólo en el vacío metodológico, sino fue moldeado por un contexto histórico marcado por transformaciones sociales profundas. Edward L. Kaplan desarrolló sus primeras ideas en la posguerra estadounidense, en un periodo donde la medicina enfrentaba el reto de analizar cohortes afectadas por mortalidad temprana, abandono y un seguimiento de observación incompleto. La censura, elemento central del método, puede interpretarse como propia de su época: movilidad poblacional intensa, pérdida de pacientes por desigualdades estructurales y prácticas clínicas aún no estandariza-

das. Así, el estimador *Kaplan–Meier* es producto de un momento histórico particular en el que urgía capturar la incertidumbre real del tiempo de supervivencia sin imponer supuestos rígidos. Lo anterior es un ejemplo claro de cómo la matemática se moldea por las necesidades, tensiones-limitaciones del tiempo y el lugar en que es concebida (12).

Estos métodos introdujeron una ruptura epistemológica en medicina: pasaron de la comparación de proporciones a la dinámica del tiempo. El riesgo dejó de ser estático y se convirtió en un proceso continuo. No obstante, su adopción también generó una nueva normalización, se asumió que la proporcionalidad de riesgos era universal cuando en muchas realidades (como las poblaciones con alta movilidad estacional o enfermedades con periodos de transición marcados) esta suposición puede no cumplirse. La revisión metodológica contemporánea de Abd ElHafeez y colaboradores (13) muestra la necesidad de actualizar y adaptar estos métodos a contextos emergentes.

Más allá del canon: aportes latinoamericanos y mexicanos, la estadística situada y ruptura de la normalización

Si bien la narrativa tradicional enfatiza los avances europeos y estadounidenses, múltiples regiones han contribuido a la matemática aplicada a la salud. La poca difusión de estas contribuciones es también un proceso de normalización, pues la historia de la ciencia se ha contado de forma lineal y geográficamente sesgada (14-18), excluyendo por origen y género. Reconocer estas aportaciones permite ampliar la comprensión del proceso científico como un fenómeno verdaderamente global.

En América Latina, el análisis epidemiológico ha generado innovaciones metodológicas de enorme valor. Aun cuando las contribuciones latinoamericanas en estadística aplicada a la salud han demostrado solidez conceptual y capacidad explicativa en múltiples especialidades, su adopción internacional ha sido históricamente menor que la de los métodos euro-estadounidenses. Esta asimetría no deriva de limitaciones metodológicas, sino de factores estructurales como la hegemonía editorial anglosajona (15,16), la centralización del financiamiento científico internacional (17), las barreras idiomáticas (18) y la temprana incorporación curricular de los métodos europeos y estadounidenses como estándares universales.

No obstante, cuando los desarrollos latinoamericanos han respondido a problemas globales, han demostrado una alta capacidad de adopción e impacto. Entre los aportes latinoamericanos más influyentes destaca la clasificación *NOVA*, desarrollada por Carlos Monteiro en Brasil, que reorganiza los alimentos según su grado de procesamiento en lugar de su composición nutricional (18). Otro ejemplo en el mismo país son las Cohortes de Pelotas de Victora, Barros y Horta quienes desarrollaron indicadores de lactancia, métricas de desigualdad como el *Slope Index* y el *Concentration Index* aplicados a salud infantil (19). En Chile, el índice multidimensional de pobreza en salud es un modelo que integra privaciones sanitarias, educativas y ambientales, útil en epidemiología, economía de la salud y análisis de desigualdades. En Colombia, el trabajo de Castillo-Salgado estandarizó indicadores de inequidad en salud (20). En Cuba, los métodos de conjugación glicoproteica desarrollados por Vérez y Campa permitieron la creación de vacunas conjugadas (21).

México ocupa un lugar central en esta cartografía metodológica y clinimétrica. Entre sus contribuciones más relevantes destacan el índice de margi-



Foto: Karla Bautista Celestino. Hospital General Cd. Madero

nación, desarrollado por Gómez de León, Tuirán y Welti (22), éste constituye uno de los indicadores estructurales más utilizados en América Latina para caracterizar desigualdades sociales y territoriales. Asimismo, las propuestas de evaluación del desempeño de los sistemas de salud (23) contribuyeron a la consolidación de marcos analíticos adoptados internacionalmente para valorar capacidades institucio-

nales y funciones sanitarias. En el ámbito de la nutrición, el Índice Mexicano de Calidad de la Dieta (24) representa una herramienta ampliamente utilizada para evaluar patrones alimentarios y monitorear la calidad de la alimentación en poblaciones mexicanas y latinoamericanas. La **Figura 2** ilustra los aportes latinoamericanos a las métricas, indicadores y marcos analíticos aplicados a la salud antes expuestos.

Figura 2. Aportes latinoamericanos a las métricas, indicadores y marcos analíticos aplicados a la salud



Los aportes latinoamericanos demuestran que la producción de indicadores, métricas y marcos analíticos para la salud no constituye una adaptación pasiva de modelos externos, sino una contribución original que ha enriquecido la comprensión de los fenómenos sanitarios desde contextos sociales, culturales y epidemiológicos diversos.

La metodología situada que se presenta con congruencia epidemiológica y social desde América Latina demuestra que desde nuestros contextos se desarrollan marcos analíticos propios que enriquecen la comprensión de los fenómenos en salud y, a la vez,

cuestionan la universalidad de los modelos tradicionales, pues desafían las normalizaciones impuestas por modelos globales al demostrar que la matemática puede y debe adaptarse al territorio (25,26).

Metodologías en perspectiva

Más allá de los modelos clásicos, la epidemiología contemporánea ha incorporado enfoques provenientes de la teoría de sistemas complejos, donde los fenómenos de salud se entienden como resultado de interacciones no lineales, adaptativas y dependientes del contexto.

Los modelos de complejidad que abarcan los sistemas dinámicos, las redes y los procesos estocásticos permiten analizar desde las transiciones epidemiológicas, considerando su trayecto no necesariamente lineal o la propagación de enfermedades ajustadas por diversos aspectos sociales y biológicos, hasta la continua interacción entre los determinantes sociales y biológicos de la salud en constante cambio. Cuando estos enfoques se analizan considerando el contexto temporal y la incertidumbre inherente, posibilitan delimitar y aplicar el conocimiento con mayor pertinencia. Los procesos estocásticos han ampliado la capacidad de modelar sistemas biológicos sujetos a incertidumbre y constituyen uno de los fundamentos matemáticos de los enfoques contemporáneos de complejidad (27) que aún se encuentran en constante evolución.

A manera de corolario, la inteligencia artificial (IA) representa una extensión, no una ruptura, con la tradición matemática en la medicina. La irrupción de la IA generativa ha transformado radicalmente la accesibilidad de los métodos estadísticos. Tanto estudiantes y profesionales de la salud pueden ejecutar modelos complejos que van desde las regresiones hasta análisis de supervivencia y métodos bayesianos, sin requerir infraestructura especializada ni software propio. Sin embargo, a la vez la generación de datos artificiales en sustitución de los registros capturados y la saturación ante grandes conjuntos de datos constituyen aún un problema que sólo los investigadores más minuciosos podrán haber corroborado. A pesar de los avances actuales, en ausencia de verificación detallada, el procesamiento estadístico de datos por IA para fines de investigación médica podría conducir a explicaciones imprecisas.

CONCLUSIONES

La historia de la medicina moderna no puede comprenderse sin reconocer el papel central de la matemática y la estadística en la construcción del conocimiento científico. Desde la distribución normal de Gauss y el Teorema del Límite Central de Laplace hasta los modelos de regresión; el análisis de supervivencia y las herramientas contemporáneas de IA, demuestran que los métodos cuantitativos han permitido describir, explicar y predecir fenómenos complejos relacionados con la salud y la enfermedad.

Esta evolución también ha estado acompañada por procesos de normalización metodológica y epistemológica. En este sentido, se reconoce que los métodos estadísticos son construcciones históricas, desarrolladas en contextos específicos y sujetas a supuestos particulares que permiten una aplicación crítica en la práctica clínica y epidemiológica.

Para los sistemas de salud contemporáneos, particularmente en contextos de transformación como IMSS Bienestar, comprender los fundamentos matemáticos y epistemológicos de la bioestadística resulta esencial para fortalecer la toma de decisiones basada en evidencia. Una formación que integre pensamiento crítico, análisis contextual, competencias cuantitativas y cualitativas permitirá interpretar con mayor pertinencia social los problemas de salud de poblaciones diversas y contribuir al diseño de intervenciones más equitativas, eficientes y adaptadas a las necesidades reales de la población.

REFERENCIAS

1. Gauss CF. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Hamburg: Perthes and Besser; 1809.
2. Laplace PS. *Théorie analytique des probabilités*. Paris: Courcier; 1812.
3. Feller W. *An Introduction to Probability Theory and Its Applications*. 2nd ed. New York: Wiley; 1957.
4. Galton F. Regression Towards Mediocrity in Hereditary Stature. *J Anthropol Inst*. 1886;15:246–63.
5. Pearson K. Contributions to the mathematical theory of evolution. *Proc R Soc Lond*. 1896;58:240–42.
6. Pearson K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. 1900 Jul 1;50(302):157–75.
7. Student (Gosset WS). The probable error of a mean. *Biometrika*. 1908;6(1):1–25.

8. Fisher, R.A. 1990. Statistical methods, Experimental design and scientific inference (ed. J.H. Bennett), Oxford: Oxford University Press
9. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:457–81.
10. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep.* 1966 Mar 1;50(3):163–70.
11. Cox DR. Regression models and life tables. *J R Stat Soc Series B.* 1972;34:187–220.
12. Stalpers LJ, Kaplan EL. Edward L. Kaplan and the Kaplan-Meier survival curve. *BSHM Bulletin: Journal of the British Society for the History of Mathematics.* 2018 May 4;33(2):109–35.
13. Abd ElHafeez S, et al. Methods to Analyze Time-to-Event Data. *Oxid Med Cell Longev.* 2021;2021:1–15.
14. Blyden J, Harding SG. *The science question in feminism revisited: methodological dominance and epistemic exclusion.* Cambridge: Polity Press; 2020.
15. Gaillard J, Arvanitis R. Research capacity in the Global South: inequality, knowledge flows, and scientific dependency. *Sci Public Policy.* 2019;46(4):645–656.
16. Montgomery SL. *Does science need a global language? English and the future of research.* Chicago: University of Chicago Press; 2013.
17. Porter TM. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton: Princeton University Press; 1995.
18. Monteiro CA, Cannon G, Lawrence M, Costa Louzada M, Machado P. *Ultra-processed foods, diet quality, and human health.* Geneva: WHO; 2019.
19. Victora CG, Barros FC, Horta BL, et al. The Pelotas birth cohort studies: 1982–2015. *Int J Epidemiol.* 2016;45(3):632–642.
20. Castillo-Salgado C, Schneider C, Loyola E, Mujica O, Roca A, Yerg T. Measuring health inequalities: Gini coefficient and concentration index. *Epidemiol Bull.* 2001 Mar;22(1):3–4.
21. Vérez-Bencomo V, Fernández-Sánchez T, Hardy E, et al. A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b. *Science.* 2004;305(5683):522–525
22. Gómez de León J, Tuirán R, Welte C. *Índice de Marginación 2015: Documento Metodológico.* Ciudad de México: Consejo Nacional de Población (CONAPO); 2016. Disponible en: <https://www.gob.mx/conapo>
23. Frenk J, González-Pier E, Gómez-Dantés O, Lezana MA, Knaut FM. Health system reform in Mexico. *Lancet.* 2006;368(9546):152–163. doi:10.1016/S0140-6736(06)69005-9.
24. Rodríguez-Ramírez S, Rivera JA, Barquera S, Campos-Nonato I, Ariza AC, et al. Dietary patterns in Mexican children and adolescents: Results from the Mexican National Health and Nutrition Survey 2012. *J Nutr.* 2016;146(9):1934S–1943S.
25. García-Olp M, Nelson C, Saiz L. Decolonizing mathematics curriculum and pedagogy: Indigenous knowledge has always been mathematics education. *Educational Studies.* 2022 Jan 2;58(1):1–6.
26. Kulago HA, Wapeemukwa W, Guernsey PJ, Black M. *Land, water, mathematics, and relationships: What does creating decolonizing and Indigenous curricula ask of us?* *Educational Studies.* 2021 May 4;57(3):345–63
27. Roberts MG, Hickson RI, McCaw JM. *How immune dynamics shape multi-season epidemics: a continuous-discrete model in one dimensional antigenic space.* *J Math Biol.* 2024 Mar 27;88(4):48. doi: 10.1007/s00285-024-02076-x.