

## REVISTA SIGMA

Departamento de Matemáticas y Estadística

*Volumen XVIII N<sup>o</sup>1(2022), páginas 1–7*

*Universidad de Nariño*

# Bioestadística: evolución y aplicaciones

Kelly Doranny Solís <sup>1</sup>  
Jhon Sebastián Oviedo <sup>2</sup>

**Abstract:** This article is a brief bibliographic review of the origins, the evolution and importance of biostatistics, as well as its applications in recent years. The advances in biostatistics have represented challenges and opportunities for the development of modern statistical methods that, in some instances, combine the computational resource, allowing a greater understanding of the biosciences.

*Keywords:* statistics, biostatistics, biosciences.

**Resumen:** Este artículo hace una breve revisión bibliográfica de los orígenes, la evolución y la importancia de la bioestadística, así como sus aplicaciones en últimos años. Los avances de la bioestadística han representado desafíos y oportunidades para el desarrollo de métodos estadísticos modernos que en algunas instancias combinan el recurso computacional, permitiendo una mayor comprensión de las biociencias.

*Palabras Clave:* estadística, bioestadística, biociencias.

---

<sup>1</sup> Maestría en Estadística aplicada, Universidad de Nariño, email: [doranny423@gmail.com](mailto:doranny423@gmail.com)

<sup>2</sup> Maestría en Estadística aplicada, Universidad de Nariño, email: [jnsebastian@hotmail.com](mailto:jnsebastian@hotmail.com)

## 1. Introducción

La investigación en las áreas que abarcan las biociencias ha tenido un notable avance durante los últimos años impulsando el desarrollo de la bioestadística que ha favorecido la comprensión de innumerables problemas de salud pública, como las enfermedades crónicas, el cáncer, el crecimiento y desarrollo humano, la relación entre la genética y el medio ambiente, entre otros, al igual que problemáticas ecológicas, zoonóticas, agrícolas y demás que involucran organismos vivos [1].

En términos generales, la bioestadística se refiere a la aplicación y desarrollo de técnicas estadísticas para las ciencias de la vida o biociencias, abarcando una amplia gama de disciplinas: medicina, bacteriología, epidemiología, genética, salud pública, farmacología, veterinaria, botánica, ecología, por nombrar algunas [1]. Los avances recientes en la investigación de estas ciencias han creado nuevos desafíos y oportunidades para esta rama de la estadística, que contribuye a la obtención de evidencias científicas fiables y de calidad [2].

En este sentido, las aplicaciones modernas de la bioestadística continúan desarrollándose en profundidad con métodos que incluyen el análisis multivariante, de supervivencia, discriminante, de regresión y con más protagonismo en los últimos años, la combinación de bioestadística e informática para el reconocimiento de patrones, la clasificación y la predicción [3].

Basado en lo anterior, esta investigación tuvo como objetivo revisar el origen, la evolución e importancia de la bioestadística junto con algunas aplicaciones. Para ello se pretende mostrar el uso de la minería de datos, el análisis de datos funcionales, los métodos de regresión, la aplicación de técnicas descriptivas, la comparación de medias y la asociación de variables en estudios del área de la salud. De la misma forma, la aplicación de técnicas estadísticas multivariadas como el análisis factorial y de agrupamiento en el ámbito agrícola.

## 2. Orígenes y evolución de la estadística y bioestadística

La literatura ha documentado ampliamente acerca del surgimiento de la estadística haciendo referencia a los esfuerzos del mundo antiguo por el registro de información, entre ellos, los balances del imperio romano de Augusto, el inventario de las posesiones de Carlo Magno, y el Domesday book que contenía cifras de la propiedad, extensión y valor de las tierras inglesas, considerados intentos tempranos de la estadística descriptiva [4], [5].

Sin embargo, sólo hasta 1654 la estadística se hace más visible con la noción de probabilidad, que en su forma más simple se remonta a Blaise Pascal, quien contribuyó con la solución de cuestiones matemáticas relacionadas con los juegos de azar, sobre cómo repartir equitativamente las apuestas en el caso de que se interrumpiera un juego y las posibilidades de obtener al menos un seis lanzando un dado cuatro veces u obtener un doble seis lanzando dos dados veinticuatro veces, estableciendo principios fundamentales de la teoría de la probabilidad [6].

Extraordinariamente, esta teoría paso a utilizarse desde su concepción para representar la incertidumbre en el conocimiento científico y constituyó un gran avance en la era de la Revolución Científica, en la cual muchos personajes reconocidos como Galileo y Newton hicieron uso de la probabilidad y la estadística [7], [8]. A partir de este periodo, las ideas estadísticas

comenzarían a ganar mayor importancia y a principios del siglo XVIII, el concepto de probabilidad tomaría mayor reconocimiento en la esencia como se conoce actualmente, siendo Pierre-Simon de Laplace el principal teórico científico de la época y un firme partidario del potencial que representa el análisis estadístico en diversos ámbitos, incluyendo la salud [9].

En el contexto sanitario, la aplicación de la estadística se remonta a inicios del siglo XIX cuando Europa lidiaba con diversas “fiebres” categorizadas como una enfermedad en sí, a la cual se concedió un sinnúmero de denominaciones que incluían fiebre séptica, puerperales, coléricas, amarilla, frenéticas y diarreica; un concepto que se mantuvo tras la revolución francesa hasta que el médico parisino Joseph Victor Broussais argumentó que todas las fiebres tenían el mismo origen, pues eran manifestaciones de la inflamación de los órganos [10].

Como consecuencia, surgen las teorías de Broussais que fueron popularizadas y aceptadas con rapidez por la comunidad médica de la época, las cuales consideraban la extracción de sangre (sangría terapéutica o flebotomía) usando sanguijuelas en la superficie del órgano inflamado un tratamiento eficaz [10]. Sin plena certidumbre del beneficio de esta práctica, la misma tuvo gran impacto en Europa a tal punto que la especie llegó a estar en peligro de extinción en la región, pues el consumo era tan elevado que sólo en Francia se importaron alrededor de 42 millones de sanguijuelas en 1833 y el consumo por año se acercó a los 100 millones [11].

Paralelo al auge de las teorías de Broussais surgieron las dudas de la validez de sus afirmaciones impulsadas por el también médico Pierre Charles Alexandre Louis, quien en busca de estimulación intelectual se involucra en el rol de asistente clínico en el Hospital La Charite de París durante varios años, en los cuales a través de la observación y registro, analizó cuidadosamente eventos clínicos, resultados de autopsias y mortalidad del lugar, que le permitieron proporcionar cautelosas generalizaciones basadas únicamente en estos hechos observados [12].

Durante este periodo, su investigación más importante se centró en pacientes con neumonía. Examinó 77 casos y los analizó según la duración del evento comparando los resultados en pacientes tratados con sangría terapéutica en la fase temprana versus la fase tardía de la enfermedad. Los hallazgos revelaron que de los pacientes que se les extrajeron sangre en los primeros 4 días, 18 de los 41 murieron, en comparación con 9 de 36 cuya extracción sanguínea fue más tarde. Con los resultados llegó a la conclusión que la sangría terapéutica era inútil en el tratamiento de la neumonía y posiblemente era perjudicial para los pacientes que atravesaban la peor fase de su enfermedad [12].

A partir de las demostraciones y contribuciones metodológicas de Louis a la observación clínica, su nombre se asocia estrechamente con la salud pública, la epidemiología, y la bioestadística durante los siglos XIX y XX [10]. En este punto, es sustancial aclarar que a pesar que la bioestadística surge y se desarrolla en el contexto sanitario, esta no sólo se encamina a este campo de aplicación. En este sentido, actualmente se reconoce la bioestadística como la aplicación de la estadística a la biociencia (o ciencia de la vida), definida como las ramas de las ciencias naturales que se ocupan de la estructura y el comportamiento de los organismos vivos, como la botánica, la biología, la medicina o las ciencias de la salud, y también encajarían la genética, la inmunología y la bioquímica [13].

En este orden de ideas, las definiciones de bioestadística la enmarcan como una rama de la estadística aplicada a las ciencias biológicas o médicas, tal como lo expresa Chiang en su artículo ¿Qué es la bioestadística?, donde especifica que se trata de una disciplina que se ocupa del desarrollo y la aplicación de la teoría y los métodos estadísticos para el estudio de los fenómenos que surgen en las ciencias de la vida [14] Algunas de las principales aplicaciones

de la bioestadística se relacionan con: diagnóstico, etiología y manejo de trastornos médicos, ensayos clínicos, epidemiología, salud comunitaria, demografía, genética médica humana; proteómica o análisis de micromatrices en genómica, estudios de redes de genes en biología, zoología, ecología, entre otras [3].

Su importancia en estas áreas radica en el diseño de los experimentos, la recolección y categorización de datos, y la organización, análisis y comparación de los resultados que luego sirven para la interpretación de los procesos y conceptos clínicos, conductuales, biológicos y biomédicos y la toma de decisiones relacionadas con estos aspectos [3]. En este sentido, los expertos en esta rama argumentan que el trabajo bioestadístico sólido requiere no solo una comprensión de las matemáticas, la probabilidad y las fuentes de sesgo, que sustentan la teoría y los métodos estadísticos, sino también amplias habilidades técnicas, incluida la informática, razón por la cual es sustancial una formación profunda para desarrollar estas destrezas junto con la aptitud necesaria para conceptualizar los problemas y desenvolverse apropiadamente entre las cuestiones y complejidad de las biociencias [15].

### 3. Aplicaciones de la bioestadística

En los últimos años las aportaciones en tecnología han abierto las puertas a innumerables investigaciones en bioestadística y en sus vertientes desde la agronomía hasta el campo de las ciencias médicas. Es precisamente la bioestadística una de las herramientas más útiles, influyentes e indispensables tanto para los profesionales de la salud, como a la comunidad académica en general que buscan explicar y contribuir al desarrollo científico. En este apartado se presentan algunas investigaciones con un alto interés en la predicción y análisis de enfermedades y tratamientos médicos, así como un estudio enfocado a la agronomía.

En el artículo de revisión “Predicción de cáncer de mama benigno y maligno mediante técnicas de minería de datos” de Chaurasia, Pal and Tiwari, los autores se centran en el desarrollo de tres modelos de predicción usando minería de datos: Naïve Bayes, RBF Red y J48 con fines de comparación de rendimiento para supervivencia del cáncer de mama en dos parámetros: benigno y pacientes con cáncer maligno. Los resultados indicaron que el algoritmo Naïve Bayes es el mejor predictor con una precisión del 97,36 % [16].

En el artículo “Una comparación de los métodos de regresión basados en la media y cuantiles para analizar los datos de ingesta dietética autoinformada” de Vidoni, Reininger y Lee, se propone la comparación del método de regresión lineal basado en la media con el método de regresión cuantil para describir el impacto de una intervención de comportamiento de salud en personas sanas y no saludables. Los hallazgos mostraron que el uso de la regresión por cuantiles indicó una relación no constante entre el índice de alimentación no saludable y los grupos de estudio en la cola superior de la distribución del índice de alimentación no saludable [17].

El artículo “Pronóstico de las tasas de mortalidad por cáncer de cerebro específicas por edad mediante modelos de análisis de datos funcionales” de Pokhrel y Tsokos, los autores combinaron metodologías de análisis de datos funcionales, estadísticas no paramétricas y pronósticos de series de tiempo, con el fin de modelar tendencias temporales de mortalidad para este tipo de tumores en grupos edad y pronosticar la función de mortalidad específica por edad. El modelo predijo un patrón decreciente de mortalidad para todos los pacientes con tumores cerebrales y del sistema nervioso central de los grupos de edad de 0 a 19 y de 20 a 64 años en los Estados Unidos durante los próximos 10 años [18].

El artículo “Concordancia entre las diferentes puntuaciones de riesgo cardiovascular en per-

sonas con artritis reumatoide y psoriasis”, de Gonzáles, Grande, Pertega, Seoane, Balboa y Veiga, los autores realizaron un estudio observacional descriptivo para determinar el riesgo cardiovascular y la concordancia entre las diferentes puntuaciones en personas con artritis reumatoide y psoriásica. La comparación de medias entre dos grupos se realizó mediante la prueba de la t de Student o la prueba de Mann-Whitney (previo estudio de la normalidad). La asociación de variables cualitativas se determinó con la prueba de chi-cuadrado o de Fisher según correspondiera. El estudio concluyó que los factores de riesgo más prevalentes fueron el sobrepeso y la obesidad, seguidos del tabaquismo y la hipertensión arterial [19].

El artículo “Aplicación de técnicas estadísticas multivariadas para el agrupamiento de materiales genéticos de cacao (*Theobroma cacao* L.)” de Righetto, Kakamura, López y dos Santos Dias, presenta un análisis factorial reteniendo los factores: características del fruto, características de la semilla y el componente del rendimiento que explica el 84,14% de la varianza acumulada. El método utilizado para el agrupamiento fue el jerárquico de Ward. Los 20 materiales genéticos de cacao fueron agrupados en cinco grupos con características distintas.

## 4. Conclusiones

Los avances en estadística a través de la historia han contribuido al desarrollo de nuevas maneras y métodos de afrontar problemas en las ciencias de la vida. La importancia de estudiar y aplicar diferentes técnicas para la explicación y predicción de fenómenos biológicos lleva a tener innumerables investigaciones que se enriquecen año a año gracias al desarrollo tecnológico y la exploración de nuevas vertientes en este campo de la ciencia.

Aplicaciones en el campo de la salud y en agricultura, desde el uso de procesos descriptivos e inferenciales, métodos multivariados de datos, minería de datos, análisis de datos funcionales y modelos lineales son un claro ejemplo de las múltiples herramientas estadísticas que son parte de las investigaciones alrededor del mundo en la solución de problemas.

## Referencias

- [1] Khurshid, A., Ageel, M. I., & Anwer, S. (2013). Evolution of (bio) statistics in medical research: Fifty eight years of “Numbering Off”. *DataCrítica: International Journal of Critical Statistics*, 4(1), 5-17. Available from: <https://www.researchgate.net/publication/242503773> 2
- [2] Santabárbara, J. (2020). Asociación entre las actitudes hacia la estadística y desempeño en bioestadística en estudiantes de medicina: un estudio metaanalítico. *FEM: Revista de la Fundación Educación Médica*, 23(5), 251-256. Available from: [www.fundacioneducacionmedica.org](http://www.fundacioneducacionmedica.org) 2
- [3] Zahir, H., Javaid, A., Rehman, R., & Hussain, Z. (2014). Statistical concepts in biology and health sciences. *Journal of Ayub Medical College Abbottabad*, 26(1), 95-97. 2, 4
- [4] Jaramillo, G. V., & Castillo, F. C. (2013). El levantamiento de datos estadísticos en la biblioteca. *Biblioteca Universitaria*, 16(1), 46-53. 2
- [5] Kendall, M. G. (1960). Studies in the History of Probability and Statistics. Where Shall the History of Statistics Begin? *Biometrika*, 47(3/4), 447-449. <https://doi.org/10.2307/2333315> 2
- [6] Monari, P. (2015). Considerations on probability: from games of chance to modern science. *Statistica*, 75(4), 345-360. <https://doi.org/10.6092/issn.1973-2201/6223> 2
- [7] Jaoudé, A. A. (2021). The Paradigm of Complex Probability and Isaac Newton’s Classical Mechanics: On the Foundation of Statistical Physics. In (Ed.), *The Monte Carlo Methods - Recent Advances, New Perspectives and Applications*. *IntechOpen*. <https://doi.org/10.5772/intechopen.98341> 2
- [8] Gnedenko, B.V., Sheĭnin, O.B. (1992). The Theory of Probability. In: Kolmogorov, A.N., Yushkevich, A.P. (eds) *Mathematics of the 19th Century*. *Birkhäuser, Basel*. [https://doi.org/10.1007/978-3-0348-5112-1\\_4](https://doi.org/10.1007/978-3-0348-5112-1_4) 2
- [9] Ögüs, E. (2017). To be Together Medicine and Biostatistics in History: Review. *Turkiye Klinikleri Journal of Biostatistics*, 9, 74-83. 3
- [10] Morabia A. (2006). Pierre-Charles-Alexandre Louis and the evaluation of bloodletting. *Journal of the Royal Society of Medicine*, 99(3), 158-160. <https://doi.org/10.1177/014107680609900322> 3
- [11] Whitaker, I. S., Rao, J., Izadi, D., & Butler, P. E. (2004). Historical Article: *Hirudo medicinalis*: ancient origins of, and trends in the use of medicinal leeches throughout history. *The British journal of oral & maxillofacial surgery*, 42(2), 133-137. [https://doi.org/10.1016/S0266-4356\(03\)00242-0](https://doi.org/10.1016/S0266-4356(03)00242-0) 3
- [12] Best, M., & Neuhauser, D. (2005). Pierre Charles Alexandre Louis: master of the spirit of mathematical clinical science. *Quality & safety in health care*, 14(6), 462-464. <https://doi.org/10.1136/qshc.2005.016816> 3
- [13] Chakrabarty, Dhritikesh. (2019). Association of Statistics with Biostatistics Research. *Biometrics & Biostatistics International Journal*. 8. 104 - 109. [10.15406/bbij.2019.08.00279](https://doi.org/10.15406/bbij.2019.08.00279) 3
- [14] Chiang, C. L., & Zelen, M. (1985). What Is Biostatistics? *Biometrics*, 41(3), 771-775. <https://doi.org/10.2307/2531297> 3
- [15] Lee, K. J., Moreno-Betancur, M., Kasza, J., Marschner, I. C., Barnett, A. G., & Carlin, J. B. (2019). Biostatistics: a fundamental discipline at the core of modern health data science. *The Medical journal of Australia*, 211(10), 444-446.e1. <https://doi.org/10.5694/mja2.50372> 4

- [16] Chaurasia, V., Pal, S., & Tiwari, B.B. (2018). Prediction of benign and malignant breast cancer using data mining techniques. *Journal of Algorithms and Computational Technology*, 12, 119 - 126. [4](#)
- [17] Vidoni, M. L., Reininger, B. M., & Lee, M. (2019). A comparison of mean-based and quantile regression methods for analyzing self-report dietary intake data. *Journal of Probability and Statistics*, 2019. [4](#)
- [18] Pokhrel, K. P., & Tsokos, C. P. (2015). Forecasting age-specific brain cancer mortality rates using functional data analysis models. *Advances in Epidemiology*, 2015. [4](#)
- [19] Gonzalez-Martin, C., Grande Morais, S., Pertega-Diaz, S., Seoane-Pillado, T., Balboa-Barreiro, V., & Veiga-Seijo, R. (2019). Concordance between the Different Cardiovascular Risk Scores in People with Rheumatoid Arthritis and Psoriasis Arthritis. *Cardiology research and practice*, 2019, 7689208. <https://doi.org/10.1155/2019/7689208> [5](#)