

Historia de la Estadística

El uso de los métodos estadísticos se remonta al menos al siglo V a. C. El historiador Tucídides en su *Historia de la Guerra del Peloponeso*² describe como los atenienses calculaban la altura de la muralla de Platea, contando el número de ladrillos de una sección expuesta de la muralla que estuviera lo suficientemente cerca como para contarlos. El conteo era repetido varias veces por diferentes soldados. El valor más frecuente (la moda en términos más modernos) era tomado como el valor del número de ladrillos más probable. Multiplicando este valor por la altura de los ladrillos usados en la muralla les permitía a los atenienses determinar la altura de las escaleras necesarias para trepar las murallas.

En el poema épico indio *Majabhárata* (libro 3: la historia del rey Nala), el rey Ritupama estimaba el número de frutas y hojas (2095 frutas y 50,00,000 hojas (5 crores)) en dos grandes hojas de un árbol Vibhitaka contándolos en un solo vástago. Este número era luego multiplicado por el número de vástagos en las ramas. Este estimado fue posteriormente verificado y se halló que estaba muy cerca del número verdadero. Con el conocimiento de este método Nala pudo subsecuentemente reconquistar su reino.

El primer escrito de estadística fue encontrado en un libro o del siglo IX titulado *Manuscrito sobre el descifrado de mensajes criptográficos*, escrito por Al-Kindi (801-873). En su libro, Al-Kindi da una descripción detallada sobre el uso de las estadísticas y análisis de frecuencias en el descifrado de mensajes, este fue el nacimiento tanto de la estadística como del criptoanálisis.^{3 4}

La Prueba del Pyx es una prueba de pureza de la moneda del Royal Mint, que ha sido llevada a cabo regularmente desde el siglo XII. La prueba en sí misma está basada en métodos de muestreo estadístico. Después de acuñar una serie de monedas —originalmente de 10 libras de plata— una moneda singular era colocada en el Pyx (una caja en la Abadía de Westminster). Después de un tiempo —ahora una vez al año— las monedas son retiradas y pesadas. Luego, una muestra de monedas retiradas de la caja es probada por pureza.

La *Nuova Crónica*, una historia de Florencia del siglo XIV escrita por el banquero florentino y oficial Giovanni Villani, incluye mucha información estadística sobre la población, ordenanzas, comercio, educación y edificaciones religiosas, y ha sido descrito como la primera introducción de la estadística como elemento positivo en la historia,⁵ aunque ni el término ni el concepto de la estadística como campo específico existía aún. Esto se demostró que era incorrecto después del hallazgo del libro de Al-Kindi sobre análisis de frecuencias.^{3 4}

Aunque era un concepto conocido por los griegos, la media aritmética no fue generalizada a más de dos valores hasta el siglo 16. La invención del sistema decimal por Simon Stevin en 1585 parece haber facilitado estos cálculos. Este método fue adoptado por primera vez en astronomía por Tycho Brahe, el que intentaba reducir errores en sus estimados de las localizaciones de varios cuerpos celestiales.

La idea de la mediana se originó en el libro de navegación de Edward Wright (*Certain errors in navigation*) en 1599 en una sección concerniente a la determinación de una localización con un compás. Wright sintió que este valor era el que más probablemente estuviera correcto en una serie de observaciones.

John Graunt en su libro *Natural and Political Observations Made upon the Bills of Mortality*, estimó la población de Londres en 1662 a través de registros parroquiales. El sabía que había cerca de 13,000 funerales al año en Londres y que de cada once familias tres personas morían por año. El estimó de los registros parroquiales que el tamaño promedio de las familias era 8 y calculó que la población de Londres era de cerca de 384,000. Laplace en 1802 estimó la población de Francia con un método similar.

Los métodos matemáticos de la estadística surgieron de la teoría de probabilidades, la cual tiene sus raíces en la correspondencia entre Pierre de Fermat y Blaise Pascal (1654). Christiaan Huygens (1657) proveyó el primer tratamiento científico sobre el tema que se conozca hasta la fecha. El libro *Ars Conjectandi* de Jakob Bernoulli (póstumo 1713) y *La doctrina de las probabilidades* (1718) de Abraham de Moivre trataron el tema como una rama de las matemáticas. En su libro, Bernoulli introdujo la idea de representar certeza completa como el número 1 y la probabilidad como un número entre cero y uno.

Galileo luchó contra el problema de errores en las observaciones y había formulado ambiguamente el principio de que los valores más probables de cantidades desconocidas serían aquellos que hicieran los errores en las ecuaciones razonablemente pequeños. El estudio formal en teoría de errores puede ser originado en el libro de Roger Cotes (*Opera Miscellanea*, póstumo 1750). Tobias Mayer, en su estudio de los movimientos de la Luna (*Kosmographische Nachrichten*, Núremberg, 1750), inventó el primer método formal para estimar cantidades desconocidas generalizando el promedio de las observaciones bajo circunstancias idénticas al promedio de los grupos de ecuaciones similares.

Un primer ejemplo de lo que posteriormente fue conocido como la curva normal fue estudiado por Abraham de Moivre, quien trazó esta curva en Noviembre 12, 1733.⁶ De Moivre estaba estudiando el número de caras que ocurrían cuando una moneda “justa” era lanzada.

En sus memorias —*Un intento por mostrar la emergente ventaja de tomar la media de un número de observaciones en astronomía práctica*— preparada por Thomas Simpson en 1755 (impreso en 1756) aplicaba por primera vez la teoría a la discusión de errores en observaciones. La reimpresión (1757) de sus memorias sostiene el axioma que errores positivos y negativos son igualmente probables, y que hay ciertos valores límites dentro de los cuales todos los errores se encuentran; los errores continuos son discutidos y se provee una curva de probabilidad. Simpson discutió varias posibles distribuciones de error. Primero consideró la distribución uniforme y después la distribución triangular discreta simétrica, seguida por la distribución triangular continua simétrica.

Ruder Boškovic en 1755 se basó en su trabajo sobre la forma de la Tierra propuesto en el libro *De litteraria expeditione per pontificiam ditionem ad dimetiendos duos meridiani gradus a PP. Maire et Boscovicli* para proponer que el verdadero valor de una serie de observaciones sería aquel que minimizara la suma de los errores absolutos. En terminología moderna este valor es la media.

Johann Heinrich Lambert en su libro de 1765 *Anlage zur Architectonic* propuso el semicírculo como una distribución de errores:

$$f(x) = \frac{1}{2}\sqrt{(1-x^2)}$$

con $-1 = x = 1$.

Pierre-Simon Laplace (1774) hizo su primer intento de deducir una regla para la combinación de observaciones desde los principios de la teoría de las probabilidades. El representó la ley de a probabilidad de errores mediante una curva y dedujo una fórmula para la media de tres observaciones.

Laplace en 1774 notó que la frecuencia de un error podía ser expresada como una función exponencial de su magnitud una vez descartado el signo.^{7 8} Esta distribución es ahora conocida como distribución de Laplace.

Lagrange propuso una distribución parabólica de errores en 1776:

$$f(x) = \frac{3}{4}(1-x^2)$$

con $-1 = x = 1$.

Laplace en 1778 publicó su segunda ley de errores en la cual notó que la frecuencia de un error era proporcional a la función exponencial del cuadrado de su magnitud. Esto fue descubierto subsecuentemente por Gauss (posiblemente en 1797) y es ahora mejor conocida como distribución normal, la cual es de importancia central en la estadística.⁹ Esta distribución fue referida como «normal» por primera vez por Pierce en 1873, quien estaba estudiando las medidas de error cuando un objeto era dejado caer sobre una superficie de madera.¹⁰ Escogió el término «normal» debido a su ocurrencia frecuente en variables que ocurrían en la naturaleza.

Lagrange también sugirió en 1781 otras dos distribuciones para errores —una distribución coseno—:

$$f(x) = \frac{\pi}{4}\cos\left(\frac{\pi x}{2}\right)$$

con $-1 = x = 1$ y una distribución logarítmica

$$f(x) = \frac{1}{2} \frac{1}{|x|}$$

con $-1 = x = 1$ donde $||$ es el --valor absoluto-- de x .

Laplace obtuvo una fórmula (1781) para la ley de facilidad de un error (un término acuñado por Joseph Louis Lagrange, 1774), pero esta conllevaba a ecuaciones inmanejables. Daniel Bernoulli (1778) introdujo el principio del máximo producto de las probabilidades de un sistema de errores concurrentes.

Laplace, en una investigación del movimiento de Saturno y Júpiter en 1787, generalizó el método de Mayer usando diferentes combinaciones lineales de un grupo de ecuaciones.

En 1802 Laplace estimó la población en Francia a 28,328,612.¹¹ Él calculó este número usando la cantidad de nacimientos del año anterior y el dato del censo de tres comunidades. Los datos de los censos de estas comunidades mostraron que tenían 2,037,615 personas y que el número de nacimientos era de 71,866. Asumiendo que estas muestras eran representativas de Francia, Laplace produjo un estimado para la población entera.

El método de los mínimos cuadrados, el cual era usado para minimizar errores en la medición de datos, fue publicado independientemente por Adrien-Marie Legendre (1805), Robert Adrain (1808), y Carl Friedrich Gauss (1809). Gauss había usado el método en su famosa predicción en 1801 de la localización del planeta enano Ceres. Las observaciones en las que Gauss basó sus cálculos fueron hechas por el monje italiano Piazzi. Posteriormente se dieron demostraciones por Laplace (1810, 1812), Gauss (1823), Ivory (1825, 1826), Hagen (1837), Bessel (1838), Donkin (1844, 1856), Herschel (1850), Crofton (1870), y Thiele (1880, 1889).

El término «error probable» (*der wahrscheinliche Fehler*) —la desviación media— fue introducido en 1815 por el astrónomo alemán Frederik Wilhelm Bessel.

Antoine Augustin Cournot en 1843 fue el primero en usar el término «mediana» (*valeur médiane*) para el valor que divide la distribución de probabilidad en dos mitades iguales.

Otros contribuyentes a la teoría de errores fueron Ellis (1844), De Morgan (1864), Glaisher (1872), y Giovanni Schiaparelli (1875).^[cita requerida] La fórmula de Peters (1856) para T , el "error probable" de una sola observación fue ampliamente usada e inspiró

tempranamente la estadística robusta (resistente a valores atípicos: ver criterio de Peirce).

En el siglo 19 los autores de la teoría estadística incluían a Laplace, S. Lacroix (1816), Littrow (1833), Dedekind (1860), Helmert (1872), Laurant (1873), Liagre, Didion, De Morgan, Boole, Edgeworth,¹² and K. Pearson.¹³ y K. Pearson.¹⁴

Gustav Theodor Fechner usó la mediana (*centralwerth*) en fenómenos sociológicos y sociológicos.¹⁵ Anteriormente había sido usado solamente en astronomía y campos relacionados.

Las primeras pruebas de la distribución normal fueron inventadas por el estadístico alemán Wilhelm Lexis en 1870. El único conjunto de datos disponible para él, en que le era posible mostrar que estaba normalmente distribuido, era la frecuencia de nacimientos.

Francis Galton estudió una variedad de características humanas —altura, edad, peso, tamaño de las pestañas, entre otras— y encontró que muchos de estos factores podían ser ajustados a una distribución normal.¹⁶

Francis Galton en 1907 entregó un artículo a la revista *Nature* acerca de la utilidad de la mediana.¹⁷ Él examinó la precisión de 787 intentos de adivinar el peso de un buey en una feria de campo. El peso real era de 1208: la mediana de todas las conjeturas fue 1198 libras. Las conjeturas fueron marcadamente no normales en su distribución.

El noruego Anders Nicolai Kiær introdujo el concepto de muestreo estratificado en 1895.¹⁸ Arthur Lyon Bowley introdujo el muestreo aleatorio en 1906. [20] Jerzy Neyman en 1934 hizo evidente que el muestreo aleatorio estratificado era en general un mejor método de estimación que el muestreo intencional (por cuota).¹⁹

El nivel de significación del 5 % parece ser introducido por Fisher en 1925.²⁰ Fisher expresó que las desviaciones que excedían dos veces la desviación estándar eran consideradas significativas. Previamente a esto las desviaciones que excedían tres veces el error probable eran consideradas significativas. Para una distribución simétrica el error probable la mitad del rango intercuantil. El cuantil superior de la distribución normal estándar está entre 0.66 y 0.67, su error probable es aproximadamente 2/3 de la

desviación estándar. Parece que el criterio de Fisher del 5% tenía sus raíces en la práctica previa.

En 1929 Wilso y Hilferty re-examinaron los datos de Pierce de 1873 y descubrieron que en realidad no estaba realmente normalmente distribuida.²¹