GAUSS Y LA ESTADÍSTICA

PILAR IBARROLA

Resumen. Este artículo, de carácter histórico, es un resumen de las principales aportaciones de Gauss a la Estadística: El método de los Mínimos Cuadrados y como consecuencia el llamado Modelo Lineal de Gauss.

Depués de considerar brevemente la polémica de Gauss con Legendre a propósito de la autoría del método de los mínimos cuadrados, se hace una exposición de dicho método, insistiendo en las aportaciones estadísticas de Gauss al mismo, distinguiendo entre la "Primera Aproximación de Gauss" (1809), en que supone la normalidad de los errores de observación y la "Segunda Aproximación de Gauss" (1821), en que restringe la clase de estimadores a las funciones lineales de las observaciones y suprime la normalidad de los errores. En la primera aproximación, el tratamiento es inferencial, en la segunda es un tratamiento de Teoría de la Decisión.

Se hablará también de Gauss como precursor de los Métodos Secuenciales, y finalmente se aludirá al posterior desarrollo del Modelo Gaussiano.

Carl Friedrich Gauss nació el 30 de Abril de 1777 en Brunswick (Alemania) en el seno de una familia humilde.

Su padre tuvo diferentes empleos, desde jardinero a maestro de obras hidráulicas, ayudante de un comerciante y tesorero de una pequeña aseguradora. El propio Gauss lo describió como digno de estima pero dominante, inculto y no refinado. Su madre fue el soporte de su devoción filial y murió con 97 años, después de vivir 22 años en casa de su hijo.

Gauss fue un niño precoz y autodidacta; sin ayuda aprendió a calcular antes que a hablar. Con tres años según una anécdota bien contrastada, corrigió un error en las cuentas de su padre. Aprendió a leer solo y en su primera clase de aritmética, a la edad de 8 años, dejó perplejo al profesor al resolver el problema de hallar la suma de los cien primeros números enteros.

En 1792 recibió una beca del Duque de Brunswick e ingresó en el Brunswick Collegium Carolinum. Estando en el Collegium, con 17 años, ya formuló, según afirmación propia, el principio de los mínimos cuadrados, autoría que fue objeto de posterior controversia según veremos. Estudió después en las Universidades de Göttingen y Helmstedt donde se doctoró en 1799.

A partir de 1807 se trasladó a Göttingen donde fue nombrado director del observatorio y permaneció hasta su muerte (el 23 de Febrero de 1855).

En 1856 su amigo Sartorius publicó una biografía que ha sido una importante fuente de información para el resto de sus biógrafos.

Gauss puede ser considerado uno de los mejores científicos de todos los tiempos; su profunda investigación y sus prolíficos resultados lo atestiguan. No obstante a veces sus resultados fueron producidos más rápidamente que publicados. Un ejemplo de ellos fue su acurada predicción, en 1801, de la localización en el firmamento de un supuesto planeta que G. Piazzi había brevemente observado y perdido en Enero de ese año. En Diciembre fue localizado el planeta Ceres, en la posición predicha por Gauss.

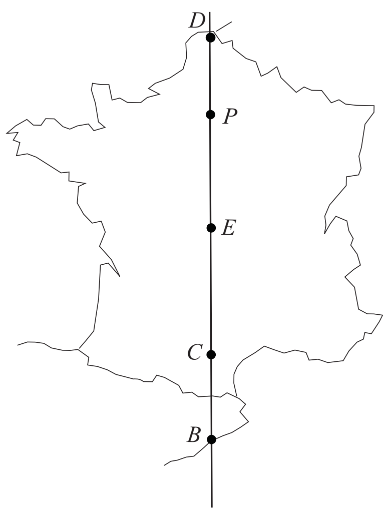
Como Gauss no hizo públicos hasta 1809 los procedimientos que había utilizado para dicha predicción (refinamiento de la teoría de la órbita y método de los mínimos cuadrados), su descubrimiento tomó un cariz sobrehumano y el personaje adquirió una fama de genio matemático y científico de primer orden.

Las principales aportaciones de Gauss a la Estadística fueron en la teoría de la Estimación: el método de los mínimos cuadrados y como consecuencia el llamado modelo lineal de Gauss.

El método de los mínimos cuadrados fue desarrollado independientemente por Gauss en Alemania, Legendre en Francia y Adrain en América. Legendre, aunque pudo no ser el primero en utilizar el método, sí que fue el primero en publicarlo (*Nouvelles méthodes pour la determination des orbites des comètes*, 1805) y fue el que le puso el nombre.

Gauss reclamó en 1806 (Monatl. Corresp. Beförd. Erd Himmelskd14, 181-186), su prioridad en el uso del método de los mínimos cuadrados (aunque no en su publicación) asegurando que hacía 12 años que venía utilizándolo y prometió publicar sus resultados más tarde. Lo hizo en 1809 en su *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, donde discute el método, menciona el trabajo de Legendre y asegura que él lo había utilizado en 1795. Legendre, a raíz de la publicación de este libro, dirigió una carta a Gauss de enhorabuena, reivindicando no obstante la autoría del método de los mínimos cuadrados. En 1820 Legendre publicó un suplemento a su memoria de 1805, atacando de nuevo a Gauss por la prioridad de los mínimos cuadrados.

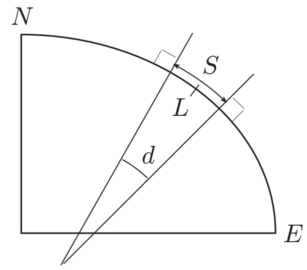
Desconociendo aparentemente el trabajo de Legendre y el de Gauss (no publicado aún), Adrain desarrolló independientemente en 1808 el método de los mínimos cuadrados y lo utilizó para resolver distintos problemas.



La polémica entre Gauss y Legendre acerca de la prioridad sobre el método de los mínimos cuadrados es famosísima en la historia de la Estadística y son muchos los científicos posteriores (Plackett 1972, Stigler 1981, Celmins 1998, etc.) que han tratado de dilucidarla, sin llegar a una conclusión definitiva.

Sin pretender entrar en profundidad en esta polémica, daremos algunas pinceladas sobre ella: En efecto, tras el ataque de Legendre, Gauss trató de probar su aplicación del método de los mínimos cuadrados anteriormente a 1805, pero no tuvo demasiado éxito, pues sus propias notas de cálculo se habían perdido y sus colegas o no recordaban discusiones con Gauss sobre el tema o no quisieron involucrarse en la disputa. Tan solo el astrónomo Olbers incluyó, en un artículo de 1816, una nota a pie de página asegurando que

Gauss le había enseñado el método de los mínimos cuadrados en 1802. Bessel publicó una nota similar en un trabajo en 1832.



En 1799 Gauss, a propósito de un trabajo sobre la medición del arco de meridiano terrestre publicado en *Allgemeine Geographische Ephemeriden*, escribe que ha utilizado "meine Methode". Surgió este trabajo a raíz de que la Academia de Ciencias Francesa decidiera en 1793 basar el nuevo sistema métrico en una unidad, el metro, igual a una 10.000.000 ésima parte del cuadrante de meridiano, distancia del polo norte al ecuador. Para ello decidieron medir el arco de meridiano que va de Dunkerque a Barcelona, pasando por París. Dividieron el arco en 4 segmentos y para cada segmento recogieron los siguientes datos: la longitud de arco S, la diferencia de latitud d y la latitud L del punto medio del arco. Los datos recogidos, de los cuales dispuso Gauss para sus cálculos, y los resultados de ajuste a los que llegó Gauss, figuran en la citada publicación.

En 1831 Shumacher escribió a Gauss, a propósito de este trabajo, sugiriéndole que repitiera los cálculos y probase que era el método de los mínimos cuadrados el empleado. Gauss se negó, alegando que su palabra era suficiente. La sugerencia de Shumacher fue recogida siglo y medio más tarde por Stigler (1981), y Aivars Celmins (1998), quienes a partir de los datos en cuestión, hicieron el ajuste por mínimos cuadrados, no llegando a los mismos resultados que Gauss.

Queda la pregunta ¿Qué método empleó Gauss? Stigler se inclina por la posibilidad de que utilizara el método de los mínimos cuadrados pero utilizando desarrollos de 2º orden. Celmins por la posibilidad de que los resultados publicados por Gauss contuvieran errores aritméticos y concluye "como dijo Gauss debemos confiar en su palabra".

Gauss y los mínimos cuadrados

Resaltemos la importancia que el propio Gauss atribuyó a este método: La primera exposición por Gauss del método de los mínimos cuadrados aparece en el libro segundo, sección 3 de su *Theoria Motus Corporum*

Coelestium (1809); se trata de la determinación de órbitas planetarias, discute la estimación de las 6 constantes o parámetros que determinan la órbita elíptica, en base a un número de observaciones $n > 6$. Comienza en el artículo 175 con "A este fin aparquemos nuestro problema particular y entremos en una discusión muy general y en una de las más fructíferas aplicaciones del cálculo a la filosofía natural".

La segunda exposición (Gauss 1821, 1823, 1826: *Theoria Combinationes Erroribus Mínimis Obnoxiae*) fue presentada en una serie de tres largos artículos a la "Royal Society of Góttingen". Aquí introduce el asunto como sigue: "El problema es ciertamente el más importante que presenta la aplicación de las matemáticas a la filosofía natural".

En 1809 Gauss plantea la estimación de las k constantes desconocidas $\theta_1, \ldots, \theta_k$ que determinan la órbita, en base a $n > k$ observaciones $y_1, \ldots, y_n$. No es posible observar los θ_i; sólo es posible observar ciertas funciones de ellos $\xi_i = \xi_i(\theta_1, \ldots, \theta_k)$. Si los ξ_i pudieran ser observados sin error, $y_i = \xi_i$, entonces $\xi_1, \ldots, \xi_n$ tendrían valores conocidos, y bastaría seleccionar k de entre ellos y despejar $\theta_1, \ldots, \theta_k$; en este caso las ecuaciones $\xi_i = \xi_i(\theta_1, \ldots, \theta_k)$ deberían ser consistentes y las $(n-k)$ no utilizadas cumplirse idénticamente con los valores θ_j despejados de las otras k. Sin embargo no es posible en la práctica observar las ξ_i sin error. Las relaciones entre las y_i y las ξ_i serán de la forma $y_i = \xi_i + \varepsilon_i$, siendo ε_i el error de observación. Con notación vectorial, $Y = \xi + \varepsilon$, siendo Y, ξ, ε los vectores (columna) de componentes $y_i, \xi_i, \varepsilon_i$.

Las ecuaciones $y_i = \xi_i$ serán ahora inconsistentes, cualquier subconjunto de k de ellas llevará a diferentes valores de $\theta_1, \ldots, \theta_k$ y en cada caso las $(n-k)$ restantes no se satisfarán.

El problema está en utilizar la totalidad de las n observaciones para obtener estimaciones óptimas $\hat{\theta}_i$ de los θ_i teniendo en cuenta la incertidumbre introducida por los errores de observación. Este es el problema que Gauss califica como "el más importante de la aplicación de las matemáticas a la Filosofía natural".

Distinguiremos entre principio de mínimos cuadrados y teoría estadística de mínimos cuadrados.

Principio de los mínimos cuadrados

El principio de mínimos cuadrados elige θ de forma a minimizar

$$(1) \qquad Q = \sum_i (y_i - \xi_i)^2.$$

Entonces θ es la solución de $\partial Q/\partial\theta_j = 0$, $j = 1, \ldots, k$, con lo que las ecuaciones de mínimos cuadrados resultan

$$(2) \qquad \sum_i (y_i - \xi_i)\frac{\partial \xi_i}{\partial \theta_j} = 0.$$

En el caso de que las ξ_i sean lineales en las θ_j, $\xi_i = \sum_j x_{ij}\theta_j$ o con notación matricial $\xi = X\theta$, siendo las x_{ij} constantes conocidas, el modelo es

$$(3) \qquad Y = X\theta + \varepsilon,$$

conocido como modelo lineal de Gauss y discutido en los textos estadísticos como regresión lineal.

Las ecuaciones de los mínimos cuadrados son

$$(X'X)\hat{\theta} = X'Y,$$

donde X' denota la matriz traspuesta de X, y el estimador de mínimos cuadrados es

$$(4) \qquad \hat{\theta} = (X'X)^{-1}X'Y.$$

Mientras que Legendre (1805) se limitó a la consideración del principio de mínimos cuadrados tal como acabamos de exponer, Gauss (1809 y publicaciones posteriores) desarrolló además la teoría estadística de mínimos cuadrados que exponemos a continuación.

Teoría Estadística de los mínimos cuadrados

Primera aproximación de Gauss (1809: *Theoria Motus Corporum Coelestum*)**.** Suponiendo que $\xi_i = y_i - \varepsilon_i$ son v.a. independientes con distribución $f(\varepsilon_i)$, la distribución conjunta de los errores de observación es

$$\Omega = \prod_i f(\varepsilon_i) = \prod_i f(y_i - \xi_i).$$

Suponiendo que todos los valores $\theta_1, \ldots, \theta_k$ son igualmente probables, la distribución a priori de θ se supone constante (distribución uniforme) y la distribución a posteriori de θ dados los valores de Y es, por el Teorema de Bayes, proporcional a Ω:

$$f(\theta_1, \cdots, \theta_k | y_1, \cdots, y_n) = k \prod_i f(y_i - \xi_i) = k\Omega,$$

siendo k una constante de normalización, que hace que las probabilidades integren la unidad.

Gauss elige el valor más probable $\hat{\theta}$ (es decir la moda de Ω) como estimador de θ: es obtenido como raíz de

$$\sum_i \frac{\partial \log f(y_i - \xi_i)}{\partial \xi_i} \cdot \frac{\partial \xi_i}{\partial \theta_j}, j = 1, 2, \ldots, k.$$

Esta aproximación es equivalente al método de la máxima verosimilitud desarrollado por Fisher en 1922.

Para aplicar el método, la forma matemática de f debe ser conocida. A tal fin, Gauss supuso que para el caso especial en que $y_i = \theta_i + \varepsilon_i$ para todo i (es decir, que hubiera un solo parámetro θ_1), el estimador de mínimos cuadrados sería la media aritmética $\hat{\theta}_1 = \bar{y}$.

En este caso la ecuación de verosimilitud resulta

$$\sum_i \frac{f'(y_i - \theta_1)}{f(y_i - \theta_1)} = 0.$$

Para que $\hat{\theta}_1 = \bar{y}$ sea solución, habrá de ser

$$\sum_i \frac{f'(y_i - \bar{y})}{f(y_i - \bar{y})} = 0,$$

cuya solución es $f(z) = ke^{-h^2 z^2}$, y que con notación moderna, tomando $h^2 = 1/2\sigma^2$ y $k = 1/\sqrt{2\pi}\sigma$, resulta la distribución normal o gaussiana para los errores.

Suponiendo que los errores siguen una distribución normal, la distribución a posteriori de θ es entonces proporcional a

$$\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-Q/2\sigma^2},$$

con $Q = \sum_i (y_i - \xi_i)^2$.

Esta probabilidad se maximiza, minimizando Q que es el principio de los mínimos cuadrados anteriormente descrito y los estimadores de $\hat{\theta}_i$ son los estimadores de mínimos cuadrados. Como dijo Gauss,

> Este principio que es extremadamente útil en todas las aplicaciones de las matemáticas a las ciencias naturales, debería ser considerado como un axioma, por el mismo motivo que tomamos la media aritmética de los valores observados de una cantidad única como el valor más verosímil de esta cantidad.

Gauss tuvo también en cuenta posibles diferencias de precisión en los y_i, y generalizó el resultado a los mínimos cuadrados ponderados $Q = \sum_i \dfrac{1}{\sigma_i^2}(y_i - \xi_i)^2$.

Gauss llegó igualmente a probar, por argumentos hoy día estándares en los textos de estadística, que $\hat{\theta}$ tiene una distribución normal multivariante de media θ y de covarianza $(X'X)^{-1}\sigma^2$.

Fue bien entrado el siglo XIX cuando la Ley normal, así bautizada por Galton, obtuvo una aceptación universal, siendo reconocida como la ley de los errores por excelencia. Su popularidad se basó en la idea de un gran número de errores elementales que se combinan para formar los errores ε_i del modelo.

Pero cuando Gauss dió su primera aproximación (1809) en su *Theoria Motus Corporum*, la ley normal era poco conocida. Gauss la descartó en su trabajo posterior principalmente porque descubrió una desigualdad aplicable a cualquier distribución continua, simétrica respecto a una moda única. Esta desigualdad es:

$$P(|X - \mu|) \geq \lambda\sigma \leq \begin{cases} 1 - \lambda/\sqrt{3} & (\lambda \leq 2/\sqrt{3}) \\ 4/9\lambda^2 & (\lambda > 2/\sqrt{3}) \end{cases}$$

y su uso con $\lambda = 2$ prueba que hay una probabilidad al menos del 89 % de que tal variable caiga dentro de dos desviaciones típicas alrededor de la media y moda. Con lo que si un estimador paramétrico se acompaña de su desviación típica, el uso de un argumento inverso permitirá sacar conclusiones sobre el rango probable de los valores del parámetro.

Laplace (1810) dio una versión rudimentaria del Teorema Central del Límite (en *Mem. Cl. Sc. Math. Phys. Inst. Fr.*) y la utilizó para demostrar que no sólo la media aritmética es ventajosa cuando la ley de los errores es normal sino también cuando el número de observaciones es grande o cuando se toma la media de resultados, basados cada uno en un número grande de observaciones y que en tal caso se debe utilizar el método de los mínimos cuadrados.

Gauss comparó su formulación (1809) con la de Laplace (1810) y concluyó que ninguna era enteramente satisfactoria introduciendo:

Segunda Aproximación de Gauss (1821,1823,1826) (*Theoria Combinationibus Erroribus Minimis Obnoxiae*, partes 1 y 2). La principal particularidad de esta segunda aproximación es que cuando θ se estima por $\hat{\theta}$, se comete un error $\theta - \hat{\theta}$ que acarrea una pérdida. El estimador $\hat{\theta}$ se elige entonces de forma a minimizar la pérdida esperada. Toma una pérdida proporcional a $(\theta - \hat{\theta})^2$ con lo que $\hat{\theta}$ se elige de forma a que minimice el error cuadrático medio $E(\theta - \hat{\theta})^2$.

Gauss supuso los errores ε_i suficientemente pequeños para que sus cuadrados y potencias superiores pudiesen ser ignorados y restringió su atención a estimadores lineales $\hat{\theta} = CY$ tales que $CX = I$ (matriz identidad $k \times k$).

Probó entonces que entre tales estimadores, el estimador de mínimos cuadrados [4] minimiza el error cuadrático medio. El error cuadrático medio mínimo es entonces $(X'X)^{-1}\sigma^2$, llegándose entonces así a los resultados de la 1ª aproximación, con la excepción de la normalidad de $\hat{\theta}$.

Gauss obtuvo también otros resultados como el siguiente:

Una observación adicional de y con x-valores correspondientes

$$x' = (x_{n+1,1}, x_{n+1,2}, \ldots, x_{n+1,k})$$

se puede incorporar al estimador de mínimos cuadrados original $\hat{\theta}$ para formar el nuevo estimador

$$\theta^* = \hat{\theta} - M(x'\hat{\theta} - Y)/(1 + W)$$

con $M = (X'X)^{-1}x$ y $W = x'(X'X)^{-1}x = x'M$. La matriz de covarianzas de θ^* es $((X'X)^{-1} - MM'/(1 + W))\,\sigma^2$, y si indicamos por Q_m el antiguo mínimo de Q el nuevo mínimo es

$$Q_m^* = Q_m + (x'\hat{\theta} - y)^2/(1 + W).$$

Este procedimiento de mínimos cuadrados recursivos, permite incorporar en el procedimiento de estimación, nuevas observaciones, obtenidas secuencialmente, sin necesidad de invertir en cada instante la nueva matriz $(X'X)$.

Como $E[Q_m] = (n - k)\sigma^2$, Gauss estimó σ por $\hat{\sigma} = \sqrt{Q_m/(n - k)}$.

Obtuvo también el error estándar de la estimación y observó que cuando las ε_i son normales, este error estándar es el error estándar de la suma de $(n - k)$ errores independientes ε_i.

La diferencia en cuanto a generalidad entre la 1ª y la 2ª aproximaciones de Gauss debe ser resaltada. La 1ª aproximación permite que $\hat{\theta}$ sea cualquier función de las observaciones, pero requiere que los errores de observación se distribuyan normalmente con media cero. La 2ª aproximación restringe el estimador $\hat{\theta}$ a las funciones lineales de las observaciones pero permite a los ε_i tener cualquier distribución con media cero y varianza finita.

Respecto a la 1ª aproximación, observemos que Gauss no dio a sus resultados una interpretación bayesiana. No habló de probabilidades o de precisión del parámetro desconocido, sino que le dio una interpretación frecuentista al hablar de la precisión del estimador $\hat{\theta}$.

El método de estimación al que llega es el de la máxima verosimilitud. Gauss obtiene el método de los mínimos cuadrados como un caso especial de máxima verosimilitud, apropiado cuando la distribución es normal, o de forma equivalente, cuando la media aritmética es el mejor estimador de posición.

Si la distribución no es normal, el método de máxima verosimilitud no tiene porqué coincidir con el método de los mínimos cuadrados, el caso más famoso es aquél en que los errores sigan una distribución de Cauchy

$$f(e) = \frac{1}{\pi}\frac{1}{1+e^2}, \quad e = y - \theta,$$

y en tal caso la media aritmética $\bar{y}$ es casi el peor estimador posible de θ, y el método de los mínimos cuadrados no produce un buen estimador pero el de máxima verosimilitud sí.

Cabe preguntarse porqué Gauss no se planteó relajar la supremacía de $\bar{y}$ e investigar otras posibles distribuciones para los errores. La razón puede estar en que su época la única distribución continua surgida experimentalmente era la normal.

No obstante en su segunda aproximación dejó libre la distribución de los errores, exigiendo sólo que tuviera media cero y desviación típica finita.

En su segunda aproximación, Gauss abandona el tratamiento "inferencial" y como comunicó en una carta a Bessel se inclina por un tratamiento de "Teoría de la Decisión".

Empieza en la primera parte (1821) de su *Theoria Combinationibus Erroribus* por comparar el problema de la estimación de un parámetro desconocido con un juego en que hay temor a perder y no hay esperanza de ganar. El error cometido se considera como pérdida que uno sufre. Tan indeseable es el juego que se mide por la pérdida esperada, suma de los productos de las posibles pérdidas por sus respectivas probabilidades. Toma como conveniente, aunque arbitrario, que la pérdida sea proporcional al cuadrado del error cometido.

Fue así Gauss un precursor de la moderna teoría de la Decisión formalizada por Wald a mediados del siglo XX.

Un aspecto desafortunado del tratamiento del modelo lineal de Gauss en los textos modernos es su referencia como teoría de Gauss-Markov, ya que la asociación con el nombre de Markov no parece justificado. Además se suele insistir en que Gauss imponía que los estimadores fuesen insesgados, requerimiento que puede conducir a resultados absurdos; es cierto que la condición $CX = I$ implica que $\hat{\theta}$ sea insesgado, pero Gauss no insistió en considerar sólo estimadores insesgados ya que propuso como estimador de σ, $\hat{\sigma} = \sqrt{Q_m/(n-k)}$, que es sesgado. La condición $CX = I$ es de hecho un criterio de consistencia que hace que si $\varepsilon_i = 0$, para todo i, las ecuaciones exactas $Y = X\theta$ lleven a que el estimador $\hat{\theta}$ coincida con θ.

Como dice Bertrand (1888, p 255) traductor oficial del trabajo de Gauss al francés “Car, sans cela, toutes les mesures etant supposeés exactes, la valeur qu'on en déduit ne le serait pas”.

También, los textos modernos no restringen el dominio de aplicación de los estimadores lineales, como hizo Gauss, suponiendo que los errores de observación son suficientemente pequeños para ignorar sus cuadrados y potencias superiores.

Posterior desarrollo del modelo Gaussiano. Respecto al posterior desarrollo del modelo lineal Gaussiano aparecen dos innovaciones primordiales: los tests de hipótesis lineales y las distribuciones exactas en el muestreo asociadas a la distribución normal de los errores.

Ya, en, 1836 Cauchy refiriéndose al modelo de Legendre, planteó la cuestión de si uno o más términos del modelo lineal podían ser descartados por no significativos, pero fue Fisher en 1922 quien introdujo en el modelo de Gauss la idea de contrastar la nulidad de un grupo de parámetros, es decir consideró la hipótesis lineal $A\theta = 0$ y propuso la técnica de hacerlo, procedimiento conocido como F-test. Ya en esa época las distribuciones χ^2 de Pearson, t de Student, F de Snedecor, también llamada F de Fisher, habían sido introducidas por los autores que les dieron nombre. Fue también Fisher en 1922 quien introdujo la utilización de variables cualitativas que dan lugar a X-vectores de componentes nulas y la elección de ellas de forma que permitan lograr un análisis posterior numéricamente sencillo, llegando así al Análisis de la Varianza.

Otras contribuciones de Gauss a la Estadística. Existen también otras contribuciones estadísticas de Gauss. Una notable (1866) es la demostración de que para una distribución normal, el estimador más preciso de σ^2 entre los estimadores que dependen de $S_k = \sum |e_i|^k$, con $e = (x'\hat{\theta} - y)$ se obtiene cuando $k = 2$.

En otro artículo (1824) Gauss utiliza el método de los mínimos cuadrados ponderados para determinar la longitud mediante un cronómetro.

Presenta una discusión sobre la estructura de los errores de observación ε_i obtenidos en el uso del cronómetro y toma la desviación típica σ_i proporcional a la raíz cuadrada del tiempo transcurrido entre observaciones. En toda la demostración aparece un híbrido entre teoría y aplicación bastante raro hoy en día.

Conclusión

Independientemente de la polémica entre Gauss y Legendre sobre la paternidad del principio de los mínimos cuadrados, fue Gauss y no Legendre quien desarrolló el método como herramienta estadística, encajándolo en un marco estadístico, introduciendo un tratamiento probabilístico de los errores de observación y planteando el modelo lineal.

Dejando aparte la controversia de utilizar el Teorema de Bayes, los cálculos de Gauss son hoy en día estándares. En gran medida la teoría de la regresión y el diseño de experimentos que forman la base de la estadística moderna, dependen de la descomposición de Q en suma de cuadrados.

Recogemos las siguientes afirmaciones de Fisher (Fisher 1970, 21-22):

> Gauss aproximó el problema de la estimación estadística con espíritu empírico, recalcando la cuestión de la estimación no sólo de las probabilidades sino de otros parámetros cuantitativos. Descubrió que, para este propósito, el método de la máxima verosimilitud era el apropiado, aunque trató de justificar el método por el principio de la probabilidad inversa [...].
>
> Gauss además perfeccionó el ajuste sistemático de las fórmulas de regresión simple y múltiple por el método de los mínimos cuadrados.

y (Fisher 1973, 88):

> El camino recto fue trazado por Gauss hace muchos años y a su método sólo le falta en su refinamiento el uso de la t-Student para muestras pequeñas.

Además el trabajo de Gauss sobre “los mínimos cuadrados recursivos” fue redescubierto en el último tercio del siglo XX y aplicado profusamente al tratamiento de series temporales. Como dice Joung (1974, 210):

> En particular el método recurrente tuvo gran importancia en el procesamiento de datos para estimar la órbita y trayectoria de la misión Apolo,

y sigue

> Uno se pregunta si Gauss podría haberse imaginado que su sencilla aproximación al análisis de observaciones serviría 160 años más tarde a contribuir al éxito del primer viaje extraterrestre del hombre.

Finalmente observemos que el filtro de Kalman, de gran aplicación actual en series temporales, no es sino un refinamiento del método recurrente de Gauss.

Clamens Schäfer, uno de sus biógrafos, escribió en Nature (128 (1931), p.341):

> No fue realmente un físico en el sentido de buscar nuevos fenómenos, sino un matemático que procuró formular en términos matemáticos exactos los resultados experimentales obtenidos por otros.

Seal (1967) escribe:

> Gauss aparece como el matemático ideal, ostentando en proporciones heroicas para una persona, las capacidades atribuidas colectivamente a la comunidad de profesionales matemáticos.

Referencias

Adrain R. (1808) Analyst 1, p. 93-109.

Bertrand J. (1888) Calcul des probabilités (2nd. Ed. 1972. Chelsea).

Celmins A. (1998) *The method of Gauss in 1799*. Statistical Science 13 , nº 2, p. 123-135.

Cauchy A. (1836) *On a new formula for solving the problem of interpolation in a manner applicable to physical investigations*. Phil. Mag 8, p. 459-468.

Dictionary of Scientific Biography. *Gauss, Carl Friedrich*, vol 5, p. 298-314.

Encyclopedia of Statistical Science. *Gauss, Carl Friedrich*, vol 3, p.305-309.

Encyclopedia of Statistical Science. *Least Squares*, vol 4, p. 593-598.

Fisher R.A. (1922) *On the mathematical foundations of theoretical Statistics*. Phil. Trans Roy.Soc. A 222, p.309-368.

Fisher R.A. (1970) Statistical Methods for Research Workers. Oliver& Boyd.

Fisher R.A. (1973) Statistical Methods and Scientific Inference. Hafner Press.

Gauss C.F. (1799) Vermischte Nachrichten nº 3. Allgemeine Geographische Ephemeridenz 4, p. 378.

Gauss C.F. (1806) Monatl. Corresp Beförd. Erd Himmelskd 14, p. 181-186.

Gauss C.F. (1809) *Theoria Motus Corporum Coelestum*, Werke 7 (traducido al inglés por C.H. Davis (1963), Dover).

Gauss C.F. (1816) *Bestimmung der Genauigkeit der Beobachtungen*. Werke 4, p. 109-117.

Gauss C.F. (1821-1823-1826) *Theoria Combinationis Erroribus Minimis Obnoxiae*. Partes 1, 2 y suplemento, Werke 4, 1-108.

Gauss C.F. (1824) *Chronometrische Langer Bestimmungen*. Astromische Nachrinten 5,p. 227.

Laplace P.S. (1810) Mém. Cl. Sc. Math. Phys. Inst. Fr., Année 1810, p. 279-247.

Laplace P.S. (1812) Théorie Analytique des Probabilités (Paris).

Legendre A.M. (1805) Nouvelles méthodes pour la determination des orbites des cometes. (Apéndice: Sur la méthode des moindres carrés).

Olbers W. (1816) *Ueber den veränderlichen Stern im Halse des Schwans.* Z. für Astronomie und verwandte Wissenschaften 2, p.181-198.

Plackett R.L. (1972) *The discovery of the method of least squares.* Biometrika 59, p. 239-251.

Sartorius W. (1856) Gauss zum Geächtniss. Leipzig

Seal H.L. (1967) *Studies in the history of Probability and Statistics XV: The historical development of the Gauss linear model.* Biometrika 54, p. 1-24.

Shäfer C. (1931) Nature 128, p.341.

Sprot D.A. (1978) *Gauss's contributions to statistics.* Historia Mathematica 5, p. 183-203.

Stigler S.M. (1981) *Gauss and the invention of least squares.* Ann. Statist. 9, p. 465-474.

Young P. (1974. *Recursive approach to time series analysis.* Bull. Inst. Math and its Applications 10, p.209-224.