

La influencia de las teorías de Newton y Darwin en el nacimiento de la Estadística Matemática

por DANIEL PEÑA SANCHEZ DE RIVERA

Departamento de Estadística
E.T.S.I. Industriales,
Universidad Politécnica de Madrid

RESUMEN

Este trabajo analiza la influencia de dos paradigmas científicos en la evolución y consolidación de la Estadística Matemática. La contrastación de la teoría de Newton es el desencadenante de la creación de la teoría de errores. La contrastación de la teoría de Darwin da lugar a la inferencia estadística. El estudio de estos períodos históricos pone de manifiesto la importancia de la interacción entre la teoría y la práctica en la creación de la Estadística Matemática y en su desarrollo posterior.

Palabras clave: Cálculo de probabilidades, Newton y la teoría de errores, la Estadística en el siglo XIX, problemas de la evolución, la herencia y la teoría de la Inferencia estadística, Estadística Matemática.

1. INTRODUCCION

La Estadística como disciplina científica ocupa un lugar singular entre el conjunto de las ciencias. La Física, la Medicina o la Sociología tienen un área sustantiva de conocimiento y cuando utilizan modelos matemáticos, los subordinan al objeto principal de hacer avanzar el conocimiento en su parcela de estudio de la realidad. El objetivo de la Matemática, en contraposición, es ampliar la concepción y generalidad de sus propias herramientas analíticas, con absoluta independencia de la posible relación entre los entes matemáticos abstractos y los fenómenos reales.

La Estadística participa de estos dos objetivos, aunque con rasgos muy peculiares. Su campo de estudio son los fenómenos aleatorios que están presentes, en mayor o menor medida, en toda actividad humana de adquisición de conocimiento empírico. La comprensión de un fenómeno aleatorio concreto implica, forzosamente, la interacción con la disciplina sustantiva, que estudia esa clase de fenómenos. Las técnicas estadísticas actúan como una gran lupa que amplifica ciertas facetas medibles del fenómeno estudiado que no son observables sin este análisis riguroso, mientras que la ciencia que estudia esa clase de fenómenos aporta el modelo conceptual en el que los hechos empíricos identificados y cuantificados encuentran su interpretación lógica. Por esta razón, la Estadística es fundamental en todo proceso de adquisición de conocimientos a través de datos empíricos y se justifica la afirmación de Mood y Graybill (1972) de que «La Estadística es la tecnología del método científico»¹.

La Estadística comparte con la Matemática el estudio lógico-deductivo de sus propias herramientas conceptuales, aunque con una importante diferencia. El criterio de calidad final en la Estadística debe ser, en definitiva, la relevancia de los instrumentos desarrollados para la comprensión de los fenómenos aleatorios, más que la generalidad y elegancia de los procedimientos empleados.

Este carácter dual de la Estadística explica que los avances fundamentales de la disciplina hayan provenido, con cierta frecuencia, de investigadores en otras áreas y que se dé la paradoja de que la Estadística se ha desarrollado, en gran parte, por el impulso de investigadores que provienen de otras disciplinas y que no pueden considerarse como «estadísticos puros», sino, sobre todo, como científicos.

Este trabajo pretende ilustrar la importancia decisiva que ha tenido en esta disciplina el proceso de interacción entre el planteamiento de nuevos problemas prácticos susceptibles de análisis estadístico y la invención de nuevas herramientas teóricas que, generalizadas posteriormente, se convierten en nuevas áreas de investigación teórica y de aplicación en otros problemas. Esta interacción entre la teoría y la práctica aparece destacadamente en las investigaciones que dan lugar a la consolidación de la Estadística Matemática actual, como disciplina, a principios de nuestro siglo. En concreto, pensamos que en la evolución de esta ciencia ha tenido un papel central la necesidad de

¹ Esta misma idea ha sido expuesta por Barnard (1972), que asigna a la Estadística el papel de «midwife» de los avances científicos. Cornfield (1975) argumenta que mientras otras ramas científicas tienen por objeto la acumulación de conocimiento sobre el mundo externo, la Estadística estudia los *métodos* para adquirir ese conocimiento, más que el conocimiento en sí y que la Estadística es «bed-fellow» de otras ciencias. El lector interesado en distintas formulaciones del papel de la Estadística con relación a otras disciplinas, encontrará un amplio espectro de opiniones en las conferencias anuales de los Presidentes de la Royal Statistical Society y The American Statistical Association, que se publican en *JRSS, A* y *JASA*, respectivamente, cada año.

generar instrumentos conceptuales operativos para contrastar dos paradigmas [en el sentido de Khun (1962)] fundamentales en la historia de la ciencia. El primero de ellos ha contribuido de forma destacada a la comprensión del mundo físico y del Universo: la teoría de Newton. El segundo ha proporcionado un marco global para interpretar la evolución de las especies en nuestro planeta: la teoría de Darwin.

La contrastación de la teoría de Newton está en la raíz de los importantes avances metodológicos en la Estadística durante el siglo XVIII y la primera mitad del siglo XIX. Los problemas de la herencia y de la selección unifican las investigaciones que conducen al nacimiento de la Estadística Matemática en Inglaterra entre finales del XIX y comienzos del XX.

Este trabajo está estructurado como sigue: En la sección dos se revisa brevemente el nacimiento del cálculo de probabilidades para concluir exponiendo su situación a comienzos del siglo XVIII. La sección tres relaciona las aportaciones del XVIII y mediados del XIX con el paradigma de Newton. La sección cuatro resume la situación de la Estadística —en el sentido que ha tenido esta disciplina durante los siglos XVIII y XIX— antes de su unificación con la teoría de la inferencia, desarrollada en Inglaterra a finales del siglo pasado.

La sección cinco resume las aportaciones más importantes que surgen relacionadas con los problemas de la evolución y la herencia planteados por Darwin, y en la sección seis se comenta brevemente la unificación de la Estadística Matemática y su expansión posterior. Finalmente, la sección siete resume algunas conclusiones especialmente destacables.

2. EL PERIODO FUNDAMENTAL DEL CALCULO DE PROBABILIDADES

Un hecho aparentemente sorprendente en la evolución del pensamiento científico es la lentitud con la que la noción actual del azar aparece en la mente humana. Según David (1962), los juegos de azar tienen una antigüedad de más de cuarenta mil años, como parece confirmar la abundante presencia del hueso astrágalo de oveja o ciervo (que constituye el antecedente inmediato del dado) en las excavaciones arqueológicas más antiguas. La utilización del astrágalo en culturas más recientes, Grecia, Egipto y posteriormente Roma, ha sido ampliamente documentada [David (1955), Todhunter (1965)]. En las pirámides de Egipto se han encontrado pinturas que muestran juegos de azar y que provienen de la primera dinastía (3.500 a. de C.), y Herodoto se refiere a la popularidad y difusión en su época de los juegos de azar, especialmente mediante la tirada de astrágalos y dados. Los dados más antiguos que se han encontrado se remontan a unos tres mil años antes de Cristo y se utilizaron tanto en el juego como en las ceremonias religiosas [Hasofer (1967), Kendall (1956)].

A pesar de esta presencia continuada de los experimentos aleatorios en las civilizaciones antiguas, los sucesos aleatorios no se consideran objeto de investigación científica hasta el Renacimiento. La razón más convincente, en mi opinión, para explicar el hecho, ha sido expuesta por Kendall (1956). En las civilizaciones antiguas, la voluntad divina se manifiesta mediante signos interpretables por los hombres. Uno de los signos que aparece con insistente repetición en culturas muy diversas es la explicación del azar mediante la voluntad divina. Los oráculos, sacerdotes o pitonisas utilizan la configuración resultante de tirar cuatro dados en Grecia y Roma para predecir el futuro y revelar la voluntad favorable o desfavorable de los dioses. Por ejemplo, en Grecia clásica y Roma la aparición de la combinación Venus (aparición de 1, 3, 4, 6 al tirar cuatro dados) era favorable y se ha descubierto en Asia Menor una completa descripción de la interpretación profética de los posibles resultados al tirar cuatro dados. Prácticas similares se han encontrado en culturas tan distantes como la tibetana (Dawid, 1955), la india (Mahalanobis, 1957) o la judía (Hasofer, 1967).

El factor común de estas prácticas religiosas, basadas en experimentos aleatorios, es que los resultados de la experimentación son impredecibles y no es posible encontrar una causa o conjunto de causas que permitan explicar el resultado. Es, por tanto, natural, en una concepción sobrenatural del Universo, atribuir el resultado a la voluntad divina. Un hecho interesante es que esta actitud mágica ante el azar se manifiesta igualmente en los niños, Piaget (1968), y que no está actualmente todavía del todo ausente en las culturas actuales.

El Renacimiento supuso un nuevo enfoque global de considerar el mundo, que indujo una observación cualitativamente distinta de muchos fenómenos naturales. En concreto, el abandono progresivo de explicaciones teológicas conduce a una consideración de los experimentos aleatorios. Sin entrar en la polémica de cuál fue el primer autor que intuyó la simetría en las tiradas de un dado, es indudable que los matemáticos italianos de comienzos del siglo XVI, habían realizado el proceso de conceptualización necesario para interpretar los resultados de experimentos aleatorios simples. Por ejemplo, Cardano, en 1526, establece, por condiciones de simetría, la equiprobabilidad de aparición de las caras de un dado a largo plazo.

Galileo (1564-1642) debe considerarse también uno de los precursores del cálculo de probabilidades por sus comentarios de respuesta [Todhunter (1965)] a un jugador que le preguntó por qué es más difícil obtener 9 tirando 3 dados que obtener 10. Galileo razonó que de las 216 combinaciones posibles equiprobables, 25 conducen a 9 y 27 a 10. Señalamos este dato porque la diferencia empírica entre ambos resultados es únicamente de $\frac{2}{216} \approx 0,01$, lo que muestra cómo a finales del siglo XVI existía un intuitivo pero preciso análisis empírico de los resultados aleatorios.

El desarrollo del análisis matemático de los juegos de azar se produce lentamente durante los siglos XVI y XVII. Una opinión muy extendida, aunque discutida [véase Pearson y Kendall (1970)], es considerar como origen del cálculo de probabilidades la resolución del problema de los puntos en la correspondencia entre Pascal y Fermat en 1654. El problema planteado a estos autores por Chevalier de Meré, un jugador empedernido de la Francia del XVII, fue cómo debería repartirse el dinero de las apuestas depositado en la mesa si los jugadores se ven obligados (presumiblemente por la Policía, ya que el juego estaba entonces prohibido) a finalizar la partida sin que exista un ganador.

El cálculo de probabilidades se consolida como disciplina independiente en el período que transcurre desde la segunda mitad del siglo XVII hasta comienzos del siglo XVIII. El primer tratado de esta disciplina es debido a Huygens, y fue publicado en 1657. Posteriormente, De Moivre² introdujo las leyes generales de adición, las funciones generatrices de probabilidad y las ecuaciones en diferencias para el cálculo de probabilidades en su tratado «De Mesura Sortis», en 1711. Dos años después, en 1713, se publica póstumamente «Ars Conjectandi», de J. Bernoulli (1654-1705), que extiende los resultados de Huygens para el problema de la duración del juego y demuestra la primera ley de los grandes números.

Estos trabajos configuran el primer paradigma del cálculo de probabilidades. El concepto de probabilidad no está todavía formalizado, aunque sí lo está el de esperanza matemática. El ámbito de aplicación se reduce a las distribuciones discretas, dentro de las cuales, la binomial ocupa un lugar central y para la que se ha establecido la ley de Bernoulli de los grandes números. Han surgido los primeros esbozos de una teoría de procesos ligados al problema de la duración de un juego³, aunque el concepto de dependencia estocástica, básico para la creación de una teoría de procesos estocásticos, no se abordará con generalidad hasta dos siglos más tarde.

3. NEWTON Y EL DESARROLLO DE LA TEORIA DE ERRORES

La obra de Newton (1642-1727) constituyó la mayor revolución científica de los siglos XVII y XVIII y su influencia en la evolución de las ciencias físicas es ampliamente

² De Moivre, estudiando la ley de los grandes números, perfeccionó los límites obtenidos por J. Bernoulli y obtuvo, de hecho, por primera vez áreas de la distribución normal. Sin embargo, la idea de una distribución continua no existía todavía, por lo que la demostración de la convergencia de la binomial a la normal es realmente debida a Laplace (Seal, 1968).

³ El problema de la duración del juego está estrechamente relacionado con el análisis de un paseo aleatorio con dos barreras absorbentes. Véase Thatcher (1957) para una historia del mismo.

conocida. En Astronomía, Newton explicó no solamente las leyes de Kepler por el principio de gravitación universal, sino que, estableció un modelo global para estudiar las relaciones entre los cuerpos estelares. En Física estableció una teoría común para explicar fenómenos que habían sido objeto de estudios fragmentarios e incompletos como péndulos, planos inclinados, mareas, etc. En Matemáticas contribuyó con Leibnitz a la creación del cálculo diferencial e integral.

Durante todo el siglo XVIII y parte del XIX, la investigación en Física y Astronomía está dirigida por el paradigma de Newton. Esta investigación se centra en: *a*) campos de observación y experimentación que la teoría de Newton señala como especialmente relevantes; *b*) contrastación de las predicciones de la teoría con los datos; *c*) extender las aplicaciones de la teoría en otros campos. Estas investigaciones van a ser de importancia fundamental en el desarrollo de la Estadística Matemática y D. Bernoulli, Simpson, Lagrange, Laplace y Gauss realizan sus aportaciones probando y contrastando el paradigma de Newton, principalmente en Astronomía.

El primer problema que se presentaba en el proceso de contrastación empírica de esta teoría en Astronomía es el tratamiento de los errores de medición. Se disponía de varias medidas independientes de una determinada magnitud física y se presentaba el interrogante de cómo combinarlas para obtener un resultado más preciso. Aunque este problema se había planteado en la Astronomía desde la antigüedad ⁴, la necesidad de comparar con exactitud los datos observados con la teoría requiere un tratamiento riguroso del mismo, que va a dar lugar a la teoría de errores.

T. Simpson obtiene en 1755 la distribución de la media aritmética en el supuesto de que la distribución de los errores de observación sigue una distribución discreta y fue el primero en considerar distribuciones continuas ⁵ [véase Plackett (1958)], aunque impuso la restricción, presente también en el trabajo de D. Bernoulli, de que la función de densidad de los errores debe ser simétrica y con rango finito. D. Bernoulli (1700-1782) introduce [D. Bernoulli (1777)] el concepto de estimador máximo-verosímil y aplica un test de significación para determinar si puede aceptarse la hipótesis de aleatoriedad en el ordenamiento de las órbitas de los planetas. Este autor introdujo también, el concepto de utilidad resolviendo la famosa paradoja de San Petersburgo, y fue pionero en la aplicación del cálculo infinitesimal al cálculo de probabilidades.

Pero el impulso fundamental de este período proviene de la obra de Pierre Simon, marqués de Laplace (1749-1827). Laplace introduce la primera definición explícita de

⁴ El lector interesado en la historia de este problema puede acudir a Plackett (1958).

⁵ Buffon arguyó haber considerado distribuciones continuas en 1733, estudiando el famoso problema de la aguja que lleva su nombre, pero sus resultados no se publicaron hasta 1777.

probabilidad y se plantea el problema general de inferencia, que resuelve redescubriendo el teorema de Bayes, que había sido publicado póstumamente en 1764 [Bayes (1764)]. Introduce el criterio de minimización del valor esperado ⁶ del error absoluto y demuestra las ventajas para algunas distribuciones de probabilidad de utilizar como estimador la mediana muestral, en lugar de la media. Introduce la ley normal de errores y demuestra el primer teorema central del límite utilizando una nueva herramienta: la función característica. Laplace plantea el problema de regresión múltiple, tomando como criterio de estimación la minimización de las desviaciones absolutas y, aunque sus procedimientos matemáticos fueron muy «ad hoc» [Stigler (1975)], sus resultados fueron sorprendentemente precisos ⁷.

La obra de Laplace va a tener una influencia fundamental en la evolución de la Estadística. En su «*Theorie Analytique des Probabilités*» (1818) presenta por primera vez una comparación entre métodos alternativos de estimación y obtiene las condiciones en las que el criterio de desviaciones absolutas proporciona estimadores mejores que los mínimos cuadrados con el criterio de menos varianza en su distribución asintótica. Como ha señalado Stigler (1973b), Laplace al considerar la distribución conjunta de dos estimadores realiza el trabajo previo que llevó a Fisher al descubrimiento del concepto de suficiencia de un estimador. Además, sus trabajos sobre la mediana muestral son los primeros estudios de distribución de estadísticos ordenados (Stigler 1973a), área en la que también debe considerarse a Laplace como pionero.

La segunda contribución fundamental de este período es debida a Gauss (1777-1855), que introduce y desarrolla el modelo lineal. Aunque la paternidad original de los mínimos cuadrados hay que atribuirlos a Legendre, que introdujo en 1805 este método como un procedimiento de ajuste astronómico, la concepción y desarrollo del modelo lineal pertenece plenamente a Gauss ⁸.

Gauss parte del modelo de regresión lineal y establece [Seal (1967)] la hipótesis de que si la distribución «a priori» sobre los parámetros es uniforme, la moda de la distribución posterior debe ser igual a la media. Esto le conduce a tomar como distribución de los errores de observación la distribución normal. Partiendo de esta formulación,

⁶ Valor esperado respecto a la distribución posterior tomando como distribución prior la uniforme.

⁷ Una excelente discusión de la teoría de Júpiter y Saturno se encuentra en Stigler (1975).

⁸ La paternidad del método de los mínimos cuadrados fue objeto de agrias polémicas entre Gauss y Legendre. Gauss reclamó haberlo descubierto en 1795, a los dieciocho años de edad. Sorenson (1970) ha descrito cómo curiosamente se ha producido una polémica similar entre Swerling y Kalman a propósito de la generalización de los mínimos cuadrados para problemas de estimación recursiva de modelos en el espacio de los estados, lo que ha conducido al algoritmo conocido como filtro de Kalman.

obtuvo la distribución posterior de los parámetros, que será máxima cuando la suma de los cuadrados de los errores sea mínima, lo que conduce inmediatamente al principio de mínimos cuadrados. Gauss estableció un procedimiento computacional eficiente para resolver el sistema de ecuaciones y demostró que si abandonamos la hipótesis de normalidad, el método de mínimos cuadrados se justifica por el teorema que actualmente se denomina de Gauss-Markov ⁹.

Durante la primera mitad del siglo XIX los matemáticos-astrónomos continúan ampliando la teoría de errores y podemos observar la aparición de problemas y métodos que van a tener gran influencia posterior. Bravais (1846), geólogo y astrónomo, es el primero en considerar la distribución normal multivariadamente como matriz de varianzas y covarianzas no diagonal, aunque como en astronomía no existía la necesidad de una medida de correlación, Bravais no intenta interpretar los términos del exponente que conducirán a Galton y Pearson al concepto y definición del coeficiente de correlación. Benjamin Pierce (1852) propone el primer criterio para rechazar «outliers» y S. Newcomb, el más famoso astrónomo americano del XIX, introduce métodos de estimación robusta que son análogos [Stigler (1973a)] a la clase M de Huber (1964) y que constituyen en la actualidad objeto de prometedoras investigaciones.

Por tanto, a mediados del siglo XIX existen ya las herramientas básicas que van a dar lugar a la Estadística Matemática actual. Sin embargo, la aplicación de estos principios va a restringirse a la Física y la Astronomía y va a tener en esta época poca influencia sobre otras áreas de conocimiento.

En particular, los avances de la teoría de errores tienen poca influencia sobre otra disciplina científica cuyo campo de estudio es el análisis cuantitativo de datos demográficos, sociales y económicos y que se conoce, desde el siglo XVII, con el nombre de Estadística.

4. LA ESTADÍSTICA EN EL SIGLO XIX

Desde la antigüedad, los Estados han recogido información sobre la población y riqueza que existía en sus dominios. Los censos romanos, los inventarios de Carlomagno de sus posesiones, etc., pueden considerarse precedentes de la institucionalización de la recogida de datos demográficos y económicos por los Estados modernos. Esta aritmética política o Estadística Descriptiva evoluciona durante los siglos XVII y XVIII tomando progresivamente un carácter más cuantitativo.

La publicación de «Observations» por Grant, en 1662, es probablemente el primer intento de aplicar un razonamiento propiamente estadístico, en el sentido actual del tér-

⁹ Este teorema fue redescubierto por Markow (1912) y atribuido a este autor hasta 1938.

mino, a datos demográficos. Grant intenta obtener conclusiones globales aplicables a una población a partir de una muestra: se planteó el problema de estimar la población inglesa en su época, estimó por primera vez tasas de mortalidad por edades y dedujo en frecuencia de nacimientos de hombres y mujeres, entre otros análisis demográficos relevantes.

El tipo de razonamiento de Grant es puramente analítico y desligado completamente del concepto de probabilidad. En la misma línea, Petty, en su «Political Arithmetic», publicado en 1690, analiza datos demográficos así como datos económicos de ingresos, educación y comercio.

Las primeras tablas completas de mortalidad fueron publicadas por Edmund Halley en 1693, que estudió el problema de los seguros de vida. Durante el siglo XVIII se produce un rápido crecimiento, principalmente en Inglaterra, de los seguros de vida y los seguros marítimos y debido en gran parte a la influencia de las ideas de Grant y Petty, se comienzan a realizar los primeros censos oficiales. El primer censo del que tenemos noticias se realiza en Irlanda en 1703. En España, el primer censo se efectúa en 1787 impulsado por el conde de Floridablanca. A comienzos del siglo XIX puede afirmarse que la casi totalidad de los países europeos, recogen información oficial mediante censos de datos demográficos, económicos, climáticos, etc. Paralelamente surgen las agencias oficiales de estadística y en 1834 se crea en Londres la Royal Statistical Society, seguida, en 1839, por la American Statistical Association.

Durante el siglo XVIII y la mayor parte del siglo XIX, la Estadística evoluciona como ciencia separada del cálculo de probabilidades y la teoría de errores. Aunque A. de Moivre y Deparcieux, entre otros, aplican el cálculo de probabilidades a datos demográficos, y Condorcet y Laplace a problemas de aritmética política, existe durante este período escasa comunicación entre ambas disciplinas. Una contribución importante hacia la síntesis de ambas disciplinas es debida a A. Quetelet (1846), que sostuvo la importancia del cálculo de probabilidades para el estudio de datos humanos. Quetelet demostró que la estatura de los reclutas de un reemplazo seguía una distribución normal e introdujo el concepto de «hombre medio». Sin embargo, la diferencia de concepción y de lenguaje entre los matemáticos-astrónomos y los estadísticos-demógrafos dificultó la interacción entre ambos grupos. La unión entre ambas corrientes, se va a producir a comienzos del siglo XX, favorecida en gran parte por los nuevos problemas teóricos y metodológicos que planteaba la contrastación empírica de la teoría de Darwin.

5. LOS PROBLEMAS DE LA EVOLUCION, LA HERENCIA Y LA TEORIA DE LA INFERENCIA ESTADISTICA

La revolución que supuso en la física Newton se produjo en la Biología por la obra de Darwin. Dos facetas importantes de esta teoría eran: a) permitía establecer pre-

dicciones sobre la evolución de poblaciones animales que, en determinadas condiciones, podrían ser contrastadas empíricamente; b) la contrastación debería ser estadística, ya que la unidad que va a sufrir la evolución es la población en su conjunto; los dos mecanismos de la selección natural, producción de variabilidad y selección mediante lucha por la existencia, tienen un atractivo inmediato desde el punto de vista estadístico. La producción de variabilidad mediante el azar entronca con el cálculo de probabilidades y la selección natural con el estudio de poblaciones y con la idea de correlación. Aquellos organismos que estén más adaptados sobrevivirán un mayor período de tiempo y dejarán mayor número de descendientes, por lo que tiene que existir una correlación entre determinadas características genéticas transmisibles y el grado de supervivencia y descendencia de los individuos de una especie.

Weldon escribía en 1900, en el primer editorial de *Biométrica*:

El punto de partida de la teoría de evolución de Darwin es precisamente la existencia de aquellas diferencias entre los individuos miembros de una raza o especie que los morfólogos niegan en su mayor parte. La primera condición necesaria para que comience un proceso de selección natural es la existencia de diferencias entre los miembros de la raza o especie, y el primer paso en una investigación de los posibles efectos de la selección sobre una característica de una raza, debe ser estimar la frecuencia con la que aparecen individuos con cierto grado de anormalidad respecto a esa carácter. La unidad no es el individuo, sino la raza, y el resultado del análisis debe ser una cuantificación que muestre la frecuencia relativa con la que ocurren las distintas clases de individuos que componen la raza.

El primero en resaltar la necesidad de acudir a métodos estadísticos para contrastar la teoría de Darwin fue Francis Galton (1822-1911). Galton, primo de Darwin, fue un hombre de profunda curiosidad intelectual que le llevó a viajar por todo el mundo y a realizar actividades tan diversas como redactar leyes para los hotentotes que gobernaban en el sur de África ¹⁰, o realizar fecundas investigaciones en Meteorología (a él le debemos el término anticiclón). La lectura de la obra de Darwin supuso una transformación radical en la vida de Galton que, casi a los cuarenta años, dedica sus esfuerzos al estudio de la herencia humana. Su trabajo principal es «Natural Inheritance», publicado en 1889 (a la edad de sesenta y siete años), donde desarrolla y amplía muchas de las ideas expuestas en trabajos anteriores. Galton estudió exhaustivamente la distribución normal e introdujo el concepto de línea de regresión comparando, como es bien conocido, las estaturas de padres e hijos. Galton comprendió certeramente la importancia de una medida adimensional de asociación entre variables, es decir, de un coeficiente de correlación, que obtuvo haciendo adimensional el coeficiente de regresión mediante la división de las variables por el recorrido semi-intercuartílico.

¹⁰ Entre Bahía de Walvis y el lago Ngami, K. Pearson (1914-1930) ha realizado una extraordinaria biografía de Galton.

La importancia de Galton radica no solamente en el nuevo enfoque que introduce en el problema de la dependencia Estadística, sino también en su influencia directa sobre Weldon, K. Pearson, R. A. Fisher y Edgeworth, entre otros. El primer departamento ¹¹ de Estadística Matemática, en el sentido actual de la palabra, fue patrocinado por él y llevó su nombre, y la revista *Biométrica* fue posible gracias a su generoso apoyo económico.

El enfoque estadístico propugnado por Galton para el estudio de los problemas de la evolución en «Natural Inheritance» es aceptado entusiásticamente por W. R. F. Weldon, (1860-1906), entonces catedrático de Zoología en la Universidad de Londres. Weldon abandona el camino de los estudios embriológicos y morfológicos como medio de contrastar las hipótesis de Darwin y comienza a investigar en la aplicación de los métodos estadísticos la biología animal. En 1893 (Weldon, 1983) escribe:

Es necesario insistir que el problema de la evolución animal es esencialmente un problema estadístico... debemos conocer *a*) el porcentaje de animales que exhiben un cierto grado de anomalía respecto a un carácter; *b*) el grado de anomalía de otros órganos que acompañan a las anomalías en uno dado; *c*) la diferencia entre la tasa de mortalidad en animales con diferentes grados de anomalía respecto a un órgano, y *d*) la anomalía de los descendientes en términos de la anomalía de los padres y viceversa.

La resolución de estos problemas requiere el desarrollo de métodos estadísticos más avanzados que los existentes y Weldon busca para ello la colaboración de un matemático y filósofo: K. Pearson (1857-1936).

La colaboración de estos dos autores y el apoyo de Galton va a constituir el impulso generador de la corriente de contribuciones que va a fundamentar la Estadística Matemática actual.

Uno de los primeros problemas en la investigación biológica, es encontrar una familia de modelos de distribución de probabilidad lo suficientemente rica para tratar las distribuciones asimétricas y marcadamente no normales que aparecen en estos casos. Para resolver este problema, Pearson construye el sistema de curvas de frecuencia que lleva su nombre, a las que llegó generalizando las propiedades de la normal y estableciendo una ecuación diferencial cuya solución general comprende, como caso particular, las familias de densidad que habían sido tratadas antes aisladamente. Pearson propuso especificar el modelo mediante la estimación de los parámetros por el método de los momentos,

¹¹ En 1906, Galton, que había fundado el Laboratorio de Eugenesia, se lo cedió a Pearson, profesor del University College. En 1911, Pearson une este laboratorio con el suyo de Biometría, para formar el primer departamento de Estadística Matemática, que fue el único existente en el mundo durante casi veinte años. En 1933, al retirarse K. Pearson, este departamento se dividió en dos. El de Eugenesia, cuyo director fue Fisher, y el de Estadística Matemática, dirigido por E. Pearson, hijo de Karl.

y en la búsqueda de un criterio objetivo de educación del modelo introdujo el contraste X^2 , que se reveló como su instrumento enormemente útil en la Estadística Teórica y Aplicada.

El segundo problema típico en el estudio de la herencia es la determinación de las relaciones entre las variables que se han estudiado previamente de forma individual. Pearson recoge la idea de Galton e introduce el coeficiente de correlación productomomento que lleva su nombre; formuló la teoría general de correlación para tres variables y estableció la unión analítica entre los métodos de correlación y regresión. Además de estas aportaciones, que sientan las bases de la metodología actual del estudio estático de dependencias estadísticas. Pearson dedicó una parte importante de sus esfuerzos a la publicación de tablas estadísticas que permitieran la utilización práctica de los nuevos métodos, con lo que contribuyó, decisivamente, a su rápida difusión.

El laboratorio de Pearson se convierte en un polo de atracción para las personas interesadas en el análisis empírico de datos. W. S. Gosset (1876-1937), que trabajaba en la firma cervecera Guinness de Dublín, fue una de las personas que acudieron a Londres a estudiar bajo el patrocinio de Pearson. Gosset se había encontrado en sus investigaciones sobre los efectos de las características de la materia prima en la calidad de cerveza final con el problema de las pequeñas muestras. No era posible económicamente, en este caso, obtener las grandes cantidades de datos que permitirían utilizar los métodos para muestras grandes desarrolladas por Pearson y su escuela. Gosset estudió la distribución de la varianza muestral S^2 en poblaciones normales, demostró que S^2 y la media muestral estaban incorreladas y obtuvo la distribución t , que lleva su nombre ¹². A continuación, Gosset realizó el primer trabajo de investigación estadística mediante el método de Monte Carlo, tomando 750 muestras aleatorias de cuatro elementos de los datos recopilados por W. R. McDonnell sobre la anchura y la longitud del dedo corazón de 3.000 delincuentes ¹³.

La importancia del trabajo de Gosset radica en el impulso que aportó al problema de obtener la distribución de un estadístico en el muestreo. Antes de Gosset, el procedimiento habitual era asociar a un estadístico muestral una medida de su error probable o de su desviación típica. Para determinar la «precisión» de esta medida de dispersión

¹² Gosset publicó sus trabajos bajo el seudónimo de Student, ya que Guinness prohibía entonces, y lo mantuvo hasta después de la segunda guerra mundial, la publicación de trabajos de su personal. Estrictamente hablando, la demostración rigurosa de la distribución del estadístico t fue debida a R. A. Fisher en 1925, ya que Gosset intuyó certeramente, pero no pudo demostrar, la independencia de la media y la varianza muestrales en poblaciones normales.

¹³ La recopilación de estos datos respondía a las teorías de Lombroso, en boga a finales del XIX, que sostenía la existencia de relación entre las características físicas de un individuo y su propensión a la delincuencia.

asociada al estadístico muestral, se requería calcular un tercer estadístico y así sucesivamente. Aunque en la práctica, los estadísticos se contentaban con la medida de dispersión —ya que el proceso podía si no continuar indefinidamente—, la solución era parcial e insatisfactoria. Con la obtención de la distribución exacta del estadístico en el muestreo, este proceso recurrente se cortaba, ya que dicha distribución recoge toda la información relevante sobre la precisión del estadístico. Más importante aún es que esta distribución permite, con un pequeño avance conceptual, el desarrollo de una teoría matemática de estimación por intervalos y de contraste de hipótesis. Gosset obtuvo también la distribución exacta del coeficiente de correlación muestral, cuando el coeficiente poblacional es cero; Student (1908) introdujo el estudio de observaciones dependientes y contribuyó, en su amistosa colaboración con Fisher, al desarrollo de la teoría de diseños de experimentos.

Las contribuciones de Galton, Weldon, Gosset y K. Pearson, conjuntamente con las de Edgeworth, que expondremos en la sección siguiente, configuran el primer período de consolidación de la Estadística Matemática entre 1890 y 1915. El segundo período, entre 1915 y 1925, está dominado por la figura de R. A. Fisher (1890-1962).

Fisher se interesó primeramente por la eugenesia, lo que le conduce, siguiendo los pasos de Galton, a la investigación Estadística. En su primer trabajo estadístico, Fisher (1915) aborda directamente el problema de estimación y sugiere el método de máxima verosimilitud como alternativa al método de los momentos de Pearson. En sus trabajos posteriores, Fisher completa la teoría de correlación generalizando los resultados de Student sobre la distribución exacta del coeficiente de correlación en el muestreo, y obteniendo la distribución del coeficiente de correlación intraclases, del coeficiente de regresión, del coeficiente de correlación parcial y del coeficiente de correlación múltiple.

Paralelamente, Fisher desarrolla una teoría general de la estimación estadística, introduce en 1920 el concepto de estadístico suficiente y demuestra, en 1925, la eficiencia asintótica del estimador máximo-verosímil¹⁴. Estos trabajos culminan con la publicación de *Statistical Methods for Research Workers*, con el que se cierra conjuntamente, en mi opinión, la fase de fundamentación de la Estadística Matemática y el período en el cual los problemas de la evolución y la herencia va a ser el impulso motor de los desarrollos teóricos.

En este libro aparece ya claramente el cuerpo metodológico básico que constituye la Estadística Matemática actual: el problema de identificar el modelo a partir de datos

¹⁴ Pratt (1976) ha estudiado cómo Edgeworth se anticipó a Fisher en la intuición de la eficiencia del estimador máximo verosímil.

empíricos ¹⁵, la deducción matemática de las propiedades del mismo (cálculo de probabilidades), la estimación de los parámetros condicionados a la bondad del modelo y la validación final del mismo mediante test de significación. Es indudable, como comentaremos en la sección siguiente, que en el más de medio siglo transcurrido desde la primera edición del libro de Fisher, la Estadística Matemática ha avanzado considerablemente y también es cierto que el rechazo de Fisher del principio de la probabilidad inversa y su énfasis en la probabilidad fiducial no han sido fructíferos. Sin embargo, es innegable la contribución de Fisher para crear el armazón lógico conceptual básico de la Estadística Matemática.

6. LA ESTADÍSTICA MATEMÁTICA

El hecho de que los problemas de la evolución y la herencia sirvan de vehículo impulsor de las aportaciones teóricas en el período 1890-1925 es importante para comprender la unificación de la Estadística Matemática, ya que estos problemas tienen una proximidad sustantiva directa con el análisis de poblaciones humanas y de datos económicos.

Una personalidad importante en este proceso de acercamiento es F. Y. Edgeworth (1845-1926), más conocido por sus trabajos como economista que como estadístico, y cuyos primeros trabajos se orientan hacia la aplicación de las Matemáticas a las ciencias sociales.

Las corrientes de críticas que despertó su libro «*Mathematical Physics. An essay in the application of Mathematics to the moral Science*», publicado en 1881, movió a Galton a escribirle manifestándole su completo apoyo. La amistad consiguiente entre ambos científicos influyó decisivamente en orientar el trabajo de Edgeworth hacia la Estadística. Este percibió rápidamente que los métodos estadísticos preconizados por Galton, para el estudio de la evolución y de la herencia genética, podían ser igualmente útiles para la investigación social y económica, y entre 1883 y 1910 realiza destacadas contribuciones a la Estadística Matemática. Entre otros trabajos, Edgeworth realizó un análisis general de la normal multivariante (Seal, 1976), contribuyó con Pearson al estudio de curvas de frecuencia (Pearson, E. S., 1967), introdujo el primer test de normalidad comparando la media y la mediana muestral, realizó el primer análisis de la varianza de una tabla de clasificación cruzada (Stigler, 1978) y, entre otros trabajos de Estadística Económica, estudió las propiedades de los números índices y la utilización de la regresión para eliminar la tendencia de series temporales.

¹⁵ Este problema ha sido durante un amplio período considerado como muy secundario, hasta que, como ha señalado Box (1979), recientemente ha vuelto a alcanzar un tono de respetabilidad al ser rebautizado como «Técnicas de análisis de datos».

El trabajo de Edgeworth contribuyó en gran parte a la difusión de los métodos estadísticos entre los investigadores sociales y a realizar la ósmosis entre la estadística tradicional del XIX con los nuevos métodos de inferencia estadística.

Una medida del acercamiento entre estas áreas es la introducción de los conceptos de inferencia estadística en las investigaciones sociológicas y económicas oficiales. Es sintomático que entre 1890 y 1925 se produce la transición entre los estudios socioeconómicos que tratan de cubrir toda la población a los estudios sistemáticos mediante muestreo [Kruskal y Mosteller (1980)]. La primera introducción del término «muestreo representativo» con relación a estas investigaciones ocurre en 1895, en la reunión del International Statistical Institute y los conceptos probabilísticos se van introduciendo lentamente en las reuniones de Instituto en los años siguientes, hasta que en la reunión de Roma, en 1925, los conceptos de muestreo han alcanzado general aceptación.

A partir de los años veinte, la evolución de la Estadística va a estar marcada por tres factores principales: *a*) una enorme expansión de las aplicaciones de las técnicas ya conocidas; *b*) la invención de nuevas metodologías de análisis; *c*) una revisión conceptual de sus fundamentos teóricos. Estos tres factores están además, como veremos, estrechamente relacionados.

El primer factor es la aplicación de los métodos estadísticos en áreas tan diversas como la Ingeniería (Control de Calidad por Shewart, estudio de procesos y flitros), la Física (teoría cinética de los gases), la antropología, la psicología o la Medicina. Fisher (1953) ha expuesto claramente este proceso de expansión. La búsqueda de respuestas a los nuevos interrogantes planteados por las diversas ciencias impulsan a su vez el desarrollo de nuevos métodos estadísticos. Los problemas en Agronomía conducen a Fisher a crear la teoría de diseños experimentales y un problema de discriminación en antropología (concretamente el estudio de cráneos, véase J. Box, 1978), lleva a Fisher a inventar el análisis discriminante y la psicología y la antropología impulsan el desarrollo de los métodos multivariantes. Los problemas de ingeniería conducen a un estudio sistemático de la teoría de procesos estocásticos, y las necesidades en el de control de procesos ¹⁶ sugieren a E. S. Pearson la consideración de hipótesis, conjuntamente con Neyman.

Es probablemente una constante en la historia de cualquier disciplina científica que los períodos de grandes avances conceptuales y metodológicos producen paralelamente una revisión de los fundamentos conceptuales de la disciplina y un intento de estable-

¹⁶ Según E. S. Pearson (1967), los trabajos de Gosset le sugirieron la idea de considerar hipótesis alternativas en los tests de significación clásicos y el término potencia de un test provino de la lectura de un trabajo de Shewart, pionero en el desarrollo de métodos de control de calidad en la industria.

cerla sobre nuevas bases más firmes. En la Estadística este proceso se produce en dos direcciones: la concepción de la probabilidad y las bases de la inferencia Estadística, problemas ambos estrechamente relacionados entre sí.

El paradigma frecuentista se asienta en la fundamentación del concepto de probabilidad por Von Mises en 1928 y tiene su desarrollo natural en la teoría que producen Neyman y E. Pearson entre 1925 y 1940. El paradigma bayesiano parte de la concepción de la probabilidad como grado de creencia racional (Keynes y Jeffreys) o subjetivo (Ramsey y De Finetti) y tiene su continuación natural en la teoría de inferencia bayesiana creada por Jeffreys (1931-1939).

Finalmente, las propiedades matemáticas de la probabilidad y la fundamentación rigurosa de la teoría de los procesos estocásticos es debida en gran parte a Kolmogoroff.

Es claro que tanto la concepción frecuentista como la bayesiana tienen sus raíces en los dos grandes periodos que hemos considerado. Jeffreys es un físico que recoge la tradición del enfoque de Laplace y Gauss y las aplicaciones de la Estadística a la Física y la Astronomía, que utilizan el enfoque frecuentista. Fisher, aunque se opuso al principio de la probabilidad inversa, con su énfasis en el estudio de la verosimilitud y con su concepto de probabilidad fiducial se encuentra de hecho más próximo al enfoque bayesiano que al clásico ¹⁷.

7. CONCLUSIONES

En las páginas anteriores se ha intentado esbozar cómo las demandas de otras disciplinas científicas juegan un papel central en el nacimiento de la Estadística Matemática y su expansión posterior. Creemos que esta visión histórica ayuda más a la comprensión de la ciencia Estadística que las interpretaciones que ponen el énfasis en la continuidad del desarrollo de los procedimientos matemáticos. Como ha señalado Box (1978), ilustrándolo con la vida de Fisher, la interacción fructífera entre la teoría y la práctica está en la raíz de los descubrimientos importantes en esta ciencia. Un argumento adicional especialmente relevante es que los científicos sobresalientes en el período que hemos estudiado han sido, prácticamente siempre, personas interesadas en una gama muy amplia de problemas y con un espectro marcadamente multidisciplinario en su aproximación a la realidad. Laplace, Gauss, Galton, Edgeworth, Pearson, Gosset y Fisher realizaron apor-

¹⁷ La probabilidad fiducial ha sido descrita por Savage (1961) como «un intento de hacer una tortilla bayesiana sin romper los huevos bayesianos». Es de señalar que los trabajos de Fisher en su último período convirtieron a muchos estadísticos al enfoque bayesiano. Véase, por ejemplo, Geisser (1980).

taciones no sólo a la estadística, sino sobre todo a la ciencia en el sentido más amplio de la palabra. Su vida y obra representan un ejemplo que merece ser seguido.

BIBLIOGRAFIA

- BARNARD, G. A.: «The Unity of Statistics». *JRSS, A.*, 135, 1, pp. 1-15, 1972.
- BAYES, T.: «An essay towards solving a problem in the doctrine of chances» (1764). Reproducido en *Biometrika*, 45, pp. 293-315, 1958.
- BOX, G. E. P.: «Science and Statistics», *JASA*, 71, 356, pp. 791-799, 1976.
- : «Some problems of Statistics and Every day life». *JASA*, 74, 365, pp. 1-4, 1979.
- BOX, J. F.: *R. A. Fisher. The life of a scientist*. Wiley, 1978.
- BRAVAIS, A.: «Sur les probabilités des erreurs de situation d'un point». *Med. Acad. Roy. Sci. Inst. France*, 9, pp. 255-332, 1846.
- BERNOULLI, D.: «The most probable choince between several discrepant observations and the most formation there from of the most likely induction» (1777). Reproducido en *Biometrika*, 48, páginas 1-18, 1961.
- CORNFIELD, J.: «A Statistician's Apology». *JASA*, 70, 349, pp. 7-14, 1975.
- DAVID, F. N.: «Dicing and gaming». *Biometrika*, 42, pp. 1-15, 1955.
- : *Games, Gods and Gambling: A History of Probability and Statistical Ideas*, Charles Griffin, 1962.
- EISENHART, C.: «Carl Friederich Gauss», en *Int. Encycl. of Soc. Science*, Macmillan, 1968.
- FISHER, R. A.: «Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population». *Biometrika*, 10, pp. 507-521, 1915.
- : «The expansion of Statistics». *J. R. Statist. Soc.*, A, 116, pp. 1-10, 1953.
- GEISSER, S.: «The contributions of Sir Harold Jeffreys to Bayesian Inferences». En *Bayesian Analysis in Econometrics and Statistics*, A. Zellner editor, Nort-Holland, 1980.
- HASOFER, A. M.: «Random Mechanisms in Talmudic literature». *Biometrika*, 54, pp. 316-21, 1967.
- HUBER, P. J.: «Robust Estimation of a location Parameter». *Annals of Mathematical Statistics*, 43, 4, pp. 1041-1067, 1964.
- JEFREYS, H.: *Scientific Inference*. Cambridge University Press, 1931.
- : *Theory of Probability*. Oxford University Press, 1939.
- KENDALL, M. G.: «The beginnings of a probability calculus». *Biometrika*, 43, pp. 1-14, 1956.
- : «Francis Ysidro Edgeworth, 1845-1926». *Biometrika*, 5, 1968.
- KRUSKAL, W., y MORTELLER, F.: «Representative Sampling, IV: The History of the concept in Statistics, 1895-1939». *Int. Stat. Reviews*, 48, pp. 169-195, 1980.
- KUHN, T.: *The Structure of Scientific Revolutions*. The University of Chicago Press, 1962.
- MAHALANOBIS, P. C.: «The foundations of Statistics». *Sankhya*, 18, pp. 183-194, 1957.

- MOOD, A. M., y GRAYBILL, F. A.: *Introducción a la Teoría de la Estadística*. Aguilar, Madrid, 1972.
- PEARSON, E. S., y KENDALL, M. G.: *Studies in the History of Statistics and Probability*. Charles Griffin, 1970.
- PEARSON, K.: *The Life, Letters and Labours of Francis Galton*, 3 vols. Cambridge University Press, 1914-1930.
- PLACKETT, R. L.: «The principle of the arithmetic mean». *Biometrika*, 45, pp. 130-135, 1958.
- PIERCE, B.: «Criterion for the Rejection of Doubtful Observations». *Astronomical Journal*, 2, pp. 161-163, 1852.
- PIAGET, J.: *Etudes sur la logique de l'enfant*. Delachaux y Niestlé, 1968.
- PRATT, J. W.: «F. Y. Edgeworth and R. A. Fisher on the efficiency of Maximum likelihood estimation». *The Annals of Statistics*, 4, 3, pp. 501-514, 1976.
- QUETELET, A.: *Theory of probability*, 1846.
- SEAL, H. L.: «The historical development of the gauss linear model». *Biometrika*, 54, pp. 1-24, 1967.
- : «Abraham de Moivre», en *Int. En cycl. of Soc. Sciences*. Macmillan, 1968.
- STIGLER, S. M.: «Simon Newcomb, Percy Daniell, and the History of Robust Estimation, 1885-1920». *JASA*, 68, 344, pp. 872-879, 1973a.
- : «Laplace, Fisher and the discovery of the concept of sufficiency». *Biometrika*, 60, 3, páginas 439-445, 1973b.
- : «Napoleonic Statistics: The work of Laplace». *Biometrika*, 62, 3, pp. 503-517, 1975.
- : «Francis Ysidro Edgeworth, Statistician». *Technical Report*, 510. University Wisconsin. Madison, 1978.
- STUDENT, X.: «The probable error of a mean». *Biometrika*, 6-1, 1908.
- THATCHER, A. R.: «A note on the early solutions of the problem of the duration of play». *Biometrika*, 44, pp. 515-518, 1957.
- TODHUNTER, I.: *History of the Mathematical Theory of Probability*, 3.^a ed., Chelsea, 1965.
- WELCH, B. L.: «Student and small sample theory». *JASA*, 53, 284, pp. 777-778, 1958.
- WELDON, W. F. R.: «On certain correlated variations in *carcinus moenas*». *Proc. Roy. Soc.*, 54, pp. 318-329, 1983.

SUMMARY

This work analyses the influence of two scientific paradigms in the evolution and consolidation of Mathematical Statistics. The contrast of Newton's theory is the unleashing of the theory of errors. The contrast of Darwin's theory gives rise to Statistical Inference. A study of these histo-

ric periods brings out the importance of the interaction between theory and practice in the creation of Mathematical Statistics and in its later development.

Key words: Probability calculus, Newton and the theory of errors, statistics in the XIX century, problems of inheritance and the theory of statistical inference, mathematical statistics.

AMS, 1970. Subject classification: 62-03.

