Inducción, deducción y decisión en las teorías estadísticas de la inferencia científica

ANDRÉS RIVADULLA
(Universidad Complutense)

1. Introducción

Después de un placentero viaje en el AVE procedente de Madrid, Andy R. Roderick llega a Sevilla, donde ha sido invitado por unos colegas a pasar unos días para visitar la exposición universal. Tras un breve descanso en el hotel, decide dar un paseo para gozar del atardecer perfumado de la ciudad hispalense. Mientras camina ensimismado por el barrio de Santa Cruz, casi tropieza con una cabina telefónica y piensa que sería una ocasión excelente para llamar a casa. Así que saca de su monedero diez piezas de cien pesetas, pero con tan mala fortuna que se le caen al suelo. La forma como lo hacen es la siguiente: ocho muestran 'cara' y dos muestran 'cruz'. Esto llama poderosamente su atención. De modo que, de vuelta en la intimidad de su habitación, intenta explicarse lo observado. He aquí el curso de sus reflexiones.

La cuestión acerca de cómo extraer conocimiento válido del mundo ha preocupado a los teóricos del conocimiento desde el origen mismo de la filosofía occidental. La respuesta de Aristóteles según la cual la *inducción enumerativa* garantiza la inferencia de verdades empíricas universales a partir de un número finito de observaciones había sido aceptada ampliamente. Y aunque algunos filósofos como Hume[1], en la primera mitad del XVIII, y Popper, en pleno siglo XX, mantuvieron la ilegitimidad lógica de la inferencia ampliativa y conservadora de la verdad, e. d. de la *inducción*, la atribución de valores de probabilidad a las conclusiones obtenidas de observaciones metódicas transformó la cuestión de la validez de la inducción en la de la posibilidad de la probabilidad inductiva. Las raíces de este cambio hay que buscarlas, empero, en la historia del cálculo de probabilidades, no en la de la filosofía.

En efecto, en la segunda mitad del siglo XVIII el clérigo inglés Thomas Bayes publicó con carácter póstumo un artículo[2] que por primera vez permitió un tratamiento matemático de la inferencia inductiva[3]. Con lo que a partir de

1. HUME, D.: *A Treatise of Human Nature*, 1739, libro I, parte III, y *An Enquiry Concerning Human Understanding*, 1748, §§ 2-7.

2. BAYES, T.: «An Essay Towards solving a Problem in the Doctrine of Chances». *Philosophical Transactions of the Royal Society of London*, 1763.

3. Aunque cincuenta años antes el matemático suizo Jakob Bernoulli –en la parte cuarta de su *Ars Conjectandi*– había intentado, ante el prudente escepticismo de Leibniz, inferir inductivamente probabilidades a partir de frecuencias observadas, su empresa no fue acompañada por el éxito. Por lo que corresponde a Thomas Bayes el honor de haber inaugurado en la teoría de probabilidades el uso propiamente inverso de la probabilidad matemática. Véase al respecto RIVADULLA, A.: *Probabilidad e Inferencia Científica*, 1991, pp. 124 y ss.

entonces el problema de la inducción dejó de ser tema exclusivo de debate en el seno de la filosofía para pasar a convertirse también en objeto de investigación matemática. Al mismo tiempo, la aplicación del cálculo de probabilidades al problema de la obtención de inferencias inciertas se convirtió en una cuestión intensamente debatida tanto entre los teóricos de la estadística matemática como entre los filósofos de la ciencia.

La pregunta acerca de cómo la probabilidad pudo despertar el interés de los matemáticos interesados en la obtención de inferencias exactas a partir de las observaciones no es difícil de contestar. Es bien conocido que, desde Pascal y Fermat, a mediados del XVII, la probabilidad se aplicaba a fin de calcular con exactitud la incertidumbre de los resultados en los juegos de azar[4]. Algunas cuestiones planteadas a Pascal por el caballero de Méré promovieron en 1654 una correspondencia entre Pascal y Fermat, publicada parcialmente en *Varia Opera Mathematica D. Petri de Fermat*, en Toulouse, 1679, y en las obras de Pascal[5]. Pero el primer libro sobre probabilidades fue *Liber de Ludo Alae*, de Gerolamo Cardano, posiblemente escrito en 1564, aunque inédito hasta 1663, cuando fue publicado en *Opera Omnia*, de Cardano. La primera obra publicada en teoría de probabilidades fue, pues, *De Ratiociniis in Ludo Alae*, publicada por Christiaan Huyghens en 1657. Con el ya mencionado *Ars Conjectandi*, de Jakob Bernoulli, donde por vez primera se formula y demuestra la *Ley* (débil) *de los grandes números*, Basilea, 1713, y la *Doctrine of Chances*, de Abraham de Moivre, Londres, 1718, se cierra el proceso de consolidación del cálculo de probabilidades como un campo matemático autónomo.

La relevancia de la probabilidad para los juegos de azar es obvia: dadas las reglas de los juegos de cartas y dados, la probabilidad de determinado resultado se sigue *directa o deductivamente*. Naturalmente el resultado en cuestión no se puede predecir con certeza, pero la probabilidad de su ocurrencia se puede calcular fácilmente. La probabilidad mide, pues, de modo preciso la incertidumbre de todos los resultados lógicamente posibles en los juegos de azar. Tras el fallido intento de Jakob Bernoulli, Thomas Bayes primero, y su eminente valedor, Pierre Simon Laplace, once años después[6], consideraron que la probabilidad podía ser aplicada también en procedimientos *inversos o inductivos*, e. d. en casos donde la probabilidad de ocurrencia del fenómeno observado es desconocida, a fin de estimarla a partir de los resul-

4. En relación a la aplicación –en la que se puede situar el origen de la estadística matemática– de la probabilidad matemática a dominios empíricos *distintos* de los juegos de azar, véase RIVADULLA, A.: «The Roots of Theoretical Statistics», 1990, y «Apriorismo y Base Empírica en los Orígenes de la Estadística Matemática», 1991.

5. Cfr. TODHUNTER, I.: *A History of Mathematical Theory of Probability*, 1865, pp. 7-8. Más recientemente esta correspondencia ha sido compilada por ABOUT, P.-J., y BOY, M., en *La correspondance de Blaise Pascal et de Pierre de Fermat*, 1983. En alemán existe también una versión de la misma en SCHNEIDER, I. (Hrsg.): *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933*, 1988.

6. LAPLACE, P. S.: «Mémoire sur la probabilité des causes par les événements». *Mém. Acad. Roy. Sci. Paris*, 1774. Reimpreso en LAPLACE: *Oeuvres*, vol. 8, 1891.

tados observados[7]. Al decir de Ronald Fisher, padre de la estadística objetivista contemporánea, tal posibilidad de un uso inverso de la probabilidad matemática cautivó mentes del calibre de un Gauss o de un Poisson. Posibilidad rechazada por estadísticos frecuentistas tales como el propio Fisher, Jerzy Neyman, Egon S. Pearson, etc., y por filósofos antiinductivistas como Karl Popper.

2. El *teorema de Bayes* y la reacción de Fisher contra el cálculo de la *probabilidad inversa*

El asunto que ocupa a Bayes en su *Essay* de 1763 es el del cálculo de la *probabilidad (chance)* de que la probabilidad de cierto suceso esté situada entre dos valores. Su solución proporciona, según Laplace[8], la respuesta a la pregunta acerca de las probabilidades de las causas y de los sucesos futuros en relación a la evidencia disponible. O, como dice Fisher[9]: «To Thomas Bayes must be given the credit of broaching the problem of using the concepts of mathematical probability in discussing problems of inductive inference, in which we argue from the particular to the general; or, in statistical phraseology, argue from the sample to the population, from which, *ex hypothesi*, the sample was drawn. Bayes put forward, with considerable caution, a method by which such problems could be reduced to the form of problems of probability.» Si bien, puntualiza Fisher, el procedimiento de Bayes se apoyaba en un error, que no pudo ser descubierto en su tiempo, ya que para los matemáticos de la época, convencidos de la naturaleza inductiva del método científico, no cabía la menor duda de que toda inferencia científica había de ser formulada en términos de probabilidad.

La reconstrucción por Fisher del teorema de Bayes[10] muestra que éste no es sino una consecuencia lógica de postular un conocimiento *a priori* acerca de una superpoblación imaginaria de la que la población investigada ha sido extraída aleatoriamente. Con lo que las supuestas inferencias de la muestra a la población correspondiente no constituyen sino consecuencias probabilísticas perfectamente deductivas. En efecto, si con P designamos la *probabilidad (chance)* de que la probabilidad p de conseguir a 'éxitos' (p. e., a veces 'cara')

7. Respecto a la relación entre Bayes y Laplace en el contexto del problema de la probabilidad inversa puede leerse GOMEZ-VILLEGAS, M. A.: «El Problema de la Probabilidad Inversa: Bayes y Laplace». *Actas del Encuentro de Lógica y Filosofía de la Ciencia. Rudolf Carnap y Hans Reichenbach in Memoriam. Madrid, 1991*. Así como RIVADULLA, A.: *Probabilidad e Inferencia Científica*, 1991, p. 126.

8. LAPLACE, P. S.: *Théorie Analytique des Probabilités*, 1820. Introduction, p. 148.

9. Cfr. FISHER, R.: «Two New Properties of Mathematical Likelihood». *Proceedings of the Royal Statistical Society of London*, 1934, p. 285.

10. Para una presentación detallada de la misma véase RIVADULLA, A.: *Probabilidad e Inferencia Científica*, pp. 130-133.

en *n* ensayos independiente se sitúe entre dos valores probabilísticos *u* y *v* previamente dados, entonces

(1) $$P = \frac{(n+1)!}{a!\,(n-a)!} \int_u^v p^a\,(1-p)^{n-a}dp,$$

caso de que conozcamos *a priori* que *p* está distribuida uniformemente[11] en la superpoblación imaginaria, de la que nuestra población de 'éxitos' y 'fracasos' ('caras' y 'cruces', 'unos' y 'ceros', etc.) ha sido extraída al azar.

Ahora bien, (1), es decir, el *teorema de Bayes*, es una consecuencia lógica del postulado conocimiento *a priori*, cuando éste actúa como superpremisa de nuestro razonamiento deductivo. Como la asunción necesaria de conocimiento *a priori* transforma la inferencia en una deducción lógica, el procedimiento de Bayes parece que no proporciona, contra la intención original, ningún medio para la estimación *inversa* de las causas o razones de los fenómenos observados. Así que Fisher lo excluye como método adecuado para aprender inductivamente de la experiencia.

Tal decisión obliga a Ronald Fisher a buscar formas alternativas de inferencia de lo particular a lo general, e. d. de las muestras aleatorias a las poblaciones de las que proceden. A este respecto Fisher desarrolla dos procedimientos ampliamente aceptados: el *método de máxima verosimilitud* para la estimación de los valores de los parámetros poblacionales, y los *test de significación de hipótesis estadísticas*. Ambos son concebidos por él como formas de razonamiento *inverso* o *inductivo*, a fin de extraer conclusiones válidas sobre la población a partir de muestras aleatorias de la misma.

Para Ronald Fisher la *verosimilitud* de que cierto parámetro poblacional posea un valor determinado es una cantidad proporcional a la probabilidad de que los datos observados sean los actualmente disponibles, si el valor paramétrico conjeturado fuese verdadero[12]. Por consiguiente, el valor cuya verosimilitud es máxima será la mejor estimación del parámetro poblacional. Tanto la *verosimilitud matemática* como la *probabilidad matemática* son medidas de *creencia racional*, si bien la primera no satisface el axioma de adición del cálculo de probabilidades[13]. Existen, pues, según Fisher[14] dos medidas diferentes

---

11. Para una explicación de los conceptos de 'ensayos independientes' y 'distribución uniforme' puede consultarse RIVADULLA, A.: *Probabilidad e Inferencia Científica*, pp. 130-132.

12. Cfr. FISHER, R.: «On the Mathematical Foundations of Theoretical Statistics», *Phil. Trans. R. Soc. of London*, 1922, p. 310 *et passim*.

13. Para una presentación accesible del cálculo de probabilidades puede consultarse RIVADULLA, A.: *Filosofía Actual de la Ciencia*, 1986, cap. 0. «Cálculo Básico de Probabilidades», así como «Cálculo axiomático de la probabilidad lógica», *Theoria*, 1992, volumen conmemorativo del XL aniversario.

14. Cfr. FISHER, R.: «Inverse Probability», *Proc. Cambr. Phil. Soc.*, 26, 1930, p. 532.

de creencia racional propias para cada caso: «Knowing the population we can express our incomplete knowledge of, or expectation of, the sample in terms of probability; knowing the sample we can express our incomplete knowledge of the population in terms of likelihood.» La probabilidad se aplica, pues, en los casos de inferencias deductivas o directas de la población a sus propias muestras aleatorias, mientras que el *método de máxima verosimilitud* permite extraer inferencias inversas o inductivas (aunque no probabilísticas) de las muestras a la población de la que han sido extraídas al azar.

Por su parte, los *test de significación* tienen como objeto valorar las desviaciones de las observaciones respecto de la *hipótesis testada*, que metodológicamente se considera verdadera. El concepto clave aquí es el de *resultado estadísticamente significativo*. Para explicarlo Fisher[15] recurre al siguiente experimento mental: Una señora afirma ser capaz de discernir, tras probar una taza de té con leche, si la leche se vertió sobre el té en la taza o el té sobre la leche. A fin de comprobar la supuesta capacidad de la señora se hace necesario diseñar un experimento en el que la probabilidad de acertar por casualidad sea muy pequeña. A tal efecto podríamos proponerle que distribuyera ocho tazas de té con leche en dos grupos de cuatro, según el ingrediente que fue vertido sobre el otro. En este experimento la probabilidad de acertar por casualidad en una única ejecución del experimento, es decir, tras probar cada una de las ocho tazas una sola vez, es ciertamente pequeña, pues al haber 70 combinaciones posibles de las ocho tazas en grupos de cuatro, la proporción de éxito es sólo del 1,4%. Supongamos, pues, que nuestra hipótesis testada es la de que la señora carece de la habilidad de que presume. Un resultado exitoso del experimento debería ser considerado estadísticamente significativo, y refutaría la hipótesis nula al nivel de significación 1,4%. Un resultado harto improbable desde el punto de vista de la hipótesis testada es, pues, *estadísticamente significativo* si se considera que su ocurrencia no puede ser debida al azar; con lo que es tanto más significativo cuanto menor es su probabilidad de acontecer. Un test de significación clasifica, previamente a la realización del experimento aleatorio, los resultados posibles del mismo en dos grupos disjuntos: el de aquellos que se estipula que se desvían significativamente de la hipótesis testada, y el de los que se conviene que concuerdan con ella. El razonamiento que subyace a la aplicación de todo test de significación es obvio: *si, en una única ejecución del experimento, se produce un resultado cuya frecuencia de aparición a la larga es muy pequeña, dada la hipótesis testada, rechazaremos la verdad de ésta.*

Fisher impone dos restricciones a las *decisiones* posibles a tomar –*aceptación* o *rechazo* de la hipótesis testada– en la aplicación de un test de significación. La primera es que *rechazo* no equivale a *refutación lógica*, ya que incluso un fenómeno muy improbable puede acontecer por casualidad, refutando

15. Cfr. FISHER, R.: *The Design of Experiments*, 1935, p. 11.

así erróneamente una hipótesis verdadera. La segunda consiste, complementariamente, en que *aceptación* no significa *verificación concluyente*. Ambas decisiones son, pues, siempre provisionales, y pueden ser corregidas en posteriores ejecuciones del experimento. Y, en todo caso, son relativas a un *nivel de significación* dado, e. d. a lo que, de entrada, consideramos harto improbable que pueda ocurrir por casualidad.

3. **El Programa Neyman-Pearson de inferencia estadística**

Según Jerzy Neyman la estadística matemática tiene como objeto el establecimiento de reglas de comportamiento, e. d. reglas que permiten seleccionar acciones previamente fijadas de acuerdo con los resultados observados de experimentos aleatorios. Un test estadístico es, pues, una regla de este tipo, ya que permite elegir entre aceptar o rechazar una hipótesis testada; y también lo es la decisión de que el valor verdadero de un parámetro poblacional, estimado por los procedimientos de la teoría de la estimación estadística confidencial, se encuentra situado entre dos valores calculados.

La teoría Neyman-Pearson del test de hipótesis estadísticas completa la fisheriana de los test de significación introduciendo la referencia a hipótesis alternativas a la testada. Así, un test estadístico debe comenzar con el establecimiento de la hipótesis testada –o *hipótesis nula*– H_0 y de la *hipótesis alternativa* H_1, que comporta que la aceptación de una de ellas supone el rechazo de la otra.

Los resultados experimentales que se desvían significativamente de la hipótesis testada constituyen la *región crítica* o *región de rechazo* de la hipótesis nula, mientras que el resto forman su *región de aceptación*. Para Neyman[16], pues, establecer el test de una hipótesis estadística equivale a fijar la región crítica correspondiente. Así, si X es una variable aleatoria binomial[17], la *fórmula de Bernoulli* nos permite calcular la probabilidad de x veces 'éxito' en n ensayos independientes, bajo el supuesto de que la probabilidad constante p de 'éxito' sea conocida:

(2) $$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

16. Cfr. NEYMAN, J.: *First Course in Probability and Statistics*, 1950, p. 265.

17. Este es el caso cuando una variable X sólo tiene dos valores posibles: 'éxito' o 'fracaso', 'cara' o 'cruz', etc. y los resultados experimentales son mutuamente independientes en el sentido de que ninguno de ellos influye sobre los demás.

Esta es la fórmula que debemos aplicar si deseamos conocer la probabilidad de obtener *a* veces 'cara' en *n* lanzamientos de una moneda supuestamente insesgada, o, lo que es lo mismo, en un lanzamiento de *n* monedas supuestamente leales.

Sea la hipótesis nula $H_0 : p = p_0$. En *n* observaciones independientes de la variable aleatoria binomial *X* hay 2^n resultados lógicamente posibles, cada uno de los cuales constituye un elemento del denominado *espacio muestral*, cuya probabilidad se calcula por medio de (2). Si nos fijamos en los valores de probabilidad más pequeños correspondientes a determinados elementos del espacio muestral y los sumamos, obtendremos el valor α que denota la probabilidad de observar determinados resultados de la región de rechazo de la hipótesis testada. α designa, pues, el *nivel de significación* del test, y al mismo tiempo la probabilidad (frecuencia a la larga) de cometer un *error del primer tipo*, e. d. la proporción de casos en que H_0 será rechazada erróneamente a la larga. En efecto, como un suceso harto improbable puede ocurrir por casualidad, y como ignoramos al mismo tiempo si H_0 es verdadera, nos veremos obligados a rechazarla en favor de su alternativa H_1. Por otra parte, cuando ésta es la verdadera puede suceder que acontezca un resultado favorable a la hipótesis testada; en tal caso rechazaremos erróneamente la hipótesis alternativa verdadera en favor de la hipótesis nula, cometiendo así un *error del segundo tipo*. La probabilidad (frecuencia a la larga) de incurrir en este tipo de error se designa por medio de ß y sólo se puede calcular tras fijar el valor de p_1 en $H_1 : p = p_1$. Una vez que las probabilidades de error α y ß han sido fijadas, las probabilidades complementarias 1 - α y 1 - ß expresan respectivamente las frecuencias a la larga de aceptación correcta de las hipótesis nula y alternativa. 1 - α y 1 - ß se denominan respectivamente *coeficiente de confianza* y *poder del test*. Un test potente es aquel cuya probabilidad de error ß es pequeña.

Por lo que respecta al otro sistema de reglas de comportamiento: la teoría de la estimación estadística del valor verdadero de parámetros poblacionales por medio de intervalos de confianza, el procedimiento más apropiado sugerido por Neyman[18] es el del cómputo de los valores límites L_1 y L_2 del valor verdadero. Estos valores, denominados *estimaciones superior e inferior* del valor paramétrico, constituyen un intervalo de confianza en el que, con una frecuencia determinada a la larga, estará situado el valor buscado. Así, si designamos el valor verdadero desconocido por medio de *T*, la teoría estadística de la estimación confidencial tiene como objetivo formular enunciados del tipo

(3) $$P(L_1 \leq T \leq L_2) = 1-\alpha,$$

18. NEYMAN, J.: «Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability», 1937.

donde 1 - α constituye, al igual que en el caso del test de hipótesis, un coeficiente de confianza. La interpretación que Neyman[19] ofrece de (3) es la de que el coeficiente de confianza expresa el porcentaje de casos en que el estadístico tendrá razón, cuando afirma que el parámetro estimado se encuentra dentro del intervalo de confianza calculado.

4. **Inducción probabilística e inducción estadística**

Según vimos en la sección 2, el método de Bayes-Laplace fracasó, en opinión de Ronald Fisher, en su empeño por mostrar, en términos de probabilidades matemáticas, el carácter inductivo de las inferencias estadísticas. Esto llevó a Fisher a desarrollar procedimientos matemáticos no probabilísticos que capacitan para *aprender inductivamente* de la experiencia. La lógica de la inferencia estadística devino, pues, la lógica inductiva por excelencia. Con ello Fisher continuó vinculado a una tradición dominante en la filosofía de la ciencia que considera a la empresa científica gobernada por el método inductivo. Esta tradición podría retrotraerse a Francis Bacon, cuyo *Novum Organum* fue publicado en 1620 y, con matices muy importantes, contaría entre sus máximos representantes a John Stuart Mill, Charles Peirce, William Whewell, el positivismo lógico, Bertrand Russell, Rudolf Carnap, Jaakko Hintikka, etc.

En tiempos de Laplace, pues, cuya *Théorie analytique des probabilités* fue publicada por vez primera en 1812, el pensamiento científico estaba dominado por la idea de que la ciencia aplica el método inductivo a fin de alcanzar conocimiento verdadero del mundo. Por contra, la afirmación fisheriana de que la moderna estadística matemática proporciona los medios para aprender inductivamente de la experiencia chocaba con corrientes importantes de la epistemología contemporánea, en particular la que personifica Karl Popper, quien reivindicando y modernizando el argumento humeano de invalidez de la inducción, presenta batalla contra cualquier forma de inferencia inductiva, inclusive la versión carnapiana de la inducción probabilística[20]. Esta circunstancia obligaba a ser escéptico respecto a la idea fisheriana de la naturaleza inductiva de la inferencia estadística. Como ésta parece consistir esencialmente en que unas pocas observaciones metódicas –las premisas singulares del razonamiento inductivo– bastan para extraer conclusiones acerca de determinadas propiedades de poblaciones eventualmente infinitas, la pregunta que se

---

19. Cfr. NEYMAN, J., *op. cit.*, p. 263.

20. En relación al problema de la probabilidad inductiva puede consultarse RIVADULLA, A.: «On Popper-Miller's Proof of the Impossibility of Inductive Probability», 1987; «Probabilidad Inductiva», 1989, y *Probabilidad e Inferencia Científica*, 1991, caps. I y II.

plantea es la siguiente: ¿quién tiene razón: los filósofos que niegan la legitimidad lógica de la inferencia conservadora de la verdad y ampliadora del contenido, inclusive la inducción probabilística, o los matemáticos que aseveran que la estadística proporciona los métodos para realizar inferencias correctas de lo particular a lo general, en terminología estadística, de muestras aleatorias a las poblaciones (supuestamente infinitas) de las que han sido extraídas?

Llegado a este punto en su argumentación, Roderick vuelve de nuevo su atención a la observación que la produjo y lleva a cabo las siguientes reflexiones finales.

Si deseo explicar satisfactoriamente el resultado observado: ocho veces 'cara' y dos veces 'cruz', debo conocer, o al menos plantear la hipótesis de que sé que esta observación constituye una muestra aleatoria de una población que sigue una ley probabilística determinada. Por ejemplo, que está distribuida binomialmente, o normalmente, o de otra forma determinada. Si asumo, pues, *a priori*, que los lanzamientos de monedas de curso legal son independientes entre sí y que la probabilidad de obtener 'cara' es constante, aunque sea desconocida, entonces la variable aleatoria: número de 'caras' en n lanzamientos de una moneda o en un lanzamiento de n monedas, es binomial. El caso más usual es aquel en que ignoro el valor de la constante p que designa la probabilidad de que ocurra 'cara' con la(s) moneda(s) de curso legal. Luego, a la vista del resultado obtenido, debo elegir entre *estimar* p o *testar* alguna hipótesis que considere viable acerca de su verdadero valor.

Si decido estimar p puedo recurrir al método fisheriano de *máxima verosimilitud*, y en tal caso, según Fisher[21], la función de verosimilitud $p^8(1 - p)^{10-8}$ correspondiente a la observación realizada habrá de ser tomada como base para una *inferencia inductiva* que permita estimar el valor verdadero de p. Esta inferencia, además, es incierta, ya que, supuesto el carácter infinito de la población de valores, no hay ninguna razón concluyente que obligue a creer que el valor conjeturado de p cuya verosimilitud es máxima coincidirá necesariamente con el valor verdadero. Lo que, parece, redunda en favor de la idea del carácter inductivo de toda inferencia construida con el método fisheriano de máxima verosimilitud.

Ahora bien, la función de verosimilitud sólo expresa una pura relación matemática (disponible una vez que el carácter binomial de la distribución de la población correspondiente ha sido admitido hipotéticamente) entre el valor conjeturado de un parámetro poblacional y los datos observacionales. Como la inferencia del valor paramétrico provisto de la máxima verosimilitud[22] es el resultado de ciertas manipulaciones matemáticas en la función de verosimili-

21. Cfr. FISHER, R.: «Inverse Probability and the Use of Likelihood», 1932, p. 259.

22. Respecto al uso de Ronald Fisher del método de máxima verosimilitud para la estimación de parámetros poblacionales, véase RIVADULLA, A.: *Probabilidad e Inferencia Científica*, p. 135.

tud correspondiente, cuando determinadas constantes numéricas se sustituyen por los valores observados, la inferencia es puramente *analítica*. Además, los datos observacionales solos, es decir, independientes de la premisa argumentativa que se ha de admitir al menos hipotéticamente, son complemente inoperantes. Luego los argumentos de verosimilitud no pueden ser ampliativos y, por consiguiente, inductivos.

El caso del test de hipótesis es aún más simple, ya que los resultados de experimentos aleatorios no añaden nada que no sea previamente conocido. Si conociéramos, como antes, o asumiéramos hipotéticamente que la población está distribuida binomialmente, sabríamos, previa aplicación de la fórmula de Bernoulli, qué probabilidad de ocurrencia tienen todos y cada uno de los resultados posibles, todos y cada uno de los elementos del espacio muestral. De esta manera tendríamos, desde el principio, un panorama completo de las decisiones a adoptar de acuerdo con las observaciones. De manera que si ocurriera un suceso de la que hubiéramos dado en considerar *región de rechazo de la hipótesis testada*, no habría acontecido nada imprevisto, pues su probabilidad de ocurrencia ya sería conocida matemáticamente. Como es sólo cuestión de *decisión* previa a la realización de experimentos declarar *harto improbables* a determinados sucesos observables, *la aceptación o rechazo de la hipótesis testada es una pura cuestión de decisión y no de inferencia, ni deductiva ni inductiva*.

El razonamiento deductivo juega ciertamente un papel importante en el cómputo de las probabilidades de los resultados lógicamente posibles, así como en el control de las probabilidades de error de primer y segundo tipos que pueden ser cometidos. Así, el resultado: ocho veces 'cara' tiene la probabilidad $P = 0{,}044$, bajo el supuesto de que H_0: $p = 0{,}5$ es verdadera; pero tendría la probabilidad $P = 0{,}121$ si $H_0 : p = 0{,}6$ fuera verdadera, y $P = 0{,}011$ si lo fuera $H_0 : p = 0{,}4$, etc. Supongamos, pues, que estoy dispuesto a aceptar que H_0: $p = 0{,}5$ es verdadera, entonces puedo construir la región de aceptación correspondiente de tal manera que contenga los resultados 2 ó 3 ó ... ó 8 veces 'cara', mientras la región de rechazo contendría los valores 0 ó 1 ó 9 ó 10 veces 'cara'. Como la suma de las probabilidades de estos resultados, calculadas por medio de (2) en relación a la hipótesis nula considerada, es 0,022, habría que fijar en = 2,2% el nivel de significación del test de $H_0 : p = 0{,}5$. Lo que esta *decisión* comporta es que *hemos de estar dispuestos a reconocer que el 2,2 % de las veces a la larga rechazaremos equivocadamente la hipótesis testada, incurriendo en un error del primer tipo*. Admitamos ahora, por contra, que la hipótesis verdadera es la alternativa $H_1 : p = 0{,}7$. Como, calculadas en relación a esta hipótesis las probabilidades de los sucesos que constituyen la región de aceptación de $H_0 : p = 0{,}5$, éstos tienen una probabilidad de acontecer del 85,1%, acabaremos aceptando erróneamente la hipótesis testada un 85,1% de las veces, *cometiendo un error del segundo tipo*.

El cómputo de las probabilidades de error es ciertamente competencia del razonamiento deductivo. Pero la inferencia concluye aquí. Pues la conducta a seguir, de acuerdo con las observaciones realizadas, cae en el dominio de la *teoría de la decisión*[23].

Finalmente, si lo que deseamos es calcular un intervalo de confianza para la estimación de nuestro parámetro poblacional desconocido *p*, tendremos que llevar a cabo, al igual que si empleáramos el método fisheriano de máxima verosimilitud, transformaciones tautológicas (matemático-probabilísticas) de la información de que disponemos acerca de la población investigada, *a fin de inferir deductivamente las estimaciones inferior y superior del parámetro buscado*. El método de la estimación estadística confidencial es, como Neyman mantiene correctamente frente a Fisher, perfectamente deductivo[24].

La conclusión general de Roderick es que en ninguno de los dos casos considerados: el test de hipótesis y la teoría de la estimación, la estadística frecuentista tiene absolutamente nada que ver con inducción. Pues cuando nos las tenemos que ver con inferencias, éstas son perfectamente deductivas, y, cuando hemos de adaptar nuestra conducta a los resultados observacionales, abandonamos el campo de la lógica y nos adentramos en el de la teoría de la decisión.

Si se puede hablar de inferencia estadística inductiva, y en qué sentido sería legítimo hacerlo, es algo que sólo cabría imaginar en el marco de la teoría bayesiana[25]. Pero ello supondría abandonar completamente el paradigma de la estadística clásica.

BIBLIOGRAFIA

ABOUT, P. J., y BOY, M.: *La correspondance de Blaise Pascal et de Pierre de Fermat: La géometrie du hasard ou le début du calcul des probabilités*. Fontenay aux Roses, ENS, 1983.

BAYES, T.: «An Essay Towards Solving a Problem in the Doctrine of Chances». *Philosophical Transactions of the Royal Society of London*, 53 and 54, 1763. Reimpreso en Pearson, E. S., y Kendall, M. G. (eds.): *Studies in the History of Statistics and Probability*. Charles Griffin, London, 1970.

FISHER, R. A.: «On the Mathematical Foundations of Theoretical Statistics». *Philosophical Transactions of the Royal Society of London*, series A, vol. 222, 1922, pp. 309-368.

FISHER, R. A.: «Inverse Probability». *Proceedings of the Cambridge Philosophical Society*, 26, 1930, pp. 528-535.

FISHER, R. A.: «Inverse Probability and the Use of Likelihood». *Proceedings of the Cambridge Philosophical Society*, 28, 1932, pp. 257-261.

FISHER, R. A.: «Two New Properties of Mathematical Likelihood». *Proceedings of the Royal Society of London*, series A, vol. 144, 1934, pp. 285-307.

---

23. Para un examen más exhaustivo de esta cuestión véase RIVADULLA, A.: «Mathematical Statistics and Metastatistical Analysis», 1991, sección 3.2.

24. Cfr. RIVADULLA, A.: «Mathematical Statistics and Metastatistical Analysis», 1991, sección 3.1.

25. Para una introducción a la misma puede consultarse RIVADULLA, A.: *Probabilidad e Inferencia Científica*, 1991, pp. 203 y ss.

FISHER, R. A.: *The Design of Experiments*. Oliver & Boyd. Edinburgh, 1935.

FISHER, R. A.: «Statistical Methods and Scientific Induction». *Journal of the Royal Statistical Society*, series B, vol. 17, 1955, pp. 69-78.

GOMEZ-VILLEGAS, M. A.: «El Problema de la Probabilidad Inversa: Bayes y Laplace». *Actas del Encuentro de Lógica y Filosofía de la Ciencia. Rudolf Carnap & Hans Reichenbach in Memoriam. Madrid, 13-15 de noviembre de 1991*. Ed. Siglo XXI. Madrid, 1993.

HUME, D.: *A Treatise of Human Nature*, 1739. Edición de L. A. Selby-Bigge. Clarendon Press. Oxford, 1973.

HUME, D.: *An Enquiry Concerning Human Understanding*, 1748. En Hume, D.: *Enquiries concerning Human Understanding and concerning the Principles of Morals*. 3.ª ed. de P. H. Nidditch. Clarendon Press. Oxford, 1975.

LAPLACE, P. S.: «Mémoire sur la probabilité des causes par les événements». *Mém. Acad. Roy. Sci. Paris*, 6, 1774. Reimpreso en Laplace, P. S.: *Oeuvres*, vol. 8, 1891.

LAPLACE, P. S.: *Théorie Analytique des Probabilités*, 3.ª ed. París, 1820.

NEYMAN, J.: «Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability», 1937. Reimpreso en Neyman, J.: *A Selection of Early Statistical Papers*. University Press. Cambridge, 1967.

NEYMAN, J.: *First Course in Probability and Statistics*. Constable and Co., Ltd. London, 1950.

RIVADULLA, A.: *Filosofía Actual de la Ciencia*. Editorial Tecnos. Madrid, 1986.

RIVADULLA, A.: «On Popper-Miller's Proof of the Impossibility of Inductive Probability». *Erkenntnis*, 27, 1987, pp. 353-357.

RIVADULLA, A.: «Probabilidad Inductiva». *Arbor*, 523-524, 1989, pp. 61-74.

RIVADULLA, A.: «The Roots of Theoretical Statistics. Probabilities of life in the second half of the 17th Century». En Díaz, A., et al. (eds.): *Structures in Mathematical Theories*. Universidad del País Vasco. San Sebastián, 1990.

RIVADULLA, A.: *Probabilidad e Inferencia Científica*. Editorial Anthropos. Barcelona, 1991.

RIVADULLA, A.: «Apriorismo y Base Empírica en los Orígenes de la Estadística Matemática». *Llull*, vol. 14, 1991, pp. 187-219.

RIVADULLA, A.: «Mathematical Statistics and Metastatistical Analysis». *Erkenntnis*, 34, 1991, pp. 211-236.

RIVADULLA, A.: «Cálculo axiomático de la probabilidad lógica». *Theoria*, 1992, volumen conmemorativo del XL aniversario.

SCHNEIDER, I. (Hrsg.): *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933. Einführungen und Texte*. Wissenschaftliche Buchgesellschaft. Darmstadt, 1988.

TODHUNTER, I.: *A History of Mathematical Theory of Probability*. Macmillan and Co. Cambridge and London, 1865.