Departament d'Estadística i I.O., Universitat de València.
Facultat de Matemàtiques, 46100–Burjassot, València, España.
Tel. 34.6.386.4314 (directo), 34.6.386.4362 (secretaría); Fax 34.6.386.4735,
Internet: jose.m.bernardo@uv.es, Web: http://www.uv.es/~bernardo/

Impreso el 6 de Octubre de 1997
Pendiente de aparecer en la *Revista de la Academia de Ciencias de Madrid*

Bruno de Finetti en la Estadística Contemporánea

JOSÉ-MIGUEL BERNARDO
Universitat de València, España

RESUMEN

Bruno de Finetti, [1906–1985], es unánimamente considerado como una de las figuras más relevantes en la estadística del siglo XX. Desde nuestro punto de vista, sus aportaciones más trascendentes para el desarrollo de la estadística contemporánea han sido (i) la formalización del concepto de probabilidad como *grado de creencia*, que permite un tratamiento riguroso del concepto de probabilidad que se deduce a partir de la teoría de la decisión; (ii) el concepto de *intercambiabilidad* que, a través de los *teoremas de representación*, permite integrar en un paradigma unificado los conceptos estadísticos frecuencialistas asociados a modelos paramétricos con el concepto de probabilidad como grado de creencia; y (iii) el desarrollo de las *funciones de evaluación*, que permiten calibrar la asignación de probabilidades y, en particular, contrastar la idoneidad de un modelo probabilístico. En este trabajo se apuntan algunos datos de su biografía, se sintetizan el contenido y las implicaciones de algunas de sus aportaciones, y se ofrece un conjunto de referencias que puede facilitar el acceso a las fuentes originales.

Palabras Clave: COHERENCIA; INFERENCIA BAYESIANA; FUNCIONES DE EVALUACIÓN; GRADOS DE CREENCIA; INTERCAMBIABILIDAD; PROBABILIDADES INTERSUBJETIVAS; PROBABILIDADES SUBJETIVAS; TEOREMAS DE REPRESENTACIÓN.

1. ELEMENTOS BIOGRÁFICOS

Bruno de Finetti nació el 13 de Junio de 1906 en Innsbruck (Austria) de padres italianos, donde su padre trabajaba como ingeniero en la construcción de la *Stubaithalbahn*. Estudió primaria y secundaria en Trento, lugar de origen de su madre, a donde se trasladó con ella tras la prematura muerte de su padre. Ingresó posteriormente en el Politécnico de Milán, pero tras dos años de estudios de ingeniería, se inscribió en la recién creada Universidad de Milán, por la que se licenció en Matemáticas en 1927. Su formación en probabilidad y estadística fué esencialmente autodidacta; se cita el *Wahrscheinlichkeitslehere* de Czuber como el texto que despertó su interés en la materia.

Desde el principio mostró su propensión hacia las matemáticas, entendidas como un instrumento para desarrollar aplicaciones específicas (en física, ingeniería, biología, economía, estadística), y como un elemento de ayuda en la profundización de cuestiones conceptuales (en

---

José Miguel Bernardo is Catedrático de Estadística en la Universidad de Valencia. Investigación financiada con el Proyecto PB93-1204 de la DGICYT, Madrid.

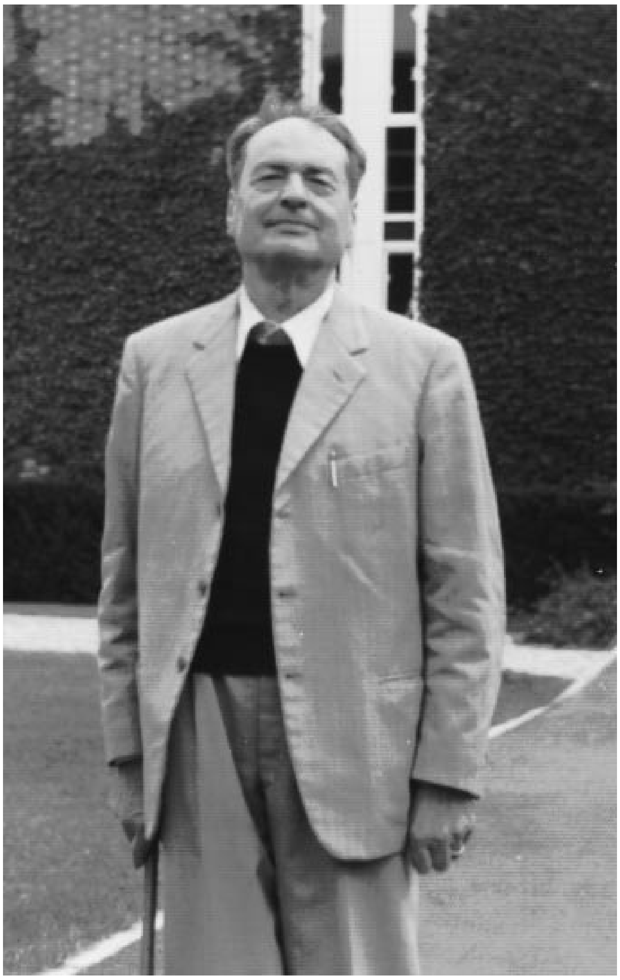
lógica, probabilidad, teoría de la ciencia), mientras que explícitamente rechazaba la visión de las matemáticas como un formalismo abstracto cerrado en sí mismo (de Finetti, 1981, p. xviii; Fürst, 1993).

Un primer trabajo sobre herencia mendeliana (de Finetti, 1926) llamó la atención de Conrado Gini, quién le propuso como director de la oficina matemática del *Instituto Centrale di Statistica* en Roma. Estos cuatro primeros años en la capital italiana, [1927–1931], constituyen la época en la que aparecen sus grandes aportaciones originales, contemporáneas de las de Kolmogorov, Levi, Fisher o Cantelli. Así, en Septiembre de 1928, presenta *Funzione Caratteristica di un Fenomeno Aleatorio* (de Finetti, 1930) ante el Congreso Internacional de Matemáticas de Bolonia, un trabajo que ya contiene su concepto de *intercambiabilidad*; en la audiencia destacan los nombres de Borel, Cantelli, Darmois, Fisher, Fréchet, Khintchine, Lévi, Neyman y Polya. En 1929 escribe *Probabilismo* (de Finetti, 1931a), un amplio ensayo crítico sobre el concepto de probabilidad como grado de creencia, en la línea del positivismo lógico de Mach. Cerrando esta etapa, publica *Sul Significato Soggetivo della Probabilità* (de Finetti, 1931b), la primera exposición sistemática de sus ideas.

En 1931 acepta una oferta de la compañia de seguros *Assicurazioni Generali*, con base en Trieste, para la que trabaja hasta 1946 en que obtiene una cátedra en la Universidad de esa misma ciudad adriática. Su estancia en Trieste, [1931–1954], coincide con el desarrollo de su imagen internacional. Su trabajo más conocido, *La Prévision, ses Lois Logiques, ses Sources Subjectives* (de Finetti, 1937), que contiene el desarrollo matemático formal de la probabilidad como grado de creencia y el teorema de representación de variables dicotómicas intercambiables, es la versión escrita del ciclo de cinco conferencias que fué invitado a dictar en el *Institut Henrí Poincaré* de Paris, en 1935. En los años cuarenta publica *Matemática Lógico-Intuitiva* (de Finetti, 1944), que constituye su aportación más conocida a la enseñanza universitaria de las matemáticas. En 1950 es invitado al segundo *Berkeley Symposium* (de Finetti, 1951), donde presenta un trabajo pionero en la teoría de métodos Bayesianos robustos. Durante 1952, es invitado a visitar Chicago; es el principio de una larga y fructífera colaboración con Leonard (Jimmie) Savage que amplificaría notablemente su proyección internacional.

En 1954 obtiene una cátedra en la Facultad de Economía de la Universidad de Roma, que sería seguida en 1961 por otra cátedra en la Facultad de Ciencias de la misma universidad, en la que permanecería hasta su jubilación en 1976. En esta segunda época romana de plena madurez se inscriben muchas de sus publicaciones más conocidas. El 1960 es invitado al cuarto *Berkeley Symposium* (de Finetti, 1961), donde introduce un tratamiento Bayesiano de las observaciones atípicas. A lo largo de los años sesenta, desarrolla una teoría para la evaluación de predicciones probabilísticas (de Finetti, 1962, 1965) a partir del uso sistemático de la *función de evaluación cuadrática*. En 1970 aparece su *opera magna*, la *Teoria delle Probabilità* (de Finetti, 1970a), síntesis de sus ideas sobre probabilidad y estadística, traducida al inglés y al alemán, y destinada, en palabras de Dennis Lindley, *to be recognized as one of the great books of the world* (Lindley, 1975). Poco después, a propuesta de Jimmie Savage, publica *Probability, Induction and Statistics* (de Finetti, 1972), una versión en inglés, —ampliada, corregida y reestructurada por el autor—, de importantes trabajos suyos originalmente publicados en italiano entre 1949 y 1967. Su conferencia invitada ante el congreso del *International Statistical Institute* celebrado en Viena en 1973 (de Finetti, 1974), constituye una lúcida descripción del papel unificador de los métodos estadísticos Bayesianos en los fundamentos y en las aplicaciones de la estadística.

Tres años más tarde, ante la *First European Conference on New Developments and Applications of Bayesian Methods*, —la primera conferencia internacional sobre métodos Bayesianos que tuvo lugar en el mundo—, celebrada en Fontainebleau en Junio de 1976, de Finetti analiza el problema de las probabilidades de orden superior (de Finetti, 1977a). El autor de estas líneas, entonces recién doctorado, tuvo en esa ocasión la oportunidad de conocer al *Maestro*, y el privilegio de compartir con él y con Dennis Lindley una larga sobremesa sobre el papel de la aditividad numerable en la teoría de la probabilidad y en la estadística, un tema todavía polémico 30 años después. Unos meses más tarde, en Noviembre de 1976, de Finetti dictaba en la Universidad de Roma su conferencia de despedida, *La probabilità: Guardarsi dalle Contraffazioni*, (de Finetti, 1976).



Bruno de Finetti en el Congreso de Fontainebleau, Junio de 1976. Foto de J. M. Bernardo.

En 1981, con ocasión del 75 cumpleaños del *Maestro*, se reeditaban sus primeros trabajos, precedidos de una nota autobiográfica (de Finetti, 1981), y la Academia de Ciencias italiana organizaba en su honor una conferencia internacional sobre el desarrollo y las aplicaciones del concepto de intercambiabilidad (Koch & Spizzichino, 1982). Poco después, ya enfermo, de Finetti publica un segundo apunte autobiográfico, *Probability and my Life* (de Finetti, 1982), que sería su último escrito. Bruno de Finetti moría en Roma el 20 de Julio de 1985.

Los años posteriores vieron aparecer varias reediciones y nuevas traducciones de sus trabajos. Sus escritos sobre la lógica de la incertidumbre fueron recogidos por Marco Mondatori (de Finetti, 1989), y sus publicaciones del periodo [1931–1936] fueron reimpresas por *l'Associazione per la Matematica Applicata* (de Finetti, 1991). Poco después, Paola Monari y Daniela Cocchi editaron una edición bilingüe, italiana e inglesa, de algunos de sus trabajos más representativos sobre probabilidad e inducción (de Finetti, 1993) y, recientemente, A. Mura ha preparado una reedición de sus trabajos sobre filosofía de la probabilidad (de Finetti, 1995).

Numerosos autores han publicado trabajos sobre la vida y sobre la obra del Profesor de Finetti. En la preparación de esta conferencia se han utilizado los de Daboni (1984), Lindley (1986, 1987), Piccinato (1986), Fürst (1993), Jeffrey (1993) y Cifarelli & Regazzini (1995). Goel & Zellner (1986) editaron un libro de contribuciones en honor a Bruno de Finetti; Daboni *et al.* (1987) organizaron en Trieste un congreso en recuerdo suyo, cuyas Actas contienen diferentes perspectivas sobre su trabajo. Regazzini (1987) y Muhere & Parmigiani (1993) han publicado estudios históricos sobre el desarrollo de la teoría de la probabilidad en los que las contribuciones del *Maestro* pueden ser situadas en su perspectiva histórica.

2. EL CONCEPTO DE PROBABILIDAD

Una de las aportaciones fundamentales de Bruno de Finetti al pensamiento estadístico contemporáneo fué la construcción de una formulación *matemática* precisa para el concepto *semántico* de probabilidad, tal como se usa en el lenguaje cotidiano en frases como 'es probable que la Unión Europea disponga de una moneda única antes del año 2005' y que, como es sabido, resulta indispensable para una toma de decisiones coherente (Savage, 1954; Bernardo & Smith, 1994, Cap. 2).

Para de Finetti, la probabilidad es necesariamente *subjetiva*: la probabilidad de un suceso mide el *grado de creencia* en su ocurrencia percibido por el sujeto en el momento que la expresa. Para cualquier suceso específico, sólo existen dos alternativas 'objetivas': el suceso es (o será) cierto o falso; sin embargo, si quiere *medirse* el grado de incertidumbre que se tiene sobre cuál de esas dos alternativas es correcta, la probabilidad es la *única* forma coherente de medida. Así, cuando una persona asigna probabilidad 0.1 al suceso $E \equiv \{$ el dígito 10^6 del número π es $0\}$, no está realizando una afirmación 'objetiva' (ese suceso es demostrablemente cierto o falso) sino expresando su incertidumbre al respecto: con la información de que dispone, los 10 dígitos le parecen igualmente probables. A la pregunta (con más de dos siglos de historia) sobre ¿qué es la probabilidad?, de Finetti contesta provocativamente que '*la probabilità non esiste*', que la probabilidad no tiene existencia 'objetiva'; no es una propiedad de la naturaleza, ni tampoco de los individuos: describe una *relación* entre *un* individuo y la naturaleza: $P_i(E \mid H)$ es una medida del grado de creencia en la ocurrencia de un suceso E, asignada por un individuo i, en unas condiciones H; en palabras de su amigo Jimmie Savage, '*probabilities are states of mind, not states of nature*' (Savage, 1981, p. 674). Sin embargo, esta actitud aparentemente solipsista es compatible con la existencia de un amplio conjunto de probabilidades sobre cuyos valores puede existir consenso, como en el caso de la probabilidad 0.1 de que el dígito 10^6 del número π sea 0 que acabamos de mencionar; el desarrollo científico se apoya precisamente en la existencia de estas probabilidades 'intersubjetivas'.

Inscrito en la tradición del positivismo lógico de Mach, Einstein y Bridgman, de Finetti apuesta por definiciones *operativas* de todos los conceptos que introduce. En particular, su *definición* de probabilidad, es totalmente operativa: $P_i(E \mid H)$ es el *máximo* 'precio' que el individuo i estaría dispuesto a pagar por un ticket de una 'rifa', jugada en las condiciones H, en la que obtendría un premio unidad si, y sólamente si, E tiene lugar. Así, una persona que afirma que la probabilidad de que una moneda lanzada al aire salga cara es 0.5, debería estar dispuesta a pagar un máximo de 50 pesetas por un ticket que le daría derecho a cobrar 100 pesetas si al lanzar la moneda se obtiene una cara. Recíprocamente, si el máximo precio que una persona estaría dispuesta a pagar por un ticket que le daría derecho a cobrar 100 pesetas si un determinado equipo de fútbol gana su próximo partido es de 70 pesetas, entonces *su* probabilidad de que tal equipo gane el partido es 0.7. Como el propio de Finetti pone de manifiesto en una nota a pié de página de la traducción al inglés de su trabajo más citado (de Finetti, 1937), es obvio que para que la definición sea realmente operativa, las cantidades involucradas deben ser *pequeñas* comparadas con los recursos de la persona que utiliza la definición; más formalmente, su función de utilidad debe ser aproximadamente lineal para esas cantidades.

De Finetti demuestra que si los sucesos $\{E_1, \ldots, E_k\}$ forman un partición, y $\{p_1, \ldots, p_k\}$ son los grados de creencia en la ocurrencia de cada uno de ellos asignados por un individuo i en las condiciones H, ésto es $p_j = P_i(E_j \mid H)$, entonces la condición necesaria y sufiente para que *no* sea posible construir un conjunto de apuestas en las que i necesariamente pierda dinero (*dutch book*) es que las p_j's constituyan una *distribución de probabilidad*, *i.e.*, que $p_j \geq 0$, y $\sum_j p_j = 1$; para evitar un comportamiento incoherente, intrinsecamente contradictorio, *los grados de creencia deben comportarse como probabilidades* (de Finetti, 1931b, 1937). Las definiciones axiomáticas de probabilidad (*e.g.*, Kolmogorov, 1933), sin relación operativa alguna con el mundo real, se convierten así en un teorema: los grados de creencia deben satisfacer las propiedades convencionalmente utilizadas para definir formalmente la probabilidad y, consecuentemente, puede definirse una medidad de probabilidad a partir de ellos. Para de Finetti, *cualquier* incertidumbre *debe* ser expresada mediante una distribución de probabilidad. Esta es, como es sabido, la característica *esencial* de la metodología estadística *Bayesiana*.

Naturalmente, el hecho de que *exista* una definición operativa, no implica que utilizarla sea en la práctica el mejor *método* de asignar probabilidades. De hecho, el *concepto* de probabilidad como grado de creencia incluye como *caso particular* a los conceptos *clásico* y *frecuencialista* (de Finetti, 1970a, Prefacio). En efecto, si las simetrías de un problema sugieren una percepción de 'equiprobabilidad' entre N casos posibles mutuamente excluyentes, entonces el sujeto asignará la misma probabilidad $1/N$ a cada uno de ellos y consecuentemente, en virtud de las leyes de la probabilidad que los grados de creencia deben obedecer, la probabilidad asociada a un suceso constituido por la unión de k de ellos será *necesariamente* k/N. Análogamente, si la información histórica de que se dispone sugiere una comparación con situaciones razonablemente parecidas, entonces la probabilidad asignada será presumiblemente establecida en base a las correspondientes frecuencias relativas; así, si una compañia de seguros estima que todos los conductores varones noveles con menos de 25 años presentan un riesgo de accidente comparable, y sus series históricas demuestran que en el pasado un 3% de ellos tuvieron de hecho un accidente en su primer año como conductores, es razonable suponer que sus analistas asignarán una probabilidad 0.03 al suceso de que un conductor varón novel de menos de 25 años tenga un accidente en su primer año como conductor. Estos son dos ejemplos sencillos de probabilidades 'intersubjetivas' en el sentido antes mencionado. Obsérvese sin embargo que, aunque pueden resultar útiles para *asignar* probabilidades, los *conceptos* clásico y frecuencialista de la probabilidad no permiten una *definición* formal de probabilidad, que resultaría circular al tener, respectivamente, que incluir conceptos como 'casos equiprobables' y 'convergencia en probabilidad'.

Consecuente con su defensa de las definiciones estrictamente operativas, de Finetti defiende la *posibilidad* de trabajar con probabilidades tan sólo *finitamente* aditivas (de Finetti, 1937, 1970a), sin exigir la condición de σ-aditividad que, en el caso contínuo, permite garantizar la existencia de funciones de densidad. Sin embargo, la notable complejidad de una teoría de la probabilidad no σ-aditiva y las paradojas de aglomeración que tal teoría lleva necesariamente consigo (la hipótesis de conglomerabilidad implica la σ-aditividad) han llevado a una aceptación pragmática practicamente universal, —incluyendo algunos trabajos del propio de Finetti—, de las probabilidades σ-aditivas. A nivel de fundamentos, sin embargo, la polémica continúa (Heath & Sudderth, 1978, 1989, Cifarelli & Regazzini, 1987; Consonni & Veronese, 1989; Lindley, 1997).

La notación convencional para las probabilidades, $P(E \mid H)$ en lugar de $P_i(E \mid H)$, omite la referencia al sujeto i que las especifica. Sin embargo, no existe ningún argumento que permita *demostrar* que dos individuos con la misma información deben, necesariamente, asignar las mismas probabilidades. Las probabilidades 'lógicas' en el sentido de Keynes (1921) o de Carnap (1950) no están matemáticamente justificadas. Sin embargo, como ya hemos apuntado, existen numerosas situaciones, —especialmente en el contexto del diseño y análisis de la experimentación científica—, en las que es posible elaborar probabilidades de consenso, probabilidades intersubjetivas que permiten la comunicación científica.

En la línea de Kim, el personaje de Kipling en el *Libro de la Selva*, y de Henri Poincaré (*la théorie des probabilités n'est que le bon sens réduit au calcul*), la teoría de la probabilidad es para de Finetti una forma de higiene mental, un modelo (en el sentido del *als ob* de Veihinger) que permite la incorporación y gestión consistente de las numerosas fuentes de información relevante de que puede disponerse en un momento dado (de Finetti, 1938b, 1963, 1964, 1967, 1968, 1977b, 1978). Sus *teoremas de representación* constituyen una de las herramientas más importantes para llevar adelante este proyecto.

3. INTERCAMBIABILIAD Y TEOREMAS DE REPRESENTACIÓN

Dentro de su programa *operativo*, de Finetti parte de la modelización de observables. Así, puesto que toda incertidumbre debe ser descrita mediante probabilidades, la información de que se dispone en las condiciones H sobre el valor de una observación futura $x \in X$ debe ser descrita mediante una densidad de probabilidad $p_x(. \mid H)$ en el sentido de que

$$P(x \in A \mid H) = \int_A p_x(x \mid H)\, dx, \quad A \subset X.$$

Si disponemos de una base de datos $\{x_1, \ldots, x_n\}$ constituida por n observaciones 'semejantes' a x, para 'predecir' el valor de x deberemos especificar su correpondiente densidad de probabilidad *condicional* $p(x \mid x_1, \ldots, x_n, H)$. Obsérvese que de Finetti no *introduce* 'parámetros' desconocidos: $p(x \mid x_1, \ldots, x_n, H)$ es una función totalmente especificada.

Suprimiendo por simplicidad en la notación los subindices correspondientes a las distintas densidades de probabilidad, el comportamiento de un *conjunto* $\{x_1, \ldots, x_n, x_{n+1}\}$ de $n+1$ observaciones en las condiciones H será descrito por su densidad de probabilidad *conjunta*, $p(x_1, \ldots, x_n, x_{n+1} \mid H)$ de forma que

$$P(\{x_1, \ldots, x_n, x_{n+1}\} \in B \mid H) = \int_B p(x_1, \ldots, x_n, x_{n+1} \mid H)\, dx_1 \ldots dx_{n+1}, \quad B \subset X^{n+1}$$

y, consecuentemente,

$$P(x_{n+1} \in A \mid x_1, \ldots, x_n, H) = \int_A p(x_{n+1} \mid x_1, \ldots, x_n, H)\, dx_{n+1}, \quad A \subset X$$

con

$$p(x_{n+1} \mid x_1, \ldots, x_n, H) = \frac{p(x_1, \ldots, x_{n+1} \mid H)}{p(x_1, \ldots, x_n \mid H)}$$

de forma que para resolver el problema de *predicción* planteado, —que de Finetti acertadamente considera el problema básico en la investigación científica—, es necesario y suficiente especificar la distribución conjunta de una sucesión cualquiera de observaciones 'semejantes'.

De Finetti empieza por proponer una *definición* formal que capture la idea básica de observaciones 'semejantes'.

Definición de intercambiabilidad. *Las variables aleatorias* $\{x_1, x_2, \ldots\}$ *de una sucesión son intercambiables (equivalenti, exchangeable, échangeable, austauschbar) si la distribución conjunta de cualquier subconjunto finito,* $p(x_1, \ldots, x_n)$*, es invariante ante permutaciones de sus índices.*

En un conjunto de variables intercambiables toda la información relevante está contenida en los *valores* de las x_i's, de forma que sus índices no proporcionan información alguna. Obsérvese que el concepto de intercambialilidad generaliza el de independencia condicional: un conjunto de observaciones independientes idénticamente distribuidas son *siempre* un conjunto de observaciones intercambiables.

Utilizando en concepto de intercambiabilidad, de Finetti demuestra su famoso teorema de representación para variables dicotómicas:

Teorema de representación para variables dicotómicas. *(De Finetti, 1930, 1937). Si* $\{x_1, x_2, \ldots, \}$*,* $x_i \in \{0, 1\}$ *son variables aleatorias intercambiables, entonces*

(i) existe $\theta = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \theta \in]0, 1[,$ *y*

(ii) existe una densidad de probabilidad $p(\theta)$*,*

tales que la densidad conjunta $p(x_1, \ldots, x_n)$ *tiene la representación integral*

$$p(x_1, \ldots, x_n) = \int_0^1 \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} p(\theta)\, d\theta$$

El teorema demuestra que un conjunto de variables aleatorias dicotómicas intercambiables *necesariamente* se comporta como una muestra aleatoria de observaciones Bernoulli, cuyo parámetro θ es el límite de su frecuencia relativa, y para el que *necesariamente* existe una distribuión inicial $p(\theta)$. Lindley & Phillips (1976) proporcionan una sencilla demostración moderna de este resultado crucial.

En el caso dicotómico, la intercambibilidad identifica las observaciones como una muestra aleatoria de un modelo probabilístico específico (Bernoulli) y garantiza la existencia de una distribución inicial sobre su parámetro. En el caso general, para variables aleatorias de cualquier rango y dimensión, la intercambiabilidad identifica las observaciones como una muestra aleatoria de *algún* modelo probabilístico y garantiza la existencia de una distribución inicial sobre el parámetro que lo describe:

Teorema de representación general. *(Hewwit & Savage, 1955; Kingman, 1978). Si* $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \}$*,* $\boldsymbol{x}_i \in X$ *son variables aleatorias intercambiables, entonces*

(i) existen una función $f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ *y un modelo probabilístico* $p(\boldsymbol{x} \mid \boldsymbol{\theta})$*,* $\boldsymbol{x} \in X$*, con* $\boldsymbol{\theta} = \lim_{n\to\infty} f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n), \quad \boldsymbol{\theta} \in \Theta,$ *y*

(ii) existe una función de densidad de probabilidad $p(\boldsymbol{\theta})$*,*

tales que la densidad conjunta $p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ *tiene la representación integral*

$$p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \int_\Theta \prod_{i=1}^{n} p(\boldsymbol{x}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})\, d\boldsymbol{\theta}$$

En una terminología convencional, el teorema demuestra que un conjunto de observaciones intercambiables pueden ser siempre *interpretadas* (de nuevo el *als ob* de Veihinger) como una muestra aleatoria de *algún* modelo probabilístico, controlado por un parámetro que no es introducido de manera arbitraria, sino que resulta *definido* como el límite de una función de las observaciones. Además, el teorema *garantiza la existencia* de una *distribución inicial* sobre el parámetro que gobierna el comportamiento del modelo. Debe subrayarse que las observaciones intercambiables $\boldsymbol{x}_i$ pueden ser *vectores* de cualquier dimensión y que, incorporando en ellas como covariables todos los elementos de información relevantes para que sus índices no sean informativos, *siempre* es posible postular la intercambiabilidad de los vectores así obtenidos.

Los teoremas de representación son resultados de teoría de la probabilidad no sujetos a polémica alguna, pero al demostrar la *existencia* de la distribución inicial constituyen una demostración de la necesidad *lógica* de la metodología Bayesiana. Obsérvese, sin embargo, que se trata tan *sólo* de teoremas de existencia: en cada problema concreto es necesario indentificar tanto el modelo como la distribución inicial apropiadas.

Como de Finetti magistralmente sintetiza (de Finetti, 1974), el concepto de probabilidad como grado de creencia proporciona, junto a los teoremas de representación, un paradigma unificado para la inferencia estadística. Dado un conjunto de observaciones intercambiables, (posiblemente vectoriales) $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, el proceso de inferencia sobre una magnitud de interés ϕ relacionada con ellas se descompone de forma natural en tres fases: especificación del modelo probabilístico, especificación de la distribución inicial relevante, y determinación de la distribución final de la magnitud de interés.

1. *Especificación del modelo*. De la intercambiabilidad de las $\boldsymbol{x}_i$'s se deduce que el conjunto de observaciones $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ constituye una muestra aleatoria de *algún* modelo probabilístico $p(\boldsymbol{x} \,|\, \boldsymbol{\theta})$, de forma que las $\boldsymbol{x}_i$'s son condicionalmente independientes dado el valor de un parámetro $\boldsymbol{\theta}$ (posiblemente vectorial), *i.e.*,

$$p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \,|\, \boldsymbol{\theta}) = \prod_{i=1}^{n} p(\boldsymbol{x}_i \,|\, \boldsymbol{\theta}),$$

donde $\boldsymbol{\theta} = \lim_{n\to\infty} f(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, $\boldsymbol{\theta} \in \Theta$, se *define* como el límite de alguna función de las observaciones. Frecuentemente, la información de que se dispone en el contexto en que las observaciones han sido obtenidas permite especificar el modelo; por ejemplo, la teoría de errores puede sugerir una distribución normal para las observaciones obtenidas en un experimento astronómico, o la teoría económica puede sugerir una distribución de Pareto para observaciones sobre la renta de los individuos. En otros casos, el modelo puede ser determinado mediante hipótesis adicionales sobre el comportamiento de las $\boldsymbol{x}_i$'s; por ejemplo, si las x_i's son variables reales intercambiables y tienen simetría esférica, entonces se trata necesariamente de observaciones normales $N(x \,|\, \mu, \sigma)$ (Smith, 1981), donde

$$\mu = \lim_{n\to\infty} \overline{x}, \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\sigma = \lim_{n\to\infty} s, \quad s^2 = \frac{1}{n}\sum_{i=1}^{n} [x_i - \overline{x}]^2.$$

Finalmente, el concepto de *intercambiabilidad parcial* (de Finetti, 1938a), y los teoremas de representación asociados a ella (de Finetti, 1970a; Diaconis and Freedman, 1980, 1984; Diaconis, Eaton and Lauritzen, 1992), permiten la construcción de sofisticados modelos *jerárquicos* que, partiendo de los trabajos pioneros de Good (1965) y Lindley & Smith (1972), constituyen hoy una de las áreas de aplicación más fecundas en la estadística contemporánea (ver *e.g.*, Good, 1980; Bernardo & Smith,1994, Cap. 4; Gelman *et al.,* 1995, Cap. 5; y referencias citadas en ellos) .

2. *Especificación de la distribución inicial.* De la intercambiabilidad de las $\boldsymbol{x}_i$'s se deduce la *existencia* de una distribución inicial $p(\boldsymbol{\theta})$ sobre los parámetros que gobiernan el modelo probabilístico. Si se dispone de información relevante sobre el valor de $\boldsymbol{\theta}$, esta información debe ser descrita por medio de $p(\boldsymbol{\theta})$; la forma de hacerlo depende necesariamente del contexto; Gatsonis *et al.* (1993, 1995, 1997) y Gelman *et al.* (1995) proporcionan numerosos ejemplos. Si no se dispone de información relevante sobre $\boldsymbol{\theta}$, o si en un contexto científico no quiere utilizarse la que se dispone por no tratarse de información intersubjetiva consensuada, entonces debe utilizarse una distribución inicial de *referencia* $\pi(\boldsymbol{\theta})$, que dependerá del modelo utilizado $p_{\boldsymbol{x}}(.\,|\,\boldsymbol{\theta})$ y de la magnitud de interés ϕ, y que se obtiene como aquella $p(\boldsymbol{\theta})$ que minimiza la influencia de la distribución inicial en la distribución final de la magnitud de interés (Bernardo, 1979, 1997; de Finetti, 1979; Berger & Bernardo, 1992; Bernardo & Ramón, 1998). Las distribuciones de referencia son un ejemplo importante de probabilidades intersubjetivas.

3. *Determinación de la distribución final de la magnitud de interés.* La magnitud de interés puede ser tanto una función de los parámetros del modelo $\phi = \phi(\boldsymbol{\theta})$ como una función de un conjunto de observaciones futuras $\phi = \phi(\boldsymbol{x}_{n+1}, \ldots, \boldsymbol{x}_{n+m})$. En ambos casos, su determinación es un problema bien definido de cálculo de probabilidades. En primer lugar, es necesario determinar, mediante el teorema de Bayes, (que da nombre al procedimiento de inferencia), la distribución final conjunta de los parámetros del modelo,

$$p(\boldsymbol{\theta}\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \frac{p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\,|\,\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)} = \frac{\prod_{i=1}^n p(\boldsymbol{x}_i\,|\,\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_\Theta \prod_{i=1}^n p(\boldsymbol{x}_i\,|\,\boldsymbol{\theta})p(\boldsymbol{\theta})\,d\boldsymbol{\theta}}\,.$$

Si la magnitud de interés $\phi = \phi(\boldsymbol{\theta})$, es una función de los parámetros del modelo, su distribución final $p(\phi\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ puede deducirse de $p(\boldsymbol{\theta}\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ como la correspondiente distribución marginal. Si es una función de observaciones futuras $\phi = \phi(\boldsymbol{x}_{n+1}, \ldots, \boldsymbol{x}_{n+m})$, entonces la distribución *condicional* $p(\phi\,|\,\boldsymbol{\theta})$ puede ser deducida a partir del modelo, y la distribución final de ϕ puede calcularse utilizando el teorema de la probabilidad total, de forma que

$$p(\phi\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \int_\Theta p(\phi\,|\,\boldsymbol{\theta})p(\boldsymbol{\theta}\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\,d\boldsymbol{\theta}.$$

Los resultados obtenidos en un análisis como el descrito son obviamente *condicionales* al modelo probabilístico y a la distribución inicial utilizadas. Es importante por lo tanto disponer de herramientas que permitan *evaluar* la asignación de probabilidades. Good (1965) y de Finetti (1962, 1965, 1970b) fueron pioneros en el uso de *funciones de evaluación* con este propósito.

4. FUNCIONES DE EVALUACIÓN

Para de Finetti, la *única* forma de describir la incertidumbre es una medida de probabilidad; consecuentemente, cualquier otra expresión es necesariamente incompleta o inadecuada. Los *test de respuesta múltiple*, tan frecuentes en nuestra sociedad, proporcionan, en su forma convencional, un buen ejemplo de expresión *incompleta* de la incertidumbre. En efecto, cuando se requiere a un alumno que señale cuál de las k respuestas posibles, —que se le asegura son mutuamente excluyentes—, es la correcta, se está violando el principio mencionado y, consecuentemente, se obtiene una información incompleta: el profesor *no puede* determinar si un alumno ha elegido la respuesta correcta porque la conocía, o porque ha elegido al azar y ha acertado. La solución obvia, consistente con el principio enunciado, es requerir al alumno que especifique una *distribución de probabilidad* sobre las respuestas posibles. De esta forma, un alumno que conoce la respuesta debería asignar una distribución degenerada, con probabilidad 1 para la respuesta correcta, mientras que un alumno que dudase entre las dos primeras respuestas debería asignar una distribución $(0.5, 0.5, 0, \ldots, 0)$, y un alumno que no sabe nada debería asignar una distribución uniforme $(1/k, \ldots, 1/k)$. Esta formulación plantea sin embargo un nuevo problema extremadamente interesante: ¿cuál debe ser la forma de *evaluar* este tipo de respuesta?, ¿cuál debe ser, en función de la respuesta correcta, la puntuación que merece un alumno que asigna una distribución de probabilidad $(p_1, p_2, \ldots, p_k)$ sobre las k respuestas posibles?

Obviamente, la función de evaluación debe ser elegida de forma que *motive* al alumno a decir la verdad, esto es que su estrategia óptima sea escribir $(1/k, \ldots, 1/k)$ si realmente no sabe nada y no 'pretender' que sabe escribiendo, por ejemplo, una distribución degenerada que asigne probabilidad uno a una respuesta escogida al azar. De Finetti (1962, 1965, 1970b) propone utilizar una *función de evaluación cuadrática* de la forma

$$u[(p_1, p_2, \ldots, p_k), i] = a(1 - \delta_i^2) + b_i, \quad a > 0, \quad b_i \in \Re$$

$$\delta_i^2 = (1 - p_i)^2 + \sum_{j \neq i} p_j^2$$

que a una 'predicción' $(p_1, p_2, \ldots, p_k)$ sobre la respuesta correcta, denotada i, le asigna una puntuación proporcional a $1 - \delta_i^2$, donde δ_i es la distancia euclidea entre la 'predicción' $(p_1, p_2, \ldots, p_k)$ y la 'predicción perfecta' que asigna probabilidad 1 a la respuesta correcta. Es fácil demostrar que, con esta función de evaluación, la estrategia óptima del alumno es decir la verdad; en efecto, su puntuación *esperada* si escribe $(q_1, q_2, \ldots, q_k)$ cuando realmente piensa $(p_1, p_2, \ldots, p_k)$ es

$$\sum_{i=1}^{k} u[(q_1, q_2, \ldots, q_k), i]\, p_i$$

que, para la función de evaluación cuadrática, se maximiza si $q_i = p_i, i = 1, \ldots, k$, y sólamente en ese caso. Si se decide calificar con 1 punto a una contestación perfecta (probabilidad uno asignada a la respuesta correcta) y con 0 puntos a la ignorancia absoluta (distribución uniforme), entonces $a = k/(k-1)$, $b = -1/(k-1)$ y la función de evaluación resulta ser

$$u[(p_1, p_2, \ldots, p_k), i] = 1 - \frac{k}{k-1}\, \delta_i^2, \quad i = 1, \ldots, k$$

con lo que la puntuación que merece el 'disparate' de asignar probabilidad 1 a una respuesta falsa (¡mucho peor que confesar ignorancia!) resulta ser $-(k+1)/(k-1)$, que tiende a -1 cuando el número de respuestas posibles tiende a infinito, pero es, por ejemplo, igual a -2 si sólamente existen tres respuestas alternativas, e igual a -3 si únicamente se trata de predecir si una afirmación es cierta o falsa.

La función de evaluación cuadrática, atribuida a Brier por McCarthy (1956), fué independientemente redescubierta por de Finetti, quién la propuso como un método de elicitación de probabilidades subjetivas, —posteriormente generalizado por Lindley (1982)—, y la utilizó frecuentemente para valorar predicciones reales sobre la quiniela italiana (de Finetti, 1982).

Una función de evaluación que motive al alumno a decir la verdad recibe el nombre de función de evaluación *propia*. La función de evaluación cuadrática no es, desde luego, la única función de evaluación propia. Por ejemplo, la función de evaluación logarítmica

$$u[(p_1, p_2, \ldots, p_k), i] = a \log p_i + b_i, \quad a > 0, \quad b_i \in \Re$$

que toma la forma $1 - (\log p_i)/(\log k^{-1})$ si se asigna 1 punto a una respuesta perfecta y 0 puntos a una distribución uniforme, es una función de evaluación propia y es, además, la única que depende *exclusivamente* de la probabilidad asignada a la respuesta correcta (Savage, 1971). Sin embargo, se ha conjeturado que la función de evaluación cuadrática es la *única* que satisface la condición de motivar una respuesta veraz *sin* exigir que la respuesta dada $(q_1, q_2, \ldots, q_k)$ sea una distribución de probabilidad; tiene además la ventaja de estar acotada, mientras que con la función de evaluación logarítmica, la puntuación que se obtiene si se asigna probabilidad 1 a una respuesta falsa es $-\infty$.

La generalización de las ideas expuestas al caso de una variable contínua, posiblemente vectorial, tiene importantes implicaciones, y puede servir como ejemplo de las poderosas implicaciones que las ideas originales del *Maestro* pueden llegar a tener.

En el caso discreto, la función de evaluación cuadrática propuesta por de Finetti para valorar una predicción $(p_j, j \in J)$ puede reescribirse en la forma

$$u[(p_1, p_2, \ldots), i] = a(2p_i - \sum_{j \in J} p_j^2) + b_i, \quad a > 0, \quad b_i \in \Re.$$

Para evaluar una predicción $p_{\boldsymbol{x}}(.\,|\,H)$ sobre el valor de $\boldsymbol{x} \in X$ en las condiciones H basta, análogamente, calcular

$$u[p_{\boldsymbol{x}}(.\,|\,H), \boldsymbol{x}] = a\Big(2\,p_{\boldsymbol{x}}(\boldsymbol{x}\,|\,H) - \int_X p_{\boldsymbol{x}}^2(\boldsymbol{x}\,|\,H)\,d\boldsymbol{x}\Big) + b(\boldsymbol{x}), \quad a > 0, \quad b \in \{X \to \Re\},$$

que sigue siendo una función de evaluación propia.

Las funciones de evaluación propias proporcionan un potente *método de contraste* para juzgar si el par formado por un modelo probabilístico y una distribución inicial sobre sus parámetros describe o no adecuadamente el comportamiento de los datos. En efecto, dado un conjunto de datos $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n\}$, un modelo probabilístico y una distribución inicial sobre sus parámetros, el valor *medio* de la evaluación que merece la predicción que puede hacerse sobre cada una de las observaciones en base a todas las demás, esto es

$$t = t(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n) = \frac{1}{n} \sum_{i=1}^{n} u[\,p_{\boldsymbol{x}}(.\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n),\ \boldsymbol{x}_i],$$

donde u es cualquier función de evaluación propia, y $p_{\boldsymbol{x}}(.\,|\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n)$ es la distribución predictiva de una nueva observación $\boldsymbol{x}$ basada en todas las observaciones disponibles excepto la $\boldsymbol{x}_i$, es una estimación por Monte Carlo del valor esperado de la evaluación que merecería una predicción sobre $\boldsymbol{x}$ basada en los datos $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_n\}$ y en el modelo y la distribución inicial asumidos. Consecuentemente, t proporciona un *estadístico de contraste global* intuitivamente razonable, una medida intrínseca de la 'bondad' del par (modelo probabilístico, distribución inicial) para explicar el comportamiento de las $\boldsymbol{x}_i$'s.

Para comparar entre sí un conjunto de modelos probabilísticos y/o distribuciones iniciales alternativos, basta calcular el valor de t que se deduce de cada uno de ellos; la mejor combinación (modelo probabilistico, distribución inicial) será aquella que minimiza el valor de t (Bernardo & Smith, 1994, Cap 6).

5. AGRADECIMIENTOS

El autor agradece a los Profesores Donato Cifarelli, Daniela Cochi, Ludovico Piccinato y Eugenio Regazzini su inestimable colaboración en el acopio de los materiales necesarios para preparar este trabajo, y al Dr. Raúl Rueda sus valiosos comentarios a una primera versión del manuscrito.

REFERENCIAS

Berger, J. O. y Bernardo, J. M. (1992). On the development of reference priors. *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). Oxford: University Press, 35–60 (con discusión).

Bernardo, J. M. (1979b). Reference posterior distributions for Bayesian inference. *J. Roy. Statist. Soc. B* **41**, 113–147 (with discussion). Reprinted in *Bayesian Inference* (N. G. Polson and G. C. Tiao, eds.). Brookfield, VT: Edward Elgar, 1995, 229–263.

Bernardo, J. M. (1997). Noninformative priors do not exist. *J. Statist. Planning and Inference* (pendiente de publicación, con discusión).

Bernardo, J. M. and Ramón, J. M. (1998). An introduction to Bayesian reference analysis: inference on the ratio of multinomial parameters. *The Statistician* **47**, (pendiente de publicación).

Bernardo, J. M. y Smith, A. F. M. (1994). *Bayesian Theory*. Chichester: Wiley.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: University Press.

Cifarelli, D. M. y Regazzini, E. (1987). Priors for exponential families which maximize the association between past and future observations. *Probability and Bayesian Statistics* (R. Viertl, ed.). London: Plenum, 83–95.

Cifarelli D. M. y Regazzini, E. (1995). De Finetti's contribution to probability and statistics. *Proc. Rimini Conference on Bayesian Robustness* (J. O. Berger *et al.* , eds.). *IMS Lectures Notes* **29**, Hayward, CA: IMS,1–64.

Consonni, G. y Veronese, P. (1989). Some remarks on the use of improper priors for the analysis of exponential regression problems. *Biometrika* **76**, 101–106.

Daboni, L. (1984). Few reflexions on the work of the Master. *Riv. Mat. Sci. Econ. Soc.* **7**, 5–13.

Daboni, L. *et al.* (1987). *Atti del Convegno 'Ricordo di Bruno de Finetti, Professore nell'Ateneo Trientino'*. Trieste: Dip. Matematica Applicata "Bruno de Finetti".

de Finetti, B, (1926). Considerazioni matematiche sull'eredità mendeliana. *Metron* **6**, 3–41. Reimpreso en de Finetti (1981), 1–41.

de Finetti, B, (1930). Funzione caratteristica di un fenomeno aleatorio. *Memorie della R. Accademia dei Lincei* **6** 86–133. Reimpreso en de Finetti (1981), 265–315.

de Finetti, B, (1931a). Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza. *Biblioteca di Filosofia* (A. Alivera, ed.). Napoli: Pezzella, 163–219. Reimpreso en de Finetti (1991), 81–143.

de Finetti, B, (1931b). Sul significato soggetivo della probabilità. *Fundamenta Matematicae* **17**, 298–329. Reimpreso en de Finetti (1991), 145–178. Reimpreso en de Finetti (1993), 31–61 (en italiano) y 291–321 (en inglés)

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7**, 1–68. Traducido como 'Foresight; its logical laws, its subjective sources', *Studies in Subjective Probability* (H. E. Kyburg and H. E Smokler, eds.). New York: Dover (1980), 93–158. Reimpreso en Kotz y Johnson (1992), 127-174, con una introducción por R. E. Barlow.

de Finetti, B. (1938a). Sur la condition d'équivalence partielle. *Actualités Scientifiques et Industrielles* **739**. Paris: Herman et Cii. Traducido en *Studies in Inductive Logic and Probability* **2** (R. Jeffrey, ed.), 1980, Berkeley: Univ. California Press, 193–206.

de Finetti, B. (1938b). Probabilisti di Cambridge. *Suppl. Statist. Nouvi Problemi Politica Storia ed Economia* **4**, 21-37. Traducido como 'Cambridge probability theorists', *Riv. Mat. Sci. Econ. Soc.* **8**, (1985), 79–91.

de Finetti, B, (1944). *Matematica Logico-Intuitiva. (Nozioni di matematiche complementari e di calcolo differenziale e integrale come introduzione agli studi di Scienze Economiche, Statistiche a Attuariali)*. Trieste: EST. Segunda ed. 1957, Roma: Cremonese. Tercera ed. 1959, Roma: Cremonese.

de Finetti, B. (1951). Recent suggestions for the reconciliation of theories of probability. *Proc. Second Berkeley Symp.* (J. Neyman ed.). Berkeley: Univ. California Press, 217–226. Reimpreso en Finetti (1993), 115–127 (en italiano) y 375–387 (en inglés)

de Finetti, B. (1961). The Bayesian approach to the rejection of outliers. *Proc. Fourth Berkeley Symp.* **1** (J. Neyman and E. L. Scott, eds.). Berkeley: Univ. California Press, 199–210.

de Finetti, B. (1962). Does it make sense to speak of 'Good Probability Appraisers'? *The Scientist Speculates: An Anthology of Partly-Baked Ideas* (I. J. Good, ed.). New York: Wiley, 257–364. Reimpreso en de Finetti (1972), 19–23.

de Finetti, B. (1963). La décision et les probabilitiés. *Rev. Roumaine Math. Pures Appl.* **7**, 405–413.

de Finetti, B. (1964). Probabilità subordinate e teoria delle decisioni. *Rendiconti Matematica* **23**, 128–131. Traducido como 'Conditional probabilities and decision theory', de Finetti (1972), 13–18.

de Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item. *British J. Math. Statist. Psychol.* **18**, 87–123. Reimpreso en de Finetti (1972), 25–63.

de Finetti, B. (1967). Logical foundations and measurement of subjective probability. *Acta Psychologica* **34**, 129–145.

de Finetti, B. (1968). Probability: interpretations. *Internat. Encyclopedia of the Social Sciences*, **12**. London: Macmillan, 496–504.

de Finetti, B, (1970a). *Teoria delle Probabilità (Sintesi introdutiva con appendice critica)* (2 vol.). Torino: Einaudi. Traducido como *Theory of Probability*. Chichester: Wiley, 1975 (vol. 1) y 1976 (Vol. 2). Traducido como *Wahrscheinlichkeitstheorie*. Wien: Oldenburg Verlag, 1981.

de Finetti, B, (1970b). Logical foundations and measurement of subjective probability. *Acta Psychologica* **34**, 129–145.

de Finetti, B. (1972). *Probability, Induction and Statistics*. Chichester: Wiley.

de Finetti, B, (1974). Bayesianism: its unifying role for both the foundations and applications of statistics *Internat. Statist. Rev.* **42**, 117–130. Reimpreso en de Finetti (1993), 205–228 (en italiano) y 467–490 (en inglés)

de Finetti, B, (1976). La probabilità: guardarsi dalle contraffazioni. *Scientia* **111**, 255-281. Traducido como 'Probability: Beware of falsifications!', *New Developments in the Applications of Bayesian Methods* (A. Aykaç and C. Brumat, eds.). Amsterdam: North-Holland(1977), 347–379. Reimpreso en *Studies in Subjective Probability* (H. E. Kyburg y H. E. Smokler, eds.). Melbourne, FL: Krieger (1980), 193–224.

de Finetti, B, (1977a). Probabilities of probabilities: a real problem or a misunderstanding? *New Developments in the Applications of Bayesian Methods* (A. Aykaç and C. Brumat, eds.). Amsterdam: North-Holland, 1–10.

de Finetti, B, (1977b). Inferenza statistica Bayesiana. *Atti Convegno su I Fondamenti dell'Inferenza Statistica*. Firenze: Parenti, 166-177. Reimpreso en de Finetti (1993), 249–260 (en italiano) y 513–424 (en inglés)

de Finetti, B. (1978). Probability: interpretations. *Internat. Encyclopedia of Statistics* (W. H. Kruskal y J. M. Tanur, eds.) London: Macmillan, 496–505.

de Finetti, B, (1979). Discussion of Bernardo (1979). *J. Roy. Statist. Soc. B* **41**, p. 135.

de Finetti, B, (1981). *Scritti* [1926–1930]. Padova: Cedam.

de Finetti, B, (1982). Probability and my life. *The Making of Statisticians* (J. Gani, ed.) Berlin: Springer, 3–12.

de Finetti, B, (1989). *La Logica dell'Incerto*. (M. Mondatori, ed.). Milano: Il Saggiatore.

de Finetti, B, (1991). *Scritti* [1931–1936]. Bologna: Pitagora.

de Finetti, B, (1993). *Probabilità e Induzione. Induction and Probability*. (P. Monari y D. Cocchi, eds.). Bologna: Clueb

de Finetti, B, (1995). *Filosofia della Probabilità*. (A. Mura, ed.). Milano: Il Saggiatore.

Diaconis, P., Eaton, M. L. y Lauritzen, S. L. (1992). Finite de Finetti theorems in linear models and multivariate analysis. *Scandinavian J. Statist.* **19**, 289–316.

Diaconis, P. and Freedman, D. (1980). De Finetti generalizations of exchangeability. *Studies Inductive Logic and Probability* (Jeffrey, R. C. ed.). Berkeley: Univ. California Press, 223–249.

Diaconis, P. and Freedman, D. (1984). Partial exchangeability and sufficiency. *Statistics: Applications and New Directions* (J. K. Ghosh and J. Roy, eds.). Calcutta: Indian Statist. Institute, 205–236.

Fürst, D. (1993). Bruno de Finetti and the teaching of mathematics. *Induction and Probability*. (P. Monari y D. Cocchi, eds.). Bologna: Clueb, 277–288.

Gatsonis, C. A., Hodges, J. S., Kass, R. E. y Singpurwalla, N. (eds.) (1993). *Case Studies in Bayesian Statistics*. Berlin: Springer.

Gatsonis, C. A., Hodges, J. S., Kass, R. E. y Singpurwalla, N. (eds.) (1995). *Case Studies in Bayesian Statistics 2*. Berlin: Springer.

Gatsonis, C. A., Hodges, J. S., Kass, R. E. y Singpurwalla, N. (eds.) (1997). *Case Studies in Bayesian Statistics 3*. Berlin: Springer.

Gelman, A., Carlin, J. B., Stern, H. y Rubin, D.B. (1995). *Bayesian Data Analysis*. London: Chapman and Hall.

Goel, P. K. y Zellner, A. (eds.) *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*. Amsterdam: North-Holland.

Good, I. J. (1965). *The Estimation of Probabilities. An Essay on Modern Bayesian Methods*. Cambridge, Mass: The MIT Press.

Good, I. J. (1980). Some history of the hierarchical Bayesian mehodology. *Bayesian Statistics* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.). Valencia: University Press, 489–519, (con discusión).

Heath, D. L. y Sudderth, W. D. (1978). On finitely additive priors, coherence and extended admissibility. *Ann. Statist.* **6**, 333–345.

Heath, D. L. y Sudderth, W. D. (1989). Coherent inference from improper priors and from finitely additive priors. *Ann. Statist.* **17**, 907–919.

Hewitt, E. and Savage, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80**, 470–501.

Jeffrey, R. (1993). De Finetti's radical probabilism. *Induction and Probability*. (P. Monari y D. Cocchi, eds.). Bologna: Clueb, 263–275.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan. Segunda ed. 1929, London: Macmillan. Reimpreso en 1962, New York: Harper and Row.

Kingman, J. F. C. (1978). Uses of exchangeability. *Ann. Prob.* **6**, 183–197.

Koch, G. y Spizzichino, F. (eds.) (1982). *Exchangeability in Probability and Statistics*. Amsterdam: North-Holland.

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer. Traducido como *Foundations of the Theory of Probability*, 1950, New York: Chelsea.

Lindley, D. V. (1975). Preface to *Theory of Probability*, traducción al inglés de *Teoria delle Probabilità* (de Finetti, 1970a). Chichester: Wiley, i–ix.

Lindley, D. V. (1982). Scoring rules and the inevitability of probability. *Internat. Statist. Rev.* **50**, 1–26 (con discusión).

Lindley, D. V. (1986). Obituary. Bruno de Finetti, 1906–1985. *J. Roy. Statist. Soc. A* **149**, p. 152.

Lindley, D. V. (1987). de Finetti, Bruno. *Encyclopedia of Statistical Sciences* **Supp.** (S. Kotz, N. L. Johnson and C. B. Read, eds.). New York: Wiley, 46–47.

Lindley, D. V. (1997). Some comments on Bayes Factors. *J. Statist. Planning and Inference* **61**, 181–189.

Lindley, D. V. y Phillips, L. D. (1976). Inference for a Bernoulli process (a Bayesian view). *Amer. Statist.* **30**, 112–119.

Lindley, D. V. y Smith, A. F. M. (1972). Bayes estimates for the linear model. *J. Roy. Statist. Soc. B* **34**, 1–41 (con discusión).

McCarthy, J. (1956). Measurements of the value of information. *Proc. Nat. Acad. Sci. USA* **42**, 654–655.

Muhere P. y Parmigiani, G. (1993). Utility and means in the 30's. *Statist. Sci.* **8**, 421–432.

Piccinato, L. (1986). De Finetti's logic of uncertainty and its impact on statistical thinking and practice. *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti* (P. K. Goel and A. Zellner, eds.). Amsterdam: North-Holland, 13–30.

Regazzini, E. (1987). Probability theory in Italy between the two wrold wars: a brief historical survey. *Metron* **44**, 5–42.

Savage, L. J. (1954). *The Foundations of Statistics.* New York: Wiley. Second edition in 1972, New York: Dover.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc.* **66**, 781–801. Reimpreso en *Studies in Bayesian Econometrics and Statistics: in Honor of Leonard J. Savage* (S. E. Fienberg and A. Zellner, eds.). Amsterdam: North-Holland, 1974, 111–156.

Savage, L. J. (1981). *The Writings of Leonard Jimmie Savage: a Memorial Collection*. Washington: ASA/IMS.

Smith, A. F. M. (1981). On random sequences with centred spherical symmetry. *J. Roy. Statist. Soc. B* **43**, 208–209.