De Finetti's Contribution to Probability and Statistics

Donato Michele Cifarelli, Eugenio Regazzini

Statistical Science, Volume 11, Issue 4 (Nov., 1996), 253-282.

Stable URL:
http://links.jstor.org/sici?sici=0883-4237%28199611%2911%3A4%3C253%3ADFCTPA%3E2.0.CO%3B2-T

Statistical Science
1996, Vol. 11, No. 4, 253–282

De Finetti's Contribution to Probability and Statistics

Donato Michele Cifarelli and Eugenio Regazzini

Abstract. This paper summarizes the scientific activity of de Finetti in probability and statistics. It falls into three sections: Section 1 includes an essential biography of de Finetti and a survey of the basic features of the scientific milieu in which he took the first steps of his scientific career; Section 2 concerns de Finetti's work in probability: (a) foundations, (b) processes with independent increments, (c) sequences of exchangeable random variables, and (d) contributions which fall within other fields; Section 3 deals with de Finetti's contributions to statistics: (a) description of frequency distributions, (b) induction and statistics, (c) probability and induction, and (d) objectivistic schools and theory of decision. Many recent developments of de Finetti's work are mentioned here and briefly described.

Key words and phrases: Associative mean, Bayesian nonparametric statistics, Bayes–Laplace paradigm, completely additive probabilities, correlation and monotone dependence, exchangeable and partially exchangeable random variables, finitely additive probabilities, gambler's ruin, Glivenko–Cantelli theorem, infinitely decomposable laws, predictive inference, prevision, principle of coherence processes with independent increments, reasoning by induction, statistical decision, subjective probability, utility function

This paper includes the text of two lectures that we had the honor to give at the Second International Workshop on Bayesian Robustness, on May 22, 1995.

0. INTRODUCTION

This is to mark the tenth anniversary of de Finetti's death on 20 July 1985. We were thus deprived of his sincere and discreet friendship, of his critical and constructive insights, of his passionate and disinterested battle for a fairer social order. He is still very much with us, though, with his impressive scientific corpus scattered in more than 290 writings.

Donato Michele Cifarelli and Eugenio Regazzini are Professors of Statistics, Istituto di Metodi Quantitativi, Università "L. Bocconi," v. Sarfatti 25, 20136 Milano. Donato Michele Cifarelli is Head of this Institute and Eugenio Regazzini is Director, Istituto per le Applicazioni della Matematica e dell'Informatica (IAMI-CNR), v. Ampère 56, 20131 Milano.

The Scientific Committee of the Second International Workshop on Bayesian Robustness thought it fitting to devote a special session to a retrospective analysis of de Finetti's scientific endeavor. We have the honor and pleasure of delivering two lectures on de Finetti's contributions to probability and statistics. In point of fact, we have partly enlarged the scope of our original plan by inserting also a survey of the Italian scientific circle which was closest to probability and statistics when de Finetti embarked on his venture into the subjects.

Recently, Barlow has written that the publication of de Finetti's paper La prévision (1937a) was the first major event that brought about the rebirth of Bayesian statistics, and that:

> Despite the growing number of papers and books which have been influenced by de Finetti, his overall influence is still minimal. Why is this? Perhaps one reason is communication. De Finetti's style of writing is difficult to understand, even for Italian mathematicians and few English-speaking mathematicians have

> really tried to interpret what he has written. [See Barlow, 1992, page 131.]

As a matter of fact, the language of Bruno de Finetti is elaborate, full of nuances and perfectly appropriate for the manifestation of his sharp way of thinking. It is then difficult to translate it into a language other than the original without losing some of its peculiar stylistic aspects or introducing some simplifications which may distort the original meaning. Even if this kind of imperfection is sometimes present in translations of de Finetti's works, yet we are grateful to the scholars who have helped bring de Finetti's contributions to the attention of the international scientific community.

We have taken upon ourselves the task of summarizing the scientific activity of de Finetti in probability and statistics, hoping to help spread his major ideas and change the picture portrayed by Barlow. In going through de Finetti's original papers, we have found it convenient not to linger on well-known works such as *Teoria delle Probabilità* (1970), but rather on a number of less known earlier papers which blazed the trail of de Finetti's scientific venture.

To get one's bearings in his copious production is no easy matter, and we are indebted to previous work, mentioned in the General References section, for a number of bibliographic hints. In particular, de Finetti's complete bibliography can be found in Daboni (1987).

Many recent developments of de Finetti's work are mentioned here and briefly described, as are the lengthier developments, in *note* within the text.

1. HISTORICAL NOTES

1.1 Essential Biography

De Finetti was born of Italian parents on 13 June 1906 in Innsbruck (Austria). At 17 he enrolled at the Polytechnic of Milan with a view to obtaining a degree in engineering. During his third year of study (in Italy engineering is a five-year course), a Faculty of Mathematics was opened in Milan. He attended a few classes there, and, following his more theoretical bent, shifted subjects. In 1927 he obtained a degree in applied mathematics. He discussed his graduation dissertation on affine geometry with Giulio Vivanti, a mathematician known for some noteworthy contributions to complex analysis. Soon afterwards, de Finetti accepted a position in Rome, at the Istituto Centrale di Statistica, presided over, at that time, by an outstanding Italian statistician: Corrado Gini. De Finetti worked there until 1931. In those years, he laid the foundations for his principal contributions to probability theory and statistics: the subjective approach to probability; definition and analysis of sequences of exchangeable events; definition and analysis of processes with stationary independent increments and infinitely decomposable laws; theory of mean values. In 1930 de Finetti qualified in a competition as a university lecturer of Mathematical Analysis.

In 1931, he moved to Trieste, where he accepted a position as actuary at the Assicurazioni Generali. Trieste was included in the Austro-Hungarian Empire until 1918. De Finetti's father and grandfather had worked as engineers there. In 1906 his father was actually engaged in the construction of new railway lines in the Innsbruck area (Upper Tyrol). During his stay in Trieste, de Finetti developed the research he had started in Rome, but he also obtained significant results in actuarial and financial mathematics as well as in economics. In addition, he was active in the mechanization of some actuarial services. It was probably this operational background which later enabled him to understand the revolutionary impact of the computer on scientific calculus. As a matter of fact, he was one of the first mathematicians in Italy able to solve problems of numerical analysis by means of a computer (cf. Fichera, 1987). In spite of his impending actuarial activity, in Trieste de Finetti also started his teaching career (mathematical analysis, actuarial and financial mathematics, probability), with a few years's stint at the renowned University of Padua. However, it was only in 1947 that he obtained his chair as full Professor of Financial Mathematics in Trieste. He had actually won it in a nationwide competition back in 1939, but at that time he was unmarried and, by a law in force in those days in Italy, bachelors were not permitted to hold any position in the public service. Such a law was cancelled by the Allies, who ruled over Trieste from the end of the war (April 1945) until 1954.

In 1954 de Finetti moved to the Faculty of Economics at the University of Rome "La Sapienza." In 1961, he changed to the Faculty of Sciences in Rome, where he was Professor of the Theory of Probability until 1976.

1.2 Probability and Statistics in Italy During the 1920s and 1930s

De Finetti did not attend any course in probability or statistics. With the exception of the University of Rome, in the 1920s no Italian faculty of mathematics included those subjects in its course of studies. Around 1925, de Finetti was attracted by probability, indirectly, through the reading of a paper on Mendelian heredity by C. Foà, a biologist. De

Finetti started research on the subject, reaching results that Foà, Vivanti and Mortara (a statistician, at that time, active in Milan) considered of interest. Mortara submitted part of de Finetti's research for publication in *Metron*, to Gini (editor and owner of that journal), who published the paper in an issue of 1926. Other parts of the same research were published in the proceedings of the Reale Accademia dei Lincei in 1927, with nomination by Foà. These are the first three items in the long list of de Finetti's scientific works.

The above-mentioned genetics problem led him to go deeper into probability and statistics. By the way, from some of his written recollections, we know that he studied probability in Czuber's *Wahrscheinlichkeitslehre* and in Castelnuovo's *Calcolo delle Probabilità e Applicazioni* (Castelnuovo, 1919). The former is well known to statisticians, while the latter is more renowned among pure mathematicians, following his outstanding contributions to geometry. In point of fact, Castelnuovo played a fundamental role in introducing probability into the culture of official Italian mathematical circles.

Thanks to the initiative of Guido Castelnuovo (1865–1952), the Faculty of Mathematics at the University of Rome, starting from 1915, introduced courses of probability and actuarial mathematics. They represented the germ of the School of Statistics and Actuarial Sciences, officially established in 1927. Castelnuovo was the first head of that school. Among its students, a few became well known for their remarkable contributions to probability theory and its applications, for example, Onicescu and Mihok. In 1935 that school merged with the Special School of Statistics imbued with Gini's line of thinking in statistics, to found the present-day Faculty of Statistics in Rome. That very same year, Castelnuovo gave up teaching. In 1938, like all Italian Jews he bore the brunt of Fascist racial laws.

The first edition of Castelnuovo's book on probability dates back to 1919. Considering the time it was conceived, this treatise looks quite accurate from the mathematical angle, and most up-to-date. In particular, it bears witness to the influence of the Russian School of St. Petersburg (Chebyshev, Markov, Lyapunov) on probability theory. Castelnuovo's stance as to the foundations of probability can be included under the heading of the *empirical approach* (empirical law of chance), the same position as Lévy's, Fréchet's and Cantelli's. Cantelli was actually Castelnuovo's most important collaborator in the drafting of the book. As far as the relationship between probability and inductive reasoning is concerned, Castelnuovo went deeper into the subject. In compliance with his approach to probability, he rejected the Bayesian formulation of the problem at issue. He proposed to deal with the question of rejecting an isolated hypothesis on the ground of its sole likelihood. In this sense, he could be considered as a forerunner of Fisher's viewpoint on inductive reasoning. In any case, Castelnuovo, because of his preference for pure mathematics, refrained from taking sides in the controversy on the foundations of probability and statistics. He disinterestedly helped not only scholars who followed the main lines of his approach, but also those, like de Finetti, who supported alternative viewpoints; see de Finetti (1976).

Francesco Paolo Cantelli (1875–1966) was the first "modern" Italian probabilist. His name is definitively linked with some fundamental results about the convergence of sequences of random variables: the very famous 0–1 law, strong laws of large numbers (Cantelli, 1917a, b), the fundamental theorem of mathematical statistics (Cantelli, 1933b) and a formulation of the law of the iterated logarithm (Cantelli, 1933a) which extends Kolmogorov's version of the same law, under suitable conditions; see Epifani and Lijoi (1995).

Cantelli's point of view about the interpretation of probability substantially agrees with Castelnuovo's. In fact, he had no genuine interest in the controversy about the foundations of probability. Cantelli's conception of science was extremely positive and operative, in compliance with his position as an actuary at a very high international standard level. Here is how Harald Cramér portrays Cantelli (Wegman, 1986):

> I should mention also another name, the Italian mathematician Cantelli, with whom we also had contact in our Stockholm probabilistic group. Cantelli, like myself, was also working as an actuary. As a matter of fact, he had been, among other things, the actuary of the pension board of what was then called the Society of Nations in Geneva. When he resigned his position, I became his successor. That gave me a contact with Cantelli which I value very much. You know his name from the Borel–Cantelli condition. He had written several very valuable papers on probability, papers which have perhaps not received quite the attention that they really do deserve. He was a very temperamental man. When he was excited, he could cry out his views with his powerful voice: a very energetic fellow.

The extroversion and the determination of Cantelli's attitudes are confirmed by Gaetano Fichera. They set off the low profile and shy attitudes of de Finetti, so as to conclude that Cantelli was the photographic negative of de Finetti; see Fichera (1987). But, although Cantelli had no inclination for philosophical speculation, he stressed the necessity for a formal definition of probability. It could be taken as the starting point for a rigorous construction of probability theory considered as a formal theory. On this subject, he formulated an *abstract theory* of probability—shortly before the publication, in 1933, of Kolmogorov's *Grundbegriffe*—whose axioms essentially coincide with those propounded by the great Russian mathematician; see Cantelli (1932) and Ottaviani (1939).

We cannot conclude these notes without quoting Cantelli's initiative leading to the publication of the *Giornale dell'Istituto Italiano degli Attuari*, in 1930. During his editorship, the *Giornale* was one of the most prestigious journals dealing with probability, statistics and actuarial mathematics, and could count on the collaboration of outstanding scholars. In particular, a number of de Finetti's writings appeared there, although none of them actually deal with the foundations of probability. An explanation for this can probably be found in Fichera's above-quoted paper, when he recollects that speaking to Cantelli about subjective probability and finitely additive probabilities was tantamount to pulling a tiger by its tail.

Passing from probability to statistics, the Italian statistical panorama during the 1920s and 1930s was dominated by Corrado Gini (1884–1965). As a matter of fact, Gini deeply influenced Italian statistics until the end of the 1960s. His cultural interests were deep and broad, so that he handed down striking applications of statistics to many aspects of the real world. He was accustomed to gearing his statistical methodology toward the specific problem involved under each circumstance. This approach—setting up instruments and techniques best befitting specific targets—was always actually oriented by an implicit intuitive overall vision. Indeed, starting from the 1940s, he concentrated on "a critical work of synthesis drawing its form and substance from the particular techniques worked out"; see Castellano (1965). According to this author, the main works of Gini's early period may be grouped under several headings, namely, the theory of means, the theory of variability and the theory of statistical relations. In particular, let us recall Gini's proposal of the *mean difference* Δ as a general measure of the variability in a frequency distribution F and of the *concentration ratio* G (obtained by dividing the mean difference by twice the arithmetic mean) as a measure of the particular aspect of variability, for nondegenerate distributions with positive support, known as concentration:

$$\Delta = \int_{\mathbb{R}^2} |x - y| \, dF(x) \, dF(y)$$

$$G = \frac{\Delta}{2\mu},$$

where

$$F(x) = 0 \quad \text{for } x < 0$$

and

$$\mu = \int_{[0, +\infty)} x \, dF(x)$$

(cf. Stuart and Ord, 1994). Gini introduced Δ in 1912 as a measure of the variability, independently of previous works of the astronomers Jordan, von Andrae and Helmert who used the mean difference for quite different reasons.

In a series of papers dating back to the quadriennium 1914–1917, Gini shaped a systematical methodology for the analysis of relations between two statistical characteristics, which brought clarity and order into a field to whose progress outstanding contributions had already been made by Bravais, Galton, Pearson and Yule.

Insofar as statistical inference is concerned, Gini deemed Fisher's approach logically weak and opposed it vigorously. Nonetheless, he refused the Neyman–Pearson approach. In the above-mentioned paper, Castellano summarizes Gini's position in the following terms:

> With penetrating logic Gini proved... how important and unavoidable are prior probabilities in any judgement on the measures deriving from a sample; he restored to induction its essential character of a conclusion based upon an experience (the facts) and an independent and preliminary a priori assumption by which the facts can be interpreted.

Gini rendered explicit this position, which basically supports the Bayes–Laplace paradigm, in 1939, on the occasion of the first meeting of the Società Italiana di Statistica, in an address titled *I pericoli della statistica*. As a matter of fact, in 1911, he had made an application of that paradigm to the study of the sex ratio at birth, after determining the conditional probability of k_2 successes in n_1 future trials given k_1 successes in n_1 past trials, when the prior distribution of the unknown Bernoulli parameter is beta. On the other hand, consistently with his objective interpretation of probability, Gini developed a skeptical attitude toward the value of statistical inferences. This, in addition to a penetrating and proud freedom of thought, prevented him from being a follower of the neo-Bayesian way of thinking. He maintains that probability is a frequency: not an ordinary frequency, nor a limiting frequency at that. Probability is the frequency of a phenomenon in the totality of cases in which that very phenomenon could occur. In any case, what is involved is a finite collective, defined as the class of all cases with respect to which we deem it fit to regulate our conduct.

The above survey should give you an idea of the studies in probability and statistics in Italy in the days when de Finetti began tackling those issues, and, consequently, draw your attention to his orthogonal position with respect to the prevailing way of thinking. In 1976, with reference to Savage's cooperation, de Finetti wrote:

> I must stress that I owe it to him if my work is no longer considered as a blasphemous but harmless heresy, but as a heresy with which the official statistical church is being compelled, unsuccessfully, to come to terms.

To complete the picture, let us now briefly analyze the connections between the small coterie of Italian scholars dealing with probability and statistics and foreign schools.

1.3 Connections with Foreign Schools

De Finetti started dealing with probability and statistics in a period of tremendous developments for the subjects. For instance, Kolmogorov (1903–1987) and Lévy (1886–1971) were making their decisive contribution to the modern theory of probability, and Fisher (1890–1962) was setting out the basic technical concepts for his new approach to statistics. Castelnuovo, Cantelli, Gini and young de Finetti got involved in this impressive cultural revival: on one side, with their first-hand scientific contributions, and, on the other side, with the opportunity they gave to distinguished foreign scholars of publishing the new results of their research in Italian scientific journals.

For example in *Metron*, the journal founded by Gini in 1920, Fisher gave striking examples of his art of calculating explicit sampling distributions, providing the distribution of F and of its logarithm Z (Fisher, 1921) and an interesting Note on Student's distribution (Fisher, 1925). Romanowsky, in two papers appearing in 1925 and 1928, respectively, determined the moments of the standard deviation and of the correlation coefficient in normal samples, and masterfully dealt with a criterion that two given samples belong to the same normal distribution.

Castelnuovo presented three important papers of Kolmogorov's for publication in *Atti della Reale Accademia dei Lincei*, from 1929 to 1932, dealing with the strong law of large numbers (Kolmogorov, 1929), the general representation of an associative mean in a discrete distribution (Kolmogorov, 1930) and the representation of the characteristic function of an infinitely decomposable law with finite variance (Kolmogorov, 1932a, b). The last two papers are closely connected with de Finetti's research. In particular, they originate from de Finetti's studies on processes with independent increments and, in turn, represent the starting point of Lévy's fundamental paper, published in 1934 in *Annali della Scuola Normale Superiore di Pisa*. This is the paper in which Lévy—in addition to the generalization of the Kolmogorov representation—states the renowned decomposition of a process with independent increments.

The publication of the *Giornale dell'Istituto Italiano degli Attuari* (*GIIA*), starting from 1930 under Cantelli's editorship, made the submission of foreign works more matter-of-fact. Let us mention Glivenko's paper, which appeared in 1933 in *GIIA*, dealing with the almost-sure uniform convergence of the empirical distribution function F_n to F, in a sequence of i.i.d. random variables with common continuous distribution function F. In the same year, *GIIA* published Kolmogorov's paper in which the asymptotic distribution of $\sup_x \sqrt{n}|F_n(x)-F(x)|$ is stated. In 1935, *GIIA* accepted Neyman's paper on the factorization theorem. An interesting result reached by Romanowsky, dealing with the asymptotic behaviour of the posterior distribution, is included in an issue of 1931. Other first-rate contributions in the domain of probability which appeared in *GIIA* in the 1930s are noteworthy: Khinchin (1932a), on stationary sequences of random events; Khinchin (1935), including the characterization of the domain of attraction of the normal law; Khinchin (1936), dealing with a strong law of large numbers; Lévy (1931a), on the strong law of large numbers; Lévy (1935), on the application of the geometry of Hilbert spaces to sequences of random variables; Cramér (1934), focussing on a version of the strong law of large numbers; Cramér (1935), providing asymptotic expansions of a distribution function in series of Hermite polynomials; Glivenko (1936), on a useful refinement of Lévy's continuity theorem; Slutsky (1934), on some basic results on stationary (in a wide sense) processes;

and Slutsky (1937), on the foundations of the theory of random functions according to a point of view similar to that in Doob (1953). This very same paper includes the original version of the celebrated criterion for the continuity of a random function due to Kolmogorov.

Let us conclude with a look at the International Congress of Mathematicians held in Bologna, in 1928. Among the statisticians and probabilists who attended that congress, let us mention B. Hostinsky, M. Fréchet, É. Borel, P. Lévy, G. Darmois, C. Gini, C. Bonferroni, G. Castelnuovo, J. Neyman, A. Lomnicki, S. Bernstein, G. Pólya, F. P. Cantelli, B. de Finetti, R. A. Fisher, E. J. Gumbel, C. Jordan, A. Khinchin, O. Morgenstern, E. Slutsky, O. Onicescu, V. Romanowsky and F. M. Urban. The Bologna meeting gave de Finetti a good opportunity to get in touch with the international mathematical community of the time. In a communication particularly appreciated by Fréchet, Khinchin and Neyman, he summarized his recent research on the representation of the law of a sequence of exchangeable events. In addition to the proceedings of the congress, let us mention a stimulating recent paper by Seneta (1992) dealing with a controversy between Cantelli and Slutsky, which occured during the Bologna Congress, over priority for the strong law of large numbers.

After sketching out the basic features of the scientific milieu in which de Finetti took the first steps of his scientific career in probability and statistics, let us review some of his main contributions.

2. DE FINETTI'S WORK IN PROBABILITY

Three topics stand out in de Finetti's contributions to probability: the foundations of probability; processes with independent increments; and sequences of exchangeable variables.

2.1 Foundations of Probability

De Finetti had examined this subject in depth ever since his early approach to probability. In fact, in 1931, he was in a position to explain his subjectivistic point of view in a nearly definitive way (cf. de Finetti, 1931a, b). Before the appearance of these papers, he had concisely stated his outlook on probability on the occasion of a polite dispute with Maurice Fréchet (Fréchet, 1930a, b, and de Finetti, 1930b, c), which originated from a paper of de Finetti's on the limit operation in the theory of probability; see de Finetti (1930a). In that paper, de Finetti, starting from the assumption that probability need not be completely additive, develops some critical remarks on the usual formulation of Cantelli's theorem, according to which:

THEOREM. *Let $(X_n)_{n\geq 1}$ be a sequence of $\{0,1\}$-valued independent and identically distributed random variables defined on the completely additive probability space $(\Omega, \mathscr{F}, P)$, such that $P(X_1 = 1) = p$. Then*

$$\lim_{N\to\infty} P\left(\sup_{n\geq N}\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| \leq \varepsilon\right)$$

$$= P\left(\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n} X_k = p\right) \tag{1}$$

$$= 1. \tag{2}$$

Statements (1) and (2) hold if P is completely additive but they need not hold if P is not completely additive. In any case, one has the following:

THEOREM. *Given $\varepsilon, \eta > 0$, there is $N_0 = N_0(\varepsilon, \eta)$ such that*

$$P\left(\sup_{N\leq n\leq N+M}\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| \leq \varepsilon\right) > 1-\eta, \quad N \geq N_0,\ M \in \mathbb{N}. \tag{3}$$

However, (3), generally speaking, does not entail

$$P\left(\sup_{n\geq N}\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| \leq \varepsilon\right) > 1-\eta, \quad N \geq N_0.$$

Moreover, equality (1) need not occur.

De Finetti explains these facts by giving an interesting example, which we report here. Let

$$\Omega = \left\{\frac{1}{2}; \frac{1}{4}, \frac{3}{4}; \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}; \dots; \frac{1}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n-1}{2^n}; \dots\right\}.$$

Writing each term of Ω in its terminating binary expansion, we obtain the following representation for the elements of Ω:

$$\begin{matrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 1 & 0 & 0 & \cdots \\ 1 & 1 & 1 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{matrix}$$

Now, for each k in $\mathbb{N}$, define $X_k : \Omega \to \{0, 1\}$ to be the kth digit of ω in Ω. Clearly,

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n} X_k(\omega) = 0 \qquad \text{for each } \omega \text{ in } \Omega \tag{4}$$

and

$$\sup_{n\geq N}\left|\frac{1}{n}\sum_{k=1}^{n}X_k(\omega)-\frac{1}{2}\right|=\frac{1}{2} \tag{5}$$

for each ω in Ω and N in $\mathbb{N}$.

Moreover, define

$$\Omega_n=\left\{\frac{1}{2},\frac{1}{4},\frac{3}{4},\dots,\frac{1}{2^n},\dots,\frac{2^n-1}{2^n}\right\},$$

$$P_n(\{\omega\})=\begin{cases}\dfrac{1}{2^n-1}, & \text{if } \omega\in\Omega_n,\\ 0, & \text{if } \omega\in\Omega\setminus\Omega_n,\end{cases}$$

for each n in $\mathbb{N}$. Then

$$P_n(X_k=1)=\begin{cases}0, & \text{if } k\geq n+1,\\ \dfrac{2^{n-1}}{2^n-1}, & \text{if } k\leq n,\end{cases}$$

$$P_n(\{X_k=1\}\cap\{X_i=1\})$$
$$=\begin{cases}0, & \text{if } \{k\geq n+1\}\vee\{i\geq n+1\}\\ \dfrac{2^{n-2}}{2^n-1}, & \text{if } \{k\leq n\}\wedge\{i\leq n\}\end{cases}\qquad k\neq i,$$

$$\vdots$$

$$P_n\left(\bigcap_{j=1}^{i}\{X_{k_j}=1\}\right)=\begin{cases}0, & \text{if } \bigvee_{j=1}^{i}\{k_j\geq n+1\},\\ \dfrac{2^{n-i}}{2^n-1}, & \text{if } \bigwedge_{j=1}^{i}\{k_j\leq n\},\end{cases}$$
$$k_1\neq k_2\neq\cdots\neq k_i$$

Let us now consider any (finitely additive) probability P on $\mathscr{P}(\Omega)$ such that $P(A)=\lim_{n\to+\infty}P_n(A)$ for every A in $\mathscr{P}_l(\Omega):=\{A\subset\Omega: \lim_{n\to+\infty}P_n(A)$ exists$\}$. Then $(X_n)_{n\geq1}$ with respect to such a P is a sequence of independent and identically distributed random variables with two values (0 and 1) and $P(X_1=1)=1/2$. Hence, (3) holds with $p=1/2$ (weak formulation of the uniform convergence in probability) in spite of (4), which states that $(\sum_{k=1}^{n}X_k/n)_{n\geq1}$ converges to 0 in the sense of pointwise convergence. Moreover, (5) shows that the usual strong formulation of the Cantelli law of large numbers does not hold with respect to P. See also Ramakrishnan and Sudderth (1988) for a similar example.

These facts merely show that a family of consistent finite-dimensional probability distributions, if complete additivity is not assumed as a requisite of probability, *does not determine* the probability of "infinitary events" of the type

$$\left\{\lim_{n\to+\infty}\sum_{k=1}^{n}X_k/n=E(X_1)\right\},$$

$$\left\{\sup_{n\geq N}\left|\frac{1}{n}\sum_{k=1}^{n}X_k-E(X_1)\right|\leq\varepsilon\right\},$$

and so on. Those very same facts reveal that the validity of strong laws cannot be ascribed to reasons of physical and objective nature, but to the *arbitrary* choice of a coherent extension to "infinitary events" of a probability law assessed on "finitary events" only.

In the same paper, de Finetti dwells upon the following circumstance. If additivity of probability is not complete, then almost-sure convergence of a sequence of random variables to X does not imply convergence in law of the same sequence to X (the previous example, in which $X\equiv0$ and the sequence converges in law to $Y\equiv1/2$). This problem, at that time, was of particular interest to him, since it was related to his research on processes with independent increments. So he stated that, in order that almost sure convergence of $(X_n)_{n\geq1}$ to X imply convergence in law of the same sequence to X, it suffices that $(X_n)_{n\geq1}$ converges in probability to X. And that is indeed a revolution when you consider the run-of-the-mill analysis of stochastic convergence! Maybe, in studying *consistency* of statistical procedures, for instance, statisticians ought to ponder these "paradoxical" conclusions.

In those months, Fréchet had submitted a paper for publication in *Metron*, summarizing the content of his 1929 and 1930 courses on stochastic convergence, from the σ-additive point of view. This circumstance made the French mathematician particularly susceptible to de Finetti's remarks. Thus, he sent a very short Note to *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, in which de Finetti's paper had been published, to say that de Finetti's examples were interesting but, on the other hand, meaningless once the extension of the additivity condition is assumed; see Fréchet (1930a). De Finetti answered Fréchet's objections (de Finetti, 1930b) by stating that the problems concerning stochastic convergence are mere signs of a deeper problem concerning the correct mathematical definition of probability. In fact, in his opinion, this definition has to adhere to the notion of probability as it is conceived by all of us in everyday life, and he preannounces the publication of *Probabilismo* and *Fondamenti per la teoria delle probabilità*, devoted to the philosophic and mathematical aspects, respectively, of the definition of probability. He maintains that one has anyway no right to make arbitrary use of the properties introduced to give a mathematical definition of probability. Indeed, these very properties have to be not only formally consistent but also intrinsically necessary with respect to a meaningful interpretation of probability. He goes on to state that the definition he intends to make is not necessarily conducive to a reduction of all admissible probabilities to the one family of σ-additive probabilities. In this light, his remarks in *Sui passaggi al limite* (de Finetti, 1930a) do make sense, unless one rejects the above definition *a priori*. Moreover, even if one should confine oneself strictly to the formal aspects of the problem, the restrictive stance referred to above would indeed lead to some odd conclusions, namely: (i) one has no right to consider a countably infinite class of pairwise disjoint events the probabilities of which have the same order of magnitude; (ii) the limit of a sequence of probability distributions need not be a probability distribution. Finally, he dwells upon the analogy between probability and measure supported by Fréchet. De Finetti shares Fréchet's opinion implying that each definition, even if of a mere mathematical nature, is more or less directly and distinctly triggered by intuition. Nevertheless, this very definition can effectively be arbitrary, provided that one confines himself to deducing purely formal conclusions from it. This is the case of the definition of measure. A different case is connected with the definition of weight, since "we cannot force a pair of scales to work according to our definition." The case of probability is finally different from the previous ones. If we, on the basis of a *convention*, state that $P(A_k) = 0$, $k \geq 1$, entails $P(\cup_{k=1}^{\infty} A_k) = 0$, then we intuitively think of $\cup_{k=1}^{\infty} A_k$ as a nearly impossible event, whereas the formal definition allows us only to conclude that 0 is the value at $\cup_{k=1}^{\infty} A_k$ of the function which we, conventionally, have called "probability."

Fréchet promptly answered de Finetti's new arguments (Fréchet, 1930b) by expressing his sympathy for de Finetti's position, at the same time, by showing reluctance to remove complete additivity as a characteristic property for a probability distribution. He resorts to some technical arguments that de Finetti, in his subsequent answer (de Finetti, 1930c), succeeds in using his own theory. See, for example, the argument employed by Fréchet to show that σ-additivity can be deduced from the "empirical" definition of probability stated in Fréchet and Halbwachs (1925).

We have lingered on the debate with Fréchet just because in it de Finetti's stance on probability foundations stands out most clearly. We are now also in a position to review the mathematical formulation of his viewpoint, following *Problemi determinati e indeterminati* and *Sul significato soggettivo* (de Finetti, 1930d and 1931b, respectively). In point of fact, in the corpus of de Finetti's works the abovementioned *Fondamenti per la teoria delle probabilità* is not listed. It was probably just a provisional title for *Sul significato soggettivo* or for *Problemi determinati e indeterminati*, or for both.

A person who wants to summarize his degree of belief in the different values of a random event E by a number p is supposed to accept any bet on E with gain $c(p - \mathbf{1}_E)$, where $\mathbf{1}_E$ denotes the indicator of E and c is any real number chosen by an opponent. Since c may be positive or negative, there is no advantage for the person in question in choosing a value p such that cp and $c(p-1)$ are both strictly positive for some c. Hence, p is an admissible evaluation of the probability of E if it meets the following *principle of coherence*: p has to ensure that there is no choice of c so that the realizations of $c(p - \mathbf{1}_E)$ are all strictly positive (or strictly negative). The concept of coherence is then extended to a (finite or infinite) class $\mathbf{E}$ of events in the following way. The real-valued function P on $\mathbf{E}$ is said to be a *probability* on $\mathbf{E}$ if, for any finite subclass $\{E_1, \dots, E_n\}$ of $\mathbf{E}$ and any choice of $(c_1, \dots, c_n)$ in $\mathbb{R}^n$, $n = 1, 2, \dots$, the gain

$$G = G(E_1, \dots, E_n; c_1, \dots, c_n) = \sum_{k=1}^{n} c_k\{P(E_k) - \mathbf{1}_{E_k}\}$$

is such that $\inf G \leq 0 \leq \sup G$, where the infimum and supremum are taken over all constituents of $\{E_1, \dots, E_n\}$.

With regard to the problem of the existence of at least a probability on a given class of events, de Finetti, in his early papers on the foundations of probability, confines himself to giving some hints based on geometrical arguments. In *Sull'impostazione assiomatica*, published in 1949, he dealt with this problem thoroughly by proving the following:

EXTENSION THEOREM. *If $\mathbf{E}_1$ and $\mathbf{E}_2$ are classes of events such that $\mathbf{E}_1 \subset \mathbf{E}_2$, and if P_1 is a probability on $\mathbf{E}_1$, then there is a probability P_2 on $\mathbf{E}_2$ such that $P_1 = P_2$ on $\mathbf{E}_1$.*

De Finetti obtained this theorem independently of the almost contemporary Horn–Tarski extension theorem which, however, is less general than de Finetti's; see Hornand Tarski (1948). De Finetti's proof is by induction, transfinite induction when $\mathbf{E}_2$ is not a countable class. Hence, his argument rests on the axiom of the choice that he adopts here as well as on other occasions, mentioning it explicitly.

Then, given a class $\mathbf{E}$ of events and an element E of this class, any p in $[0, 1]$ represents a coherent assessment on $\mathbf{E}_1 := \{E\}$, so that, in view of the extension theorem, there exists at least one probability P on $\mathbf{E}$ such that $P(E) = p$.

After defining probability, de Finetti, in *Sul significato soggettivo*, proves that the usual rules of the calculus of probability *are necessary for the coherence of P on* $\mathbf{E}$. More precisely, he states that:

THEOREM. *If P is a probability on $\mathbf{E}$, then:*

(π_1) $A \in \mathbf{E} \Rightarrow P(A) \in [0, 1]$;
(π_2) $\Omega \in \mathbf{E} \Rightarrow P(\Omega) = 1$;
(π_3) $A_1, \ldots, A_n \in \mathbf{E}$, $\cup_{k=1}^{n} A_k \in \mathbf{E}$ *and* $A_k \cap A_j = \emptyset$ *for* $1 \le k \ne j \le n$

$$\Rightarrow \quad P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k).$$

In particular, he deduces the following:

CHARACTERIZATION OF P ON AN ALGEBRA. *If $\mathbf{E}$ is an algebra, then $P: \mathbf{E} \to \mathbb{R}$ is a probability on $\mathbf{E}$ if and only if P is nonnegative and finitely additive, with $P(\Omega) = 1$.*

Hence, a probability coherent in de Finetti's sense need not satisfy the condition of σ-additivity. On the other hand, one obtains:

COHERENCE IS PRESERVED IN A PASSAGE TO THE LIMIT. *Let $(P_n)_{n \ge 1}$ be a sequence of probabilities on $\mathbf{E}$ and let $\mathbf{L} = \{E \in \mathbf{E}: \lim_{n\to+\infty} P_n(E)$ exists$\}$. Then, if $\mathbf{L} \ne \emptyset$, one has that $\tilde{P}: \mathbf{L} \to \mathbb{R}$ defined by $\tilde{P} = \lim_{n\to+\infty} P_n$ is a probability on $\mathbf{L}$.*

In connection with the principle of coherence let us mention that de Finetti's paper *Sul significato soggettivo* includes the first precise formulation of a system of axioms of a purely qualitative nature concerning the comparison between events. Such a system of axioms defines a *qualitative* (or *comparative*) *probability structure*. In that very same paper, de Finetti considers the problem of the existence of a *quantitative* subjective probability which agrees with a given probability structure. De Finetti's approach to qualitative probability played an important role in Savage's fundamental contribution to the foundations of statistics; see Savage (1954).

Insofar as the concept of conditional probability is concerned, de Finetti thinks of a *conditional event* $E|H$ as a logical entity which is true if $E \cap H$ is true, false if E is false and H is true, void if H is false (e.g., see de Finetti, 1937a). However, to the best of our knowledge, he refrained from providing a systematic theory of conditional probabilities. He just confined himself to mentioning the problem in *Teoria delle Probabilità* (de Finetti, 1970, volume 2; page 399 of the English translation), where he maintains that:

> it is necessary to strengthen the condition of coherence by saying that *the evaluations conditional on H must turn out to be coherent conditional on H* (i.e., under the hypothesis that H turns out to be true...). Although this strengthening of the condition of coherence might seem obvious...there are several forms of strengthening conditions....

NOTE. A concise review of the work covering the strengthening of the condition of coherence can be found in Berti and Rigo (1989). We recall only one of the propositions following from the abovementioned condition.

CHARACTERIZATION OF CONDITIONAL PROBABILITIES. *Let $\mathbf{E}$ and $\mathbf{H}$ be two algebras of events such that $\mathbf{H} \subset \mathbf{E}$, and let $P(\cdot|\cdot\cdot)$ be a real-valued function defined on $\mathbf{E} \times \mathbf{H}^0$, where $\mathbf{H}^0 = \mathbf{H} \setminus \{\emptyset\}$. Then P is a (coherent) conditional probability on $\mathbf{E} \times \mathbf{H}^0$ if and only if:*

(γ_1) *$P(\cdot|H)$ is a probability on $\mathbf{E}$, for each H in $\mathbf{H}^0$;*
(γ_2) *$P(H|H) = 1$ for all H in $\mathbf{H}^0$;*
(γ_3) *$P(A \cap B|C) = P(A|B \cap C)P(B|C)$ whenever C and $B \cap C$ belong to $\mathbf{H}^0$ and A and B are in $\mathbf{E}$.*

Conditions (γ_1)–(γ_3) are introduced in the form of axioms for conditional probability in de Finetti (1949). In view of the definition of conditional event, $P(E|\Omega)$ must coincide with the probability of E, so that we will write $P(E)$ in place of $P(E|\Omega)$. Hence, de Finetti's theory of probability is a particular case of the theory of coherent conditional probability.

In spite of the absence of a general theory, de Finetti, in *Sulla proprietà conglomerativa* (1930e), singled out an interesting circumstance which can occur in connection with a coherent conditional distribution: *the absence of conglomerability.* Let $P(\cdot|\cdot\cdot)$ be a conditional probability on $\mathbf{E} \times \mathbf{H}^0$ according to the previous characterization, and let $\pi = \{H_1, H_2, \ldots\}$ be a denumerable partition of Ω in $\mathbf{H}^0$. Given any E in $\mathbf{E}$, if π is finite, then

$$P(E|\Omega) = \sum_{k \ge 1} P(E|\Omega \cap H_k) P(H_k|\Omega)$$

[from (γ_1) and (γ_3)]

and, consequently,

$$\inf_k P(E|H_k) \leq P(E) \leq \sup_k P(E|H_k).$$

De Finetti wonders whether this property, which he names *conglomerability*, holds even if π is infinite. He gives a few examples which, contrary to intuition, show that P need not be conglomerable. The second example can be restated in the following terms. Let $\mathbf{E}$ be the power set of $\mathbb{N}^2$ and let $\mathbf{H}_1 = \{\{n\} \times \mathbb{N}, \mathbb{N} \times \{m\} : n \in \mathbb{N}, m \in \mathbb{N}\}$. Then, for any H in $\mathbf{H}_1$ and E in $\mathbf{E}$, put

$$P_1(E|H) = 0 \text{ or } 1$$

according to $E \cap H$ is finite or infinite.

In view of a coherence criterion due to Rigo (1988) (see Berti, Regazzini and Rigo, 1990 and 1996, for some generalizations and applications), it is immediate to obtain that P_1 is a coherent conditional probability on $\mathbf{E} \times \mathbf{H}_1$. Hence, there exists at least one extension of P_1 to a conditional coherent probability on $\mathbf{E} \times \mathbf{E}^0$, say P. Since, for $E = \{(n, m) \in \mathbb{N}^2 : n \geq m\}$,

$$P_1(E|\{n\} \times \mathbb{N}) = P_1(E^c|\mathbb{N} \times \{n\}) = 0 \quad n \in \mathbb{N},$$

it is clear that P cannot be conglomerable on E and on E^c with respect to the partitions $\{\{n\} \times \mathbb{N} : n \in \mathbb{N}\}$ and $\{\mathbb{N} \times \{n\} : n \in \mathbb{N}\}$, respectively. By the way, the same example was later given by Lévy (Lévy, 1931b) and, in modern literature, it is known as Lévy's paradox. In de Finetti's paper *Sull'impostazione assiomatica* (1949), one can find an example of nonconglomerability with respect to a partition whose elements have nonnull probabilities. Conglomerability is connected with disintegrability of probability measures. It has been studied in recent years by Dubins and other authors (i.e., Dubins, 1975, and Schervish, Seidenfeld and Kadane, 1984). In modern Bayesian literature, the phenomenon of nonconglomerability has come up in the so-called *marginalization paradoxes*.

The above-mentioned paper *Sull'impostazione assiomatica* (de Finetti, 1949), in addition to a complete account of subjective approach to probability based on the principle of coherence, includes a detailed analysis of the differences between that approach and the axiomatic theory of Kolmogorov.

A noteworthy analysis of the structure of a finitely additive probability is worked out in *La struttura delle distribuzioni* (de Finetti, 1955a), where de Finetti, independently of previous works covering the same topic (Bochner, 1939; Sobczyk and Hammer, 1944a, 1944b), provides a decomposition for finitely additive probabilities into the *discrete* component (masses concentrated at the points of a countable set), the *agglutinated* component (nonconcentrated in the previous sense, but concentrated on ultrafilters which constitute a countable class), and the *continuous* component (which can be indefinitely subdivided). Moreover, he decomposes the continuous component into the *condensed* (singular) subcomponent and the *diffuse* (absolutely continuous) subcomponent, with respect to a given continuous finitely additive measure. Finally, he proves a Radon–Nikodým theorem for finitely additive measures. Let us mention here the method for proving these statements. The proof of the decomposition is based on a suitable *coefficient of divisibility*, whereas the Radon–Nikodým theorem is proved through an extension of the well-known notion of Lorenz–Gini concentration function.

NOTE. In recent years de Finetti's extension of the notion of concentration function has been used as a measure of robustness of Bayesian statistical methods and in order to obtain a new proof for "exact versions" of the Radon–Nikodým theorem; see Cifarelli and Regazzini (1987), Fortini and Ruggeri (1994) and Berti, Regazzini and Rigo (1992).

We conclude this subsection by explaining another important aspect of de Finetti's contribution to the foundations of probability: the treatment of the concept of expectation of a random quantity. The main feature of this treatment is that our author suggests a definition on the basis of which the expectation of a random quantity can be evaluated without assessing its probability distribution. An early hint of this idea was included in a course of 1930 (de Finetti, 1930f), but the first fully fledged attempt to make this idea precise appeared later, in 1949. As to an exhaustive exposition, it came up for publication only in 1970 with *Teoria delle Probabilità*. Starting from the concrete meaning of expectation of a random quantity, de Finetti extends the betting scheme, recalled at the beginning of this present subsection, to random quantities, in order to define the concept of prevision of a *bounded* random quantity. Hence, given any class $\mathbf{K}$ of bounded random quantities and a real-valued function $\mathbb{P}$ on $\mathbf{K}$, one says that $\mathbb{P}$ is a *prevision* on $\mathbf{K}$ if, for every finite subclass $\{X_1, \ldots, X_n\}$ of $\mathbf{K}$ and for every n-tuple $(c_1, \ldots, c_n)$ of real numbers, one has

$$\inf \sum_{k=1}^{n} c_k \{\mathbb{P}(X_k) - X_k\} \leq 0 \leq \sup \sum_{k=1}^{n} c_k \{\mathbb{P}(X_k) - X_k\}.$$

In other words, $\mathbb{P}$ is a prevision if and only if there is no (finite) betting system which makes uniformly strictly negative the gain of the person who adopts $\mathbb{P}$ as a system of unit-prices for bets on the elements of $\mathbf{K}$. It is clear that the probability of an event E must coincide with the prevision of $\mathbf{1}_E$ and that the theory of probability is included in that of prevision. The same argument used to prove de Finetti's extension theorem can prove the following:

PROPOSITION. *If* $\mathbf{K}_i$, $i = 1, 2$ *are classes of bounded random quantities with* $\mathbf{K}_1 \subset \mathbf{K}_2$, *and if* $\mathbb{P}_1$ *is a prevision on* $\mathbf{K}_1$, *then there is a prevision* $\mathbb{P}_2$ *on* $\mathbf{K}_2$ *such that* $\mathbb{P}_1 = \mathbb{P}_2$ *on* $\mathbf{K}_1$.

In view of this proposition, given any class $\mathbf{K}$ of bounded random quantities there exists at least one prevision on $\mathbf{K}$. Moreover, if $\mathbf{B}$ designates the class of all bounded random quantities, then:

PROPOSITION. *The function* $\mathbb{P}$ *is a prevision on* $\mathbf{K} \subset \mathbf{B}$ *if and only if* $\mathbb{P}$ *can be extended as a positive, linear functional* $\mathbb{P}'$ *on* $\mathbf{B}$, *such that* $\inf X \le \mathbb{P}'(X) \le \sup X$ *for each* X *in* $\mathbf{B}$.

This proposition shows that prevision meets all the properties of the expectation, with the exception of the continuity property. Moreover, if one considers each element of $\mathbf{B}$ as a bounded function from Ω into $\mathbb{R}$, then:

PROPOSITION. The function $\mathbb{P}$ *is a prevision on* $\mathbf{K} \subset \mathbf{B}$ *if and only if there is a probability* π *on the class of all subsets of* Ω *such that*

$$\mathbb{P}(X) = \int_\Omega X \, d\pi \qquad (X \in \mathbf{K})$$

with the integral thought of as an abstract Stieltjes integral.

Apropos of the integral representation of a prevision, let's also recall the paper, of a purely mathematical character, *Sulla teoria astratta* (de Finetti, 1955b) as well as the paper *Sull'integrale di Stieltjes–Riemann* (de Finetti, 1935a) written with M. Jacob. In the latter, our author deals with the problem of the evaluation of $\mathbb{P}(X)$ when one knows the probability distribution function of X.

NOTE. In more recent years, de Finetti's theory of prevision has been extended to conditional bounded random variables (cf. Holzer, 1985, and Regazzini, 1985, for a review) and to random elements taking values in a Banach space (cf. Berti, Regazzini and Rigo, 1994).

2.2 Processes with Independent Increments

In 1929, de Finetti started on new research regarding functions with random increments. The crisis of determinism and of the causality principle introduced a novelty into the scientific method. This basically boils down to the replacement of classical logic with probability calculus. Rigid laws which state that a certain fact is bound to occur in a certain way are being replaced by probabilistic or statistical laws stating that a certain fact can occur depending on a variety of ways governed by probability laws. De Finetti's pioneer research on random functions aimed at preparing an analytic apparatus intended for the "translation" of deterministic laws of physics into probabilistic laws. Then, given a scalar quantity whose temporal evolution is described by function $X = X(\lambda)$, $\lambda \ge 0$, assume that the values taken by $X(\lambda)$ are known for $\lambda \le \lambda_0$ and consider the (conditional) increment

$$(X(\lambda) - X(\lambda_0) | X(u),\, u \le \lambda_0), \quad \lambda > \lambda_0.$$

Insofar as the probability distribution function $\phi(\cdot)$ of such an increment is concerned, de Finetti considers the following cases:

(a) $\phi(\cdot)$ is independent of $X(u)$ for every u in $[0, \lambda_0]$;
(b) $\phi(\cdot)$ is independent of $X(u)$ for every u in $[0, \lambda_0)$;
(c) $\phi(\cdot)$ dependent on the "history" of X on $[0, \lambda_0]$.

This classification is inspired by Volterra's classification for deterministic laws of physics and de Finetti, following Volterra's classification, calls ϕ *known* if case (a) occurs, *differential* if case (b) comes true and *integral* in case (c). Note that Kolmogorov's attitude toward the role played by stochastic processes in the formulation of physical laws was very close to de Finetti's. See Kolmogorov (1931), where he deals with case (b). In *Sulle funzioni a incremento aleatorio* (de Finetti, 1929a), de Finetti deals with the problem of characterizing the probability distribution of $X(\lambda)$ in case (a). If $X(0) \equiv 0$ and ϕ_λ and ψ_λ denote the probability distribution function and characteristic function of $X(\lambda)$, respectively, then $\{\psi_{1/n}(\cdot)\}^n$ is the characteristic function of the sum of n independent increments, identically distributed according to the law of $X(1/n) - X(0)$. If $\lim_{n\to+\infty}\{\psi_{1/n}(\cdot)\}^n$ exists, then de Finetti calls it the *derivative law* of the distribution of $X(\lambda)$ and denotes it by $\psi_0'(\cdot)$. Hence, if $\log \psi_\lambda(\cdot)$ is differentiable with respect to λ, one has

$$\begin{aligned}\psi_0'(\cdot) &= \lim_{n\to+\infty}\{\psi_{1/n}(\cdot)\}^n \\ &= \exp\left\{\frac{\partial}{\partial\lambda}\log\psi_\lambda(\cdot)\right\}\Bigg|_{\lambda=0}\end{aligned}$$

and

$$\psi_\lambda(\cdot) = \exp\left\{\int_0^\lambda \log \psi_\nu^*(\cdot)\,d\nu\right\}$$

with

$$\psi_\lambda^*(\cdot) = \exp\left\{\frac{\partial}{\partial\lambda}\log\psi_\lambda(\cdot)\right\}.$$

In particular, if ψ_λ^* is independent of λ (de Finetti, in this case, calls ϕ_λ *known and fixed*), one obtains $\psi_\lambda^*(\cdot) = \psi_0'(\cdot)$ and

$$\psi_\lambda(\cdot) = \exp\{\lambda \log \psi_0'(\cdot)\} = \{\psi_0'(\cdot)\}^\lambda,$$

which entails $\psi_1(\cdot) = \psi_0'(\cdot)$ and

$$\psi_\lambda(\cdot) = \{\psi_1(\cdot)\}^\lambda.$$

In modern literature, these processes are known as *processes with homogeneous independent increments*, and ψ_1 is called an *infinitely decomposable characteristic function* [indeed, $\psi_1 = (\psi_{1/n})^n$]. At this point, de Finetti considers the particular case in which

$$\log\psi_0'(\cdot) = imt - \frac{\sigma^2}{2}t^2, \quad m \in \mathbb{R},\ \sigma^2 > 0,$$

which gives

$$\psi_\lambda(t) = \exp\left\{i\lambda mt - \frac{\lambda\sigma^2}{2}t^2\right\}$$

(i.e., X is a Brownian motion) and sketches a proof for nowhere differentiability of X, a result generally ascribed to Paley, Wiener and Zygmund (1933).

In *Sulla possibilità di valori eccezionali* (de Finetti, 1929b), he considers processes with homogeneous independent increments and shows that ψ_λ is continuous whenever $X(\lambda)$ is continuous on $[0, +\infty)$ and $X(\lambda)$ is different from $c\lambda$. The proof is based on an interesting geometrical argument. Moreover, the examples chosen to emphasize the relevance of the assumption of continuity for X are also noteworthy: the Poisson process and the compound Poisson process. In fact, de Finetti's method enables him to deduce the characteristic properties of these processes in a way that is quite innovative with respect to the original arguments propounded by Lundberg (1903). Coming back to a continuous process X with known and fixed law, de Finetti, in *Integrazione delle funzioni* (de Finetti, 1929c), deals with the problem of determining the law of the integral

$$S = \int_0^\lambda X(u)\,du = \lim_{n\to+\infty} S_n,$$

where

$$S_n = \frac{\lambda}{n}\sum_{h=1}^n X\left(\frac{h}{n}\lambda\right).$$

In order to determine the probability distribution of S_n, he rewrites S_n as

$$S_n = \lambda X(0) + \frac{\lambda}{n}\sum_{h=1}^n (n-h+1) \cdot \left\{X\left(\frac{h}{n}\lambda\right) - X\left(\frac{h-1}{n}\lambda\right)\right\}$$

and, from the properties of X, the characteristic function of S_n, φ_{S_n}, satisfies

$$\log\varphi_{S_n}(t) = ict\lambda + \frac{\lambda}{n}\sum_{h=1}^n \log\psi_1\left((n-h+1)\frac{\lambda}{n}t\right)$$

when $X(0) \equiv c$. Hence,

$$\lim_{n\to+\infty}\varphi_{S_n}(t) = ict\lambda + \frac{1}{t}\int_0^{\lambda t}\log\psi_1(u)\,du := \log\varphi(t).$$

In a σ-additive framework, this fact suffices to state that φ is the characteristic function of S, while, in de Finetti's finitely additive context, the previous argument states that $(S_n)_{n\geq 1}$ has a limiting distribution but this need not be the distribution of S; see Section 2.1. This would be the case if $(S_n)_{n\geq 1}$ converged to S in probability. On the other hand, in order to check on such a condition, we need some information about the distribution of (S, S_n) for each n. De Finetti, by way of example of the above result, states that the characteristic function of S equals

$$\exp\left\{it\left(c\lambda + \frac{m}{2}\lambda^2\right) - \frac{\sigma^2\lambda^3}{3!}t^2\right\}$$

when $\log\psi_1(t) = imt - (\sigma^2/2)t^2$. This result is a prelude to analogous results stating the probability distribution of particularly significant functionals of the Wiener–Lévy process.

In two subsequent papers dated 1930 and 1931, he considered the problem of singling out all the characteristic functions ψ such that ψ^λ is a characteristic function for every $\lambda > 0$. In fact, this is equivalent to characterizing the characteristic function of $X(\lambda)$, X being any process with homogeneous independent increments, and is equivalent to characterizing the class of all infinitely decomposable characteristic functions. In the above-mentioned papers (de Finetti, 1930h, 1931c), de Finetti obtains the well-known theorem:

THEOREM. *The class of infinitely decomposable laws coincides with the class of distribution limits of finite convolutions of distributions of Poisson type.*

Kolmogorov and Lévy took these papers as a starting point for their renowned representations for infinitely decomposable characteristic functions; see Kolmogorov (1932a, b) and Lévy (1934).

Apart from the brillant treatment included in the second volume of *Teoria delle Probabilità* (de Finetti, 1970), de Finetti returned to the theory of stochastic processes with independent increments only once,with a short paper dated 1938, *Funzioni aleatorie* (de Finetti, 1938b). In this work, he assumes a critical attitude toward his previous papers. In fact, in the above-mentioned papers, he considered the probability of transcendental (= infinitary) conditions while, in his opinion, only conditions depending on a finite number of values of $X(\cdot)$ are susceptible of "concrete" probability evaluations. The paper in question does not actually include new results, but rather a few suggestions conducive, on one hand, to the notion of characteristic functional and, on the other hand, indicative of a way to introduce the probability that $X(\cdot)$ satisfies certain conditions of an analytical nature. For instance, one cannot possibly speak of the probability distribution of X' at the point t_1, but one can at least consider the distribution of $\{X(t_1 + h) - X(t_1)\}/h$ for each h. Consequently, one can consider the limit expression of that distribution as $h \to 0$ and interpret the limiting distribution as an approximation of the law of the difference quotient for h sufficiently small to justify the approximation.

2.3 Sequences of Exchangeable Random Variables

With regard to the connections between the subjectivistic viewpoint and the objectivistic one which, in a different way, characterizes the classical approach (based on the notion of equally probable cases) and the frequentistic approach, de Finetti considers these approaches as procedures "that have been thought to provide an objective meaning for probability," but he immediately specifies that these same procedures are not necessarily conducive to the existence of an objective probability. On the contrary, by drawing that conclusion, one encounters well-known difficulties, which do disappear only when one does not seek to eliminate but, on the contrary, one seeks to make the subjective element more precise. Thus, the classical definition of probability, based on the relation of the number of favorable cases with the number of possible cases, can be justified immediately: indeed, if there is a complete class on n incompatible events, and if *they are judged equally probable*, then by virtue of the theorem of total probability each of them will necessarily have the probability $p = 1/n$. The analysis of the frequentistic point of view is more complex. De Finetti proposed to break the analysis down into two phases and explained their subjectivistic foundations. The first phase deals with the relations between evaluations of probabilities and the prevision of future frequencies; the second concerns the relationship between the observation of past frequencies and the prevision of future frequencies. As regards the first phase, let us consider a sequence of events $E_1, E_2, \ldots$ relative to a sequence of trials and suppose that, under the hypothesis H_N stating a certain result of the first N events, a person considers equally probable the events $E_{N+1}, E_{N+2}, \ldots$. Then, denoting by $\tilde{f}_{H_N}$ the prevision of the random relative frequency of occurrence of the n events $E_{N+1}, \ldots, E_{N+n}$, conditional on H_N, the well-known properties of a prevision yield

$$p_{H_N} = \tilde{f}_{H_N},$$

where p_{H_N} indicates the probability of each $E_{N+1}, E_{N+2}, \ldots$ conditional on H_N. Hence, by estimating $\tilde{f}_{H_N}$ *via* the observation of past frequencies, one obtains an evaluation of p_{H_N}. But when is it permissible to estimate $\tilde{f}_{H_N}$ in such a manner? This is the problem of the second phase. De Finetti's answer is: when the events considered are supposed to be elements of a stochastic process whose probability law, conditional on large samples, admits, as prevision of the future frequency, a value approximately equal to the frequency observed in these samples. Since the choice of the probability law governing the stochastic process is subjective, the prevision of a future frequency based on the observation of those past is naturally subjective. De Finetti shows that the procedure is perfectly admissible when the process is *exchangeable*, that is, when only information about the number of successes and failures is relevant, irrespective of exactly which trials are successes or failures (cf. de Finetti, 1937a).

So far as we know, the notion of exchangeable events dates to a communication by Jules Haag at the *International Mathematical Congress* held in Toronto (August 1924), published in 1928. In addition to the definition, Haag exhibits an incomplete proof of the representation theorem. Independently of Haag's work, de Finetti introduced exchangeable events as a probabilistic characterization of a *random phenomenon*, that is, a phenomenon which can be repeatedly observed under homogeneous environmental conditions. He argued that a correct probabilistic translation of such an empirical circumstance leads us to think of the probability of m successes and $(n - m)$ failures in n trials as invariant with respect to the order in which successes and failures alternate, whatever n and m may be. More precisely, in a communication at the above-mentioned International Mathematical Congress held in Bologna, de Finetti defines a sequence $(E_n)_{n\geq 1}$ of events to be *equivalent* [nowadays, the more expressive and unambiguous word *exchangeable*, proposed by Pólya (cf. de Finetti, 1938c, 1939a) or Fréchet (cf. de Finetti, 1976), is used; in fact, even the term *symmetric*, used by Hewitt and Savage, admits ambiguity] if, for every finite permutation π, the probability distribution of $(\mathbf{1}_{E_1}, \dots, \mathbf{1}_{E_n}, \dots)$ is the same as the probability distribution of $(\mathbf{1}_{E_{\pi(1)}}, \dots, \mathbf{1}_{E_{\pi(n)}}, \dots)$. Contrary to Haag, de Finetti rigorously obtains the expression for the more general exchangeable distribution as a consequence of the following:

STRONG LAW OF LARGE NUMBERS. *If $(E_n)_{n\geq 1}$ is a sequence of exchangeable events, then*

$$P\left(\sup_{1\leq p\leq k}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{E_i} - \frac{1}{n+p}\sum_{i=1}^{n+p}\mathbf{1}_{E_i}\right| \leq \varepsilon\right) > 1-\eta$$

holds for every $k = 1, 2, \dots$ and for every strictly positive ε, η, provided that $n \geq N_0 = N_0(\varepsilon, \eta)$.

In its turn, this result entails:

THEOREM. *If $(E_n)_{n\geq 1}$ is a sequence of exchangeable events, then there is a probability distribution function F, whose support is included in* $[0, 1]$, *such that*

$$P\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{E_i} \leq x\right) \to F(x), \quad n \to +\infty,$$

at each continuity point x of F.

Moreover, under the usual condition of complete additivity for P, the first statement implies the existence of a random variable θ such that

$$P\left(\sup_{n\geq m}\left|\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{E_1} - \theta\right| \leq \varepsilon\right) \to 1, \quad m \to +\infty, \varepsilon > 0,$$

and F is the probability distribution function of θ.

Starting from these basic results, one easily obtains the following:

REPRESENTATION THEOREM. *If $(E_n)_{n\geq 1}$ is a sequence of exchangeable events, and if $(x_1, \dots, x_n) \in \{0, 1\}^n$, then*

$$P(\mathbf{1}_{E_1} = x_1, \dots, \mathbf{1}_{E_n} = x_n) = \sum_{M=s_n}^{N-n+s_n} \frac{\binom{N-n}{M-s_n}}{\binom{N}{M}} P\left(\sum_{i=1}^{N}\mathbf{1}_{E_i} = M\right)$$

where $s_n = \sum_{i=1}^{n} x_i$, $n \leq N$. Hence,

$$P(\mathbf{1}_{E_1} = x_1, \dots, \mathbf{1}_{E_n} = x_n) = \lim_{N\to+\infty}\sum_{M=s_n}^{N-n+s_n} \frac{\binom{N-n}{M-s_n}}{\binom{N}{M}} P\left(\sum_{i=1}^{N}\mathbf{1}_{E_i} = M\right) = \int_{[0,1]} t^{s_n}(1-t)^{n-s_n}\,dF(t),$$

F being the limiting distribution of

$$\left(\sum_{i=1}^{N}\mathbf{1}_{E_i}/N\right)_{N\geq 1}.$$

The previous line of proof reproduces an argument, recalled in *La prévision* (de Finetti, 1937a), due to Khinchin (1932b). As a matter of fact, de Finetti's original argument in his communication at the Bologna International Mathematical Congress and, more at length, in *Funzione caratteristica* (de Finetti, 1930g, 1932a), are of a merely analytical nature (Fourier–Stieltjes transform). In any case, he was able to justify the evaluation of $\tilde{f}_{H_N}$ *via* past frequencies, thanks to the the following:

ASYMPTOTIC THEOREM FOR PREDICTIVE DISTRIBUTIONS. *If $(E_n)_{n\geq 1}$ is a sequence of exchangeable events, then*

$$P\left(\sup_{N\leq n\leq N+p}\left|P(E_{n+k}|\mathbf{1}_{E_1}, \dots, \mathbf{1}_{E_n}) - \sum_{i=1}^{n}\frac{\mathbf{1}_{E_i}}{n}\right| \leq \varepsilon\right) > 1-\eta$$

for every p and k in $\mathbb{N}$ and for every strictly positive ε and η, provided that $N \geq N_0(\varepsilon, \eta)$.

In three short Notes published in *Atti della Reale Accademia dei Lincei*, in 1933, de Finetti extended the previous statements to sequences of exchangeable random quantities. In particular, he extended the strong law of large numbers to exchangeable random variables with finite fourth moment and he explained an argument to extend the representation theorem. Apropos of this latter point, further developments are explained in de Finetti (1937a), where the representation theorem is formulated in the following terms.

REPRESENTATION THEOREM. *Let $(X_n)_{n\geq 1}$ be a sequence of exchangeable random variables and let $\mathbb{F}$ denote the class of the probability distribution functions on $\mathbb{R}$ equipped with the Lévy metric. Then there is a probability μ on the class of all subsets of $\mathbb{F}$ such that*

$$P(X_1 \leq x_1, \dots, X_n \leq x_n) = \int_{\mathbb{F}}\prod_{k=1}^{n} V(x_k)\mu(dV)$$

for every x_k in $\mathbb{R}$, $k = 1, \dots, n$, $n \in \mathbb{N}$.

It has to be stressed that in de Finetti's formulation of the representation theorem, in conformity with the principle of coherence, P is thought of as finitely additive. Hence, that very same formulation need not be equivalent to the usual proposition "the X_n's are exchangeable if and only if they are conditionally i.i.d., given a random quantity θ," which holds if P is σ-additive.

NOTE. This form of de Finetti's theorem is quite widely known. It has been extended to random elements with values in a compact Hausdorff space by Hewitt and Savage (1955) and to random elements with values in a standard Borel space by Aldous (1985). On the other hand, Dubins and Freedman (1979) give an example to show that the usual formulation is false for general random elements; see also Freedman (1980) and Dubins (1983). Bühlmann (1960) found the appropriate representation for processes with exchangeable increments and Freedman (1962a, b, 1963) began to work on more general types of symmetry. Freedman (1962a) and Diaconis and Freedman (1980a) dealt with the characterization of discrete chains with a law which is invariant with respect to the order of transitions. In the first paper, one proves that, if the chain is stationary, then it must be a mixture of Markov chains. In the second paper, the authors reach the very same conclusion under a suitable condition of recurrence. Daboni (1975, 1982) obtained characterizations of mixtures of particular Markov processes starting from exchangeability of interarrivals times, and used such a characterization to prove two well-known Hausdorff and Bernstein theorems concerning completely monotone functions. Significant extensions of the above-mentioned results to general state spaces, and/or to continuous time, can be found in Freedman (1984) and Kallenberg (1973, 1974, 1975, 1982).

According to Diaconis (1988):

> de Finetti's theorem presents a real-valued exchangeable process as a mixture over the set all measures. Most of us find it hard to meaningfully quantify a prior distribution over so large a space. There has been a search for additional restrictions that get things down to a mixture of familiar families parametrized by low dimensional Euclidean parameter spaces.

In this setting, in addition to Freedman (1963), let us mention Dawid's characterization of the mean-zero p-dimensional covariance mixture of normal distributions and Smith's characterization of location–scale mixtures of univariate normal distributions (cf. Dawid, 1978 and Smith, 1981). The multidimensional extension of Smith's theorem has been given by Diaconis, Eaton and Lauritzen (1992). These results characterize those models that are compatible with a certain choice of data reduction. This, in turn, involves sufficiency and is linked to de Finetti's reconstruction of the Bayes–Laplace paradigm; see Section 3. Diaconis and Freedman (1984) provide a framework for characterizing mixtures of processes in terms of their symmetry properties and sufficient statistics. As an application, mixtures of the following kinds of processes are characterized: coin-tossing processes (de Finetti); sequences of independent identically distributed normal variables; sequences of independent, identically distributed integer-valued generalized exponential variables. Extensions of this last characterization are included in Küchler and Lauritzen (1989) and Diaconis and Freedman (1990). Other interesting de Finetti style theorems can be found in Lauritzen (1975), Ressel (1985), Accardi and Lu (1993) and Björk and Johansson (1993). To conclude this short review of recent literature related to de Finetti's representation theorem, let us mention central limit theorems: by Bühlmann (1958), by Blum, Chernoff, Rosenblatt and Teicher (1958), by Klass and Teicher (1987) and by Eaton, Fortini and Regazzini (1993). The limiting distribution of $\sum_1^n X_i/n$ has been determined by Cifarelli and Regazzini (1990, 1993) under the assumption that μ is a Ferguson–Dirichlet distribution. Recently, Berti and Rigo (1994) have provided a Glivenko–Cantelli style generalization of the asymptotic theorem for predictive distributions.

During his subsequent scientific career, de Finetti went back to exchangeability on two occasions: first, to deal with the so-called degenerate sequences of exchangeable random events; second, to consider the problem of extending finite exchangeable sequences. In *Gli eventi equivalenti e il caso degenere* (1952), he recalls that the E_n's are exchangeable if and only if

$$P(E_{n+1}|\mathbf{1}_{E_1} = x_1, \dots, \mathbf{1}_{E_n} = x_n) = p^{(n)}\Big(\sum_1^n x_k\Big)$$

for every n and $(x_1, \dots, x_n)$ in $\{0, 1\}^n$, and

$$p^{(n)}\Big(\sum_1^n x_k\Big) q^{n+1}\Big(\sum_1^n x_k + 1\Big) = q^{(n)}\Big(\sum_1^n x_k\Big) p^{n+1}\Big(\sum_1^n x_k\Big)$$

where $q^{(n)} = 1 - p^{(n)}$, and he designates as *degenerate* any exchangeable distribution such that $p_r^{(n)} = 0$ or 1, for some r and n. Then, after some interesting considerations of a geometrical nature, he considers the problem of representing a degenerate exchangeable sequence and connects this problem with the use of improper priors for the Bayesian analysis of the Bernoulli parameter; see also Daboni (1953). As for the problem of extending sequences of exchangeable random events, it is well known that a finite exchangeable process may or may not be the initial segment of some other exchangeable sequence of more steps. Then, in *Sulla proseguibilità* (1969), de Finetti suggests a geometric approach to deal with the following problem: is a given n-exchangeable process extendible? In the affirmative, how many steps can it be extended for, while preserving exchangeability? If the given n-exchangeable process is r-extendible, can we characterize the r-exchangeable process with respect to which the given process is the initial n-segment?

NOTE. De Finetti's geometric viewpoint was followed by Crisma (1971, 1982), Spizzichino (1982) and Wood (1992). A very clear exposition of such approach is given in Diaconis (1977).

In a communication at the *Colloque de Genève* in 1937, *Sur la condition d'équivalence partielle* (de Finetti, 1938a), de Finetti proposes the following extension of the notion of exchangeability. He considers k sequences of random events $(E_n^{(i)})_{n\geq 1}$, $i = 1, \ldots, k$, and he defines the resulting array of events to be *partially exchangeable* if the probability law of that array is the same as the law of $\{(E^{(1)}_{\pi_1(n)})_{n\geq 1}, \ldots, (E^{(k)}_{\pi_k(n)})_{n\geq 1}\}$, where $\pi_1, \ldots, \pi_k$ are arbitrary finite permutations. After explaining the meaning of this condition in connection with the inductive problem, he states the representation theorem for the most general law for an array of partially exchangeable events.

NOTE. In more recent years, finite exchangeable sequences have been studied by Diaconis and Freedman (1980b), who analyzed the behavior of the variation distance between the distribution of an n-exchangeable process and the closest mixture of i.i.d. random variables. Those very same authors give similar theorems for particular versions of de Finetti's theorem (like normal location, or scale parameters, mixtures of Poisson, geometric and gamma etc.) and for exponential families; see Diaconis and Freedman (1987, 1988). Diaconis, Eaton and Lauritzen (1992) consider a finite sequence of random vectors and give symmetry conditions on the joint distribution which imply that it is well approximated by a mixture of normal distributions.

As far as partial exchangeability is concerned, let us recall that contemporary scholars tend to use that term in a broad sense; see Aldous (1981) and Diaconis and Freedman (1980a). In the latter paper, for instance, partial exchangeability designates invariance with respect to the order of transitions. As a matter of fact, Petris and Regazzini (1994), following de Finetti (1959), analyze the strict relationships between that invariance and de Finetti's partial exchangeability of the array of "subsequent states;" see also Zabell (1995). Partial exchangeability, in the sense of Diaconis and Freedman, for finite sequences has been considered by Zaman (1984, 1986). Central limit theorems for partially exchangeable arrays are proved in Fortini, Ladelli and Regazzini (1994). Von Plato (1991) and Scarsini and Verdicchio (1993) have dealt with the extendibility of finite partially exchangeable random elements.

Finally, let us mention that in April 1981 an International Conference on "Exchangeability in Probability and Statistics" was held in Rome. The proceedings of the conference, edited by Koch and Spizzichino, include 8 general lectures and 21 invited contributions, in addition to 2 articles by de Finetti and Fürst, respectively.

A full survey of classical work on exchangeability can be found in Aldous (1985). Diaconis (1988) can be thought of as a useful updating supplement to the Aldous monograph.

2.4 Other Contributions to Probability Theory and Its Applications

We devote this subsection to four papers of de Finetti for which it would be difficult to find suitable pigeon holes. The first, *Resoconto critico del colloquio* (de Finetti, 1938c), represents an account of a conference, held in Geneva, October 1937, and coordinated by Fréchet, devoted to the theory of probability. The author provides a concise but clear-cut summary of each of the 16 lectures actually delivered at the Geneva conference, reporting in addition both the discussions originated by each lecture and some personal remarks. As a result, we are presented with a stimulating picture of the state of probability theory at the end of a particularly productive period. We list the lecturers and the subjects dealt with in Appendix 1.

The second paper we want to draw to your attention is *La teoria del rischio* (de Finetti, 1939b), which deals with the risk theory from Lundberg's standpoint. We mention it in the present section because of its interesting connections with a fundamental identity later proved by Wald; see Wald (1944), Section 3. Assume that G_1 and G_2 are strictly positive integers, and consider two players who start with bankrolls G_1 and G_2, respectively. Then consider a sequence of random variables $(\xi_n)_{n\geq 1}$ with their ranges in $\{-1, 1\}$: if $\xi_i = 1$ (-1, respectively), we suppose that the second player pays one unit to the first (the first pays the second, respectively). Then $S_k := \xi_1 + \cdots + \xi_k$ can be interpreted as the amount won by the first player from the second after k turns. If

$$\tau = \inf\{n \geq 1: S_n \leq -G_1 \text{ or } S_n \geq G_2\},$$

then $P_1 = P(S_\tau = -G_1)$ [$P_2 = P(S_\tau = G_2)$, respectively] represents the probability of ruin for the first player (the second, respectively) and, under suitable conditions of fairness, one has

$$P_1 + P_2 = 1, \quad P_2 G_2 - P_1 G_1 = 0.$$

In particular, these conditions are satisfied if the ξ's are independent and identically distributed with $P(\xi_1 = 1) = P(\xi_1 = -1) = 1/2$. As a matter of fact, de Finetti considers more general situations in which one, for example, has $E(\xi_{n+1}|\xi_1, \ldots, \xi_n) = 0$ for every $n \geq 1$, and he deals with the problem of determining the probability of ruin when the game is unfair, by resorting to a trick which dates back to De Moivre; see De Moivre (1711) and Thatcher (1970). Assume that the ξ's are independent and identically distributed with $P(\xi_1 = a) = p$ in $(0, 1)$ and $P(\xi_1 = b) = q = 1 - p$. Then attach value $(\exp(\alpha\xi_i) - 1)$ to the gain ξ_i and choose α in such a way that the resulting gain is fair that is, choose $\alpha_0 \neq 0$ for which $E(\exp(\alpha_0\xi_i) - 1) = 0$. Afterwards, de Finetti proves that the sign of α_0 is opposite to the sign of $E(\xi_i)$ and that $E(\exp(\alpha_0 \sum_{i=1}^n \xi_i)) = 1$ for every n. Hence, from the well-known formula for fair games

$$P_1(\exp(-\alpha_0 G_1) - 1) + P_2(\exp(\alpha_0 G_2) - 1) = 0,$$
$$P_1 + P_2 = 1,$$

one has

$$P_1 = \exp(\alpha_0 G_1)\frac{\exp(\alpha_0 G_2) - 1}{\exp[\alpha_0(G_1 + G_2)] - 1},$$
$$E(\exp(\alpha_0 S_\tau)) = 1.$$

Our author dwells upon the meaning of α_0 as a "index of risk" when $E(\xi_1) > 0$ ($\alpha_0 < 0$) and $G_2 = +\infty$ (the case, e.g., of an insurance company). In this case, one can write

$$P_1 = \exp(-G_1/B), \quad B = -1/\alpha_0,$$

and B represents the value of the company's initial capital for which the probability of ruin equals $1/e$, so that de Finetti calls it *riskiness level*.

In *Trasformazione di numeri aleatori* (de Finetti, 1953a), he proves that, given two continuous and proper probability distribution functions on $\mathbb{R}$, say F and G, there is a random variable on $\mathbb{R}$ with its range in $[0, 1]$ whose probability distribution function is uniform on $[0, 1]$ both under F and under G. This implies, for instance, that two statisticians, with different continuous posterior distributions for a real-valued parameter, can still find a common ground on which their respective opinions do coincide.

To conclude, we refer to the subject matter of the first paper taken into consideration in the present section: convergence of sequences of random variables. As mentioned in Section 1.3, the *GIIA*, in 1933, published a paper in which Glivenko proved the almost-sure convergence of $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)|$ to 0 (as $n \to +\infty$), where F_n represents the empirical distribution function associated with a random sample from the *continuous* probability distribution function F. The extension of Glivenko's theorem to whichever probability distribution function F is generally ascribed to Cantelli. In fact, in a paper published the same year in *GIIA*, Cantelli claims that the problem of possible discontinuity points of F can be dealt with by resorting to the definition of random variable he had proposed ever since 1916. In point of fact, according to Cantelli, such a definition ought to imply that the jump at x of F represents the probability concentrated at x. Actually, it is not clear how a definition of random variable could produce the above phenomenon which, instead, has to be ascribed to the properties of the probability distribution of each element of the random sample. For example, if this distribution is σ-additive on the Borel class in $\mathbb{R}$, then any jump of F takes on just that meaning, but this is not necessarily true in the finitely additive case. The correct interpretation of the jumps of F in connection with the so-called fundamental theorem of mathematical statistics is included in a paper by de Finetti, *Sull'approssimazione empirica* (de Finetti, 1933b), where the above-mentioned extension of Glivenko's theorem is thoroughly stated. This paper, which precedes Cantelli's, includes also some interesting hints at how to investigate that very same problem in the case of the existence of possible adherent masses at a discontinuity point of the distribution function. In a nutshell, de Finetti suggests that, in order to obtain a proposition whose validity is independent of the nature of the jumps of F, one should use Lévy's metric as a substitute for the uniform metric. Moreover, he replaces Glivenko's original argument with a simpler procedure which makes the most of Cantelli's strong law of large numbers in Bernoulli trials. Therefore, even if research on the convergence of the empirical distribution function, in the early 1930s drew inspiration from Cantelli's works on convergence of random sequences, the "fundamental theorem of mathematical statistics" ought to be ascribed, strictly speaking, to Glivenko and de Finetti.

3. DE FINETTI'S WORK IN STATISTICS

For expository convenience, we will split de Finetti's contributions to statistics into two subsections: the former devoted to description of frequency distributions, the latter to inductive reasoning.

3.1 Description of Frequency Distributions

The early de Finetti paper in this field deals with Gini's mean difference. In *Calcolo della differenza media* (de Finetti, 1930i), de Finetti and Paciello stated the identity

$$\text{(mean difference =)} \int_{\mathbb{R}^2} |x-y|\, dF(x)\, dF(y) = 2\int_{\mathbb{R}} F(x)\{1-F(x)\}\, dx,$$

and, subsequently, in *Sui metodi proposti* (1931d), de Finetti discussed different methods for the computation of the mean difference by showing the superiority of the so-called de Finetti–Paciello method.

That very same year saw the publication of one of the most important papers by de Finetti in the field of the analysis of statistical data: *Sul concetto di media* (de Finetti, 1931e). The starting point of this new investigation was a definition of mean given two years before by Chisini, an Italian distinguished geometer who was de Finetti's teacher at the University of Milan. According to such a definition, given n quantities $x_1, \ldots, x_n$ the concept of mean of $x_1, \ldots, x_n$ is meaningful and operationally determined when there is a new quantity z which depends on $x_1, \ldots, x_n$,

$$z = f_n(x_1, \ldots, x_n),$$

so that one can think of a mean as a value $\bar{x}$ which, if each x_i is replaced by $\bar{x}$, does not alter the value of z at $(x_1, \ldots, x_n)$. As an example, assume that the natural increase in a human population, in three consecutive years, is 1.2%, 0.2% and 0.8%. What is the mean increase in that triennium? If z is taken to be the total increase in the population of the same period, then the mean $\bar{x}$ has to satisfy

$$1.012 \cdot 1.002 \cdot 1.008 = (1+\bar{x})^3,$$

which yields $\bar{x} \simeq 0.00729$. De Finetti extends Chisini's definition of a mean to distribution functions: given a class $\mathbb{F}$ of frequency (or probability) distribution functions on $\mathbb{R}$ and a real-valued function f on $\mathbb{F}$, a mean of F in $\mathbb{F}$, with respect to the evaluation of f, is any number $\xi = \xi(F)$ such that

$$f(F) = f(D_\xi),$$

where D_x denotes the probability distribution function which degenerates at x. Assume that $\mathbb{F} = \mathbb{F}[A, B]$ is the class of all distribution functions whose support is included in $[A, B]$ with $-\infty < A < B < +\infty$, and that $m: \mathbb{F} \to \mathbb{R}$ is defined by $f(D_{m(F)}) = f(F)$ for every F in $\mathbb{F}$. In other words, m is a mean on $\mathbb{F}[A, B]$ with respect to the evaluation of f. De Finetti considers three properties which, though they might be optional, are anyway significant for a general mean with respect to some specific classes of problems. Thus, he says that m is *reflexive* if $m(D_x) = x$; m is *strictly increasing* if, for any pair of elements of $\mathbb{F}$, say F and G, such that $F(x) \le G(x)$ for every x in $\mathbb{R}$ and $F \neq G$, one has $m(F) > m(G)$; m is *associative* if for every t in $[0, 1]$ and F, G, F^*, G^* such that $m(F) = m(F^*)$ and $m(G) = m(G^*)$, one has $m(tF + (1-t)G) = m(tF^* + (1-t)G^*)$. The meaning of the second property is clear when $\mathbb{F}$ denotes the class of all random gains taking values in $[A, B]$. Indeed, in that case, $F(x) \le G(x)$ for every x in $\mathbb{R}$ and $F \neq G$ means that, for each x in $\mathbb{R}$, the probability under F of a gain greater than x is not lower than the probability of the same event under G (it is indeed greater for some x). Nowadays, this condition is identified with a well-known definition of stochastic dominance. Finally, associativity states that a mean remains unchanged in spite of changes in some part of the distribution, provided these do not alter the corresponding partial mean. In connection with these properties, de Finetti proves a significant extension of a theorem independently proved by Kolmogorov (1930) and Nagumo (1930):

REPRESENTATION THEOREM FOR AN ASSOCIATIVE MEAN (DE FINETTI–KOLMOGOROV–NAGUMO]. *Suppose that $m: \mathbb{F}[A, B] \to \mathbb{R}$ is a reflexive, strictly increasing and associative mean. Then there is a function ϕ, continuous and strictly increasing in the closed interval $[A, B]$, for which*

$$(*) \quad m(F) = \phi^{-1}\Big(\int_{\mathbb{R}} \phi(x)\, dF(x)\Big), \quad F \in \mathbb{F}[A, B].$$

Moreover, ϕ is uniquely determined up to linear transformations. Conversely, if m is defined by (∗), for a ϕ with the properties stated, then it satisfies reflexivity, strict monotonicity and associativity.

This formulation is drawn from the classic book by Hardy, Littlewood and Pólya (1934), who follow the guidelines of de Finetti's original proof. The 1931 paper includes a differential criterion which proves of interest in stating whether a mean is associative. Moreover, denoting mean (*) by m_ϕ, de Finetti proves that $m_{\phi_1} \geq m_{\phi_2}$ on $\mathbb{F}[A, B]$ if and only if $\phi_2 \circ \phi_1^{-1}$ is convex. Finally, he characterizes the m_ϕ which are homogeneous or translative. A mean m_ϕ is said to be *homogeneous* (*translative*, respectively) if $m_\phi(F_a) = a \cdot m_\phi(F)$ for every $a > 0$ [$m_\phi(F_a) = a + m_\phi(F)$ for every a, respectively] and F in $\mathbb{F}[A, B]$, with $F_a(x) = F(x/a)$ [$F_a(x) = F(x - a)$, respectively] for every x in $\mathbb{R}$.

CHARACTERIZATION OF HOMOGENEOUS AND TRANSLATIVE MEANS. *m_ϕ is homogeneous if and only if ϕ is any linear transformation of $\phi_1(x) = x^m$, $m \in \mathbb{R} \setminus \{0\}$, or of $\phi_2(x) = \log x$. Correspondingly, m_ϕ is translative if and only if ϕ is any linear transformation of $\phi_3(x) = \exp\{cx\}$, $c \in \mathbb{R} \setminus \{0\}$, or of $\phi_4(x) = x$. Therefore, the arithmetic mean is the sole m_ϕ mean which is both homogeneous and translative.*

The paper *A proposito di correlazione* (de Finetti, 1937b) originates from a debate on the use and misuse of the correlation coefficient which came to the fore during the 23rd Session of the International Statistical Institute (ISI) (London, 1934). De Finetti, thanks to his in-depth knowledge of metric and, in particular, pre-Hilbert spaces, deals with correlation from a geometrical viewpoint. First, he considers the set of all random quantities X, with $E(X^2) < +\infty$, as a metric space with distance between X and Y given by the standard deviation of $(X - Y)$. In such a case, X and Y are thought of as coincident if and only if $P(|Y - X - a| > \varepsilon) = 0$ for every $\varepsilon > 0$ and for some constant a. This metric space can be viewed as a pre-Hilbert space with inner product $\sigma(X)\sigma(Y)r(X, Y)$, where σ denotes standard deviation and r correlation coefficient. After stressing some interesting descriptive properties on the basis of geometrical properties of the space in question, de Finetti deals with some special cases and tackles the problem raised in the ISI debate: can the correlation coefficient be considered as a measure of monotone dependence between two numerical properties of a population? The answer is positive if and only if the marginals F_1 and F_2 are such that $F_1(x) = F_2(ax + b) = 1 - F_2(-ax + \beta)$ for every x and for some $a > 0$ and b, β in $\mathbb{R}$. Finally, in connection with the problem of defining true measures of monotone dependence, de Finetti proposes the following two indices for a bivariate probability distribution function with density function φ:

$$c_1 = \int_\Gamma \varphi(x, y)\varphi(\xi, \eta)\, dx\, dy\, d\xi\, d\eta - 1/2,$$

$$c_2 = 2\sigma(X)\sigma(Y)r(X, Y),$$

with $\Gamma = \{(x, y, \xi, \eta) \in \mathbb{R}^4 : (\xi - x)(\eta - y) > 0\}$. After clarifying that c_1 and c_2 may have opposite signs, de Finetti mentions an analogy between c_1 and c_2: c_1 takes into consideration the signs of $(\xi - x)(\eta - y)$ only, while c_2 considers the value of that product, too. Clearly, c_2 is the covariance of φ up to a multiplicative constant, while c_1, surprisingly, is the well-known Kendall's τ; see Kendall (1938). Hence, de Finetti introduced this index into statistical literature one year before Kendall did, so that one might designate it as *de Finetti–Kendall index* (*of monotone dependence*). To tell the truth, the *Encyclopedia of Statistical Sciences* informs us that this index had already been discussed around 1900 by Fechner, Lipps, and Deuchler and more theoretically in the 1920s by Esscher and Lindeberg.

De Finetti's line of reasoning in descriptive statistics is explained in *La matematica nelle concezioni statistiche* (de Finetti, 1943), where he introduces a natural distinction with reference to the descriptive power of means and statistical indices with respect to concrete problems. In the light of the above distinction, these indices have a *significant* value if they embody the actual statement of a problem; otherwise they have a mere *indicative* value when there is no given stated problem to be answered. For example, a mean has significant value if it is assessed according to Chisini's definition since, in this case, it meets the precise requirement of a rational statement. Passing from means to indices, de Finetti claims that Chisini's line of reasoning relating to means can be extended to other statistical indices in order to obtain significant information with reference to important aspects of a distribution. He gives some hints at a treatment for large classes of variability indices, and quotes his paper on the correlation coefficient. Finally, he mentions an interesting aspect of the construction of statistical indices in case of indices with a mere indicative value. He, indeed, points out that some important properties of a distribution have a more or less precise qualitative meaning which allows us to set up a *partial* ordering in the class of all distributions. For instance, the usual concept of concentration of a positive and transferable character leads us to state that distribution F is no more concentrated than distribution G when the (Lorenz–Gini) concentration curve of F

lies above the concentration curve of G. Then, two elements are not comparable if their concentration curves intersect. Generally speaking, any statistical index defines a complete ordering in the class of all distributions so that, if a "natural" partial ordering exists as a consequence of the definition of the property taken into consideration, then an index ought to agree with such a partial ordering. In a previous paper (de Finetti, 1939c), de Finetti had linked these ideas to lattice theory, as explained in a booklet by Glivenko; see Glivenko (1938). It seems to us that today's theory of stochastic dominance, maybe unconsciously, is very much imbued with principles clearly reminiscent of de Finetti's viewpoint on the description of frequency distributions. A course length treatment which develops this approach can be found in Regazzini (1987).

We conclude this subsection by mentioning a further contribution to descriptive statistics by de Finetti: the extension of Lévy's concept of dispersion to multivariate distributions; see de Finetti (1953b).

Despite his interest in descriptive statistics, on several occasions de Finetti doubted the very validity of any statistical argument taken disjointly from induction. In his opinion, correct inductive procedures are necessary in order that descriptive methods may rationally embody all available information with respect to the purpose of a given statistical data analysis. Hence, it is high time indeed we reviewed his work in the field of statistical inference.

3.2 Induction and Statistics

Before getting to the heart of the matter, that is, de Finetti's outlook on statistical inference, let us mention some specific cases considered by our author. During his stay at the *Istituto Centrale di Statistica*, he worked together with Gini on the prospective growth of the Italian population and, on that occasion, he dealt with the problem of determining the parameters of a logistic curve. The conclusions drawn from this study can be found in de Finetti (1931f).

In *Sulla legge di probabilità degli estremi* (de Finetti, 1932b), he considers the problem of determining the distribution of $M_n = \max(X_1, \dots, X_n)$ when $X_1, \dots, X_n$ are i.i.d. random variables. In particular, he deals with the problem of "estimating" M_n by the median ξ_n of the probability distribution function of M_n. If the probability distribution function F of each X_i is continuous and strictly increasing, then ξ_n is the unique root of the equation $F(\xi_n) = 1/\sqrt[n]{2}$. So, after providing the condition under which $|M_n - \xi_n| \to 0$ in probability (as $n \to +\infty$), he states that ξ_n can be considered as a good estimator of M_n, for n sufficiently large. In order to evaluate the order of magnitude of $|M_n - \xi_n|$, he considers the case in which F is normal and proves that $\xi_n|M_n - \xi_n|$ has a limiting distribution F_0 (as $n \to +\infty$), precisely: $F_0(x) = (1/2)^{\exp(-x)}$, $x > 0$. Some specialists in the field of order statistics acknowledge the pioneer nature of de Finetti's paper; see David (1981).

In a very short communication at the 23rd Meeting of the Società Italiana per il Progresso delle Scienze (Naples, 1934), he tackled the problem which is now known as nonparametric estimation of a cumulative distribution function; see de Finetti (1935b). In point of fact, what we have is merely a statement of the problem within a proper Bayesian framework. On one hand, he assumes that $\phi: \mathbb{R}^{n+1} \to \mathbb{R}$ represents a density function for the law of the random vector $(X_1, \dots, X_n, X)$ and identifies the problem at issue with the one of determining the distribution of X, after observing $X_1 = x_1, \dots, X_n = x_n$. Apropos of this solution, he recalls that, if $(X_1, \dots, X_n, X)$ is the initial segment of a sequence of exchangeable random variables, then the cumulative distribution corresponding to ϕ is close to the empirical distribution of the x_i's, provided that n is sufficiently large. On the other hand, de Finetti thinks of the "unknown" distribution as a realization of a random function whose trajectories are probability distribution functions. Hence, after assigning the probability distribution of that random function and the conditional probability law of data given that very same distribution function, one will use the Bayes theorem to obtain the posterior distribution of the random function. Within this framework, the unknown distribution is estimated by some suitable "typical value" of the posterior distribution (mean, median, etc.). In our opinion, the latter of the two above-mentioned approaches is valuable on two accounts. In the first place, by trying another tack, he clearly envisaged the Bayesian nonparametric analysis of statistical problems: please note that in statistical literature this came into the fore 40 years later, thanks to Ferguson (1973). In the second place, in the paper in question, de Finetti's view stood out most clearly.

3.2.1 *Probability and induction*. Anyway, this view is thoroughly expounded in *La prévision* (de Finetti, 1937a) and in *La probabilità e la statistica* (de Finetti, 1959). Let us then confine ourselves to reporting only the most significant aspects. In spite of the long lapse of time between the two works, the basic train of thought expressed throughout is always quite consistent. True enough, the 1959 lectures do include also a critical survey of the major stances of the time on statistical inference, but in 1937 all the "core" had already been clearly formulated.

According to de Finetti's interpretation, statistical inference is a special case of *reasoning by induction*, where "*reasoning by induction*" means learning from experience. However:

> to speak of inductive *reasoning* means to attribute a certain validity to this way of learning, to view it not as an upshot of quirky psychological reaction, but as a mental process in its own right susceptible of analysis, interpretation and justification. When it comes to induction, the tendency to overestimate reason, often to the point of excluding *tout court* any other factor, can prove a very detrimental bias. Reason is invaluable as a supplement to other intuitive faculties, but should never be a substitute for them....
>
> A consequence of this bias is the elevation of deductive reasoning to the status of a standard, though actually all non-tautological truths are based on something else. Thus inductive reasoning is generally considered as something on a lower level, warranting caution and suspicion. Worse still, attempts to give it dignity try to change its nature making it seem like something that could almost be included under deductive reasoning.
>
> In point of fact, there are often attempts to explain induction without even introducing the term "probability", or else trying to wrest the term from its every-day significance as a measure of a degree of belief attributed to the various possible alternatives. [*From the English translation of La probabilità e la statistica in de Finetti*, 1972, *pages* 147–148.]

It is what routinely occurs in all statistical frequentistic procedures, when the very nature and meaning of probabilistic reasoning is indeed basically altered, for example: attempting to express its conclusions as predictions, and speaking of such things as the possibility of verification or of agreement with experience; speaking of the "impossibility" of highly improbable events; and resorting in some way to those interpretations which many consider necessary in order to give status to the calculus of probability by suppressing all trace of its peculiar task, which is to deal with uncertainty. In more general terms, by postulating the existence of natural laws with more or less narrowly predetermined characteristics, you can even pass off inductive reasoning as actually deductive. All the conceptions which view the determinism as the necessary postulate of science itself are basically founded on this assumption, and, consequently, they reject the very core of inductive reasoning.

With these considerations as a starting point, de Finetti proceeds very much along the lines of Hume's idea of causality. When we regard any events as causally connected, all we do and can observe is that they frequently and uniformly go together, and the impression or idea of the one event brings with it the idea of the other. A habitual association is set up in the mind, and as in other forms of habit, so in this one, the working of the association is felt as compulsion. This feeling is the only discoverable impressionable source of the idea of causality. In other words, when we assert a causal connection between any two objects, the sole experience or evidenced necessitation is not in them but in the habituated mind. That is where mathematics comes in handy: just to make precise what can indeed be made precise. As a matter of fact, in mathematics, too, inductive reasoning cannot be ruled out. Far from it, in the creative moment it does prove necessary. Nobody could set out to prove a theorem if this very theorem could not have some kind of likelihood. Lévy used to say that if you want to get somewhere, you must anyway envisage your destination (insight) before starting off on your actual journey (logic).

Let us now set aside all the considerations mentioned in passing to concentrate on the role of mathematics in the formulation of inductive reasoning. In order to tackle induction from the mathematical angle it is imperative one should secure a mathematical instrument suitable for mastering the logic of uncertainty. Such an instrument does exist, and it is the theory of probability viewed in a broad sense as in de Finetti's subjectivistic approach. According to de Finetti, the role of probability within the framework of inductive logic boils down to the following: pinpointing how the evaluation of probability relative to future events "must change" following the result of events actually observed. The term "must change" is not to be interpreted as "must be corrected". The probability of occurrence of a conditional event is nothing but a coherent extension of a given probability law and not a "better evaluation of the probability of an unconditional event" (cf. de Finetti, 1970, Chapter 4). In any case, this "change," properly interpreted, is the mathematical equivalent of "learning by experience." Consequently, inductive logic boils down either to the principle of compound probabilities or to its more sophisticated variant which passes under the name of the theorem of Bayes.

In this sense, de Finetti is Bayesian even if, apropos of this qualification, some cautions are in order both in connection with the original Bayes-Laplace paradigm and with respect to the neo-Bayesian practice that paradigm is imbued with. Indeed, in *La prévision* he strongly criticizes the usual interpretation of the Bayes–Laplace paradigm and, in *La Probabilità e la statistica*, he does not mince words and speaks of "reconstruction of the classical formulation according to the subjectivistic viewpoint" with reference to the introduction of exchangeability and partial exchangeability into inductive logic.

A first circumstance to be mentioned here regards the nature of the condition of exchangeability, in other words how to translate the empirical vague idea of *analogous* events, or, more generally, *analogous* random elements into probabilistic terms. In the last analysis, such an assumption stems from a *subjective* judgment, and the solution that the problem of induction finds through the asymptotic theorem for predictive distributions (Section 2.3) is obviously subjective:

> but in itself perfectly logical, while on the other hand, when one pretends to *eliminate* the subjective factors one succeeds only in *hiding* them... more or less skilfully, but never in avoiding a gap in logic. It is true that in many cases—as for example in the hypothesis of exchangeability—these subjective factors never have too pronounced an influence, provided that the experience be rich enough; this circumstance is very important, for it explains how in certain conditions more or less close agreement between the predictions of different individuals is produced, but it also shows that discordant opinions are always legitimate." [*From the English translation of La prévision in Kyburg and Smokler*, 1980, *page* 107.]

A second circumstance which distinguishes de Finetti's position from the usual Bayesian statement of statistical inference is his caution towards the usual interpretation in terms of "random elements independent and identically distributed conditional on a random parameter." In order to explain this view, de Finetti resorts to two examples, at once quite simple and yet representative of situations of doubtless practical meaning: the sequence of drawings with replacement from an urn containing both white and nonwhite balls according to an unknown composition; the sequence of tosses of a same coin. The sequence of events $(E_n)_{n\geq 1}$ (with E_n the drawing of a white ball in the nth drawing, E_n is heads in the nth toss) can be viewed as exchangeable in both cases. Consequently, the representation theorem formulated in Section 2.3 holds that from a formal viewpoint the law of the sequence is just a mixture of Bernoullian processes. Again, from a merely formal viewpoint, exchangeable events correspond to those that are commonly viewed as independent with constant but unknown probability θ, where θ is distributed according to the mixing distribution which appears in de Finetti's representation theorem. Why then start out on a new definition and go to great lengths just to prove something well known, we may wonder. Here is the answer. The old definition cannot possibly be stripped of its metaphysical character. In its light, you are bound to admit that, beyond the probability law corresponding to your subjective judgment (the exchangeable law) there must "exist" another law, quite unknown, which corresponds to something real. The diverse hypotheses which might then be made on this unknown law, according to which events would no longer be dependent, but actually stochastically independent, would in turn constitute events whose probability should then be duly evaluated. To de Finetti's way of thinking, that is to say within the framework of his subjectivistic approach, all this does not actually make any sense. In the case of the urn of unknown composition, we could no doubt speak of the probability of the composition of the urn as well as of the conditional probability given each composition. In point of fact, the composition is empirically verifiable. On the other hand, if you play heads or tails and toss an "unfair" coin, you cannot infer that this very unfairness has a given influence over the "unknown probability" and hence view this unfairness as an observable hypothesis. In fact, this "unknown probability" cannot possibly be defined, and all the hypotheses one would like to introduce in this way have no objective meaning.

The difference between the two previous cases is quite substantial and cannot be neglected. One cannot "by analogy" recover in the second case the reasoning which was valid for the urn, for this reasoning no longer applies in the second case. In other words, it is plainly impossible to stick to the usual fuzzy interpretation. Instead, if one wants to go about it in a rigorous way, one should tackle both cases by consistently starting from what they do have in common, that is, the similarity of the environmental conditions in which both the successive drawings and the successive tosses take place. This is the very same similarity which is conducive to the adoption of the probabilistic condition of exchangeability. This way it is exchangeability which presides over inductive reasoning in the presence of the classical conditions which characterize the statistical version of such a kind of reasoning. Moreover, as we have seen in Section 2.3, all reasonings regarding events must be extended to random elements taking values in arbitrary spaces.

The message from these thought-provoking remarks is clear enough: statistical inference must confine itself to considering objective hypotheses on something that can be really observed, at least in principle. Hence, probabilistic models, introduced to formalize statistical problems ought first and foremost to consider finitary events but they ought not to fix *compulsory* rules in connection with the extension of probabilities from finitary to infinitary events. In fact, while it is reasonable to believe that one is able to represent, in some acceptable way, one's judgment of probability on observable facts, it is illusory to think that one might sensibly express any judgment about something which has no empirical meaning, like infinitary events. Hence, any attempt to lend objectivity or any empirical interpretation to entities which originate from any extension of a probability to infinitary events (through purely mathematical arguments) is generally in conflict with the difficulty of justifying a special status (from a logical standpoint) for any extension of the type quoted above. Diaconis (1988) is very clear on this point and he deserves to be cited directly:

> de Finetti's alarm at statisticians introducing reams of unobservable parameters has been repeatedly justified in the modern curve fitting exercises of today's big models. These seem to lose all contact with scientific reality focusing attention on details of large programs and fitting instead of observation and understanding of basic mechanism. It is to be hoped that a fresh implementation of de Finetti's program based on observables will lead us out of this mess.

For these reasons, de Finetti's representation theorem, in view of statistical interpretations, ought to be enounced according to its original "weak" formulation (Section 2.3) or, in other words, one ought to avoid the seemingly more enticing strong formulation "the observable X_n's are exchangeable if and only if they, conditional on a random element θ, are independent and identically distributed". Indeed, the strong formulation could lead to justifying inference on θ even when θ is not observable. To strengthen this position, let us recall that the existence of θ, in general, is a mere mathematical statement, due to some particular extension of a family of coherent finite-dimensional distributions to a complete law for the entire sequence $(X_n)_{n\geq 1}$. Example 3.1 in Regazzini and Petris (1992) shows that there are sequences of exchangeable events whose corresponding frequency of success does not converge in any well-specified stochastic sense. These remarks lead us to focus on de Finetti's insistence on the importance of the weak approach (based on finite additivity) to probability. In general, one thinks that the dilemma "finite versus complete additivity" concerns mathematics or philosophy, so that there are "subjectivistic Bayesian statisticians" who just think of the issue as simply trivial. In spite of this, the previous remarks show that the very meaning of statistical inference may depend on the solution envisaged.

Up to now we have dealt with statistics as inductive reasoning on analogous observable random elements, which justifies the assumption of exchangeability when it comes to those random elements. In this case, induction is independent of the order of observations. Apropos of this, de Finetti, in *La prévision*, points out:

> One can indeed take account not only of the observed frequency, but also of regularities or tendencies toward certain regularities which the observations can reveal. Suppose, for example, that the n first trials give alternately a favourable result and an unfavourable result. In the case of exchangeability, our prediction of the following trial will be the same after these n trials as after any other experience of the same frequency of 1/2, but with a completely irregular sequence of different results; it is indeed the absence of any influence of the order on the judgments of a certain individual which characterizes, by definition, the events that he will consider "exchangeable". In the case where the events are not conceived of as exchangeable, we will, on the other hand, be led to modify our predictions in a very different way after n trials of alternating results than after n irregularly disposed trials having the same frequency of 1/2; the most natural attitude

> will consist in predicting that the next trial will have a great probability of presenting a result opposite to that of the preceding trial.
>
> It would doubtless be possible and interesting to study this influence of order in some simple hypotheses, by some more or less extensive generalization of the case of exchangeability, and some developments tied up with that generalization.... [*From the English translation in Kyburg and Smokler*, 1980, *page* 106.]

At that time, a comprehensive survey of the issue was still to be carried out. A first generalization to be mentioned is the definition of partial exchangeability by de Finetti himself; see Section 2.3. In *La probabilità e la statistica*, he emphasizes the role of this generalization in inductive reasoning and gives a good piece of advice on how to fit Markov dependence into partial exchangeability.

NOTE. The trail blazed by de Finetti and leading to reconsidering the Bayes–Laplace paradigm has been followed by a small group of scholars, as shown in Section 2.3; see also Diaconis (1988). With reference to the Italian part of this group, first and foremost let us mention the book by Daboni and Wedlin (1982), who reformulate the basic statistical results according to de Finetti's approach. Other Italian scholars have considered characterizations of statistical models starting from exchangeability condition combined with a new notion of predictive sufficiency; see Campanino and Spizzichino (1981), Cifarelli and Regazzini (1982) and Spizzichino (1988). In any case, the last-mentioned subject is connected with the previously cited works by Diaconis and Freedman.

3.2.2 *de Finetti, the objectivistic schools and the theory of decisions.* De Finetti's attitude toward the so-called objectivistic schools was extremely critical, although he thought of the rise of objectivistic conceptions as a reaction to unacceptable and discredited formulations of the Bayes–Laplace paradigm, based on mysterious "unknown probabilities." In his opinion:

> Fisher's rich and manifold personality shows a few contradictions. His common sense in applications on one hand and his lofty conception of scientific research on the other lead him to disdain the narrowness of a genuinely objectivistic formulation, which he regarded as a "wooden attitude." He professes his adherence to the objectivistic point of view by rejecting the errors of the Bayes–Laplace formulation. What is not so good here is his mathematics, which he handles with mastery in individual problems but rather cavalierly in conceptual matters, thus exposing himself to clear and sometimes heavy criticism. From our point of view it appears probable that many of Fisher's observations and ideas are valid provided we go back to the intuitions from which they spring and free them from the arguments by which he thought to justify them. On the other hand, the coherent development of a theory of statistical induction on rigidly objectivistic bases is the specific characteristic of the Neyman–Pearson school. There, probability means nothing but "frequency in the long run," and this also includes moments of distraction for the sake of convenience. To accept either a hypothesis or an estimate does not mean to attribute to it any kind of probability or plausibility. Such acceptance is a mechanical act based not on a judgment of its actual validity but on the frequency of the validity of the method from which it is derived". [*From the English translation of La probabilità e la statistica in de Finetti*, 1972, *pages* 171–172.]

De Finetti met Neyman at the Geneva Conference in 1937. On that occasion, Neyman delivered a lecture on statistical estimation. De Finetti's comment on Neyman's lecture and Neyman's answer to de Finetti's comment illustrate their disagreement on the role of Bayes' theorem; see de Finetti (1939a).

Insofar as the theory of statistical decision is concerned, de Finetti interpreted the advent of the Wald decision theoretical approach to statistical inference as "the most decisive involuntary contribution toward the erosion of the objectivistic positions," in connection with the basic statement that admissible decisions are Bayesian. In fact, he viewed that theory as a modern development of Bernoulli's concept of moral expectation, quite suitable for furthering a reconciliation between different viewpoints inspired, on one hand, by de Finetti's reconstruction of the Bayes–Laplace paradigm and, on the other, by the neo-Bernoullian view of statistical problems. Insofar as we know, he did not obtain new results in decision theory, but he did produce an interesting interpretation of the Bernoulli principle of the moral expectation, based on the de Finetti–Kolmogorov–Nagumo theorem for associative means (Section 3.1). Indeed, by assuming that all random gains take values in $[A, B]$, we can identify the class of all these gains with $\mathbb{F}[A, B]$. Then, if one assumes that a person $\mathscr{P}$ is able to associate a real number $m(F)$ (called the value of F) with each F in $\mathbb{F}[A, B]$, in such a way that it is indifferent for $\mathscr{P}$ to possess F or $m(F)$, then m: $\mathbb{F}[A, B] \to \mathbb{R}$ is a mean. If, in comparing elements of $\mathbb{F}[A, B]$, $\mathscr{P}$ clings to the following rules:

(a) $F = D_\alpha \Rightarrow m(F) = \alpha$;

(b) m is strictly monotone and associative (for the economic meaning of these concepts see Section 3.1);

then, from the aforementioned theorem, one has

$$m(F) = \phi^{-1}\left(\int_{[A,\,B]} \phi(x)\, dF(x)\right),$$

where ϕ is continuous and strictly increasing. This function—the *utility function of money*—can be determined by iterating the following experiment where, without loss of generality, it is assumed that $\phi(A) = A$ and $\phi(B) = B$. Hence $\mathscr{P}$, in order to determine ϕ on $[A, B]$, can assess the value g_1 of a random gain which takes the values A and B with probability 1/2 and 1/2, respectively. Then

$$g_1 = \phi^{-1}\left(\tfrac{1}{2}\phi(A) + \tfrac{1}{2}\phi(B)\right),$$

which yields

$$\phi(g_1) = \tfrac{1}{2}(A + B).$$

Iterating this process, $\mathscr{P}$ will assess the value g_2 of a random gain which takes the values g_1 and B with probability 1/2 and 1/2, respectively. In this way, one determines $\phi(g_2)$ and so on.

The function

$$F \to \int_{[A,\,B]} \phi\, dF$$

is known as *preferability index* or *expected utility*. In a sense, the representation theorem for associative means, according to de Finetti's approach, allows us to reach the main conclusion of the von Neumann–Morgenstern theory. De Finetti's line of reasoning about the theory of statistical decisions, as opposed to that of other scholars holding different views, is clearly expounded in an interesting survey by Piccinato (1986), who is particularly attentive to all developments in the field. De Finetti was also interested in group decision-making. His results in this area have recently been taken up by Barlow, Wechsler and Spizzichino (1988), who suggest an interesting generalization on the matter.

With reference to decision-making, we must not forget that it was in this field of investigation that de Finetti encountered Leonard J. Savage. Their relationship was mutually stimulating. Savage guided de Finetti through objectivistic methods and helped him gain access to wider statistical circles; de Finetti urged Savage to go deeper into the new and exciting perspectives by the neo-Bayesian and neo-Bernoullian thought.

When expounding his thought, de Finetti was always uncompromising. He actually never spared accurate criticism against the views he did not share. Yet he was also eager to find some common ground for all statistical views. The title of one of his papers, "Recent suggestions for the reconciliation of theories of probability" (de Finetti, 1951), is enlightening. Let us quote from it:

> From the subjective standpoint, no assertion is possible without *a priori* opinion, but the variety of possible opinions makes problems depending on different opinions interesting.

And this sounds, indeed, as a wholehearted welcome to our conference on Bayesian robustness.

APPENDIX 1: GENEVA CONFERENCE LECTURES, OCTOBER 1937

- Fréchet (Paris): *On some recent advances in the theory of probability*;
- Steffensen (Copenhagen): *Frequency and probability*;
- Wald (Wien): *Compatibility of the notion of "Kollektiv"*;
- Feller (Stockholm): *On the foundations of probability*;
- Neyman (London): *Probabilistic treatment of some problems in mathematical statistics*;
- de Finetti (Trieste): *On the condition of partial exchangeability*;
- Heisenberg (Leipzig): *Probabilistic statements in the wave quantum theory*;
- Hopf (Leipzig): *Statement of mechanics problems according to measure theory*;
- Pólya (Zürich): *Random walk in a road network*;
- Hostinsky (Brno): *Random change–fluctuations of the number of objects in a compartment*;
- Onicescu (Bucharest): *Sketch of a general theory for chains with complete bounds*;
- Dodd (Austin): *Certain coefficients of regression or trend associated with largest likelihood;*
- Cramér (Stockholm): *On a new central limit theorem*;

• Obrechkoff (Sofiya): *Difference equations with constant coefficients, the Poisson distribution and the Charlier series*;
• Steinhaus (Lvov): *Theory and applications of stochastically independent functions;*
• Lévy (Paris): *Arithmetic of probability distributions.*

As a matter of fact, other thought-provoking contributions were on the agenda, namely:

• Bernstein (Leningrad): *Theory of stochastic differential equations*;
• Cantelli (Rome): *About the definition of random variable;*
• Glivenko (Moscow): *About the strong law of large numbers in a functions space;*
• Jordan (Budapest): *Critical aspects of the correlation theory from a probabilistic viewpoint;*
• Kolmogorov (Moscow): *Random functions and their applications*;
• von Mises (Istanbul): *Generalization of classical limit theorems*;
• Romanowsky (Tashkent): *Some new results in the theory of Markov chains;*
• Slutsky (Moscow): *Some statements in the theory of random functions with a continuous spectrum*.

Unfortunately these last-mentioned lecturers did not turn up. However, the complete texts or at least summaries of the lectures were published in *Actualités Scientifiques et Industrielles*, including a slightly modified French translation of de Finetti's *Resoconto Critico del Colloquio*; see de Finetti (1939a).

ACKNOWLEDGMENTS

We are grateful to Gianfranco Barbieri, Ilenia Epifani and Antonio Lijoi for having helped us, in various ways, with the draft of the present paper. Extremely useful comments were provided by two anonymous referees. This work was partially supported by the MURST (40%, 1994, "Inferenza Statistica: basi probabilistiche e sviluppi metodologici") and by Università Bocconi (Ricerca di base 1994–1995).

REFERENCES

ACCARDI, L. and LU, Y. G. (1993). A continous version of de Finetti's theorem. *Ann. Probab.* **21** 1478–1493.

ALDOUS, D. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** 581–598.

ALDOUS, D. (1985). Exchangeability and related topics. *Ecole d'Eté de Probabilités de Saint Flour XIII. Lecture Notes in Math.* **1117**. Springer, Berlin.

BARLOW, R. E. (1992). Introduction to de Finetti (1937) foresight: its logical laws, its subjective sources. In *Breakthroughs in Statistics* **1**. *Foundations and Basic Theory* (S. Kotz and N. L. Johnson, eds.) 127–133. Springer, New York.

BARLOW, R. E., WECHSLER, S. and SPIZZICHINO, F. (1988). De Finetti's approach to group decision making. In *Bayesian Statistics 3* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 1–15. Oxford Univ. Press.

BERTI, P., REGAZZINI, E. and RIGO, P. (1990). De Finetti's coherence and complete predictive inferences. Technical Report 90.5, CNR-IAMI, Milano.

BERTI, P., REGAZZINI, E. and RIGO, P. (1992). Finitely additive Radon–Nikodým theorem and concentration of a probability with respect to a probability. *Proc. Amer. Math. Soc.* **114** 1069–1085.

BERTI, P., REGAZZINI, E. and RIGO, P. (1994). Coherent prevision of random elements. Technical Report 94.15, CNR-IAMI, Milano.

BERTI, P., REGAZZINI, E. and RIGO, P. (1996). Well-calibrated, coherent forecasting systems. *Theory Probab. Appl.* To appear.

BERTI, P. and RIGO, P. (1989). Conglomerabilità, disintegrabilità e coerenza. Serie Ric. Teor. 11, Dip. Statist., Univ. Firenze.

BERTI, P. and RIGO, P. (1994). A Glivenko–Cantelli theorem for exchangeable random variables. *Statistics and Probability Letters*. To appear.

BJÖRK, T. and JOHANSSON, B. (1993). On theorems of de Finetti type for continous time stochastic processes. *Scand. J. Statist.* **20** 289–312.

BLUM, J., CHERNOFF, J., ROSENBLATT, M. and TEICHER, H. (1958). Central limit theorems for interchangeable processes. *Canad. J. Math.* **10** 222–229.

BOCHNER, S. (1939). Additive set functions on groups. *Ann. Math.* **40** 769–799.

BÜHLMANN, H. (1958). Le problème "limite centrale" pour les variables aléatoires échangeable. *C. R. Acad. Sci. Paris* **246** 534–536.

BÜHLMANN, H. (1960). Austauschbare stochastische Variabeln und ihre Grenzwertsätze. *Univ. California Publications in Statistic* **3** 1–35.

CAMPANINO, M. and SPIZZICHINO, F. (1981). Prediction, sufficiency and representation of infinite sequences of exchangeable random variables. Technical report, Inst. Mathematics "G. Castelnuovo," Univ. Roma.

CANTELLI, F. P. (1917a). Sulla probabilità come limite delle frequenze. *Atti della Reale Accademia Nationale dei Lincei, Serie V, Rendiconti* **26** 39–45.

CANTELLI, F. P. (1917b). Su due applicazioni d'un teorema di G. Boole alla statistica matematica. *Atti della Reale Accademia Nationale dei Lincei, Serie V, Rendiconti* **26** 295–302.

CANTELLI, F. P. (1932). Una teoria astratta del calcolo delle probabilità. *Giorn. Istit. Ital. Attuari* **3** 257–265.

CANTELLI, F. P. (1933a). Considerazioni sulla legge uniforme dei grandi numeri e sulla generalizzazione di un fondamentale teorema del signor Lévy. *Giorn. Istit. Ital. Attuari* **4** 327–350.

CANTELLI, F. P. (1933b). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Istit. Ital. Attuari* **4** 421–424.

CASTELLANO, V. (1965). Corrado Gini: a memoir. *Metron* **24** 3–84.

CASTELNUOVO, G. (1919). *Calcolo delle Probabilità e Applicazioni*, 1st ed. Soc. Ed. Dante Alighieri, Milano.

CHISINI, O. (1929). Sul concetto di media. *Periodico di Matematiche* **9** 105–116.

CIFARELLI, D. M. and REGAZZINI, E. (1982). Some considerations about mathematical statistics teaching methodology suggested by the concept of exchangeability. In *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, eds.) 185–205. North-Holland, Amsterdam.

CIFARELLI, D. M. and REGAZZINI, E. (1987). On a general definition of concentration function. *Sankhyā Ser. B* **49** 307–319.

CIFARELLI, D. M. and REGAZZINI, E. (1990). Distribution functions of means of a Dirichlet process. *Ann. Statist.* **18** 429–442. [Correction: *Ann. Statist.* (1994) **22** 1633–1634.]

CIFARELLI, D. M. and REGAZZINI, E. (1993). Some remarks on the distribution functions of means of a Dirichlet process. Technical Report 93.4, CNR-IAMI, Milano.

CRAMÉR, H. (1934). Su un teorema relativo alla legge uniforme dei grandi numeri. *Giorn. Istit. Ital. Attuari* **5** 3–15.

CRAMÉR, H. (1935). Sugli sviluppi asintotici di funzioni di ripartizione in serie di polinomi di Hermite. *Giorn. Istit. Ital. Attuari* **6** 141–157.

CRISMA, L. (1971). Alcune valutazioni quantitative interessanti la proseguibilità di processi aleatori scambiabili. *Rend. Istit. Mat. Univ. Trieste* **3** 96–124.

CRISMA, L. (1982). Quantitative analysis of exchangeability in alternative processes. In *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, eds.) 207–216. North-Holland, Amsterdam.

DABONI, L. (1953). Considerazioni geometriche sulla condizione di equivalenza per una classe di eventi. *Giorn. Istit. Ital. Attuari* **16** 58–65.

DABONI, L. (1975). Caratterizzazione delle successioni (funzioni) completamente monotone in termini di rappresentabilità delle funzioni di sopravvivenza di particolari intervalli scambiabili tra successi (arrivi) contigui. *Rend. Mat.* **8** 399–412.

DABONI, L. (1982). Exchangeability and completely monotone functions. In *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, eds.) 39–45. North-Holland, Amsterdam.

DABONI, L. and WEDLIN, A. (1982). *Statistica*. Utet, Torino.

DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

DAWID, A. P. (1978). Extendibility of spherical matrix distributions. *J. Multivariate Anal.* **8** 567–572.

DE FINETTI, B. See separate list following the References section.

DE MOIVRE, A. (1711). De mensura sortis. *Philosophical Transactions of the Royal Society (London)* **27** 213–264.

DIACONIS, P. (1977). Finite forms of de Finetti's theorem on exchangeability. *Synthèse* **36** 271–281.

DIACONIS, P. (1988). Recent progress on de Finetti's notions of exchangeability. In *Bayesian Statistics 3* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 111–125. Oxford Univ. Press.

DIACONIS, P., EATON, M. L. and LAURITZEN, S. L. (1992). Finite de Finetti theorems in linear models and multivariate analysis. *Scand. J. Statist.* **19** 289–315.

DIACONIS, P. and FREEDMAN, D. (1980a). de Finetti's theorem for Markov chains. *Ann. Probab.* **8** 115–130.

DIACONIS, P. and FREEDMAN, D. (1980b). Finite exchangeable sequences. *Ann. Probab.* **8** 745–764.

DIACONIS, P. and FREEDMAN, D. (1984). Partial exchangeability and sufficiency. In *Proceedings of the Indian Statistical Institute Golden Jubilee International Conference on Statistics: Applications and New Directions* (J. K. Ghosh and J. Roy, eds.) 205–236. Indian Statistical Institute, Calcutta.

DIACONIS, P. and FREEDMAN, D. (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 394–423.

DIACONIS, P. and FREEDMAN, D. (1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. *J. Theoret. Probab.* **1** 381–410.

DIACONIS, P. and FREEDMAN, D. (1990). Cauchy's equation and de Finetti's theorem. *Scand. J. Statist.* **17** 235–250.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

DUBINS, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *Ann. Probab.* **3** 89–99.

DUBINS, L. E. (1983). Some exchangeable probabilities which are singular with respect to all presentable probabilities. *Z. Wahrsch. Verw. Gebiete* **64** 1–6.

DUBINS, L. E. and FREEDMAN, D. (1979). Exchangeable processes need not be mixtures of independent identically distributed random variables. *Z. Wahrsch. Verw. Gebiete* **48** 115–132.

EATON, M. L., FORTINI, S. and REGAZZINI, E. (1993). Spherical symmetry: an elementary justification. *Journal of the Italian Statistical Society* **2** 1–16.

EPIFANI, I. and LIJOI, A. (1995). Some considerations on a version of the law of the iterated logarithm due to Cantelli. Technical Report 95.21, CNR-IAMI, Milano.

FERGUSON, T. S. (1973). A bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

FERGUSON, T. S., PHADIA, E. G. and TIWARI, R. C. (1992). Bayesian nonparametric inference. In *Current Issues in Statistical Inference: Essays in Honour of D. Basu* (M. Ghosh and P. K. Pathak, eds.). IMS, Hayward, CA.

FICHERA, G. (1987). Interventi invitati al Convegno "Ricordo di Bruno de Finetti Professore nell'Ateneo triestino." *Atti* 26–31. Dip. Matematica Applicata alle Scienze Economiche Statistiche e Attuariali, Univ. Trieste Pub. No. 1.

FISHER, R. A. (1921). On "probable error" of a coefficient of correlation deduced from a small sample. *Metron* **1** 3–32.

FISHER, R. A. (1925). Applications of "Student's distribution." *Metron* **5** 90–120.

FORTINI, S., LADELLI, L. and REGAZZINI, E. (1996). A central limit problem for partially exchangeable random variables. *Theory Probab. Appl.* **41** 353–379.

FORTINI, S. and RUGGERI, F. (1994). On defining neighbourhoods of measures through the concentration function. *Sankhyā Ser. A* **56** 444–457.

FRÉCHET, M. (1930a). Sur l'extension du théorème des probabilités totales au cas d'une suite infinie d'événements. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* **63** 899–900.

FRÉCHET, M. (1930b). Sur l'extension du théorème des probabilités totales au cas d'une suite infinie d'événements. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* **63** 1059–1062.

FRÉCHET, M. (1930c). Sur la convergence "en probabilité." *Metron* **8** 3–50.

FRÉCHET, M. and HALBWACHS, M. (1925). *Le calcul des Probabilités à la Porteé de Tous.* Dunod, Paris.

FREEDMAN, D. (1962a). Mixtures of Markov processes. *Ann. Math. Statist.* **33** 114–118.

FREEDMAN, D. (1962b). Invariance under mixing which generalize de Finetti's theorem. *Ann. Math. Statist.* **33** 916–923.

FREEDMAN, D. (1963). Invariants under mixing which generalize de Finetti's theorem: continuous time parameter. *Ann. Math. Statist.* **34** 1194–1216.

FREEDMAN, D. (1980). A mixture of independent identically distributed random variables need not admit a regular conditional probability given the exchangeable σ-field. *Z. Wahrsch. Verw. Gebiete* **51** 239–248.

FREEDMAN, D. (1984). De Finetti's theorem in a continuous time. Technical Report 36, Dept. Statistics, Univ. California, Berkeley.

GINI, C. (1911). Considerazioni sulle probabilità a posteriori ed applicazioni al rapporto dei sessi nelle nascite umane. *Studi economico-giuridici dell'Università di Cagliari* anno III.

GINI, C. (1939). I pericoli della statistica. *Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **5** 1–44.

GLIVENKO, V. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Istit. Ital. Attuari* **4** 92–99.

GLIVENKO, V. (1936). Sul teorema limite della teoria delle funzioni caratteristiche. *Giorn. Istit. Ital. Attuari* **7** 160–167.

GLIVENKO, V. (1938). *Théorie Générale des Structures*. Hermann, Paris.

HAAG, J. (1928). Sur un problème général de probabilités et ses diverses applications. In *Proceedings of the International Congress of Mathematicians* 659–674. Univ. Toronto Press.

HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1934). *Inequalities*. Cambridge Univ. Press.

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

HOLZER, S. (1985). On coherence and conditional prevision. *Boll. Un. Mat. Ital. C (6)* **1** 441–460.

HORN, A. and TARSKI, A. (1948). Measures on Boolean algebras. *Trans. Amer. Math. Soc.* **64** 467–497.

KALLENBERG, O. (1973). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrsch. Verw. Gebiete* **27** 23–36.

KALLENBERG, O. (1974). Path properties of processes with independent and interchangeable increments. *Z. Wahrsch. Verw. Gebiete* **28** 257–271.

KALLENBERG, O. (1975). Infinitely divisible processes with interchangeable increments and random measures under convolution. *Z. Wahrsch. Verw. Gebiete* **32** 309–321.

KALLENBERG, O. (1982). Characterizations and embedding properties in exchangeability. *Z. Wahrsch. Verw. Gebiete* **60** 249–281.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93.

KHINCHIN, A. (1932a). Sulle successioni stazionarie di eventi. *Giorn. Istit. Ital. Attuari* **3** 267–272.

KHINCHIN, A. (1932b). Sur les classes d'événements équivalentes. *Mathematiceskii Sbornik, Recueil. Math. Moscou* **39** 40–43.

KHINCHIN, A. (1935). Sul dominio di attrazione della legge di Gauss. *Giorn. Istit. Ital. Attuari* **6** 371–393.

KHINCHIN, A. (1936). Su una legge dei grandi numeri generalizzata. *Giorn. Istit. Ital. Attuari* **7** 365–377.

KLASS, M. and TEICHER, H. (1987). The central limit theorem for exchangeable random variables without moments. *Ann. Probab.* **15** 138–153.

KOLMOGOROV, A. N. (1929). Sur la loi des grands nombres. *Atti della Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti* **9** 470–474.

KOLMOGOROV, A. N. (1930). Sur la notion de la moyenne. *Atti della Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti* **12** 388–391.

KOLMOGOROV, A. N. (1931). Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104** 415–458.

KOLMOGOROV, A. N. (1932a). Sulla forma generale di un processo stocastico omogeneo. Un problema di Bruno de Finetti. *Atti della Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti* **15** 805–808.

KOLMOGOROV, A. N. (1932b). Ancora sulla forma generale di un processo omogeneo. *Atti R. Acc. Lincei, Serie VI, Rendiconti* **15** 866–869.

KOLMOGOROV, A. N. (1933a). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Istit. Ital. Attuari* **4** 92–99.

KOLMOGOROV, A. N. (1933b). *Grundbegriffe der Wahrscheinlichkeitsrechnung. Ergebnisse der Mathematik und ihrer Grenzgebiete* **3**. Springer, Berlin. [English translation (1956) *Foundations of the Theory of Probability*, 2nd ed. Chelsea, New York.]

KÜCHLER, V. and LAURITZEN, S. L. (1989). Exponential families extreme point models, and minimal space–time invariant functions for stochastic processes with stationary and independent increments. *Scand. J. Statist.* **15** 237–261.

LAURITZEN, S. L. (1975). General exponential models for discrete observations. *Scand. J. Statist.* **2** 23–33.

LAURITZEN, S. L. (1988). *Extremal Families and Systems of Sufficient Statistics. Lecture Notes in Statist.* **49**. Springer, New York.

LÉVY, P. (1931a). Sulla legge forte dei grandi numeri. *Giorn. Istit. Ital. Attuari* **2** 1–21.

LÉVY, P. (1931b). Sur quelques questions de calcul des probabilités. *Prace Matematyezno Fizyczne* (*Warsaw*) **39** 19–28.

LÉVY, P. (1934). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Sup. Pisa* **3** 337–366; **4** 217–235.

LÉVY, P. (1935). Sull'applicazione della geometria dello spazio di Hilbert allo studio delle successioni di variabili casuali. *Giorn. Istit. Ital. Attuari* **6** 13–28.

LUNDBERG, F. (1903). Approximerad framställning av sannolikhetsfunktionen. Återförsäkring av kollektivrisker. Thesis, Uppsala.

NAGUMO, M. (1930). Über eine Klasse von Mittelwenten. *Japan J. Math.* **7** 71–79.

NEYMAN, J. (1935). Su un teorema concernente le cosiddette statistiche sufficienti. *Giorn. Istit. Ital. Attuari* **6** 320–334.

OTTAVIANI, G. (1939). Sulla teoria astratta del calcolo delle probabilità proposta dal Cantelli. *Giorn. Istit. Ital. Attuari* **10** 10–40.

PALEY, R. E., WIENER, N. and ZYGMUND, A. (1933). Note on random functions. *Math. Z.* **37** 647–668.

PETRIS, G. and REGAZZINI, E. (1994). About the characterization of mixtures of Markov chains. Unpublished manuscript.

PICCINATO, L. (1986). De Finetti's logic of uncertainty and its impact on statistical thinking and practise. In *Bayesian Inference and Decision Techniques* (P. K. Goel and A. Zellner, eds.) 13–30. North-Holland, Amsterdam.

RAMAKRISHNAN, S. and SUDDERTH, W. D. (1988). A sequence of coin toss variables for which the strong law fails. *Amer. Math. Monthly* **95** 939–941.

REGAZZINI, E. (1985). Finitely additive conditional probabilities. *Rend. Sem. Mat. Fis. Milano* **55** 69–89.

REGAZZINI, E. (1987). Argomenti di statistica. Parti 1–2. Dip. Matematica "F. Enriques," Univ. Milano.

REGAZZINI, E. and PETRIS, G. (1992). Some critical aspects of the use of exchangeability in statistics. *Journal of the Italian Statistical Society* **1** 103–130.

RESSEL, P. (1985). De Finetti-type theorems: an analytic approach. *Ann. Probab.* **13** 898–922.

RIGO, P. (1988). Un teorema di estensione per probabilità condizionate finitamente additive. *Atti della XXXIV Riunione Scientifica della Società Italiana di Statistica* **2** 27–34.

ROMANOWSKY, V. (1925). On the moments of standard deviation and of correlation coefficient in samples from normal. *Metron* **5** 3–46.

ROMANOWSKY, V. (1928). On the criteria that two given samples belong to the same normal population. *Metron* **7** 3–46.

ROMANOWSKY, V. (1931). Sulle probabilità a "posteriori." *Giorn. Istit. Ital. Attuari* **2** 493–511.

SAVAGE, L. J. (1954). *The Foundations of Statistics.* Wiley, New York.

SCARSINI, M. and VERDICCHIO, L. (1993). On the extendibility of partially exchangeable random vectors. *Statist. Probab. Lett.* **16** 43–46.

SCHERVISH, M. J., SEIDENFELD, T. and KADANE, J. B. (1984). The extent of non-conglomerability of finitely additive probabilities. *Z. Wahrsch. Verw. Gebiete* **66** 205–226.

SENETA, E. (1992). On the history of the strong law of large numbers and Boole's inequality. *Historia Mathematica* **19** 24–39.

SLUTSKY, E. (1934). Alcune applicazioni dei coefficienti di Fourier all'analisi delle funzioni aleatorie stazionarie. *Giorn. Istit. Ital. Attuari* **5** 435–482.

SLUTSKY, E. (1937). Alcune proposizioni sulla teoria delle funzioni aleatorie. *Giorn. Istit. Ital. Attuari* **8** 183–199.

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

SOBCZYK, A. and HAMMER, P. C. (1944a). A decomposition of additive set functions. *Duke Math. J.* **11** 839–846.

SOBCZYK, A. and HAMMER, P. C. (1944b). The ranges of additive set functions. *Duke Math. J.* **11** 847–851.

SPIZZICHINO, F. (1982). Extendibility of symmetric probability distributions and related bounds. In *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, eds.) 313–320. North-Holland, Amsterdam.

SPIZZICHINO, F. (1988). Symmetry conditions on opinion assessment leading to time-transformed exponential models. In *Accelerated Time Testing and Experts' Opinions* (D. V. Lindley, C. A. Clarotti, eds.) 83–97. North-Holland, Amsterdam.

STUART, A. and ORD, K. (1994). *Kendall's Advanced Theory of Statistics* **1**. *Distribution Theory*. Edward Arnold, London.

THATCHER, A. R. (1970). A note on the early solutions of the problem of the duration of play. In *Studies in the History of Statistics and Probability* (E. S. Pearson and Sir M. Kendall, eds.) **1** 127–130. Griffin, London.

VON PLATO, J. (1991). Finite partial exchangeability. *Statist. Probab. Lett.* **11** 99–102.

WALD, A. (1944). On cumulative sums of random variables. *Ann. Math. Statist.* **15** 283–296.

WEGMAN, E. J. (1986). Some personal recollections of Harald Cramér on the development of statistics and probability. *Statist. Sci.* **1** 528–535.

WOOD, G. R. (1992). Binomial mixtures and finite exchangeability. *Ann. Probab.* **20** 1167–1173.

ZABELL, S. L. (1995). Characterizing Markov exchangeable sequences. *J. Theoret. Probab.* **8** 175–178.

ZAMAN, A. (1984). Urn models for Markov exchangeability. *Ann. Probab.* **12** 223–229.

ZAMAN, A. (1986). A finite form of de Finetti theorem for stationary Markov exchangeability. *Ann. Probab.* **14** 1418–1427.

GENERAL REFERENCES CONCERNING DE FINETTI'S WORK

(1987). *Atti del Convegno "Ricordo di Bruno de Finetti Professore nell'Ateneo triestino."* Dip. Matematica applicata alle scienze economiche statistiche e attuariali. "Bruno de Finetti" Pub. No. 1.

DABONI, L. (1987). Bruno de Finetti. *Boll. Un. Mat. Ital. A* 283–308.

GANI, J., ed. (1982). *The Making of Statisticians*. Springer, New York.

GOEL, P. K. and ZELLNER, A., eds. (1986). *Bayesian Inference and Decision Techniques. Essays in Honor of Bruno de Finetti.* North-Holland, Amsterdam.

KOCH, G. and SPIZZICHINO, F., eds. (1982). *Exchangeability in Probability and Statistics*. North-Holland, Amsterdam.

KOTZ, S. and JOHNSON, N. L., eds. (1992). *Breakthroughs in Statistics* **1**. *Foundations and Basic Theory*. Springer, New York.

KYBURG, H. E., JR. and SMOKLER, H. E., eds. (1980). *Studies in Subjective Probability*. Krieger, Malabar, FL.

(1987). *La Matematica Italiana tra le Due Guerre Mondiali*. Milano, Gargnano del Garda, 8–11 ottobre 1986. Pitagora Ed., Bologna.

LINDLEY, D. V. (1989). De Finetti, Bruno. In *Encyclopedia of Statistical Sciences* (Supplement) 46–47. Wiley, New York.

REGAZZINI, E. (1987). Probability theory in Italy between the two world wars. A brief historical review. *Metron* **44** 5–42.

VON PLATO, J. (1994). *Creating Modern Probability*. Cambridge Univ. Press.

CITED WORKS OF DE FINETTI

Articles

(1929a). Sulle funzioni a incremento aleatorio. *Atti. Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **10** 163–168.

(1929b). Sulla possibilità di valori eccezionali per una legge ad incrementi aleatori. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **10** 325–329.

(1929c). Integrazione delle funzioni a incremento aleatorio. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **10** 548–552.

(1930a). Sui passaggi al limite nel calcolo delle probabilità. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* **63** 155–166.

(1930b). A proposito dell'estensione del teorema delle probabilità totali alle classi numerabili. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* **63** 901–905.

(1930c). Ancora sull'estensione alle classi numerabili del teorema delle probabilità totali. *Rendiconti Reale Istituto Lombardo di Scienze e Lettere* **63** 1063–1069.

(1930d). Problemi determinati e indeterminati nel calcolo delle probabilità. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **12** 367–373.

(1930e). Sulla proprietà conglomerativa delle probabilità subordinate. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **12** 278–282.

(1930f). Introduzione matematica alla statistica metodologica. Riassunto di un corso di lezioni per laureati, Istituto Centrale di Statistica, Roma.

(1930g). Funzione caratteristica di un fenomeno aleatorio. *Atti Reale Accademia Nazionale dei Lincei, Mem.* **4** 86–133.

(1930h). Le funzioni caratteristiche di legge istantanea. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **12** 278–282.

(1930i). Calcolo della differenza media. *Metron* **8** 89–94 (with U. Paciello).

(1931a). Probabilismo. Saggio critico sulla teoria della probabilità e sul valore della scienza. *Logos* (*Biblioteca di Filosofia*) 163–219. Perrella, Napoli.

(1931b). Sul significato soggettivo della probabilità. *Fundamenta Mathematicae* **17** 298–329.

(1931c). Le funzioni caratteristiche di legge istantanea dotate di valori eccezionali. *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rend.* **14** 259–265.

(1931d). Sui metodi proposti per il calcolo della differenza media. *Metron* **9** 47–52.

(1931e). Sul concetto di media. *Giorn. Istit. Ital. Attuari* **2** 369–396.

(1931f). Calcoli sullo sviluppo futuro della popolazione italiana. *Annuali di Statistica* **10** 3–130 (with C. Gini).

(1931g). Le leggi differenziali e la rinunzia al determinismo. *Rendiconti del Seminario Matematico dell'Universitá di Roma* **7** 63–74.

(1932a). Funzione caratteristica di un fenomeno aleatorio. In *Atti del Congresso Internazionale dei Matematici* 179–190. Zanichelli, Bologna.

(1932b). Sulla legge di probabilità degli estremi. *Metron* **9** 127–138.
(1933a). Classi di numeri aleatori equivalenti. La legge dei grandi numeri nel caso dei numeri aleatori equivalenti. Sulla legge di distribuzione dei valori in una successione di numeri aleatori equivalenti (Three articles). *Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti* **18** 107–110; 203–207; 279–284.
(1933b). Sull'approssimazione empirica di una legge di probabilità. *Giorn. Istit. Ital. Attuari* **4** 415–420.
(1935a). Sull'integrale di Stieltjes–Riemann. *Giorn. Istit. Ital. Attuari* **6** 303-319 (with M. Jacob).
(1935b). Il problema della perequazione. In *Atti della Società Italiana per il Progresso delle Scienze (XXIII Riunione, Napoli, 1934)* **2** 227–228. SIPS, Roma.
(1937a). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7** 1–68. [English translation in *Studies in Subjective Probability* (1980) (H. E. Kyburg and H. E. Smokler, eds.) 53–118. Krieger, Malabar, FL.]
(1937b). A proposito di "correlazione." *Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **3** 41–57.
(1938a). Sur la condition de "équivalence partielle." In *Actualités Scientifiques et Industrielles* **739** 5–18. Hermann, Paris.
(1938b). Funzioni aleatorie. In *Atti del I Congresso UMI* 413–416. Zanichelli, Bologna.
(1938c). Resoconto critico del colloquio di Ginevra intorno alla teoria delle probabilità. *Giorn. Istit. Ital. Attuari* **9** 1–40.
(1939a). Compte rendu critique du colloque de Génève sur la théorie des probabilités. In *Actualités Scientifiques et Industrielles* **766**. Hermann, Paris.
(1939b). La teoria del rischio e il problema della "rovina dei giocatori." *Giorn. Istit. Ital. Attuari* **10** 41–51.
(1939c). Indici statistici e "teoria delle strutture." In *Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **5** 71–74. S.A.T.E., Ferrara.
(1943). La matematica nelle concezioni e nelle applicazioni statistiche. *Statistica* **3** 89–112.
(1949). Sull'impostazione assiomatica del calcolo delle probabilità. Aggiunta alla nota sull'assiomatica della probabilità (Two articles). *Annali Triestini* **19** sez. II 29–81; **20** sez. II 3–20.
(1951). Recent suggestions for the reconciliation of theories of probability. *Proc. Second Berkeley Symp. Math. Statist. Probab.* 217–225. Univ. California Press, Berkeley.
(1952). Gli eventi equivalenti e il caso degenere. *Giorn. Istit. Ital. Attuari* **15** 40–64.
(1953a). Trasformazioni di numeri aleatori atte a far coincidere distribuzioni diverse. *Giorn. Istit. Ital. Attuari* **16** 51–57.
(1953b). Sulla nozione di "dispersione" per distribuzioni a più dimensioni. In *Atti del IV Congresso UMI* 587–596. Cremonese, Roma.
(1955a). La struttura delle distribuzioni in un insieme astratto qualsiasi. *Giorn. Istit. Ital. Attuari* **18** 12–28. [English translation in B. de Finetti, *Probability, Induction and Statistics* (1972) 129–140. Wiley, New York).]
(1955b). Sulla teoria astratta della misura e dell'integrazione. *Ann. Mat. Pura Appl.* **40** 307–319. [English translation in B. de Finetti, *Probability, Induction and Statistics* (1972) 115–128. Wiley, New York.]
(1959). La probabilità e la statistica nei rapporti con l'induzione, secondo i diversi punti di vista. *Atti corso CIME su Induzione e Statistica* Varenna 1–115. Cremonese, Roma. [English translation in B. de Finetti, *Probability, Induction and Statistics* (1972) 147–227. Wiley, New York.]
(1969). Sulla proseguibilità di processi aleatori scambiabili. *Rend. Istit. Mat. Univ. Trieste* **1** 53–67.
(1976). La probabilità: guardarsi dalle falsificazioni. *Scientia* **111** 255–281. [English translation in *Studies in Subjective Probability* (1980) (H. E. Kyburg and H. E. Smokler, eds.) 193–224. Krieger, Malabar, FL.]

Books

(1970). *Teoria delle Probabilità*. Einaudi, Torino. Two volumes. [English translation, *Theory of Probability* (1975) Wiley, New York, two volumes.]
(1972). *Probability, Induction and Statistics*. Wiley, New York.
(1981). *Scritti* (1926–1930). Cedam, Padova.
(1991). *Scritti* (1931–1936). Pitagora Ed., Bologna.
(1993). *Probabilità e Induzione* (*Induction and Probability*) (P. Monari and D. Cocchi, eds.). Clueb Ed., Bologna.
(1995). *Filosofia della probabilità* (A. Mura, ed.). Theoria Il Saggiatore, Milano.