Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 1–17

The unsung de Finetti's first paper about exchangeability

FEDERICO BASSETTI – EUGENIO REGAZZINI

Abstract: *It is a singular fact that the first and pithy de Finetti's essay on exchangeability has not earned the same reputation as that of others of his papers about the same subject. In fact, this paper contains, on the one hand, all the main results on sequences of exchangeable events, together with the right subjectivistic interpretation of the role they play in the study of the connections between probability and frequencies. On the other hand, the paper makes use of mathematical methods abandoned, immediately after its publication, by de Finetti himself. The center of this methods is the so–called characteristic function of a random phenomenon. Independently of the destiny of the paper, we think that, apart from its undoubted historical value, it contains ideas susceptible of interesting new developments. Therefore, we have deemed it suitable to give here a detailed and faithful account of its content, for the benefit of the colleagues who are not in a position to understand Italian. Moreover, to emphasize the value of the paper at issue, we develop de Finetti's brief hint to the extendibility of exchangeable sequences of events, to obtain a new explicit necessary and sufficient condition of an algebraic nature.*

1 – Introduction

Bruno de Finetti (1906–1985) is regarded as the founder of the theory of *sequences* of *exchangeable random variables* or *random exchangeable sequences* for short. His first important article about this subject dates back to 1930 (see [6]). It appears as a *Memoria*, published in the proceedings of the *Accademia*

Key Words and Phrases: *Characteristic function of a random phenomenon – De Finetti's contributions to probability – Exchangeable events – Extendibility of exchangeable sequences.*
A.M.S. Classification: 60G09, 60-03, 01A60

dei Lincei, jointly presented to the *Accademia* by two of the most outstanding Italian scientists of the time: Guido Castelnuovo (1865–1952) and Tullio Levi–Civita (1873–1941). It is clear, from his biography, that de Finetti had examined the problem of finding a probabilistic description and interpretation for *random phenomena* – those which can be repeatedly observed under homogeneous environmental conditions – ever since his early approach to probability, a couple of years before his degree in mathematics, obtained in 1927 at the University of Milan. In fact, he presented a summary of the *Memoria* in a pithy communication at the *International Mathematical Congress* held in Bologna, in September 1928. The text of such a communication was published, in 1932, in the sixth volume of the proceedings of the *Congress*. See [11].

The present paper aims at giving a precise idea of the content of the *Memoria* and, especially, of the methods used therein, since they are different from those employed in later de Finetti's contributions to the same subject. This analysis is split into six points which form Section 2. Some new developments of de Finetti's original methods are sketched in Section 4. The intermediate Section 3 reviews a paper by Jules Haag (1882–1953) in which, so far as we know, the concept of exchangeable events had been introduced and studied for the first time. A comparison between this paper and de Finetti's *Memoria* clarifies the complete independence of the two papers, and it convinces of the prominent merits of de Finetti in this field.

2 – Characteristic function of a random phenomenon

In view of the homogeneity of the environmental conditions which distinguishes random phenomena (with *equivalent* trials) from other types of phenomena, de Finetti points out 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. Accordingly, he defines a sequence $(E_n)_{n\geq 1}$ of events to be *equivalent* if, for every *finite* permutation π, the probability distribution of $(\mathbb{I}_{E_1}, \mathbb{I}_{E_2}, \dots)$(1) is the same as the probability distribution of $(\mathbb{I}_{E_{\pi(1)}}, \mathbb{I}_{E_{\pi(2)}}, \dots)$. So, if $\omega_k^{(n)}$ denotes the probability that the random phenomenon, taken into consideration, comes true k times in m trials, one gets

$$\omega_k^{(m)} = \sum_{h=k}^{n-m+k} \omega_h^{(n)} \frac{\binom{h}{k}\binom{n-h}{m-k}}{\binom{n}{m}} \tag{1}$$

whenever $1 \leq m \leq n$ and $k = 0, \dots, m$.

(1) For any event E, $\mathbb{I}_E$ will stand for its indicator.

Nowadays, the term *equivalent* is replaced by the more expressive and unambiguous word *exchangeable*, proposed perhaps by Pólya (cf.[13], [14]) or by Fréchet (cf. [16]).

From (1) de Finetti derives a difference–differential equation for the probability generating function and, consequently, for the characteristic function, of the frequency of success in n trials. Such a characteristic function and its limit, as $n \to +\infty$, in the case of an infinite sequence, becomes the center of de Finetti's treatment of exchangeability.

From a methodological viewpoint, the use of characteristic functions joins the *Memoria* to the contemporaneous de Finetti's studies on processes with stationary independent increments, based on the analysis of the *derivative law* defined in terms of the characteristic function ψ_λ of the λ–th coordinate of the process; [5]. See also [24].

As already recalled in the first section, the application of the characteristic functions method to the study of exchangeable sequences is a peculiarity of the *Memoria*. In point of fact, on Khinchin's advice, in all subsequent papers on exchangeable random elements, de Finetti uses more direct tools such as probability distribution functions, moments, and so on. We have experimented that the original de Finetti approach has some remarkable merits with respect to some important problems like, for instance, concrete assessment of finitary exchangeable laws and extendibility of exchangeability. So, we believe that an accurate and faithful account of that approach could come in handy to all scholars who are unable to read Italian scientific literature.

2.1 – Fundamental recurrence relation.

Our description starts with Author's remark that (1), for $m = n-1$, reduces to

$$n\omega_k^{(n-1)} = (n-k)\omega_k^{(n)} + (k+1)\omega_{k+1}^{(n)}$$

with $\omega_0^{(0)} = 1$. Thus, for any sequence of N exchangeable events, he deduces the difference–differential equation

$$n\Omega_{n-1}(z) = n\Omega_n(z) + (1-z)\Omega_n'(z) \tag{2}$$

valid for $n = 1, \ldots, N$ and any complex number z, where Ω_n is the probability generating function defined by

$$\Omega_n(z) := \sum_{h=0}^{n} \omega_h^{(n)} z^h \qquad (n = 1, 2, \ldots, N;\ z \in \mathbb{C}) \tag{3}$$

with $\Omega_0(z) \equiv 1$.

Firstly, (2) is used to prove the identity

$$\frac{1}{m!}\Big(\frac{d^m\Omega_n}{dz^m}\Big)(1) = \binom{n}{m}\omega_m^{(m)}$$

and, consequently, to write

$$\Omega_n(1+z) = \sum_{h\geq 0} \binom{n}{h} \omega_h^{(h)} z^h. \tag{4}$$

At this stage, de Finetti defines the characteristic function of the frequency[2] $(\sum_{i=1}^{N} \mathbb{I}_{E_k}/N)$

$$t \mapsto \Psi_N(t/N) := \Omega_N(e^{it/N}) \qquad (t \in \mathbb{R})$$

to be the *characteristic function of the* (finite) *class* $\{E_1, \dots, E_N\}$ *of exchangeable events.* Clearly, such a function characterizes the probability distribution of the random vector $(\mathbb{I}_{E_1}, \mathbb{I}_{E_2}, \dots, \mathbb{I}_{E_N})$. Notice that this distribution is also determined by the sole knowledge of the probabilities $\omega_h^{(h)}$, $h = 0, 1, \dots, N$, with $\omega_0^{(0)} = 1$. To see this, combine (3) with (4).

From a practical viewpoint, the following proposition – that the Author states in Section 35 of the *Memoria* – may be useful.

PROPOSITION 1. *Any sequence* $(\tilde{\omega}_h^{(N)})_{h=0,\dots,N}$*, satisfying*

$$\tilde{\omega}_h^{(N)} \geq 0 \quad (h = 0, \dots, N) \quad \textit{and} \quad \sum_{h=0}^{N} \tilde{\omega}_h^{(N)} = 1,$$

generates a unique exchangeable law, for the class of events $\{E_1, \dots, E_N\}$*, according to which the probability that a random phenomenon comes true* k *times in* n *trials* $(1 \leq n \leq N, k = 0, \dots, m)$ *is given by*

$$\sum_{h=k}^{N-n+k} \tilde{\omega}_h^{(N)} \frac{\binom{h}{k}\binom{N-h}{n-k}}{\binom{N}{n}}. \tag{5}$$

Indeed, consider the partition defined by

$$A_h := \{\sum_{k=0}^{N} \mathbb{I}_{E_k} = h\} \qquad h = 0, 1, \dots, N$$

in a probability space such that $\tilde{\omega}_h^{(N)}$ is the probability of A_h. If the event A_h occurs, then h white balls along with $(N-h)$ black balls are placed into an urn. Now, consider an individual who just assesses the quantities $\tilde{\omega}_h^{(N)}$ as the

[2] Throughout the paper, the term frequency is used to designate what other authors call relative frequency.

probabilities for the events A_h and who randomly draws n balls from the urn ($n \leq N$), without replacement. So, if he sees the $N(N-1)\dots(N-n+1)$ possible outcomes as equally probable, whatever n may be, then the probability that the sample contains exactly k white balls is given by (5). In other words, any N–exchangeable $\{0,1\}$ sequence is a mixture of hypergeometric N–sequences.

After establishing these basic elementary facts, de Finetti moves on to the analysis of infinite sequences of exchangeable events. Such analysis is focused on the study of the pointwise limit of the characteristic function $\Psi_N(t/N)$, as $N \to +\infty$. As a matter of fact, in all later writings on exchangeability, he will consider a different approach, based on a law of large numbers for exchangeable sequences. As already mentioned, he adopter this approach following a suggestion of Alexander Khinchin (1984–1969), he met on the occasion of the *Congress* of Bologna. See [12], [20] and [21].

2.2 – Representation theorem

Given an infinite sequence $(E_n)_{n\geq 1}$ of exchangeable events, consistently with the previous notation define $\omega_h^{(h)}$ to be the probability of $E_1 \cap \cdots \cap E_h$, for $h = 1, 2, \dots$, and set

$$\Omega(1+z) := \sum_{h\geq 0} \omega_h^{(h)} \frac{z^h}{h!} \qquad (z \in \mathbb{C}).$$

In Section 6 of the *Memoria* de Finetti proves the following preliminary:

PROPOSITION 2. *For any strictly positive a and ϵ there is an integer $N_1 = N_1(a, \epsilon)$ such that*

$$\sup_{|z|\leq a} |\Omega(1+z) - \Omega_n(1+z/n)| \leq \epsilon \qquad (n \geq N_1).$$

Then, he uses this fact to prove a more important statement concerning the limiting behavior of the characteristic function $\Psi_n(t/n)$, as $n \to +\infty$:

PROPOSITION 3. *For every $\tau > 0$ and $\epsilon > 0$, there is $N_2 = N_2(\epsilon, \tau)$ such that*

$$\sup_{|t|\leq \tau} |\Psi(t) - \Psi_n(t/n)| \leq \epsilon$$

holds true for every $n \geq N_2$ and

$$\Psi(t) := \Omega(1+it) = \sum_{h\geq 0} \omega_h^{(h)} \frac{(it)^h}{h!} \qquad (t \in \mathbb{R}).$$

It is important to note that Proposition 2 and Proposition 3 are valid uniformly with respect to Ω and Ψ, respectively. In other words, given ϵ, a and τ, N_1 and N_2 do not depend on Ω and Ψ, respectively. See next Subsection 2.5 for a different situation apropos of the connection between frequency and predictive distribution.

So, *if one assumes that* Ψ *is a characteristic function* (see the next subsection), then the corresponding random variable must take values in $[0,1]$, with probability one. Moreover, if Φ is the corresponding probability distribution function, since Ψ can be extended as an entire function, one gets

$$\Psi(t) = \int_{[0,1]} e^{it\xi} d\Phi(\xi), \qquad \omega_h^{(h)} = \int_{[0,1]} \xi^h d\Phi(\xi) \quad (h = 0, 1, \dots).$$

This, in turn, combined with (4), gives

$$\Omega_n(1+z) = \int_{[0,1]} (1+z\xi)^n d\Phi(\xi)$$

and

$$\begin{aligned}\sum_{h\geq 0} \omega_h^{(n)} z^h = \Omega_n(z) &= \int_{[0,1]} (1-\xi+z\xi)^n d\Phi(\xi)\\ &= \sum_{h\geq 0} \binom{n}{h} z^h \int_{[0,1]} \xi^h (1-\xi)^{n-k} d\Phi(\xi).\end{aligned}$$

This encompasses the celebrated de Finetti's *representation theorem*, viz.:

Proposition 4. *The events* $(E_n)_{n\geq 1}$ *are exchangeable if and only if there is a probability distribution function* Φ *supported by* $[0,1]$ *such that the probability of* $\{\mathbb{I}_{E_1} = x_1, \dots, \mathbb{I}_{E_n} = x_n\}$ *is given by*

$$\int_{[0,1]} \xi^{\sigma_n} (1-\xi)^{n-\sigma_n} d\Phi(\xi)$$

for every $(x_1, \dots, x_n)$ *in* $\{0,1\}^n$ *for which* $x_1 + \cdots + x_n = \sigma_n$*, and for every* $n = 1, 2, \dots$. *Moreover,* Φ *is the limit* (*in the sense of weak convergence of probability distributions*) *of the probability distribution function* Φ_n *of the frequency of success in the first* n *trials, as* $n \to +\infty$.

2.3 – Important remark

The previous argument is based on the presumption that the limit, Ψ, of Ψ_n is a characteristic function. Nowadays, the validity of such an assertion is proved in any good probability textbook. On the contrary, the reference books at de Finetti's disposal – [1] and [22] – although they were superb, they contained the form of the continuity theorem according to which "if Ψ_n converges to a *characteristic function*, uniformly on any compact interval, then ...". Clearly, the argument in Subsection 2.2, apart from the fact that $t \mapsto \Omega(1+it)$ is the limit – uniform on any compact interval – of a sequence of characteristic functions, does not give further indications about the fact that the limit is a characteristic function. So, to complete the proof of the representation theorem, de Finetti was obliged to check whether the above–mentioned limiting condition was enough to assert that $t \mapsto \Omega(1+it)$ was a characteristic function. He deferred the solution of the problem to the Appendix of the *Memoria*, where he proved the desired completion of the continuity theorem – perhaps for the first time – consistently with the fact that he was dealing with *finitely* (i.e., not necessarily *completely*) *additive* distributions of general real–valued random variables. In point of fact, he explicitly assumes that the sequence of distributions corresponding to $(\Psi_n)_{n\geq 1}$ is tight.

2.4 – Strong law of large numbers

In the following Sections 11 and 12, de Finetti deals with the extension of Cantelli's strong law for frequencies of Bernoulli trials to frequencies of more general exchangeable trials. Define the random frequency $\bar{f}_n$ of success in the first n trials of a random phenomenon characterized by an infinite sequence $(E_n)_{n\geq 1}$ of exchangeable events,

$$\bar{f}_n := \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}_{E_k},$$

and consider the sequence $(\bar{f}_n)_{n\geq 1}$. The main result de Finetti achieves apropos of the latter sequence is *a mutual* form of the strong law of large numbers for $(\bar{f}_n)_{n\geq 1}$ that, consistently with the admissibility of simply additive probability distributions, he states correctly in the following "finitary" style.

Proposition 5. *Given strictly positive numbers ϵ and θ, there is a positive integer $N := N(\epsilon, \theta)$ such that the probability of the event*

$$\bigcap_{j=1}^{k} \{|\bar{f}_n - \bar{f}_{n+j}| \leq \epsilon\}$$

turns out to be uniformly (with respect to $k = 1, 2, \dots$) greater that $1-\theta$, whenever $n \geq N$.

Apparently de Finetti was aware of the fact that, in a framework of completely additive probability distributions on subsets of a sample space, the above proposition holds with $k = +\infty$, and that one can assert the existence of a random number f^*, with probability distribution function Φ, that can be viewed as the almost sure limit of $(\bar{f}_n)_{n\geq 1}$. But he had at least three good reasons, from his viewpoint, to be uninterested in the "strong" formulation of his strong law of large numbers. These reasons, briefly mentioned in many points of the *Memoria*, are discussed in a more systematic way in a few contemporaneous de Finetti's papers such as [7], [8] and [9]. It is worth recalling them here, in a three-point summary. (i) Logically speaking, it is unjustifiable to speak of an *infinite* sequence of trials of a random phenomenon: The number of the trials could be arbitrarily great but, in any case, finite. (ii) Without the assumption of complete additivity and with no reference to a sample space, there is no possibility of deducing the existence of a limiting random quantity from the sole mutual convergence of a given sequence. (iii) de Finetti deduces the whole theory of probability from a very natural condition having an obvious meaning – the so-called *condition of coherence* – and shows that *complete* additivity is not necessary for a quantitative measure of probability to be coherent. See [10].

The strong law of large numbers for the frequency of success in a sequence of exchangeable events represents the last issue dealt with in Chapter 1 of the *Memoria*. Chapter 2 contains the definitions of some operators on the set of all characteristic functions, with the intention of providing a rigorous, systematic presentation, in Chapter 3, of asymptotic properties of the posterior distribution and of the merging of the predictive distribution with the frequency of success in past trials of a given random phenomenon. In view of the purely instrumental function of Chapter 2, here we jump to the more important Chapter 3.

2.5 – Probability and experience: Posterior and predictive distributions

At the time of the draft of the *Memoria*, de Finetti was unfamiliar with techniques of statistical inference, and it's amazing how he was, nevertheless, able at picking out the essence of the inductive reasoning and the tools to deal with it, from a coherent mathematical standpoint. In his view of these subjects, exchangeability is a means to study and understand the role played by the knowledge of data, gathered from experience, with regards to the evaluation of probability. In particular, he aims at clarifying how exchangeability can be employed to provide with a basis the common belief that prevision of new facts rests on the analogy with past observed facts. In the case of a random phenomenon, this belief leads to assume, although with caution, past frequency as an approximate value for probability. So, in Section 27, de Finetti provides a new rigorous description of the asymptotic behavior of the posterior distribution[(3)], and makes use of this statement to show the merging of the predictive

[(3)] In point of fact, he was unaware of [26], where a strictly related problem is studied.

distribution with frequency in past trials. Apropos of the former, he considers *infinite sequences* of exchangeable trials of a random phenomenon, generating convergent sequences of frequencies. More precisely,

PROPOSITION 6. *Let* θ_0 *be any point in the intersection of* $(0,1)$ *with the support of the distribution function,* Φ*, of the random phenomenon. Then the posterior distribution, given* $\{\bar{f}_n = \frac{\sigma_n}{n}\}$*, converges weakly to the point mass* δ_{θ_0}*, whenever* $\sigma_n/n \to \theta_0$ *as* $n \to +\infty$*, i.e.*

$$\lim_{n\to+\infty} \frac{\int_{\theta_0-\epsilon}^{\theta_0+\epsilon} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)} = 1 \qquad (\epsilon > 0).$$

Whence, as for the conditional probability of $\{E_{n+k}\}$ *given* $\{\bar{f}_n = \frac{\sigma_n}{n}\}$*, viz.*

$$\frac{\int_{[0,1]} \theta^{1+\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)},$$

one obtains that, for any $\epsilon > 0$*, there is* $N = N(\epsilon, \Phi)$*, such that*

$$\Big|\frac{\int_{[0,1]} \theta^{1+\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n}d\Phi(\theta)} - \frac{\sigma_n}{n}\Big| \le \epsilon$$

holds true for every $n \ge N$*.*

In Section 28, de Finetti explains the "relative" value of this proposition. Indeed, in view of the dependence of N on Φ, it does not allow a quantitative statement about the approximation of frequency to probability, independently of a complete *a priori* knowledge of the characteristic function of the random phenomenon.

Chapter 3 ends with a brief mention of the use of posterior distribution in the problem of hypothesis–testing: the sole explicit hint to a statistical technique, contained in the *Memoria.*

2.6 – Classes of exchangeable events and extension of exchangeability

The main issue dealt with in the last chapter (Chapter 4, including Sections 31–36) is extendibility of exchangeability, from a finite sequence to a "longer" sequence of events. The problem can be formulated in the following terms: *Given positive integers* n *and* k*, establish conditions on the characteristic function of a random phenomenon of* n *exchangeable events in order that they may constitute the initial* n*–segment of a random phenomenon of* $(n+k)$ *exchangeable events.*

To solve this problem, de Finetti starts from (2), viewed as a first–order linear differential equation in the dependent variable Ω_{n+1}. Since the one–parameter family of solutions of this equation is

$$\Omega_{n+1}(z) = (1-z)^{n+1}\{(n+1)\int_0^z \Omega_n(x)(1-x)^{-(n+2)}dx + c\}, \tag{6}$$

then (6) can be combined with Proposition 1 to obtain

PROPOSITION 7. *$t \mapsto \Omega_n(e^{it/n})$ is the characteristic function of the initial n-segment of a sequence of $(n+1)$ exchangeable events if and only if the constant c in (6) can be determined in such a way that all the coefficients of the polynomial (of degree $(n+1)$), defined by the right–hand side of (6), are nonnegative.*

Analogously, to solve the problem for some $k > 1$, one can start from (6) with $(n+2)$ in the place of $(n+1)$, consider it as an equation in the dependent variable Ω_{n+2} and, finally, substitute Ω_{n+1} with its expression in the right–hand side of (6). So, by an obvious recursive argument, de Finetti states that Ω_{n+k} can be written as

$$\Omega_{n+k}(z) = F(z) + C_1(1-z)^{n+1} + \cdots + C_k(1-z)^{n+k} \tag{7}$$

F being a polynomial, whose coefficients are completely determined by Ω_n. Then:

PROPOSITION 8. *$t \mapsto \Omega_n(e^{it/n})$ is the characteristic function of the initial n–segment of a sequence of $(n+k)$ exchangeable events if and only if the constants $C_1, \ldots, C_k$ can be determined in such a way that all the coefficients of the polynomial (of degree $(n+k)$), defined by the right–hand side of (7), are nonnegative.*

Forty years later, de Finetti came back to the problem from a new standpoint, of a geometrical nature (see [15]), followed by some Authors such as [3], [4], [17] and [27].

In Section 4 of the present paper, we will resume the original analytical de Finetti's argument, by providing an explicit form for F in (7). Our goal is to reformulate the necessary and sufficient condition in Proposition 8 in the guise of a system of linear inequalities.

De Finetti gives a complete solution of the extendibility problem when $k = +\infty$, i.e.: *To establish conditions on $t \mapsto \Omega_n(e^{it/n})$ in order that it can be viewed as characteristic function of the first n trial of a random phenomenon of infinite exchangeable events.* Resting on the representation (see Subsection 2.2) according to which $\omega_h^{(h)}$ is the h–moment of the probability distribution function

Φ of the random phenomenon, via the Hamburger solution of the problem of moments (see, e.g., [25]), de Finetti was able to state:

Proposition 9. *In order that* $t \mapsto \Omega_n(e^{it/n})$ *may be the characteristic function of the initial* n*–segment of an infinite sequence of exchangeable events it is necessary and sufficient that all the roots of a distinguished polynomial, depending on* n*, belong to the closed interval* $[0,1]$*. The polynomial* (*in* ξ) *is*

$$Det \begin{pmatrix} 1 & \xi & \xi^2 & \dots & \xi^k \\ \omega_0^{(0)} & \omega_1^{(1)} & \omega_2^{(2)} & \dots & \omega_k^{(k)} \\ \dots & \dots & \dots & \dots & \dots \\ \omega_{k-1}^{(k-1)} & \omega_k^{(k)} & \omega_{k+1}^{(k+1)} & \dots & \omega_{2k-1}^{(2k-1)} \end{pmatrix}$$

if $n = 2k-1$*, while it is*

$$Det \begin{pmatrix} 1 & \xi & \xi^2 & \dots & \xi^k \\ \omega_1^{(1)} & \omega_2^{(2)} & \omega_3^{(3)} & \dots & \omega_{k+1}^{(k+1)} \\ \dots & \dots & \dots & \dots & \dots \\ \omega_k^{(k)} & \omega_{k+1}^{(k+1)} & \omega_{k+2}^{(k+2)} & \dots & \omega_{2k}^{(2k)} \end{pmatrix}$$

if $n = 2k$*.*

3 – Haag's contribution to exchangeability

To our knowledge, Haag was the first Author to study sequences of exchangeable events. He publicized his conclusions during the *International Mathematical Congress* held in Toronto, August, 1924. His communication appeared in Vol. 1 of the *Proceedings*, published in 1928, the very same year of the already mentioned Bologna *Congress*. See [18]. It is highly likely that de Finetti was in the dark about the Haag contribution until the 1950s, when Edwin Hewitt and Leonard J. Savage mentioned it in a famous paper about exchangeability. See [19].

It is convenient to pause here and consider what Haag really did. In the first six brief sections, he deals with finite sequences of exchangeable events and furnishes a detailed account of the expressions of the $\omega_k^{(n)}$ both in terms of $\omega_h^{(h)}$ and in terms of $\omega_0^{(h)}$, for $h = 1, 2, \dots, n$. In Section 7, Haag attains an early version of the representation theorem, but via a rather incomplete argument. He considers a sequence of exchangeable trials with a frequency of success σ_n/n converging to x as $n \to +\infty$. By the way, Haag does not hint at any form of law of large numbers, so that the reader is not able to judge whether the convergence assumption expresses an extraordinary or, instead, a common fact. By resorting to the Stirling formula, and assuming that $f(x)dx$ provides, for some continuous function f defined on $(0,1)$, an asymptotic value for the probability that σ_n/n belongs to $(x, x+dx)$, as $n \to +\infty$, he shows that

$$\sqrt{2\pi x(1-x)}\frac{(x^x(1-x)^{1-x})^n}{\sqrt{n}}f(x)$$

represents an approximate value – for great values of n – of $\omega_{\sigma_n}^{(n)}/\binom{n}{\sigma_n}$. At this stage, by means of a heuristic argument based on formal elementary computations, he concludes that

$$\binom{p+q}{p}x^p(1-x)^q\frac{1}{n}f(x) \tag{8}$$

gives an approximate value for the probability of the event "The limiting frequency belongs to $(x, x+1/n)$ and, simultaneously, one gets p successes in $n = p+q$ trials". So, the probability of p successes in $(p+q)$ trials can be represented as limit (as $n \to +\infty$) of a sum of terms like (8), i.e.

$$\binom{p+q}{p}\int_0^1 x^p(1-x)^q f(x)dx.$$

The *assumption* that the frequency converges to a random variable, weakens the validity of the Haag argument, and emphasizes the difference between his standpoint and de Finetti's stance. Indeed, de Finetti reckons that convergence of frequency must be proved, whilst it is evident that Haag is assuming the validity of some type of empirical law which postulates convergence of frequency. So, while de Finetti introduces exchangeability to explain the role of frequency in evaluating probability – within the framework of a rigorously subjectivisitc or, on depending on taste, axiomatic conception – Haag misses out on these fundamental aspects. Moreover, while de Finetti shows to have an extraordinarily advanced command of the right mathematical apparatus to deal with probabilistic problems, Haag does not go beyond the use of the elementary combinatorial calculus. In point of fact, the final part of his paper, intitled *Applications*, includes a review of classical problems solvable by means of elementary combinatorics.

4 – Some new developments on extendibility

As anticipated in Subsection 2.6, in the last part of this paper we follow de Finetti's ideas, explained in that very same subsection, to obtain new necessary and sufficient conditions for extendibility of a given finite–dimensional exchangeable distribution. These conditions are of an algebraic nature, differently from the above–mentioned conditions derived in the frame of a geometrical approach. Taking (7) as a starting point, one first determines an explicit form for F, i.e.

$$\begin{aligned}F(z) = F(z;n,k) = z^k \frac{(n+k)!}{n!} \int_{(0,1)^k} \Big(\prod_{j=1}^{k} t_j^{j-1} \Big) \{1 - z(1 - t_1 \cdots t_k)\}^n \cdot \\ \cdot\, \Omega_n \left(\frac{t_1 \cdots t_k z}{1 - z(1 - t_1 \cdots t_k)} \right) dt_1 \dots dt_k \qquad (k = 1, 2, \dots).\end{aligned} \tag{9}$$

To prove the validity of this representation, first note that (9) with $k = 1$ is consistent with (6). Then, to complete the proof, use (6), with n replaced by $(n+k)$, and proceed by mathematical induction with respect to k.

Now, substitute expression (3) into (9) to obtain

$$\begin{aligned}F(z) &= \frac{(n+k)!}{n!} \sum_{l=0}^{n} \omega_l^{(n)} z^{l+k} \int_{(0,1)^k} (t_1 \cdots t_k)^l \cdot \\ &\qquad (1 - z(1 - t_1 \cdots t_k))^{n-l} t_2 t_3^2 \dots t_k^{k-1} dt_1 \dots dt_k \\ &= \frac{(n+k)!}{n!} \sum_{l=0}^{n} \omega_l^{(n)} z^{l+k} \frac{1}{\Gamma(k)} \int_0^1 x^l (1-x)^{k-1} \{1 - z(1-x)\}^{n-l} dx \\ &\qquad \text{(see, for example, 3.3.5.11 in [23])} \\ &= \frac{(n+k)!}{n!(k-1)!} \sum_{l=0}^{n} \omega_l^{(n)} \sum_{h=0}^{n-l} \binom{n-l}{h} z^{h+l+k} (-1)^h B(l+1, k+h)\end{aligned}$$

with $B(\alpha, \beta) := \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$. Whence, from (6),

$$\begin{aligned}\Omega_{n+k}(z) = \sum_{i=0}^{k-1} (-1)^i z^i \sum_{j=1 \vee (i-1)} \binom{n+j}{i} C_j + \sum_{i=k}^{n+k} (-1)^i z^i \Big\{ \sum_{j=1 \vee (i-n)} \binom{n+j}{i} C_j \\ + \frac{(n+k)!}{n!(k-1)!} (-1)^k \sum_{l=0}^{i-k} \omega_l^{(n)} (-1)^l \binom{n-l}{n-i+k} B(l+1, i-l) \Big\}.\end{aligned}$$

Then, Proposition 8 can be restated as

Proposition 10. *$t \mapsto \Omega_n(e^{it/n}) := \sum_{h=0}^{n} \omega_h^{(n)} e^{iht/n}$ is the characteristic function of the initial n–segment of a sequence of $(n+k)$ exchangeable events if* *and only if the constants* $C_1, \ldots, C_k$ *can be determined in such a way that*

$$0 \leq \rho_i := (-1)^i \sum_{j=1\vee(i-n)}^{k} \binom{n+j}{i} C_j \qquad i = 0, \ldots, k-1$$

$$\begin{aligned} 0 \leq \rho_i := (-1)^i \Big\{ & \sum_{j=1\vee(i-n)} \binom{n+j}{i} C_j \\ & + \frac{(n+k)!}{n!(k-1)!} (-1)^k \sum_{l=0}^{i-k} \omega_l^{(n)} (-1)^l \binom{n-l}{n-i+k} B(l+1, i-l) \Big\} \\ & i = k, \ldots, n+k. \end{aligned} \tag{10}$$

Moreover, if this system of linear inequalities is consistent, then for each of the solutions $(C_1, \ldots, C_k)$*, the vector* $(\rho_0, \ldots, \rho_{n+k})$ *represents an exchangeable assessment for* $(\omega_0^{(n+k)}, \ldots, \omega_{n+k}^{(n+k)})$*, consistent with the initial segment* $(\omega_0^{(n)}, \ldots, \omega_n^{(n)})$*.*

The research of conditions for consistency of systems like (10) originated a wealth of literature on the subject. Here, we propose a solution derived from [2]. In matrix form, (10) becomes $Ax \leq b$ where

$$A = [a_{ij}]_{1 \leq i \leq n+k+1, 1 \leq j \leq k}, \quad b' = (b_1, \ldots, b_{n+k+1}), \quad x' = (C_1, \ldots, C_k),$$

with

$$a_{ij} := (-1)^i \binom{n+j}{i-1}, \qquad b_1 = 0, \ldots, b_k = 0,$$

$$b_i = (-1)^{i+k-1} \frac{(n+k)!}{n!(k-1)!} \sum_{l=0}^{i-1-k} (-1)^l \omega_l^{(n)} \binom{n-l}{n-i+k+1} B(l+1, i-1-l)$$

$$i = k+1, \ldots, n+k+1.$$

Since, as it is easy to show, the rank of A is k, Theorem 3 in [2] yields

PROPOSITION 11. $\Omega_n(e^{it/n}) := \sum_{h=0}^{n} \omega_h^{(n)} e^{iht/n}$ *is the characteristic function of the initial* n*–segment of a sequence of* $(n+k)$ *exchangeable events if, and only if, there exist* $1 \leq i_1 < i_2 < \cdots < i_k \leq n+k+1$ *such that*

$$\begin{pmatrix} a_{i_1 1} & \ldots & a_{i_1 k} \\ \ldots & \ldots & \ldots \\ a_{i_k 1} & \ldots & a_{i_k k} \end{pmatrix}$$

has a nonvanishing determinant Δ, *and*

$$\frac{1}{\Delta}\, det \begin{pmatrix} a_{i_1 1} & \dots & a_{i_1 k} & b_{i_1} \\ \dots & \dots & \dots & \dots \\ a_{i_k 1} & \dots & a_{i_k k} & b_{i_k} \\ a_{j1} & \dots & a_{jk} & b_j \end{pmatrix}$$

turns out to be nonnegative for every $j = 1, \dots, n + k + 1$.

In the particular case of $k = 1$, *this necessary and sufficient condition reduces to require that* $\{\omega_h^{(n)} : h = 0, \dots, n\}$ *satisfies*

$$\max\{\beta_i : \textit{for any even integer} \le n+1\} \le \min\{\beta_i : \textit{for any odd integer} \le n+1\}$$

where

$$\beta_i = \sum_{l=0}^{i-1} (-1)^l B(l+1, n-l+1)\omega_l^{(n)} \qquad (i = 1, \dots, n+1).$$

In fact, this result could be obtained by direct inspection of (10) with $k = 1$.

To conclude, let us remark that Proposition 11 is susceptible of interesting geometrical interpretations that one can deduce directly from the above–mentioned Cernikov paper. It would be interesting to compare them with the geometrical arguments in [15] and developed by other Authors already mentioned in Subsection 2.6.

Acknowledgements

The authors are very grateful to the organizing committee of the *Bruno de Finetti Centenary conference* (Rome, November 2006) and to referees for their time, comments and suggestions.

REFERENCES

[1] G. Castelnuovo: *Calcolo delle Probabilità*, Vol.1. 2nd ed. Zanichelli, Bologna, 1925

[2] S. N. Cernikov: *Systems of linear inequalities*, Amer. Math. Soc. Transl., **26** (1963), 11-86.

[3] L. Crisma, L.: *Alcune valutazioni quantitative interessanti la proseguibilità di processi aleatori scambiabili*, Rend. Ist. Mat. Univ. Trieste, **3** (1971), 96-124.

[4] L. Crisma: *Quantitative analysis of exchangeability in alternative processes.*, pp. 207–216 In *Exchangeability in probability and statistics*, G.Koch F.Spizzichino (eds.), North-Holland, Amsterdam, 1982.

[5] B. de Finetti: *Sulle funzioni a incremento aleatorio*, Atti Reale Accademia Nazionale dei Lincei. Rend. Serie VI, **10** (1929), 163-168.

[6] B. de Finetti: *Funzione caratteristica di un fenomeno aleatorio*, Atti Reale Accademia Nazionale dei Lincei. Mem. Serie VI., **4** (1930), 86-133.

[7] B. de Finetti: *Sui passaggi al limite nel calcolo delle probabilità*, Rendiconti Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 155-166.

[8] B. de Finetti: *A proposito dell'estensione del teorema delle probabilità totali alle classi numerabili*, Rendiconti Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 901-905.

[9] B. de Finetti: *Ancora sull'estensione alle classi numerabili del teorema delle probabilità totali*, Rendiconti del Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 1063-1069.

[10] B. de Finetti: *Sul significato soggettivo della probabilità*, Fundamenta Mathematicae **17** (1931), 298-329.

[11] B. de Finetti: *Funzione caratterisitca di un fenomeno aleatorio*, Atti del Congresso Internazionale dei Matematici, (Bologna, 3–10 Settembre, 1928), Zanichelli, Bologna, pp. 179-190.

[12] B. de Finetti: *La prévision : ses lois logiques, ses sources subjectives*, Ann. Inst. H. Poincaré, **7** (1937), 1-68.

[13] B. de Finetti: *Resoconto critico del colloquio di Ginevra intorno alla teoria delle probabilità*, Giorn. Istit. Ital. Attuari, **9** (1938), 1-40.

[14] B. de Finetti: *Compte rendu critique du colloque de Génève sur la théorie des probabilités*, In *Actualités Scientifiques et Industrielles* Vol. 766 Hermann, Paris, 1939.

[15] B. de Finetti: *Sulla proseguibilità di processi aleatori scambiabili*, Rend. Ist. Mat. Univ. Trieste, **1** (1969), 53–67.

[16] B. de Finetti: *La probabilità: guardarsi dalle contraffazioni!*, Scientia, **111** (1976), 255-303.

[17] P. Diaconis: *Finite forms of de Finetti's theorem on exchangeability*, Synthese, **36** (1977), 271-281.

[18] J. Haag: *Sur un probléme général de probabilités et ses diverses applications*, In *Proceedings of the International Congress of Mathematicians* (Toronto, August 11-26, 1924) **1** (1928), 659–674, Univ. Toronto Press.

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

[20] A. Khinchin: *Sur les classes d'événements équivalentes*, Mat.Sb. **39** (1932), 40-43.

[21] A. Khinchin: *Remarques sur les suites d'événements obéssants á la loi des grands nombres*, Mat.Sb., **39** (1932), 115-119.

[22] P. Lévy: *Calcul des Probabilités*, Gauthier–Villars, Paris, 1925.

[23] A. P. Prudnikov – Y. A. Brychkov – O. I. Marichev: *Integrals and series.*, vol. 1 Gordon & Breach Science Publishers, New York, 1926.

[24] E. Regazzini: *de Finetti's contribution to the theory of random functions*, *International Journal of Approximate Reasoning* **47** (2008), 387–393.

[25] J. A. Shohat – J. D. Tamarkin: *The Problem of Moments*, American Mathematical Society, Providence, R.I., 1943

[26] R. von Mises: *Fundamentalsaetze der wahrscheinlick keitsrechnung*, Math. Zeit., **4** (1919), 1-97.

[27] G. R. Wood: *Binomial mixtures and finite exchangeability*, Ann. Probab., **20** (1992), 1167-1173.

Lavoro pervenuto alla redazione il 20 dicemnvre 2007
ed accettato per la pubblicazione il 27 dicembre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo degli Autori:

Eugenio Regazzini – Dipartimento di Matematica – via Ferrata 1 – 27100 Pavia, Italy
E-mail: eugenio.regazzini@unipv.it

Federico Bassetti – Dipartimento di Matematica – via Ferrata 1 – 27100 Pavia, Italy
E-mail: federico.bassetti@unipv.it

Supported in part by MIUR grant 2006/134526.