UNIVERSIDAD DEL CEMA
Buenos Aires
Argentina

Serie
DOCUMENTOS DE TRABAJO

Área: Matemática

ON THE EXISTENCE OF THE TRUE VALUE OF A PROBABILITY. PART 2: THE REPRESENTATION THEOREM AND THE ERGODIC THEORY

Alberto H. Landro

Julio 2011
Nro. 458

www.cema.edu.ar/publicaciones/doc_trabajo.html
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding <jae@cema.edu.ar>

On the existence of the true value of a probability.
Part 2: The representation theorem and the ergodic theory

Alberto H. Landro[1]

Abstract: Some authors, basing their ideas on the exchangeability property, on the postulates of the representation theorem and on its interpretation in the ambit of ergodic theory, believed to find a counterexample to the subjectivist model through the theoretical justification of the existence of an objective probability. As a proof of the inconsistency of this reasoning, the representation theorem allows to assert that the convergence of the relative frequency on a "true value" of the probability is only a metaphysical illusion motivated by an asymptotic behaviour of the personal assessments of initial probabilities, leading to intersubjective assignment. With regard to the ergodic theory, its assimilation to the propensity model allows the demonstration of its metaphysical character and the resulting subjectivity in the assignment of probabilities.

1.-Introduction

According to the statement in the first part of this work, even though the classical, frequentist, logic and propensity interpretations, based on a deterministic conception, obtained objectivist definitions of the probability associated with inference structures defined by rules of interpretation, explicit or implicit, which define their identificatory model role of the true value of the probability, it should be noted that: **i)** the classical model suffers from unavoidable failures primarily related to its purely deductive character which prevents the contrast of its agreement with the observable results of the phenomena; **ii)** the frequentist model is based on the collective properties assumed as axioms that are not rigorously demonstrable; **iii)** the logical model considers probabilities as assimilable to logical relationships in an ambit consisting of abstract ideas and **iv)** the propensity model, despite Fetzer's changes to Popper and Miller's proposal, cannot avoid its metaphysical character.

On the contrary, it is further noted that the subjectivist model, which postulates the validity of probability assessments mismatched with each other whenever they meet the condition of coherence, leads to the conclusion that there is no truth about the probability, that there is no true probability but infinite versions of the same probability.

In addition, it is mentioned that some authors, basing their ideas on the exchangeability property, on the principles of the representation theorem and on its interpretation in the ambit

[1] These viewpoints are personal and do not necessarily represent the position of Universidad del CEMA.

of ergodic theory, which allows to show that the probability distribution of a stochastic process in certain conditions, can be defined in terms of averages in the time domain, raised the possibility of the existence of objective probabilities that have a physical meaning and therefore are not liable to subjective interpretation.

The claim of this second part is to demonstrate the inconsistency of this proposal from a subjectivist approach to the property of exchangeability and the representation theorem and the propensity nature (and therefore metaphysical) of the results of the ergodic decomposition theorem.

2.-The exchangeability property

The first attempts to define the exchangeability condition are related to the property of equiprobability to sequences of binomial events (that admit alternatives E and $\overline{E}$) having the same measure, n, with the same number, x, of results E.

Let a sequence of binomial events be bearing the form $\left\{E_{iq}, E_{i2}, \ldots, E_{ix}, \overline{E}_{i,x+1}, \ldots, \overline{E}_{in}\right\}$ that admits $\binom{n}{x}$ possible arrangements in which the probability of occurrence of the outcome E in the j-th. repetition of the i-th permutation, $p\left(E_{ij}\right) = p \quad \left(i = 1,2,\ldots,\binom{n}{x}; j = 1,2,\ldots,n\right)$, is unknown. It is said that the events E_i are exchangeable if the probability distribution of the variable $E^{(X,n)} \in \{0,1,2,\ldots,n\}$, representing the number of results E to be obtained in the sequence of n trials:

$$p\left(E^{(0,n)}\right) = 1 + \sum_{j=1}^{n}(-1)^{j}\binom{n}{j}p\left(E_1 \cap E_2 \cap \ldots \cap E_j\right)$$

$$p\left(E^{(x,n)}\right) = \binom{n}{x}\sum_{j=0}^{n-x}(-1)^{j}\binom{n-x}{j}p\left(E_1 \cap E_2 \cap \ldots \cap E_j\right) =$$

$$= \binom{n}{x}\sum_{j=0}^{n-x}(-1)^{j}\binom{n-x}{j}p\left(E^{(x+j,x+j)}\right) =$$

$$= (-1)^{n-x}\binom{n}{x}\Delta^{n-x}p\left(E^{(x,x)}\right) \qquad \left(x = 1,2,\ldots,n\right)$$

(where $\Delta^{n-x}(\bullet)$ denotes the $(n-x)$-th difference of the sequence)[2], is defined by a function of the arguments x and n exclusively and is therefore independent of the order in which the results E_i and $\overline{E}_i$ occur. A particular example of this type of exchangeable probabilities is the one arising from the application of Laplace's rule of succession. Suppose that in the succession $E^{(x,n)}$, according to the principle of insufficient reason, there is no reason to suppose that a value of p might be more likely than another, in other words, that the random variable p, representing the probability "a priori" of the event E occurrence, is uniformly distributed in the interval $[0,1]$. The probability "a posteriori" "*that the result of the* $(n+1)$*-th trial of succession is* E ", $E_{x+1}^{(n+1)}$, conditioned by the occurrence of $E^{(x,n)}$ according to the Bayesian conditioning scheme (which is the basis for the rule of succession), is defined the bye the following expression:

$$p\left(E_{x+1}^{(n+1)} / E_i^{(x,n)}\right) = \frac{p\left(E^{(x,n)} \cap E_{x+1}^{(n+1)}\right)}{p\left(E_i^{(x,n)}\right)} = \frac{x+1}{n+1}\frac{p\left(E^{(x+1,n+1)}\right)}{p\left(E^{(x,n)}\right)} =$$

$$= \frac{(x+1)p\left(E^{(x+1,n+1)}\right)}{(n-x+1)p\left(E^{(x,n+1)}\right)+(x+1)p\left(E^{(x+1,n+1)}\right)} =$$

$$= \frac{x+1}{(n-x+1)\dfrac{p\left(E^{(x,n+1)}\right)}{p\left(E^{(x+1,n+1)}\right)}+n+2} = f(x,n)$$

which exclusiveley depends on the values of x and n.

By using similar reasoning, other authors suggest assimilating the exchangeability to the equality of probabilities conditioned by the same number of repetitions of the event or considering as interchangeable those sequences of events for which the relative frequency $\frac{x}{n}$ of the event E is a sufficient estimator of their occurrence probability (in other words, those sequences for which $\frac{x}{n} \xrightarrow{p} p\left(E^{(x,n)}\right)$, when $n \to \infty$).

[2] Given a numeric succession $\{y_0, y_1, \ldots\}$, its k order difference assumes the form $\Delta^k y_h = (-1)^k \sum_{j=0}^{k} \binom{k}{j} (-1)^j y_{h+j} \quad (k \ge 1)$ and $\Delta^0 y_h = y_h$.

Nevertheless, all these attempts of essentially objectivist definitions fail to characterize exchangeability as a condition that goes beyondthe stochastic independence and seem to suggest that it is a formal property and, apparently, completely independent from any approach to the concept of probability[3]. It was the subjective interpretation of this property which led to a rethinking of not only the fundamentals of the probability. theory and the inference methods but also of the more general problem of induction[4].

From a subjective approach, the property of exchangeability allows to solve some controversies regarding the relationships between probability and frequency and the Bayes' theorem, which constitutes the theoretical foundation of the learning process based on experience that allows to coherently relate assignments of probabilities for different information sets, transforming statistical inference in a particular case of inductive reasoning[5].

Because of its reference to the assignation of probability for a single event, this subjective interpretation of Bayes' theorem as a special case of reasoning by induction is rejected by the objectivist conceptions which, by the postulation of the existence of objective laws having

---

[3] The arguments discussed so far make up what might be called the "prehistory" of exchangeability, which extends from its implicit consideration by the writers of the eighteenth century (the first formal application of this property to the "*coincidences problem*" in the "*game of thirteen*" is in Montmort, P.R. (1708)) to the derivation of the rule of succession proposed by Ramsey, F.P. (1931) presumably from the "*postulate of the permutation*" by his teacher W.E. Johnson (see Zabell, S.L. (1989)), including a first explicit mention by Jules Haag at the International Congress of Mathematics, Toronto, 1924. What could be considered the "history" of exchangeability in this context begins with the publications of de Finetti, B. (1928) (1930a), in which-to avoid the metaphysical assumptions stemming from the objectivist interpretation of probability which make up the basis of previous attempts- he based his reasoning on a rigorous definition of the concept of "*random phenomenon*" as one whose repetitions generate a succession of interchangeable events.

[4] de Finetti: "*Predictions cannot be but probable. No matter how grounded a prediction may seem, we can never be absolutely certain that the experience will not disprove it, (...) This is why sciences are but applications of probability calculus (...) So the most important issues associated with the meaning and value of a probability are no longer isolated issues in a particular branch of mathematics and they assume the role of fundamental gnoseological problems*".

[5] Bayes' theorem postulates that the probability of an event E occurrence, conditioned by hypothesis H is proportional to the product of the probability of E by the probability of H conditioned by E or, accordingly, that the probability of E conditioned by H is modified in the same direction and at the same magnitude as the probability of H conditioned by E (see de Finetti (1931a)(1931b)).

more or less restrictive features that govern the behavior of phenomena, transform inductive reasoning into deductive one[6]. On the other hand, even if the " objectivist anathemas " are put aside, this theorem by itself cannot justify, generally speaking, the assimilation of the probability of an event occurrence to its relative frequency (in other words, it does not solve the induction problem proposed by Hume), not due to a failure of the method but due to the essential requirements of the subjectivist interpretation which condition such assimilation to a coherent assignation of the initial probability and to the compliance with the exchangeability condition.

Let us then consider, M urns containing red and blue balls in unknown proportions $p_i = \frac{N_i^{(r)}}{N_i}$ $\left(i = 1,2,\dots, M\right)$ and $1 - p_i = \frac{N_i^{(a)}}{N_i}$, respectively. Then given an urn, M_i, if n withdrawals are made at random with replacement, the probability od obtaining j red balls and $\left(n - j\right)$ blue balls conditioned by the assumption of a partition defined by the composition of the urns (in other words, the probability that the variable relative frequencyof the result $E^{(r)}$ conditioned by such partition takes the value $\frac{j}{n}$) is given by:

$$p\left(E^{(j,n)} / M_i\right) = p\left(\theta_i^{(r)} = \frac{N_i^{(r)}}{N_i}\right) = \binom{n}{j} p_i^j \left(1 - p_i\right)^{n-j}$$

The probability of obtaining j red balls and $\left(n - j\right)$ blue balls when n withdrawals with

---

[6] de Finetti (1959): "*Speaking of inductive reasoning means attributing some value to this form of learning from experience, not considering it as the result of a peculiar psychological reaction, but as a mental process amenable to analysis, interpretation and justification. When this occurs, the tendency to overestimate the reasoning to the point of excluding any other factors may cause detrimental bias. The reason is an invaluable supplement to other intuitive faculties, but never a substitute for them (...). Since the non-tautological truths are based on more than reason, one consequence of this bias is the rise of inductive reasoning to a standard status. Thus, inductive reasoning is generally regarded as something of a lower level that generates caution and suspicion. And still worse, inductive reasoning is subject to being dignified by trying to change its nature by assimilating it to something that could almost be considered deductive reasoning. In fact, there are often attempts to explain induction without considering the term 'probability' or trying to remove it from its everyday meaning, assimilable to a measure of the degree of belief attributed to the different possible alternatives*".

repacement are made from an urn selected at random will be defined, then, by a mix of binomial distributions in which the weighings are given by the probabilities of the various urns, in the following way[7]:

$$p\left(E^{(j,n)}\right) = \sum_{i=1}^{M} \pi_i \, p\left(E^{(j,n)} / M_i\right) = \binom{n}{j} \sum_{i=1}^{M} \pi_i \, p_i^j \left(1 - p_i\right)^{n-j}$$

(where π_i $\left(i = 1,2,\ldots, M\right)$ denotes the initial probability that the hypothesis $H_i = \frac{N_i^{(r)}}{N_i}$ $\left(i = 1,2,\ldots, M\right)$ is true and p_i $\left(i = 1,2,\ldots, M\right)$ denotes the probability of drawing a red ball, conditioned by the hypothesis H_i, with a distribution function $F_M\left(x\right)$ $\left(x \in \left[0,1\right]\right)$. As it will be discussed in Section 4, in accordance with the principles of the repreentation theorem it can be assured that, given a sequence $\left\{E_1, E_2, \ldots, E_n, \ldots\right\}$ of exchangeable events, this mix of binomial probabilities defining the probability function $p\left(E^{(j,n)}\right)$ generated by this sequence, is unique.

It should be noted that in order to avoid ambiguity, de Finetti (1930c) calls "*event*" a strictly defined unique case and uses a more general notion of "*random phenomenon*", which considers an event as the result of "*each trial repeated under homogeneous boundary conditions for this phenomenon*". In other words, a *"random phenomenon"* deFinettiano is one which repetitions generate a sequence of exchangeable events. However, given that from a subjective approach events are unique in time and space, the hypothesis that the repetitions of an event are part of the same phenomenon is just a conjecture, which leads to the inevitable conclusion that the condition of exchangeability admits an interpretation of purely subjective characteristics.

Translating de Finetti's definitions to Carnap's nomenclature, we can say that the necessary and sufficient condition for the probabilities distribution corresponding to the "*state descriptions*", $\left\{E_{i1}^{(r)}, E_{i2}^{(r)}, \ldots, E_{ij}^{(r)}, E_{i,j+1}^{(a)}, \ldots, E_{in}^{(a)}\right\}$ $\left(i = 1,2,\ldots,\binom{n}{j}\right)$, to be symmetrical in the strict sense and, therefore, interchangeable, is that the are isomorphic. The disjunction of all the isomorphic state descriptions is calle a "*structural description*" by Carnap, so that the condition of exchangeability is comparable to that equiprobability of all disjunctions of a

[7] In the same way as in the case analyzed in previous pages, the expression in terms of finite differences of this equation gives the successive effect of the observed frequencies over the probabilities "a priori".

structure description[8].

From the expression given above, from a purely formal point of view and according to an objectivist approach, the condition of exchangeability is assimilated to that of stochastic independence conditioned by a constant but unknown probability (in other words, conditioned by a probability defined by a random variable). This leads to the conclusion that every sequence of independent events with constant probability is interchangeable.

de Finetti (1934) considered that this interpretation of exchangeability is incorrect to the extent that, in terms of a sujective approach, the idea of unknown probability is meaningless, that the probability varies from trial to trial according to the experience acquired by the observer and that it is the limit frequency, often improperly assimilated to an objective probability, which is unknown and with a distribution that varies according to the results of the realizations of the event[9]. Independence would exist if it were certainly known that the hypothesis H_i is true. Otherwise, since the same results can occur under various hypotheses, the probability of the extractions' results is conditioned by these assumptions and the probability of the truth of each hypothesis depends on the observed frequencies, the independence of events cannot be assumed. It should be noted that, from a subjective approach, the condition of exchangeability and its generalizations are not linked with causal independence, but with "*probabilistic independence*"[10], which means "*not merely the absence of any causal relationship but also the absence of any influence on our view on the probability assignment*"[11].

---

[8] An alternative way to avoid conceptual ambiguities of the notion of event so frequently found in the literature, may be obtained by substituiting the term "*events*" or "*experiments*" by "*individuals*" classified in k mutually exclusive and collectively exhaustive classes, according to a set of predicates $\{Q_1, Q_2, ..., Q_k\}$ .Each individual is assimilable, in this case, to random withdrawal of the domain $\{E_1, E_2, ..., E_n, ...\}$.

[9] As noted in the previous section, although the relative frequency can be considered as a particular value of the variable $E^{(j,n)}$ and, therefore, it can define a singular *"verifiable"* event, it should be considered that this interpretation ofexchangeability as comparable to independence conditioned by a random variable does not belong to the case when $n \to \infty$.

[10] It should be noted that the condition of stochastic independence (or *"probabilistic"* according to de Finetti's nomenclature) does not define a relationship among events, but it is a subjective property of the joint probabilities function.

[11] From Savage, J. (1954) and Hewitt, E., Savage, L.J. (1955) a strong international interest in the subjective interpretation of exchangeability was generated in the scientific ambit, particularly about the work of authors such as Kingman, J.F.K.

The exchangeability treated so far involves cases in which the experiment is indefinitely repeatable and is known as *"unlimited"*. de Finetti (1937) also defined a type of "*limited*" or finite exchangeability, corresponding to those sequences of events (which he called "*equivalent*") that cannot continue beyond a given limit (the typical example would be that of withdrawals without replacement).

de Finetti (1933a)(1933b)(1933c) extended the definition of exchageability to the random variables. It is said that a sequence of random variables $X = \{X_1, X_2, \ldots, X_n, \ldots\}$ is exchangeable if, for any set of n elements $\{X_1, X_2, \ldots, X_n\}$ from X, it is verified that, for each value θ belonging to a set ofhypothesis Θ over X, the X_i $(i = 1,2,\ldots,n)$ have the same conditional probability distribution:

$$p\left\{\left[\left(X_{i1} \le x_{i1}\right) \cap \left(X_{i2} \le x_{i2}\right) \cap \ldots \cap \left(X_{in} \le x_{in}\right)\right] / \theta\right\} =$$
$$= F\left[\left(x_{i1}, x_{i2}, \ldots, x_{in}\right) / \theta\right] = F\left(x_1 / \theta\right) F\left(x_2 / \theta\right) \ldots F\left(x_n / \theta\right)$$

From this expression one can conclude that the variables compound an infinite exchangeable sequence are positively correlated[12] .

Note that, similar to the previous case, the concept of exchangeability of random variables can also be interpreted as equivalent to the concept of independence conditioned by a partition of hypotheses and in addition, it is possible to distinguish between limited and unlimited exchangeability.

Obviously, if the variables $X_1, X_2, \ldots, X_n$ are independent andidentically distributed, in other words, such that $F_i(x) = F(x)$ $(i = 1,2,\ldots,n)$, then they are exchangeable. As it will be discussed in Section 4, generally speaking, de Finetti's theorem of representation shows that for each property inherent in the iid sequences, there is a similar property for exchangeable sequences[13]. Conversely, given a finite exchangeable sequence, it is possible to establish some

---

(1978), Koch, G.; Spizzichino, F. (eds.)(1982), Daboni, L.; Wedlin, A. (1982), Scozzafava, R. (1982)(1990), Aldous, D.J. (1985), Viertl, R. (ed.)(1987), Diaconis, P. (1987), Regazzini, E. (1988)(1991), Petris, G.; Secchi, P. (1989), Cifarelli, D.M.; Muliere, P. (1989), Piccinato, L. (1991), Gilio, A. (1992), Coletti, G.; Dubois, D.; Scozzafava, R. (eds.)(1994), Berti, P.; Rigo, P. (1994), Petris, G.; Regazzini, E. (1994), Bernardo, J.M.;Smith, A.F.M. (1994).

[12] See Section 4.

[13] See Chatterji, S.D. (1974a)(1974b), Kingman, J.F.C. (1978).

inequalities that link it to a sequence of iid variables. Let, for example, a set $X=\{X_1,X_2\}$ of exchangeable variables and two increasing functions $g_1(X)$ and $g_2(X)$, then it will be verified that $\left[g_1(x_1)-g_1(x_2)\right]\left[g_2(x_1)-g_2(x_2)\right]\geq 0$ (for all x_1 and x_2) and, therefore, that $E\left\{\left[g_1(X_1)-g_1(X_2)\right]\left[g_2(X_1)-g_2(X_2)\right]\right\}\geq 0$. Then, if the expected values exist, it can be concluded that the condition of exchangeability implies that $E\left[g_1(X_1)g_2(X_1)\right]\geq E\left[g_1(X_1)g_2(X_2)\right]$. Note that, if verified that $X_1 \perp X_2$, the above relationship would be reduced to Chebyshev's inequality, $E\left[g_1(X_1)g_2(X_2)\right]\geq E\left[g_1(X_1)\right]E\left[g_2(X_2)\right]$.

On the other hand, given a sequence of random variables $X=\{X_1,X_2,\ldots\}$ so that $E\left(\left|X_j\right|\right)<\infty$, it can be said that there is a non-random sequence of integers, $1\leq j_1<j_2<\ldots$, so that the limit of the sub-sequence $X_k^*=\{X_1,X_2,\ldots,X_k\}=\{x_{j_1},x_{j_2},\ldots,x_{j_k}\}$, $\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}X_j^*$, exists almost with certainty. Now then, it is shown that any finite sequence $X_n=\{X_1,X_2,\ldots,X_n\}$ dense (That is, such that $\lim_{\lambda\to\infty}\sup_n p\left(\left|X_n\right|\geq\lambda\right)=0$) contains a sub-sequence X_k^* that is asymptotically exchangeable, in other words, such that, when $k\to\infty$, the sequence $\{X_k^*,X_{k+1}^*,X_{k+2}^*,\ldots\}$ converges in distribution to an exchangeable sequence.

3 .- The convergence to intersubjectivity

As discussed in the previous section, from the application of the rule of succession it is obtained that, in the case of exchangeable events, the probability that the outcome of the $(n+1)$-th. extraction would be E, $E_{x+1}^{(n+1)}$, conditioned by the occurrence of $E^{(x,n)}$ is defined by a function of the arguments x and n exclusively, in the following way:

$$p\left(E_{x+1}^{(n+1)}/E^{(x,n)}\right)=\frac{x+1}{(n-x+1)\dfrac{p\left(E^{(x,n+1)}\right)}{p\left(E^{(x+1,n+1)}\right)}+x+2}=f(x,n)$$

Then, when $n\to\infty$, the convergence shall be verified towards the equiprobability of the events "*to obtain* x *results* E *in* n *trials*" and "*to obtain* $x+1$ *results* E *in* $n+1$ *trials*", in other words, whatever the coherent value assigned to initial probabilities, it will be verified that

$\frac{p\left(E^{(x+1,n+1)}\right)}{p\left(E^{(x,n)}\right)} \to 1$ and, therefore, that $p\left(E_{x+1}^{(n+1)} / E^{(x,n)}\right) \to \frac{x+1}{n+2} \approx \frac{x}{n}$ (it should be noted that, in de Finetti's nomenclature, the convergence implies a "*practical certainty*", not an "*absolute*" or logic "*certainty*"[14].

This result leads to two conceptually opposing postulations. From an objetivist point of view, it can be interpreted as a confirmation that, for sequences of exchangeable events, the true value of probability $p\left(E_{x+1}^{(n+1)} / E^{(x,n)}\right)$ exists and is asymptotically identifiable by the relative frequency of the outcome E. As discussed in the next section, from a subjectivist approach, it means recognozing that this postulated "true value" of the probability is not but a metaphysical illusion, that the fact that convergence to relative frequency is verified for any coherent value assigned to the probability $p\left(E_{x+1}^{(n+1)}\right)$ is due to the successive transformation of the assignment of initial probabilities because of the increase of empirical information, through the "*Bayesian conditioning*", which creates an intersubjective assignment easily confused with an objective probability.

With respect to this convergence to intersubjectivity, it should be noted that if the general assumptions adopted by observers about the nature of the event considered are correct (in a broad sense), then de Finetti's modification scheme of the probability $p\left(E_{x+1}^{(n+1)}\right)$ using a Bayesian conditioning process is an aceptable justification, but, if the evaluation of the boundary conditions that define the characteristics of the event is wrong, then all subjective probabilities conditioned by evidence will be inappropiate. For the final probabilities to be reasonable, it would the be necessary to have more drastic changes in $p\left(E_{x+1}^{(n+1)}\right)$ than the ones allowed by the very conservative scheme of Bayesian conditioning[15], a circunstance that would

[14] In fact, this convergence should be interpreted as an in-probability convergence. For convergence in the sense of the analysis, exchangeability is not a sufficient condition. According to a particular case of general theorem proposed by Gaifman, H. (1971), the necessary condition for such convergence is that the proportion of red balls will be distributed in accordance with a function that assigns a zero probability to the hypothesis that this share is in the range $(0,1)$ or assign a non-null value to the probability that the hypothesis falls within a range (a,b) (for $0 \leq a < b \leq 1$).

[15] A representative example of the transformation of the initial probabilities using Bayesian conditioning scheme is given by the aforementioned rule of succession,

be in flagrant contradiction with de Finettiano's principle of "*reduction to exchangeability*"[16].

It should further be noted that, if the condition of exchangeability is not verified, this transformation scheme of initial probabilities (subjective), it would generate sequences of probabilities totally unrelated to reality. This restriction could be solved by considering a broader set of hypotheses that could include, for example, speculations about chaotic behavior[17], but this alternative would render meaningless the process of Bayesian conditioning.

4.-The representation theorem

From the definition of unlimited exchangeability, the representation theorem allows to solve the essential problem of the theory of probability, which consists of interpreting its relationships with inductive reasoning through the aforementioned Bayesian scheme of learning from experience. This limit theorem shows how a probability law for exchangeable events formally approaches a law for independent events in which the probability "a priori"

---

used to address D. Hume's problem of induction about the assessment of the probability that the sun appears tomorrow once again (known in the literature as the "*problem of dawn*"). Considering the records of the last 5000 years, it is found that the sun rose every morning for 1,826,250 days. According to the rule of succession being $x = n = 1.,826,250$, the probability of dawn breaking tomorrow will be approximately equal to 0.9999994. Now, suppose that tomorrow the sun will not rise, it will be verified that $x = n - 1$ and that $n = 1,.826,251$ and, therefore, in accordance with the conservative "spirit" of Bayesian conditioning, the probability that the sun appears the next day will be approximately equal to 0.9999989. This decrease of 0.00005% in the assessment of the probability of a new dawn after entering information that wills certainly generate such a state of confusion that will lead to assume that the sun will never appear again determined by the rule of succession, looks, at least, inappropiate.

[16] This confirms the pragmatism-relativism and the non-axiomatic nature attribuited to exchangeability by de Finetti (see Section 6).

[17]. Consider, for example, the sequences originated in the "game of red or blue" proposed by Feller, W. (1950) or from the chaotic clock by Albert, M. (1992)(1999). Popper, K.R. (1957) used the game of red or blue to criticize what he called the "simple induction rule" and, consequently, de Finetti's reduction to exchangeability and the 1983 issue of the same work, to prove (it could be said, in vain) the impossibility of the existence of an inductive logic (see Gillies, DA (1996)).

of the occurrence of the result E is given by its relative frequency[18].

Let consider a scheme that consist in the randomly selection of a number $X \in [0,1]$, that represents the proportion of red balls in the $M \to \infty$ urns, according to a unique distribution function $F_X(x) = \lim_{M\to\infty} F_M(x)$ $(0 \le x \le 1)$, then the probability $p\left(E^{(j,n)}\right) = \sum_{i=1}^{M} \pi_i \, p\left(E^{(j,n)} / M_i\right)$ is transformed in:

$$p\left(E^{(j,n)}\right) = \binom{n}{j} \int_0^1 x^i (1-x)^{n-i} f_X(x) dx = \int_0^1 p\left(E^{(x,n)} / x\right) dF_X(x) = E\left[p\left(E^{(xj,n)} / x\right)\right]$$

(where $p\left(E^{(j,n)} / x\right)$ represents the "*credibility*"[19] that the observer attributes to the conditioning hypothesis x and $F_X(x)$ is a derivable function)[20]. This representation theorem statement (proposed by de Finetti (1937a))[21] shows that, if the probabilities "a priori" are distributed according to a unique distribution function $F_X(x)$ then the events $E^{(j,n)}$ are interchangeable and that the reverse implication is also true: given a probability distribution $p\left(E^{(j,n)}\right)$, exchangeable for every n positive integer, there is only one such distribution function $F_X(x)$ satisfying the representation formula, in other words, there is just one mix of binomial probabilities that defines it[22].

---

[18] Note that although the notion of exchangeability and the limit theorems consequently shown describe the process of learning from experience, the limits of the distributions that can be said to represent this Bayesian conditioning are not strictly defined.

[19] Link, G. (1980).

[20] Note that the above realtionship holds for all x with the same function $F_X(x)$.

[21] In addition to the first definition of exchangeable events, in 1928 Haag, J. had already proposed an incomplete demonstration of this theorem.

[22] Curiously, this theorem was not fully appreciated until late 1940's. Fréchet, M. (1943) makes a casual reference to it as "*a formula obtained by Khinchin and de Finetti*". In this regard it is worth mentioning Kyburg, H.E., Smokler, H.E. (eds.)(1964): "*In a sense the mos important concept in the subjective theory of probability is the one of 'exchangeable events'. Until the introduction of this concept (by de Finetti (931)) the subjective theory of probability was not but a philosophical*

Generalizing the above result, the probability that in the following m trials h results E y $n-h$ results $\overline{E}$ occur, if in the first n trials j results E and $n-j$ results $\overline{E}$ occurred, it will be as follows:

$$p\left(E_{x+h}^{(n+m)} / E^{(x,n)}\right) = \frac{\binom{m}{j}}{\binom{n+m}{x+h}} \frac{p\left(E_{x+h}^{(n+m)}\right)}{p\left(E_x^{(n)}\right)} =$$

$$= \frac{\int_0^1 x^h (1-x)^{m-h} x^j (1-x)^{n-j} dF_X(x)}{\int_0^1 x^j (1-x)^{n-j} dF_X(x)} = \int_0^1 x^h (1-x)^{m-h} dF_X^*(x)$$

this proves that, given the information on the outcome of the first n trials, the following m trials are exchangeable, but such that, for large enough m values, the probability $dF_X(x)$ that the relative frequency assumes a value in the range of the interval $[x, x+dx]$ is modified in a directly proportional way to $x^j(1-x)^{n-j}$, becoming $dF_X^*(x)$. So the probability on the peak $x^* = \frac{j}{n}$ grows when n increases indefinitely, which asymptotically leads to the assumption of independence and equiprobability of the trials with equal probability to the observed frequency. As discussed in the previous section, from a subjetivist interpretation, this result justifies the reasons that lead to associate -inteasubjectively- the probability of certain events to the frequency observed in similar events, replacing the unknown objective probabilities and the concept of independence by subjective probabilities and exchangeability[23].

---

curiosity. None of those for whom the probability theory was a subject of knowledge or application paid much attention to it. It shoul be noted that the introduction of the 'equivalence' or 'symmetry' or exchangeability' concept, as known at present, provided the link between the notion of subjective probability and the classical problem of inductive inference".

[23] Note that de Finetti's representation theorem demonstration is based on the application of the definition of characteristic function to exchangeable probabilities. According to the original notation, de Finetti obtained that $p\left(E^{(x,n)}\right) = \binom{n}{x} \int \theta^x (1-\theta)^{n-x} d\Phi = \binom{n}{x} \int (x-n\theta)\theta^{x-1}(1-\theta)^{x-h-1}\Phi(\theta)d\theta$. Thus, once the characteristic function is defined and its limit is calculated, the probabilities $p\left(E^{(x,n)}\right)$ remain uniquely determined (the basics of this demonstration method can be

Taking into account that the definition analyzed in Section 3:

$$p\left(E_{x+1}^{(n+1)} / E^{(x,n)}\right) = \frac{p\left(E^{(x+1,n+1)}\right)}{p\left(E^{(x,n)}\right)} = \frac{\int_0^1 x x^j (1-x)^{n-j}\, dF_X(x)}{\int_0^1 x^j (1-x)^{n-j}\, dF_X(x)} = \int_0^1 x\, dF_X^*(x)$$

(where $F_X^*(x) = \alpha \int_0^x y^j (1-y)^{n-j}\, dF_Y(y)$ and $\alpha = \dfrac{1}{\int_0^1 x^j (1-x)^{n-j}\, dF_X(x)}$) leads to the postulation of a biunivocal correspondence between the probabilities $p\left(E_{x+1}^{(n+1)} / E^{(x,n)}\right)$ and $p\left(E^{(x,n)}\right)$, it can be concluded that the probability $p\left(E_{x+1}^{(n+1)} / E^{(x,n)}\right)$ univocally determines the function $f_X(x)$[24].

On the other hand, every exchangeable sequence $X = \{X_1, X_2, \ldots, X_k\}$ has the selection property, whereby, given a non-random sequence of integers of the forma $1 \leq j_1 < j_2 < \ldots < j_k$, it is verified that $F_{X_{j_1}, X_{j_2}, \ldots, X_{j_k}}\left(x_{j_1}, x_{j_2}, \ldots, x_{j_k}\right) = F_{X_1, X_2, \ldots, X_k}\left(x_1, x_2, \ldots, x_k\right)$. However, since for a finite sequence of variables, $\{X_1, X_2, \ldots, X_k\}$, the selection property only provides information on $k-1$ marginal variables, the inverse implication is not necessarily verified. Only if the sequence of variables is infinite, the implication is verified in both ways[25]. This shows that the selection property (apparently not as strict as the exchangeability property) is a sufficient condition to ensure the principles of the representation theorem[26].

---

found in Poincaré, H. (1896), Borel, E. (1909) y Castelnuovo, G. (1925-28)).

[24] From these predictive probabilities de Finetti (1952) proposed the exchangeability "*degenerate case*" as the one in which such probabilities assume the values 0 or 1 (see Daboni. L. (1953)).

[25] See Dacunha-Castelle, D. (1974).

[26] With respect to the conceptual aspects of this theorem, it would be mentioned the contributions of Braithwaite, R.B. (1957), Good, I.J. (1965), Jeffrey, R.C. (1965), Hintikka, J. (1970), Spielman, S. (1976)(1977), Skyrms, B. (1980), Suppes, P.; Zanotti, M. (1980), von Plato, J. (1981), Suppes, P. (1984), Dawid, A.P. (1985), Zabell, S.L. (1988), Howson, C.; Urbach, P. (1989), Mura, A. (1989)(1992), Gorlino, P. (1990), Sahlin, E. (1993), Weschler, S. (1993).

From the previous results we can conclude that the representation theorem is only verified for infinite exchangeable sequences. Note, for example, that infinite exchangeable sequences negative correlations between the variables may appear that, therefore, cannot be conditionally independent or identically distributed. Let a sequence $X = \{X_1, X_2, \ldots, X_n\}$ be of exchangeable variables such that $E(X_i) < \infty$ $(i = 1,2,\ldots,n)$, $\sigma^2(X_i) = \sigma^2$ and $\rho(X_i, X_j) = \rho$ $(i \neq j, i, j = 1,2,\ldots,n)$, then $\sigma^2\left(\sum_{j=1}^{n} X_j\right) = \sum_{j=1}^{n}\sigma^2 + \sum_{i \neq j}\rho\sigma^2\left[1 + (n-1)\rho\right] > 0$ will be verified. So $\rho \geq -\frac{1}{n-1}$ (the inequality is verified if, for example, the variables X_j are distributed according to a symmetric multinomial function).

This necessary condition of infinity for the succession of exchangeable variables, paradoxically contradicts de Finetti's considerations (1931a)(1931b)(1970) with respect to the axiom of numerable additivity[27] about the fact that, although it is reasonable to accept the representativity of a judgement on the behavior of an observable fact, from a positivist position it is unrealistic to assume that it is possible to judge something that has no empirical meaning, such as the infinite domain events[28]. Moreover, this leads to the conclusion that, according to a subjetive interpretation, the representatation theorem does not allow an explanation in terms of parametric estimation and, therefore, within the scope of inference, de Finetti's result must be assumed according to his weakest formulation (objectivist), whereby the necessary and sufficient condition for the events $E^{(j,n)}$ to be exchangeable is that, conditioned by a random element x, the joint probability distribution for any finite sequence should be the same.

5- Partial exchangeability

As discussed in the previous section, the condition of exchangeability should be considered as a limiting case in which the analogy between the events is "*absolute*" or "*complete*". However, given a sequence of events $\{E_1, E_2, \ldots, E_n, \ldots\}$, divided into k classes, in the

---

[27] See, Landro, A.H. (2010a).

[28] de Finetti (1970): "*The assumption of countable additivity (...) is now commonly accepted; even though it was not originated with Kolmogorov's axioms (1933), it was systematized in these axioms. Its success is largely due to the mathematical convenience of turning the calculation of probabilities into a simple transfer of modern measurement theory (...) No one has provided any real justification of countable additivity (one which goes beyond its consideration simply as a 'natural extension' of finite additivity)*".

applications there are usually cases in which, while events can be regarded as "*of the same kind*", so that within each class they can be considered exchangeable from the point of view of probability, they present some observable differences among different classes. It is said that this sequence is partially exchangeable of order k if there exists a unique probability, $p\left(E^{(n_1,n_2,\ldots,n_k)}\right)$, that in n trials $n_1,n_2,\ldots,n_k$ results E occur (where the values $n_1,n_2,\ldots,n_k$ denote the number of trials made on each of the classes, such that $n_1+n_2+\ldots+n_k=n$)[29]. Then, the probability that in $n_1+n_2+\ldots+n_k$ events $j_1,j_2,\ldots,j_k$ results E and $h_1=n_1-j_1, h_2=n_2-j_2,\ldots,h_k=n_k-j_k$ results $\overline{E}$ occur, will be:

$$p\left(E^{(j_1,\ldots,j_k,n_1,\ldots,n_k)}\right)=(-1)^{h_1+h_2+h_k}\binom{n_1}{j_1}\ldots\binom{n_k}{j_k}\Delta_1^{h_1}\Delta_2^{h_2}\ldots\Delta_k^{h_k}\,p\left(E^{(j_1,\ldots,j_k,j_1,\ldots,j_k)}\right)$$

where:

$$\Delta_i^{h_i}\,p\left(E^{(j_1,\ldots,j_k,j_1,\ldots,j_k)}\right)=p\left(E^{(j_1,\ldots,j_{i-1},j_i+1,j_{i+1},\ldots,j_k,j_1,\ldots,j_{i-1},j_i+1,j_{i+1},\ldots,j_k)}\right)-$$
$$-\,p\left(E^{(j_1,\ldots,j_k,j_1,\ldots,j_k)}\right)\ (i=1,2,\ldots,k)$$

and,

$$\Delta_1^{h_1}\Delta_2^{h_2}\ldots\Delta_k^{h_k}\,p\left(E^{(j_1,\ldots,j_k,j_1,\ldots,j_k)}\right)=$$
$$=\sum_{i_1=0}^{h_1}\sum_{i_2=0}^{h_2}\cdots\sum_{i_k=0}^{h_k}\binom{h_1}{i_1}\binom{h_2}{i_2}\ldots\binom{h_k}{i_k}p\left(E^{(j_1+i_1,\ldots,j_k+i_k,j_1+i_1,\ldots,j_k+i_k)}\right)$$

For a given vector $[n_1\quad n_2\quad\ldots\quad n_k]$, the $(n_1+1)(n_2+1)\ldots(n_k+1)$ values $p\left(E^{(j_1,\ldots,j_k,n_1,\ldots,n_k)}\right)$ $(0\le j_i\le n_i)$ define the probabilities of the $(n_1+1)(n_2+1)\ldots(n_k+1)$ possible combinations of the k relative frequencies. Let $F_{n_1,n_2,\ldots,n_k}(x_1,x_2,\ldots,x_k)$ $(0<x_i<1)$ the distribution function of the k variables that represent the probability that $\frac{j_i}{n_i}\le x_i$ $(i=1,2,\ldots,k)$. In the same way

[29] Note that the sequence of probabilities $\left\{p\left(E^{(n,n)}\right)\right\}$ defined in Section 2 is replaced, in this case, by a k-dimensional sequence having the form $\left\{p\left(E^{(n_1,n_2,\ldots,n_k,n_1,n_2,\ldots,n_k)}\right)\right\}$.

as in version of the representation theorem analyzed in the previous section, when $(n_1, n_2, \ldots, n_k) \to \infty$, it is verified that $F_{n_1, n_2, \ldots, n_k}(x_1, x_2, \ldots, x_k) \to F_{X_1, \ldots, X_k}(x_1, x_2, \ldots, x_k)$. So it will be:

$$p\left(E^{(j_1, \ldots, j_k, n_1, \ldots, n_k)}\right) =$$

$$= \binom{n_1}{j_1} \ldots \binom{n_k}{j_k} \int_0^1 \ldots \int_0^1 x_1^{j_1} (1 - x_1)^{n_1 - j_1} \ldots x_k^{j_k} (1 - x_k)^{n_k - j_k} dF_{X_1, \ldots, X_k}(x_1, \ldots, x_k) =$$

$$= \int_0^1 \ldots \int_0^1 x_1^{j_1} (1 - x_1)^{n_1 - j_1} \ldots x_k^{j_k} (1 - x_k)^{n_k - j_k} f_{X_1, \ldots, X_k}(x_1, \ldots, x_k) dx_1 \ldots dx_k$$

(assuming that, in this last expression, the function $F_X(x)$ admits a density function such that $f_{X_1, \ldots, X_k}(x_1, \ldots, x_k) = \dfrac{\partial^k F_{X_1, \ldots, X_k}(x_1, \ldots, x_k)}{\partial x_1 \partial x_2 \ldots \partial x_k}$). In particular, if the events are exchangeable (for example, if classes are indiscernible from the point of view of the probability), the numbers $p\left(E^{(n_1, \ldots, n_k)}\right)$ depend exclusively on $n = n_1 + n_2 + \ldots + n_k$, the operators $\Delta_1, \Delta_2, \ldots, \Delta_k$ are all equal, $\dfrac{1}{\binom{n_1}{j_1}\binom{n_2}{j_2} \ldots \binom{n_k}{j_k}} p\left(E^{(j_1, \ldots, j_k, n_1, \ldots, n_k)}\right)$ is a function of $n = n_1 + n_2 + \ldots + n_k$ and $j = j_1 + j_2 + \ldots + j_k$ exclusively and the function $F_{X_1, \ldots, X_k}(x_1, x_2, \ldots, x_k)$ is concentred on the line $x_1 = x_2 = \ldots = x_k$. On the contrary, if the classes are independent, then $p\left(E^{(n_1, n_2, \ldots, n_k)}\right) = p_1\left(E^{(n_1)}\right) p_2\left(E^{(n_2)}\right) \ldots p_k\left(E^{(n_k)}\right)$, the differences $\Delta_1, \Delta_2, \ldots, \Delta_k$ operate individually on the probabilities $p_1\left(E^{(n_1)}\right), p_2\left(E^{(n_2)}\right), \ldots, p_k\left(E^{(n_k)}\right)$ and, consequently, it is obtained that $p\left(E^{(j_1, \ldots, j_k, n_1, \ldots, n_k)}\right) = p_1\left(E^{(j_1, n_1)}\right) p_2\left(E^{(j_2, n_2)}\right) \ldots p_k\left(E^{(j_k, n_k)}\right)$. That is, it is verified that $F_{n_1, \ldots, n_k}(x_1, \ldots, x_k) = F_{n_1}(x_1) F_{n_2}(x_2) \ldots F_{n_k}(x_k)$ and, in the limit, that $F_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k) = F_{X_1}(x_1) F_{X_2}(x_2) \ldots F_{X_k}(x_k)$.

Suppose we have the information that, in $n = n_1 + n_2 + \ldots + n_k$ trials in the k classes, $j = j_1 + j_2 + \ldots + j_k$ results E and $h = h_1 + h_2 + \ldots + h_k$ results $\overline{E}$ occurred and let the events be:

i) $A_i = \left\{E_{i,1}^{(\ell)}, \ldots, E_{i,j_i}^{(\ell)}, \overline{E}_{i,1}^{(\ell)}, \ldots, \overline{E}_{i,k_i}^{(\ell)}\right\}$ $\left(i = 1,2,\ldots,k\,; \ell = 1,2,\ldots,\binom{n_i}{j_i}\right)$, which represent the sequence of results that occurred in the ℓ-th permutation of trials made inclass i $(= 1,2,\ldots,k)$

and **ii)** $A = A_1 A_2 \ldots A_k$. The probability that in class i a result E_i occurs in the $(n+\ell)$-th trial $(\ell \geq 1)$, conditioned by the occurrence of A, will be:

$$p\left(E_i^{(n+\ell)} / A\right) = \frac{p\left(A \cap E_i^{(n+\ell)}\right)}{p(A)} = \frac{\dfrac{p\left(E^{(j_1,\ldots,j_{i-1},j_i+1,j_{i+1},\ldots,j_k,n_1,\ldots,n_{i-1},n_i+1,n_{i+1},\ldots,n_k)}\right)}{\binom{n_1}{j_1}\cdots\binom{n_{i-1}}{j_{i-1}}\binom{n_i+1}{j_i+1}\binom{n_{i+1}}{j_{i+1}}\cdots\binom{n_k}{j_k}}}{\dfrac{p\left(E^{(j_1,\ldots,j_k,n_1,\ldots,n_k)}\right)}{\binom{n_1}{j_1}\cdots\binom{n_k}{j_k}}} =$$

$$= \frac{\int_0^1 \cdots \int_0^1 x_1^{j_1}(1-x_1)^{h_1} \ldots x_{i-1}^{j_{i-1}}(1-x_{i-1})^{h_{i-1}}\, x_i x_i^{j_i}(1-x_i)^{h_i} \ldots x_k^{j_k}(1-x_k)^{h_k}\, dF_{X_1,\ldots X_k}(x_1,\ldots,x_k)}{\int_0^1 \cdots \int_0^1 x_1^{j_1}(1-x_1)^{h_1} \ldots x_k^{j_k}(1-x_k)^{h_k}\, F_{X_1,\ldots,X_k}(x_1,\ldots,x_k)} =$$

$$= \int_0^1 \cdots \int_0^1 x_i\, dF^*_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$$

where:

$$F^*_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \alpha \int_0^{x_1}\int_0^{x_2} \cdots \int_0^{x_k} y_1^{j_1}(1-y_1)^{h_1} \ldots y_k^{j_k}(1-y_k)^{h_k}\, dF_{X_1,\ldots,X_k}(x_1,\ldots,x_k) =$$

$$= \alpha \int_0^{x_1}\int_0^{x_2} \cdots \int_0^{x_k} y_1^{j_1}(1-y_1)^{h_1} \ldots y_k^{j_k}(1-y_k)^{h_k}\, f_{X_1,\ldots,X_k}(x_1,\ldots,x_k) dx_1 \ldots dx_k =$$

$$= \int_0^1 dx_1 \ldots \int_0^1 dx_{i-1} \int_0^1 x_i dx_i \int_0^1 dx_{i+1} \ldots \int_0^1 dx_k\, f^*_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$$

(assuming that $F_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$ it admits a density function $f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)$)[30], and

[30] The interpretation of the exchangeability condition as one that allows the translation of the vague notion of "*analog random items*" to a probabilistic language is based on a subjective concept and the solution to the problem of induction through the asymptotic theorem for predictive distributions is obviously subjective. "*... A perfectly logical solution in itself to the extent that when one seeks to eliminate all subjective factors the only achievement, with varying degrees of success, is to hide them, but never avoiding a gap in logic. It is true that in many cases in which the*

$$f^*_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \alpha\, x_1^{j_1}(1-x_1)^{h_1}\ldots x_k^{j_k}(1-x_k)^{h_k}\, f_{X_1,\ldots,X_k}(x_1,\ldots,x_k)\,.$$

Likewise, the probability that in $m_1 + m_2 + \ldots + m_k$ additional trials $j_1 + j_2 + \ldots + j_k$ results E and $h_1 + h_2 + \ldots + h_k$ results $\overline{E}$ occur, conditioned by A, will be:

$$p\left(E^{(n_1+m_1, n_2+m_2, \ldots, n_k+m_k)} / A\right) = \frac{p\left(A \cap E^{(n_1+m_1, \ldots, n_k+m_k)}\right)}{p(A)} =$$

$$= \int_0^1 \ldots \int_0^1 x_1^{j_1}(1-x_1)^{h_1}\ldots x_k^{j_k}(1-x_k)^{h_k}\, dF^*_{X_1,\ldots,X_k}(x_1,\ldots,x_k) =$$

$$= \int_0^1 x_1^{j_1}(1-x_1)^{h_1}\, dx_1 \int_0^1 x_2^{j_2}(1-x_2)^{h_2}\, dx_2 \ldots \int_0^1 x_k^{j_k}(1-x_k)^{h_k}\, dx_k\, f^*_{X_1,\ldots X_k}(x_1,\ldots,x_k)$$

As in the case of exchangeability it can be concluded that, conditioned by the n trials result, the remaining trials are partially exchangeable, but its limit distribution, $F_X(x)$, is modified in direct proportion to $\alpha(x_1, x_2, \ldots, x_k)$ and the probabilities in the environment of the values of the observed frequencies, $\left(x_1^* = \frac{j_1}{n_1}, x_2^* = \frac{j_2}{n_2}, \ldots, x_k^* = \frac{j_k}{n_k}\right)$, grow as $(n_1, n_2, \ldots, n_k) \to \infty$, leading to an intersubjective situation that considers the independence of the trials and the assimilation of probabilities to the observed frequencies in each class[31].

6 .- The principle of reduction to exchangeability

If in the representation theorem it is assumed, in particular, that the variable x is uniformly distributed in the interval $[0,1]$, it will be verified that $p\left(E^{(x,n)}\right) = \binom{n}{x} B(x+1, n-x+1) = \frac{1}{n+1}$

experience is important- such as in the exchangeability hypothesis-these subjective factors never have a pronounced influence. Whilst this is very important to explain how, under certain conditions, a more or less close agreement is produced between predictions that different individuals produce, it also shows that differing views are always legitimate" (de Finetti (1937a)).

[31] Note that, as in the case of exchangeability, this convergence to intersubjectivity depends basically on incomplete knowledge of a qualitative nature of the function $F_X(x)$.

$(x = 0,1,2,\ldots,n)$ (where $B(\bullet,\bullet)$ denotes a Beta function). This corollary demonstrates the formal linkage between the concepts of exchangeability and independence. As it has already been mentioned, the fact that the relationship is purely formal implies that the equations formulated from an objectivist viewpoint from the condition of stochastic independence can be interpreted from a subjectivist point of view from subjective probabilities and exchangeability. This reinterpretation led de Finetti to formulate the principle of *"reduction to exchangeability"* in order to eliminate the *"nebulous" "metaphysical"* concepts of objective probability and independence (essential in the formulation of the law of large numbers and of the limit central theorem) precisely in favour of the subjective probability and exchangeability.

From a diametrically opposite conceptual position, the objectivist authors propose what might be called a principle of reduction to independence, which considers that the exchangeability assumption would only be applicable in situations of objective independence, but this application would be redundant since the problem could be addressed simply by using the concepts of independence and objective probability. On the other hand, if the situation were of non-independence, the use of the exchangeability condition would lead to erroneous results and, therefore, should be avoided.

de Finetti replied to these objections on the grounds that, considering that exchangeable events are formally equivalent to equally probable events of objective independence, different subjective forms equivalent of non-independent events can be introduced (and, in particular, of Markov's chains, giving rise to events that could be called exchangeable-Markovian)[32]. So, instead of assuming the classical exchangeability hypothesis, different forms of Markovian exchangeability can be considered assigning each one an initial (subjective) probability.

It is obvious that (contrary to de Finetti's views) this generalization leads to such great complications that make it inapplicable. It should be noted that, in accordance with this extension, it would be necessary to consider all possible types of dependence that could occur in a sequence of events and assign each of them a probability "a priori". Obviously, it would be impossible to ensure that "all" possible forms of dependence had been taken into account in each case, not to mention assuming that it is possible to assign individual probabilities "a priori" to each of them.

This restriction led to the erroneous conclusion that, as from an objectivist interpretation, the analysis of a sequence of events just requires the consideration of only one initial hypothesis: independent events with constant probability with no need to consider "a priori" any alternative hypothesis of dependency or variable probabilities; the only way to avoid

[32] For a detailed treatment of Markovian processes, see Landro, A.H.; González, M.L. (2009).

complications associated with the widespread of reduction to exchangeability is from an objective approach.

In response to these considerations, it should be noted that since statistical tests allow only a methodological falsation, it is impossible to strictly verify the accuracy of the initial hypothesis of "*proven objective independence*", which shows its subjective nature and, therefore, invalidates the alleged redundancy of the exchangeability condition.

7 .- The ergodic theory

7.1 .- The decomposition theorem

As discussed, according to the principle of reduction to exchangeability, de Finetti considers interchangeable probabilities as subjective ones and binomial probabilities as objective probabilities that should be eliminated. The question to be analyzed in this section is whether this principle applies to more general classes of events in which the principle of insufficient reason is not allowed as an assumption.

Given that from an objectivist interpretation, the concept of ergodic probability means the greatest generalization of the stochastic independence property, we can consider the representation of stationary probabilities using unique mixtures of ergodic probabilities as the most general expression of the principle of exchangeability reduction. However, the uniqueness of mixtures condition creates situations in which the role of ergodic probabilities distribution $F_X(x)$ has a physical meaning to which many authors have assigned an objective nature, which could be interpreted as an exception to that principle.

Ergodic theory begins with Boltzmann, L.'s attempt (1868) to represent the probability distribution of a stochastic process in terms of averages in the time domain[33]. Given a stochastic process, continuous in the domain of states and continuous in the time domain and a function $g(X)$ in the domain $\Omega(X)$ of states, the expected value of the function g in the domain Ω is defined by $\int_\Omega g(x) f_X(x) dx$, where $f_X(x)$ denotes a density function. Likewise, let a region $A \in \Omega(X)$ be, replacing the function g by an indicator function of A, I_A, then $\int_\Omega I_A(x) f_X(x) dx = p(A)$ will be verified. Let, on the other hand, $T(t, X)$ the law that determines the trajectory of the process be $\{X(t)\}$, so that if $X(0) = x$, then $X(t) = T(t, x)$.

[33] This development known as "*abstract theory of dynamic phenomena*" derives from the process behaviour interpretation in terms of classic mechanics.

Suppose that this transformation is measure-preserving (or natural), that is, such that $M\left[T(t,A)\right]=M(A)$ (where $M(\bullet)$ denotes the class of all probability measures)[34]. The average in the time domain of the function g is given by $m\left[g(x)\right]=\lim_{t\to\infty}\frac{1}{t}\int_0^t g\left[T(y,x)\right]dy$. Substituting the function g by an indicator function I_A, this average (assimilable to a law of large numbers) can be interpreted as the average time to stay in the region A by the process, which was in the state x at time 0. The fundamental problem of ergodic theory, then, is to show that the average of the process in the time domain exists (stationarity condition), that is unique (that is, independent from the initial state of the process) and that the average in the states domain can be calculated as an average in the time domain, $\int_\Omega g(x)f_X(x)dx=\lim_{t\to\infty}\frac{1}{t}\int_0^t g\left[T(y,x)\right]dy$ (that is, to show that the probability $p(A)$ can be objectively expressed exclusively as the time average to stay by the process in the region A).

Let, in particular, a space Ω be formed by infinite binomial sequences $E=\{E_1,E_2,\ldots,E_n,\ldots\}$, let $x_n(E)$ the n-th element of a sequence of E and let a transformation T be similar to the one mentioned above, so $x_n(TE)=x_{n+1}(E)$ $(n=1,2,\ldots)$, resulting in $x_{n+1}(E)=x_1\left(T^nE\right)$. This transformation represents the repetition of the binomial sequence trial $E=\{E_1,E_2,\ldots,E_n,\ldots\}$ and its application provides the result to be achieved in the next repetition, that is, the sequence $E,TE,T^2E,\ldots$ defines a realization or trajectory of the process. It is said that a probability $p(\bullet)$ on Ω is a stationary measure if for every set $A\subset\Omega$, it is verified that $p\left(T^{-n}A\right)=p(A)$ $(n=1,2,\ldots)$, that is, if the probabilities functions on finite sequences of events remain invariant over time.

Suppose that, in particular, $E_i=1$ and $\overline{E}_i=0$ $(i=1,2,\ldots)$ represent the possible outcomes of a binomial phenomenon, the stationarity property is a necessary and sufficient condition to ensure that the limit of the relative frequency of the outcome E, $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n E_i$, exists and that, given an integrable function, the limit of the average in the time domain, $\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n g\left(T^iE\right)=E(g)=m(g)$, also exists and is independent of E. This show that all the realizations of the process will have the same asymptotic properties with a probability equal

[34] If $T(x,t)$ were invertible, in addition, it would be verified that $M(A)=M\left[T^{-1}(t,A)\right]$.

to 1. The ergodicity property implies, then, that an ergodic process has, in its entirety, the same asymptotic properties, that is, it cannot be separated into parts and guarantees the uniqueness of the limits of the relative frequencies of the outcome E [35].

The ergodic decomposition theorem[36] considers the relationship between the general concepts of ergodicity and stationarity and shows that every stationary sequence admits an integral representation in terms of stationary probabilities or, accordingly, a unique decomposition weighted in ergodic parts (that is, with identical asymptotic properties for each of the parts, but different for the different parts), each of which can be defined as a subset of the set of infinite sequences and whose weighs are given by the average time spent by the process in each part[37]. Suppose a non-ergodic transformation $p(\bullet)$, then, given an invariant set A, it will be verified that $p(A) > 0$. In addition, suppose there are two sets A_1 and $A_2 = \Omega - A_1$ containing no non-trivial invariant subsets, then the conditional probabilities $p(\bullet / A_1)$ and $p(\bullet / A_2)$ will be ergodic an its integral representation will be $p(B) = p(A_1)p(B / A_1) + p(A_2)p(B / A_2)$. Since $p(A_1) + p(A_2) = 1$ and $0 < p(A_1), p(A_2) < 1$, it can be asserted that $p(B)$ is defined by a unique mixture of ergodic probabilities, where $p(A_1)$ and $p(A_2)$ represent the weighs. In general, it can be asserted that, given a descomposition $\Im$, the integral over $\Im$ of the ergodic probabilities $p_D(E)$ corresponding to the partitions D weighted by a function $\mu(D)$, define a stationary probability $p(E) = \int_{\Im} p_D(E)\, d\mu(D)$.

As a corollary of the above results, it is obtained that the sequences of exchangeable events

---

[35] As a corollary of this statement one can conclude that the laws of large numbers are special cases of the ergodic theorem and that, consequently, the basis for the frequentist interpretation of the probability is not the particular condition of independence, but the most general property of ergodicity of the sequence of repetitions generated by the random phenomenon.

[36] Due to Koopman, B. (1930), von Neumann, J. (1932a)(1932b)(1932c) and Birkhoff, G.D. (1931). Khinchin, A. (1932a)(1932b)(1933) and Hopf, E. (1932)(1936)(1937) proposed a purely probabilistic formulation of this theorem, in terms of a measure-preserving transformations.

[37] It should be noted that, in the case of simple events, the asymptotic properties are characterized by the limits of relative frequencies that converge to the respective probabilities within each part, but are different for different parts and that the total probability is given by the weighted mixture of these limits, where the weighs are given by the partitions' measures.

are stationary, but with variable asymptotic properties due to the conditioning of the probabilities and also that sequences of independent events are associated with an ergodic measure. Then, since the Bernoullian measures satisfy the first law of large numbers and the ergodicity property implies a limit behavior identical for all sequences, we can conclude that the Bernoullian probabilities are ergodic and, therefore, that the classic representation theorem is a particular case of the ergodic decomposition theorem, in which the probability p defines an ergodic component formed by the set of all sequences whose relative frequency converge to p [38].

In the ambit of ergodic theory the representation theorem can be interpreted as follows: while the limit of relative frequency is unknown, the condition of exchangeability (that is, of stationarity) guarantees its existence. In terms of a subjectivist approach, this "ignorance" can be characterized by a mixture of its possible values in which the weighs of different hypotheses emerge as personal assignations of assumed true values of the probabilities through the Bayesian conditioning process and allow to obtain an intersubjective assessment of the probability.

Now then, if the process is such that it admits a decomposition due to the existence of ergodic partition which depends exclusively on the properties of the law that determines the trajectory of the process, the weighs of the decomposition assume a physical meaning, which led some authors to postulate that their distribution did not depend on the subjective characteristics of Bayesian conditioning. But, given the axiomatic nature of the premises on which the decomposition theorem is based on and, in contradiction to this proposition, it can be concluded that its acceptance is of a purely subjective nature.

On the other hand, keeping in mind that propensity is a physical property of the experiment design based on the assumption of a given structural behavior of the phenomenon and that its interpretation of the probability concept is referred to individual events and considering a definition of propensities based on the dynamic properties of a phenomenon that has a "real" trajectory and, therefore, supports an explanation in terms of classical mechanics, it was shown that it is possible to obtain an analytical justification of the propensity interpretation of ergodic theory. This leads to the conclusion that the ergodic theory is affected by the same metaphysical characteristics of the propensity model in which the assignment of probabilities is inevitably subjective.

7.2 .- The method of arbitrary functions

[38] The rigorous proof that the representation theorem is a special case of the ergodic theorem is due Ryll-Nardzewski, C. (1957). In addition, see Freedman, D. (1962) and Dynkin, E.B. (1978).

As an alternative to justify the existence of a unique objective definition of probability in the ambit of dynamic phenomena that support deterministic laws that determine its trajectory, Poincaré, H. (1896) and von Smoluchowski, M. (1918) introduced the method of arbitrary functions, which was widespread in the mathematical aspects by Fréchet, M. (1952), who showed that the sufficient condition for its implementation is to have a continuous of trajectories of the process and a density function on this set which, under certain conditions attached to it, will be asymptotically transformed into a single final distribution.

Hopf, E. (1934)(1936), in accordance with the principle by which the schemes based on the realization of a probability distribution "*... cannot determine the true origin of the laws of probability, since they are based on them*" proposed a new interpretation of Poincare and von Smoluchowski's definitions based on the conjecture which claims that the analytical justification of the method of arbitrary functions is exclusively found in ergodic theory (in which, as seen in the previous section, the assumption of the existence of an initial distribution is replaced by that of absolute continuity and the condition of independence is seen as a byproduct of the continuity property), that in some cases the ergodic nature of a dynamic phenomenon can be deductively determined and, therefore, that the method of arbitrary functions allows, in these cases, the definition of objective probabilities[39].

The theoretical foundation of Hopf's conjecture was provided by Sinai, Y. (1977), who showed that, if the domain of states is subject to a finite partition into macro-states, the succession of macro-states form a Bernoullian process, that is, a sequence in which the macro-states are stochastically independent and their probabilities are determined exclusively by the law governing the process trajectory.

Nevertheless, it should be noted that Hopf's conjecture and the results obtained by Sinai refer exclusively to phenomena whose behavior support an explanation in terms of classical mechanics. On the other hand if, as in the factic phenomena, the absolute continuity condition is not verified then, given the inevitable presence of random factors, the unique realization of the process cannot be regarded as a necessary consequence of its deterministic representation. It should be noted that, according to Khinchin, A. (1954), a necessary condition for the convergence of relative frequencies is that the expected values of functions in the domain of the states are independent of the initial distribution and that this condition does not hold for singular distributions, so that for the case of unique trajectories, the ergodic theorems postulates, and consequently, the laws of large numbers linking relative frequencies with probabilities, are not verified.

[39] However, it should be noted that the subjectivist interpretation of the absolute continuity condition leads to Savage, L.J.'s conjecture (1973) whereby the probability functions can only be derived from other probability functions.

8. Conclusions

As a counterexample to de Finetti's "motto" which postulates that "*the probability does not exist*", some authors have proposed a form of strict identification of the true value of a probability based on the exchangeability property, on the postulates of the representation theorem and on its interpretation in the ambit of ergodic theory.

As proof of the inconsistency of this proposal, it was firstly demonstrated that, from a subjectivist interpretation of the representation theorem it can be concluded that the statement about the convergence of the relative frequency on a "true value " of the probability, in the cases of Bernoullian phenomena, is not but a metaphysical illusion, which is actually attributable to an asymptotic behavior of the individual assessments of the initial probabilities according to a Bayesian conditioning scheme, which leads to an intersubjective assignment easily confused with an objective probability.

With regard to the results obtained by Koopman, von Neumann, Birkhoff, Khinchin and Sinai related to the principles of the ergodic partition theorem and the method of arbitrary functions, their foundation on axiomatic nature premises (stating that the phenomena support a deterministic explanation in terms of classic mechanics and satisfy the condition of absolute continuity) allowed the conclusion that their acceptance is exclusively of a subjective nature.

Finally, as an immediate consequence of the analytical justification of the propensionalist interpretation of ergodic theory, it was demonstrated that it is affected by the same metaphysical characteristics of the propensity model in which the assignment of probabilities is inevitably subjective.

References

- Accardi, L.; Lu, Y.G. (1993): "A continuous version of de Finetti's theorem". Annals of Statistics, vol. 21.
- Albert, M. (1992): "Die Falsifikation statisticher Hypothesen". Journal for General Philosophy of Science, vol. 23.
- Albert, M. (1999): "Bayesian learning when chaos looms large". Economic Letters, vol. 65.
- Aldous, D.J. (1985): "*Exchangeability and related topics*". Springer-Verlag.
- Bernardo, J.M.; Berger, J.O.; Dawid, A.P.; Smith, A.F.M. (eds.)(1992): "*Bayesian statisics*". Oxford University Press.
- Bernardo, J.M.; Smith, A.F.M. (1994): "*Bayesian theory*". Wiley.
- Berti, P.; Rigo, P. (1994): "A Glivenko-Cantelli theorem for exchangeable random variables". Working Papers, nº 52, Università degli Studi de Firenze.
- Billingsley, P. (1965): "*Ergodic theory and information*". Wiley.
- Birkhoff, G.D. (1931): "Proof of the ergodic theorem". Proceedings of the National

Academy of Sciences, vol. 17.
- Boltzmann, L. (1868): "Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten".Wissenchaftliche Abhandlungen, vol. 1.
- Borel, E. (1965): "*Eléments de la théorie des probabilités*".Hermann, 1909. Traducción al inglés, "*Elements of the theory of probability*", Prentice-Hall.
- Braithwaite, R.B. (1957): "On unknown probabilities". En Körner, S. (ed.).
- Carnap, R.; Jeffrey, R.C. (eds.)(1971): "*Studies in inductive logic and probability*". University of California Press.
- Carnap, R. (1977): "*Two essays on entropy*". University of California Press.
- Castelnuovo, G. (1925-28): "*Calcolo delle probabilità*". Zanichelli.
- Chatterji, S.D. (1974a): "A subsequence principle in probability theory". Bulletin of the American Mathematical Society, vol. 80.
- Chatterji, S.D. (1974b): "A principle of subsequences in probability theory". Advances in Mathematics, vol. 13.
- Cifarelli, D.M.; Muliere, P. (1989): "*Statistica Bayesiana*". G. Iuculano.
- Cifarelli, D.M.; Regazzini, E. (1996): "de Finetti's contribution to probability and statistics". Statistical Science, vol. 11.
- Cohen , E.G.D.; Thirring, W. (eds.)(1973): "*The Boltzmann equation*". Springer-Verlag.
- Coletti, G.; Dubois, D.; Scozzafava, R. (eds.)(1994): "*Mathematical models for handling partial knowledge in artificial intelligence*". Plenum Press.
- Daboni, L. (1953): "Aspetti di una interpretazione geometrica per la probabilità di eventi equivalenti". Rendiconti di Matematica della Università di Parma, vol. 4.
- Daboni, L.; Sigalotti, L. (1987): "Processi stocastici". Atti del Convegno "Ricordo de Bruno de Finetti, professore nell'Ateneo Triestino".
- Daboni, L. Wedlin, A. (1982): "*Statistica. Un'introduzione all'impostazione neo-Bayesiana*". UTET.
- Dacunha-Castelle, D. (1974): "Indiscernability and exchangeability in L^p spaces". Proceedings Seminar on Random Series, Convex Sets and Geometry of Banach spaces, Aarhus.
- Dawid, P. (1985): "Probability symmetry and frequency". British Journal for the Philosophy of Science, vol. 36.
- de Finetti, B. (1928): "Funzione caratteristica di un fenomeno aleatorio". Atti del Congresso Internazionale di Matematici, Bologna, Zanichelli.
- de Finetti, B. (1930a): "Funzione caratteristica di un fenomeno aleatorio". Memorie della Reale Accademia Nazionale dei Lincei, vol. 4, fasc. V.
- de Finetti, B. (1930b): "*Fondamenti logici del ragionamento probabilistico*". B.U.M.I..
- de Finetti, B. (1930c): "Funzione caratteristica di un fenomeno aleatorio". Memorie della Reale Accademia Nazionale dei Lincei, vol. 4, fasc. V.
- de Finetti, B. (1930d): "Problemi determinati e indeterminati nel calcolo delle probabilità". Rendiconti della RealeAccademia Nazionale dei Lincei, vol. 12.
- de Finetti, B. (1931a): "*Probabilismo: Saggio critico sulla teoria della probabilità e sul valor della scienza*". Biblioteca di Filosofia, Perrella.
- de Finetti, B . (1931b): "Sul significato soggettivo della probabilità" . Reeditado en de

Finetti, B. (1991).
- de Finetti, B. (1933a): “Classi di numeri aleatori equivalenti”. Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti, vol. 18.
- de Finetti, B. (1933b): “La legge dei grandi numeri nei caso dei numeri aleatori equivalenti”. Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti, vol. 18.
- de Finetti, B. (1933c): “Sulla legge di distribuzione dei valori in una successione di numeri aleatori equvalenti”. Atti Reale Accademia Nazionale dei Lincei, Serie VI, Rendiconti, vol. 18.
- de Finetti, B. (1934): “Indipendenza stocastica ed equivalenza stocastica”. Atti della Società Italiana per il Progreso delle Scienze (XXII riunione), vol. 2.
- de Finetti, B. (1937a): “La prévision: ses lois logiques, ses sources subjectives”. Anales del’Istitut Henri Poincaré, vol. 7. Traducción como “Foresight its logical laws, its subjective sources”. En Kyburg, H.E.; Smokler, H.E. (eds.), 1964.
- de Finetti, B. (1937b): “Sur la condition d’equivalence partielle”. Actualités Scientifiques et Industrielles, vol. 739. Traducido en Jeffrey, R. (1980).
- de Finetti, B. (1952): “Gli eventi equivalenti e il caso degenere”. Giornale dell’Istituto Italiano degli Attuari, vol. 15.
- de Finetti, B. (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.
- de Finetti, B. (1969): “Sulla proseguibilità di processi aleatori scambiabili”. Rendiconti dell’Istituto di Matematica della Università di Trieste, vol. 1.
- de Finetti, B. (1970): “*Teoria della probabilità*”. Einaudi.
- de Finetti, B. (1991): “*Scritti (1931-1936)*”. Pitagora Editrice.
- Diaconis, P. (1987): “Application of the method of moments in probability and statistics”. En Landau, H.J. (ed.).
- Diaconis, P.; Freedman, D. (1980): “Finite echangeable sequences”. Annals of Statistics, vol. 8.
- Dubucs, J.P. (ed.)(1993): “*Philosophy of probability*”. Kluwer.
- Dynkin, E.B. (1978): “Sufficient statistics and extreme points”. Annals of Probability, vol. 32.
- Feller, W. (1950): “*Introduction to probability theory and its applications*”. Wiley.
- Frankfurt, H.G. (1962): “Philosophical certainty”. Philosophical Review, vol. 71.
- Frechét, M. (1943): “*Les probabilités associées a un systéme d’événements compatibles et dependants, cas particuliers et applications*”. Hermann.
- Frechét, M. (1952): “Méthode des fonctions arbitraires, théorie des événements en chaine dans le cas d’un nombre fini de’etats possibles”. Gauthier-Villars.
- Gaifman, H. (1971): “Applications of de Finetti’s theorem to inductive logic”. En Carnap, R.; Jeffrey, R.C. (eds.).
- Garbolino, P, (1990): “La nuova arte di pensare: Elementi di logica dell’inferenza induttiva”. Annali dell’Università di Ferrara, Serie Filosofia, nº 14.
- Gilio, A. (1992): “C_0-coherence and extensions of conditional probabilities”. En Bernardo, J.M.; Berger, J.O.; Dawid, A.P.; Smith, A.F.M. (eds.).
- Gillies, D.A. (1996): “*Artificial intelligence and scientific method*”. Oxford University Press.

- Goel, P.K.; Zellner, A. (eds.)(1986): "*Bayesian inference and decision techniques with applications*". North-Holland.
- Good, I.J. (1965): "*The estimation of probabilities; An essay on modern Bayesian methods*". MIT Press.
- Haag, J. (1928): "Sur un problème général de probabilités et ses diverses applications". Proceedings of the International Congress of Mathematicians. Universidad de Toronto Press.
- Hewitt, E.; Savage, J. (1955): "Symetric measures on Cartesian products". Transactions of the American Mathematical Society, vol. 80.
- Hintikka, J. (1970): "Unknown probabilities, Bayesianism and de Finetti's representation theorem". Boston Studies in the Philosophy of Science, vol. 8.
- Hintikka, J. (ed.)(1977): "*Rudolf Carnap, logical empiricist*". Reidel.
- Hopf, E. (1932): "On the time average problem in dynamics". Proceedings of the National Academy of Sciences, vol. 18.
- Hopf, E. (1934): "On causality, statistics and probability". Journal of Mathematics and Physics, vol. 13.
- Hopf, E. (1936): "Ueber die Bedeutung der willkürlichen Funktionem für die Wahrscheinlichkeitstheorie". Jahresbericht der Deutsche Mathematiker-Vereinigung, vol. 46.
- Hopf, E. (1937): "*Ergodentheorie*". Springer.
- Howson, C.; Urbach, P. (1989): "*Scientifica reasoning: The Bayesian appproach*". Open Court.
- Jaynes, E.T. (1986): "Some applications and extensions of the de Finetti representation theorem". En Goel, P.K.; Zellner, A. (eds.).
- Jeffrey, R.C. (1965): "*The logic of decision*". University of Chicago Press.
- Jeffrey, R. (1980): "Studies in inductive logic and probability". University of California Press.
- Khinchin, A. (1932a): "Remarques sur les suites d'événements obeissant à la loi des grandes nombres". Matematicheski Sbornik, vol. 39.
- Khinchin, A. (1932b): "Zu Birkhoffs Lösung des Ergodemproblems". Mathematische Annalen, vol. 107.
- Khinchin, A. (1932c): "Sulle successioni stazionarie di eventi". Giornale dell'Istituto Italiano degli Attuari, vol. 3.
- Khinchin, A. (1933a): "*Asymptotische gesetze der Wahrscheinlichkeitsrechnung*". Springer.
- Khinchin, A. (1933b): "Sur les classes d'événements équivalents". Matematischeski Sbornik, vol. 39.
- Khinchin, A. (1938): "Zur Methode der willkürlichen Funktionen". Mathematischeski Sbornik, vol. 45.
- Khinchin, A. (1954): "Die Method der willkürlichen Funktionen und der Kampf gegen den Idealismus in der Wahrscheinlichkeitsrechnung". Swjewissenschaft-Naturwissenschaftliche Abteilung, vol. 7.
- Kingman, J.F.K. (1978): "Uses of exchangeability". Annals of Probability, vol. 6.

- Koch, G.; Spizzichino, F. (eds.)(1982): "*Exchangeability in probability and statistics*", North-Holland.
- Kolmogorov, A.N. (1933): "Grundbegriffe der Wahrscheinlichkeitsrechnung". Springer. Traducción al inglés, "*Foundations of the theory of the probability*", Chelsea, 1950.
- Körner, S. (ed.)(1957): "*Observation and interpretation. A symposium of philosophers and physicists*". Butterworth.
- Koopman, B. (1930): "Birkhoff on dynamical systems". Bulletin of the American Mathematical Society, vol. 26.
- Kyburg, H.E.; Smokler, H.E. (eds.)(1964): "*Studies in subjective probability*". Wiley.
- Landau, H.J. (ed.)(1987): "*Moments in mathematics*". American Statistical Society.
- Landro, A.H. (2010a): "*Acerca de la probabilidad*". Ediciones Cooperativas.
- Landro, A.H. (2010b): "Acerca de la existencia del verdadero valor de una probabilidad. Parte 1: Determinismo vs. aleatorismo". Jornadas de la AAEP.
- Landro, A.H.; González, M.L. (2009): "*Elementos de econometría de los fenómenos dinámicos*". Ediciones Cooperativas.
- Link, G. (1980): "Representation theorems of the de Finetti type for (partially) symmetric probability measures". En Jeffrey, R.C..
- Milne, P. (1997): "Bruno de Finetti and the logic of conditional events". British Journal of Philosphy of Science. Vol. 48.
- Mura, A. (1989): "Concezione logico-trascendentale della probabilità e induzione". Annali dell'Università di Ferrara. Serie Filosofia, nº 6.
- Mura, A. (1992): "*La sfida scettica: Saggio sul problema logico dell'induzione*". ETS.
- Naddeo, A. (ed.)(1987): "*Italian contributions to the methodology of statistics*", S.I.S, Cleup.
- Petris, G.; Regazzini, E. (1994): "About the characterization of mixtures of Markov chains. The recurrence case". Riunione Scientifica su Inferenza Statistica, Udine.
- Petris, G.; Secchi, P. (1989): "Sul teorema di rappresentazione di de Finetti". Metron, vol. XLVII.
- Piccinato, L. (1991): "Questioni critiche en diversi paradigmi inferenziali". Atti del Convegno S.I.S., Cagliari.
- Poincaré, H. (1894): "Sur la théorie cinétique des gaz". Revue Générale des Sciences Pures et Appliquées, vol. 5.
- Poincaré, H. (1896): "*Calcul des probabilités*". Gauthier-Villars.
- Popper, K.R. (1957): "Probability magic or knowledge out of ignorance". Dialectica, vol. 11.
- Racugno, W. (1982): "*Rassegna di metodi statistici ed applicazioni*". Pitagora.
- Ramsey, F.P. (1931): "*The foundation of mathematics and other logical essays*". Routledge-Kegan.
- Regazzini, E. (1987): "de Finetti's coherence and statistical inference". Annals of Statistics, vol. 15.
- Regazzini, E. (1988): "Argomenti di statistica". Quaderni di Matematica, Milán.
- Regazzini, E. (1991): "Coherence, exchangeability and statistical models (de Finetti's stance revisited)". Atti del Convegno S.I.S., Cagliari.
- Ryll-Nardzewski, C. (1957): "On stationary sequences of random variables and the de Finetti's equivalence". Colloquium Mathematicum, vol. 4.

- Sahlin, E. (1993): “On higher order beliefs”. En Dubucs, J-P- (ed.).
- Savage, J. (1954): “*The foundations of statistics*”. Dover.
- Savage, J. (1973): “Probability in science: a personality account”. En Suppes, P. (ed.).
- Scozzafava, R. (1982): “Introduzione elementare alla scambiabilità”. En Racugno, W. (ed.).
- Scozzafava, R. (1990): “Probabilità condizionate: de Finetti o Kolmogorov”. Scritti in omaggio a L. Daboni, Lint.
- Shimony, A. (1977): “Carnap on entropy”. En Hintikka, J. (ed.).
- Sinai, Y. (1973): “Ergodic theory”. En Cohen, G.D.; Thirring, W. (eds.).
- Sinai, Y. (1976): “*Introduction to ergodic theory*”. Princeton University Press.
- Skyrms, B. (1980): “*Causal necessity: A pragmatic investigation of the necessity of laws*”. Yale University Press.
- Skyrms, B. Harper, W.L. (eds.)(1988): “*Causation, chance, credence*”. Kluwer.
- Spielman, S. (1976): “Exchangeability and the certainty of objective randomness”. Journal of Philosophical Logic, vol. 5.
- Spielman, S. (1977): “Physical probability and Bayesian statistics”. Synthése, vol. 36.
- Suppes, P. (1973): “*Logic, methodology and philosophy of science*”. Nolrth-Holland.

- Suppes, P. (1980): “*Studies in the foundations of quantum mechanics*”. Philophy of Science Association.
- Suppes, P. (1984): “*La logica del probabile, un approccio Bayesiano alla razionalità*”. CLUEB.
- Suppes, P.; Zanotti, M. (1980): “A new proof of the impossibility of hidden variables using the principles of exchangeability and identity of conditional distributions”, En Suppes, P. (ed.).
- Viertl, R. (ed.)(1987): “*Probability and Bayesian statistics*”. Plenum Press.
- von Neumann, J. (1932a): “A proof of the quasi-ergodic hypothesis”. Proceedings of the National Academy of Sciences, vol. 18.
.- von Neumann, J. (1932b): “Zur Operatorenmethode in der klassischen Mechanik”. Annals of Mathematics, vol. 33.
- von Neumann, J. (1932c): “Physical applications of the ergodic hypothesis”. Proceedings of the National Academy of Sciences, vol. 18.
- von Plato, J. (1981): “On partial exchangeability as a generalization of symmetry principles”. Erkenntnis, vol. 16.
- von Plato, J. (1982a): “Probability and determinism”. Philosophy of Science, vol. 49.
- von Plato, J. (1982b): “The generalization of de Finetti’s representation theorem to stationary probabilities”. Proceedings of the Biennial Meeting of the Philosophy of Science Association.
- von Plato, J. (1983): “The method of arbitrary functions”. British Journal for the Philosophy of Science, vol. 34.
- von Smoluchowski, M. (1918): “Über den Begriff des Zufalls und den Ursprung der Warhrscheinlichkeitsgesetze in der Physik”. Die Naturwissenschaften.
- Wechsler, S. (1993): “Exchangeability and predictivism”. Erkenntnis, vol. 38.
- Zecchin, M. (1987): “de Finetti’s work on probability and statistics”. En Naddeo, A. (ed.).

- Zabell, S.L. (1988): “Symmetry and its discontents”. En Skyrms, B.; Harper, W.L. (eds.).