Rendiconti di Matematica, Serie VII
Volume 37, Roma (2016), 199 – 218

Guido Castelnuovo, pioneer of the theory and applications of probability in Italy

EUGENIO REGAZZINI*

Abstract: *We describe the work of Guido Castelnuovo (1865-1952) in the theory of probability and its applications to statistics and physics.*

1 – Introduction

In the preface to the Castelnuovo *Opere Matematiche* [17], after a brief but vivid account of his work in geometry between 1885 and 1906, to be considered as one of the most significant moments of the renowned Italian school, the editors point out that probability played a central role in the second part of his scientific career. The present paper centres around the talk given by the author at the meeting *Guido Castelnuovo: un ricordo a 150 anni dalla nascita* (Rome, 5th November 2015) and originates from a couple of papers ([56, 57]) devoted to the history of probability in Italy between the two world wars. The work of Castelnuovo in the field is here expounded through the study of some salient aspects of his book on the theory of probability (Section 2). Attention is also paid to his stance on inductive statistical reasoning (Section 3) and on applications to physics (Section 4). To conclude, his efforts to spurring and developing probabilistic and statistical studies, within the Italian scientific community, are highlighted (Section 5).

---

*Also affiliated with CNR-IMATI, Milano, italy.
Work partially supported by MIUR-2012AZS52J-003.

Key Words and Phrases: *Castelnuovo – Empirical law of chance – Kinetic theory of gases and Maxwell law – Law of (a vector of) errors of observation – Problem of moments – Second limit-theorem, Theory of dispersion*
A.M.S. Classification: 6001.

2 – Theory of probability

Octav Onicescu (1892-1983), one of his most renowned students in Rome, described Castelnuovo's passion for probability in this way: "Le désir de tout connaitre, de tout ordonner à sa manière, de tout rendre clair et bien exprimé mathématiquement dominait sa conversation. Ce fut peut-être cela qui a décidé de son Oeuvre dans le domaine des Probabilités"[1]. See [54]. In his talk at the above-mentioned meeting in Rome, Enrico Arbarello expressed an even more fascinating conjecture according to which Castelnuovo's interest in probability might have been stimulated by his research in enumerative geometry (1889-1890) aiming at determining the number – assumed to be finite – of solutions of a certain problem; see [17, Vol. I].

His earliest involvement with the teaching of probability took place in the academic year 1915-1916. Those lectures, together with his close scientific relationship with Francesco Paolo Cantelli (1875-1966), were at the basis of the first edition of the treatise *Calcolo delle Probabilità* (CdP, from now on) in 1919. As to the organization of the subject, Castelnuovo adopted the same structure as in the renowned French treatises [2, 3] and [55], with a remarkable novelty consisting in making the most of certain works of a few distinguished Russian mathematicians in the school of Pafnuti L. Chebyshev (1821-1894), that had been generally neglected, up to that time, by western writers. In fact, Castelnuovo regarded those works as the major contribution to probability after Laplace. Most of all, he recognized therein both the adoption of an appropriate language and the introduction of suitable new entities particularly consonant with his own idea of probability in connection with its use in the analysis of real phenomena. This opinion can be linked with his firm belief that probability was destined to become an unavoidable constituent of the modern scientific reasoning. In fact, one of the main features of CdP is represented by the effort to reach a satisfactory definition of probability in order to frame both theoretical results and applications into a logically rigorous system. The position of our Author can be classed under the *empirical approach*, as opposed both to any subjectivistic intrusion, and to the *asymptotical approach*, like that of Richard von Mises (1883-1953). In fact, the latter considers infinite sequences of trials (collectives) and aims at defining probability as limiting value of the frequency, which, from the Castelnuovo viewpoint, is devoid of any empirical value [15]. A reason why CdP deserves to be still read is that it contains one of the most transparent and in-depth explanations of the empirical approach, still shared by many contemporary probabilists and statisticians. Besides Cantelli, among the contemporaries following an analogous way of thinking, it is worth mentioning Maurice Fréchet (1878-1973) and Paul Lévy (1886-1971). In order to build his theory, he initially

[1]The desire of knowning everything, of ordering everything, of rendering everything clear dominated his conversation. Maybe this is what decided of his Work in the field of Probability.

observes that in every mathematical theory applied to real phenomena, we must consider two moments or trends: *Il primo indirizzo, astratto o formale, opera coi processi della logica deduttiva sopra concetti ai quali si richiede soltanto di soddisfare a certe condizioni compatibili, formulate dagli assiomi o postulati teorici. Il secondo indirizzo, concreto, tien conto dei significati fisici che si possono attribuire a quei concetti, e viene collegato al precedente da postulati empirici esprimenti che i concetti stessi o le loro proprietà astratte trovano una interpretazione approssimata in enti o fenomeni del mondo esterno. Solo in grazia di questo secondo indirizzo la scienza pura può ricevere applicazioni*[2]. See [11, p. XVIII].

He then observes that also in the theory of probability these moments appear and liken this theory to an experimental science. In relation to the abstract building of probability theory, Castelnuovo agrees with the so-called classical approach according to which probability is the ratio of the favourable to the possible elementary cases, and there is a correspondence between logical operations on classes and arithmetical operations on probabilities. But, even if this approach could give us a coherent and complete theory, he observes, it would be insufficient for those who want to employ probability theory to formulate previsions about random events. *Ora nessun ragionamento teorico può dimostrare che mille estrazioni eseguite da un'urna contenente una palla bianca e una palla nera daranno palle dei due colori presso a poco nella stessa proporzione. È questo un fatto di carattere sperimentale che occorre enunciare sotto forma di postulato empirico, affinchè le proposizioni astratte sulle probabilità trovino applicazioni*[3]. (See [11, p. XIX]).

Concretely, the mentioned postulate is an explanation of the Castelnuovo idea that the relation between frequency and probability is of empirical character, an idea open, *e.g.*, to the criticism of authors adhering to the subjectivistic approach. Through this line of reasoning, he introduces an "empirical postulate", the so-called

---

[2] *The former trend, abstract and formal, works through the deductive logic processes regarding concepts which must only satisfy some compatible conditions, formulated by axioms. The latter trend, concrete, emphasizes the physical meaning attributed to those concepts and is connected with the preceding one through empirical postulates showing that concepts themselves, or their abstract properties, find an approximative interpretation either in terms of objects or in terms of phenomena of the external world. It is in view of this trend that pure science can be applied.*

Italic type is used for passages drawn from the second edition of CdP, which came out in two volumes in 1926 and 1928, respectively. They are translated from Italian by the present author.

Compared to the first, the second edition shows a few additions only, concerning the two-dimensional Gaussian law and its impact on the theory of statistical regression, as well as a few refinements of the proof of the central limit theorem suggested by the recent, excellent, analogous treatise by Lévy [49]. See [12].

[3] *Now there is no theoretical argument able to demonstrate that a thousand drawings from an urn, containing one white ball and one black ball, will give us nearly the same proportion of both colours. This is a fact whose character is experimental and it is necessary to enounce it with the form of an empirical postulate, if we want to find applications of the abstract statements of the theory of probability.*

empirical law of chance [postulato empirico del caso], according to which (see [11, p. XIX]):

> *Se un evento ha una probabilità constante p in ogni prova, e se esso si verifica m volte in n prove, il rapporto, frequenza, m/n dà un valore approssimato della probabilità p; e l'approssimazione è ordinariamente tanto migliore, quanto maggiore è il numero n delle prove*[4].

Castelnuovo is aware of the limitations of the resulting theory, since the formulation of the empirical law of chance relies on vague and ambiguous expressions. Hence, he further stresses that it is impossible to compare such a law to the postulates of geometry. On the other hand, it is possible to compare it with propositions of this kind: *le proprietà della retta valgono, in modo approssimato, per i fili tesi*[5] (see [11, p. XX]), to which we have resort, albeit implicitly, when we are dealing with actual applications of geometry. However, in geometry it is possible to fix an upper bound for the error we make transferring a geometric property from the abstract field to that of applications. On the contrary, in the theory of probability it is impossible to act in this way, since the error concerning the assessment of a probability *via* a frequency cannot be predetermined. According to this, the empirical law of chance cannot be considered as a base for a logical foundation of the theory of probability. In any case, Castelnuovo's viewpoint forces him to fix the meaning of the Bernoulli law of large numbers which some scholars thought sufficient to explain the decrease of $|p - m/n|$ as n increases. He stresses that the Bernoulli theorem − within the classical theory of probability − does not in itself imply that there is a necessarily high frequency of cases in which $|p - m/n|$ remains below a fixed value. On the other hand, such an implication is justified by the empirical law of chance which, consequently, cannot be replaced by the Bernoulli law of large numbers. Here, he is referring to Jacob Bernoulli (1654-1705) and, especially, to his posthumous work [1]. On the contrary, according to Castelnuovo, it is such an empirical law which makes Bernoulli theorem interesting on the side of applications which require the shift from the abstract idea of probability to the experimental concept of frequency.

Obviously, the law of chance restricts the domain of application of the probability theory to those situations where we can determine

> *esperienze atte a dare, mediante determinazioni di frequenze, quanti si vogliano valori approssimati della probabilità di cui tratta il problema. Questo criterio avrebbe condotto a respingere senz'altro, come insufficienti, gli enunciati di varie questioni sulle probabilità geometriche che condussero ai più*

[4] *If an event has a constant probability p in each trial and if it comes true m times in n trials, the frequency m/n gives us an approximative value of the probability p; the higher the trials' number is, the better the approximation ordinarily is.*

[5] *The properties of the straight line can be applied, approximatively, to a tight wire.*

> *strani paradossi. Così non è lecito domandare la probabilità che una corda tracciata a caso in un cerchio sia minore del raggio, finchè non si precisino le modalità con cui vengono segnate le corde*[6].

As to the mathematical content, in the eyes of the development of the subject from the Thirties, CdP seems fatally outdated, but with reference to the time it appeared, it must be considered as an advanced transparent, well-written and rigorous presentation of the main achievements of the theory until the beginning of the Twenties. As an example, Chapter 3 gives the definitions of real-valued random variable [*variabile casuale*] and random vector (*sistema di variabili casuali*) meant as entities that may take just one value within a well-specified set of possible values, depending on which of the events forming a partition of the certain event comes true. Moreover, a definition of probability law for random variable and random vector is given. In the same chapter, expectation [*valor medio*, $M(\cdot)$ in symbols] is introduced and analysed so that it may be used as a positive linear operator to make precise, among other things, the definition of new entities such as moments, mean deviations, etc., and the formulation of the fundamental Bienaymé-Chebyshev inequality. Section 36 is devoted to the *Bernoulli weak law of large numbers*, and hints at more recent studies by Cantelli [5, 6] and [9] are made to mention (without proof) the strong law, and to stress the importance of the concept of convergence in probability both of a sequence of random variables, and of the supremum, from a suitable index forth, of their deviations from a given "center". Also Sections 20-30 on random variables and mean values are influenced – as admitted by the Author himself – by the concise treatment made by Cantelli in some of his papers, such as [4] and [5]. A distinguished part of CdP is that devoted to the *central limit theorem*, that is, Chapter 8 and Appendix to Volume II: a real appreciation of the contribution to the subject by Chebyshev starting from 1879, Markov (1856-1922) and Lyapunov (1857-1918) in [18], [53], and [52] respectively. Chapter 8 contains a concise, rigorous and effective presentation of the main results, with a few hints at their proof. The Appendix to Volume II includes the complete proof of the following proposition (Markov), well-known as a version of the *second limit-theorem*, which, by matching together the results in the Appendix to Volume II in Section 2 of Art. I, pp. 144-145, and Section 16 of Art. II, p. 187, reads: *Let* $X_1, \dots, X_n, \dots$ *be independent real-valued random variables such that* $M(X_n) = 0$ *for every* n, $\sigma_n^2 := \sum_{i=1}^n M(X_i^2)$ *and* $m_n^{(r)} := \sum_{i=1}^n M(|X_i|^r)$ *are finite for all positive integers* n

[6] *Experiences able to provide assessments of the involved probabilities, with the desired degree of approximation, by means of an actual sequence of frequencies. This criterion would have led to consider as inadequate the formulations of several problems, involving geometric probabilities; in fact, these problems have often been the cause of strange paradoxes. Thus, we cannot ask probability that the chord of a circle be shorter than the radius of the same circle, if we do not prescribe a set of rules to trace chords.*

and r; moreover, assume that $\lim_{n\to+\infty} \frac{m_n^{(r)}}{\sigma_n^r} = 0$ *for* $r = 3, 4, 5, \ldots$. *Then, the probability that* $X_1 + \cdots + X_n$ *belongs to* $(\alpha\sigma_n\sqrt{2}, \beta\sigma_n\sqrt{2})$ *converges to* $\int_\alpha^\beta e^{-x^2} dx/\sqrt{\pi}$, *as* $n \to +\infty$, *for every* $-\infty < \alpha < \beta < +\infty$. In proving the theorem, according to the Chebyshev approximation method based on moments, Castelnuovo makes the most of his algebraic cleverness to get useful simplifications and, at the same time, to point out interesting connections between different mathematical aspects. In particular, one notices his skill in dealing with the problem of moments jointly with the Chebyshev inequalities concerning the difference between a given probability distribution function and a second approximating discrete probability distribution function with n jumps and the same first $2n$ moments $m_0, \ldots, m_{2n-1}$ as the former. Castelnuovo will return to the problem of moments on other occasions (see, *e.g.*, [13, 14]), with an in-depth analysis of the Hamburger solution [37, 38, 39]. This is revisited by relying on arguments of an algebraic nature, that Castelnuovo considers more direct and elementary than the original analytical methods used by Hamburger. Coming back to the central limit theorem, Castelnuovo (Sections 21-25 of Art. II of the Appendix to Volume II shows how the use of the characteristic function, most notably of the Lévy continuity theorem combined with convergence of moments, is conducive to the "optimal" version of the second limit-theorem. Finally, Sections 24 and 25 present versions of the central limit theorem proved by Lyapunov [52] and Lindeberg [50] and [51], respectively. But, whilst the description of the Lyapunov contribution is complete, even if without proof, that of Lindeberg's work does not mention the main result based on the assumption generally referred to as *Lindeberg condition*, a feather in the cap of this distinguished Finnish mathematician.

3 – Inductive logic and Statistics

The analysis of the relationship between probability and inductive reasoning has been for years at the centre of the unsettled controversy among scholars regarding the meaning and applications of probability theory. Castelnuovo himself went deeply into this subject in Chapters IX and X of CdP. In conformity to his approach to probability, he rejected the Bayesian formulation of the problem. He proposed to deal with the question of rejecting an isolate hypothesis on the ground of the sole likelihood. In this sense, he could be considered as a forerunner of the Anglo-American School. To resume the reasoning followed by Castelnuovo in order to support his own thesis in relation to the elementary, but fundamental, problem of estimating the probability of success in a Bernoulli scheme, one can start from CdP, p. 197.

Da un'urna che contiene $\boldsymbol{a}$ *palle, tra bianche e nere, in proporzione incognita, si sono ottenute, mediante n successive estrazioni (riponendo via via la palla estratta nell'urna),* ν $(\leq n)$ *palle bianche. Qual'è la probabilità che* α $(\leq \boldsymbol{a})$ *palle dell'urna siano bianche? Qui* ν/n *è la frequenza constatata dell'evento; si vuol sapere quale plausibilità abbia l'ipotesi che sia* $\alpha/\boldsymbol{a}$ *la probabilità incognita dell'evento stesso.*
Si ragiona ordinariamente così. L'urna in questione può considerarsi come scelta a caso fra più urne, contenenti ciascuna $\boldsymbol{a}$ *palle con tutte le composizioni possibili, urne che possiamo indicare coi simboli* $(0, \boldsymbol{a})$ $(1, \boldsymbol{a}-1)$, ..., $(\boldsymbol{a}, 0)$ *i quali ci dicono quante palle bianche e nere sono ivi racchiuse. Siano* ω_0, ω_1, ... $\omega_{\boldsymbol{a}}$, *le probabilità "a priori" che la scelta sia caduta sul primo, secondo, ..., ultimo tipo delle dette urne. La probabilità che, essendo stata scelta un'urna* $(i, \boldsymbol{a}-i)$, *siano sortite, in n estrazioni,* ν *palle bianche è*

$$p_i = \binom{n}{\nu}\left(\frac{i}{\boldsymbol{a}}\right)^{\nu}\left(\frac{\boldsymbol{a}-i}{\boldsymbol{a}}\right)^{n-\nu}.$$

Qual'è la probabilità "a posteriori" che l'urna prescelta (cioè quella da cui le estrazioni avvennero realmente) abbia la composizione $(\alpha, \boldsymbol{a}-\alpha)$*? La formola di* Bayes *ci dà la risposta*

$$P_\alpha = \frac{\omega_\alpha p_\alpha}{\omega_0 p_0 + \omega_1 p_1 + \cdots + \omega_{\boldsymbol{a}} p_{\boldsymbol{a}}}\ ^{7}.$$

At the same time, the empirical approach prevents Castelnuovo from sharing Bayesian inferential approach; in other words, the interpretation of probability given by

[7] *From an urn containing* $\boldsymbol{a}$ *balls, which are black and white according to an unknown proportion, n balls are drawn one after the other with replacement and* $\nu(\leq n)$ *white balls are obtained. Find the probability that* α $(\leq \boldsymbol{a})$ *balls in the urn will be white. Here* ν/n *is the observed frequency, and we want to estimate the plausibility of the hypothesis* $\alpha/\boldsymbol{a}$ *pertinent to the unknown probability of success.*

Normally, one can argue in this way. The urn in question can be considered as chosen at random among several urns, each of which contains $\boldsymbol{a}$ *balls according to one of the possible compositions:* $(0, \boldsymbol{a}), (1, \boldsymbol{a}-1), \ldots, (\boldsymbol{a}, 0)$*... Let* $\omega_0, \omega_1, \ldots, \omega_{\boldsymbol{a}}$ *be the "a priori" probabilities of choosing an urn of the first, second and so on kind of urn. Then, the probability of obtaining* ν *balls in n drawings, provided that* $(i, \boldsymbol{a}-i)$ *represents the composition of the selected urn, is given by*

$$p_i = \binom{n}{\nu}\left(\frac{i}{\boldsymbol{a}}\right)^{\nu}\left(\frac{\boldsymbol{a}-i}{\boldsymbol{a}}\right)^{n-\nu}.$$

What is the "a posteriori" probability that the urn from which we have actually drawn the n balls, ν *of which turn out to be white, will have the composition* $(\alpha, \boldsymbol{a}-\alpha)$*? Bayes' theorem provides the desired answer*

$$P_\alpha = \frac{\omega_\alpha p_\alpha}{\omega_0 p_0 + \omega_1 p_1 + \cdots + \omega_{\boldsymbol{a}} p_{\boldsymbol{a}}}.$$

Castelnuovo is incompatible with the described procedure. In fact, Castelnuovo's criticism does not concern the mathematical validity (from the point of view of the Laplace definition) of Bayes' theorem, but rather its application in statistics which, consistently with the empirical point of view, requires the replacement of the unique urn from which balls are drawn with a series of fictitious urns, among which we imagine that the one in question has been chosen. With reference to this, Castelnuovo maintains (CdP, pp. 199-200)

> *La illegittimità della sostituzione appare più chiara se, in luogo dell'estrazione a sorte, ricorriamo a qualche altro semplice giuoco di azzardo.*
> *Una moneta è lanciata* 100 *volte; testa appare* 45 *volte; qual'è la probabilità che la moneta sia "giusta" o favorisca "croce"? – Imitando il procedimento di poc'anzi dovremmo ritenere la moneta scelta a caso fra un numero di monete,* 9 *ad es., le quali presentino le probabilità* 1/10, 2/10,..., 9/10 *di lasciare apparire una determinata faccia; e in base al risultato dell'esperienza dovremmo decidere la probabilità a posteriori che la scelta sia caduta sulla prima, seconda, ... od ultima di queste monete. Ora, per il modo come le monete sono costruite, sappiamo che le ipotesi che si scostano sensibilmente dalla media sono assolutamente da scartare, ed è quindi assurdo prenderle in considerazione.*
> *Il procedimento è certo difettoso e non merita fiducia. Ma è giusto riconoscere che la forma stessa dell'enunciato del problema lo suggerisce; le critiche vanno dunque dirette all'enunciato che è indeterminato, anzi, direi quasi, privo di senso. Chiedere la probabilità di un evento (che qui è la composizione di un'urna) non ha senso, sotto l'aspetto logico, se non si precisano le classi degli eventi possibili e degli eventi favorevoli, delle quali classi deve far parte l'evento in questione. La composizione dell'urna non ha di per sé una probabilità; l'acquista soltanto quando venga messa in rapporto con altre composizioni possibili che dovranno riguardarsi come ugualmente probabili o no. Tutto ciò l'enunciato deve dichiararlo esplicitamente; altrimenti il problema non è determinato.*
> *Alla stessa conclusione arriviamo se, ricorrendo al concetto empirico di probabilità (o, meglio, frequenza), ci proponiamo di determinare sperimentalmente le probabilità delle varie composizioni di urne, dalle quali sortono, in cento estrazioni,* 45 *palle bianche. Dovremmo operare sopra un certo numero di urne contenenti ciascuna ad es.* 10 *palle, tra bianche e nere, in proporzioni diverse. Ripetuta poi un numero grandissimo di volte la serie di cento estrazioni dalle varie urne, dovremmo tenere conto solo di quei sorteggi che hanno dato* 45 *palle bianche, ed esaminare quante volte uno di questi sorteggi sia avvenuto da un'urna con* 5 *palle bianche e* 5 *nere, o* 4 *palle bianche e* 6 *nere, ecc. È evidente che la risposta sperimentale dipenderà dal numero delle urne aventi la prima, o la seconda ... composizione.*

> *Anche di qua vediamo dunque che il problema non ha senso finchè non si forniscano altri dati intorno all'urna donde vengono fatte le estrazioni*[8].

At the end of this first taste of Castelnuovo's way of thinking about inductive logic, there is no disguising the fact that it is deeply affected by the reduction of "mathematical" probability to the Laplace philosophical stance on probability that, by raising a mere method of evaluation to the status of definition, is open to dispute because of its patent circularity. It is worth recalling these obscurities have been at the origin of the reaction of the young Bruno de Finetti (1906-1985) to the empirical approach, by inventing his subjectivistic theory [25] and indicating how to connect the two approaches through the concept of exchangeable trials [24]. One of the most serious, even if expected at this stage, statements contained in Section 79 of CdP, is the denial of validity of the deduction, through the Bayes theorem, of the (conditional) probability of future events given the frequency of success in a certain number of past trials. Even in this case, the Castelnuovo argument is heavily affected by the fact that the correct answer has not a sensible "representation" within the Laplacian setting, where the concept itself of conditional probability turns out to be deprived of its novel and stimulating conceptual content. From an objectivistic viewpoint, a criticism of Castelnuovo's position on the meaning of the theorems of Bernoulli and Bayes and on their uses as tools of inductive logic came from Corrado

[8] *The illegitimacy of the substitution appears more comprehensible in relation to some other simple game of chance, different from the drawing of balls from an urn. A coin is tossed one hundred times and Heads appears 45 times; what is the probability that the coin is fair or, on the contrary, favours Tails? Imitating the procedure described above, we should act as though the coin were chosen at random among a certain number of coins, for example 9, and as though* $1/10, 2/10, \ldots, 9/10$ *were the probabilities of obtaining Heads on one toss of the first, the second, ..., the 9th coin. Moreover, basing ourselves on the result of this experience, we should assess the posterior probability of choosing the first, the second, ..., the last of these coins... The procedure is certainly defective and unreliable. But we must recognize that the terms of the problem themselves suggest the adoption of this procedure; hence, the terms have to be criticized since they are undetermined and do not make sense. To assess the probability of an event does not make sense from a logical point of view, unless we determine the classes of possible and favourable events containing the event at issue...*

We reach the same conclusion if, according to the empirical concept of probability – it would be better to speak here of frequencies – we want to experimentally determine the probabilities of the different compositions of urns from which we obtain 45 white balls through one hundred drawings. Through our procedure, we should choose a certain number of urns, each one of which contains, for example, 10 balls among white and black ones, in different proportions. After repeating many times the sequence of one hundred drawings from each urn, we should consider those drawings, only, which have produced 45 white balls and we should consider the number of drawings of this type from an urn with 5 white balls and 5 black balls, or 4 white and 6 black, etc... It is obvious that the experimental answer depends on the number of urns endowed with a given composition. Examining this point, we observe that the problem does not make sense if we do not provide further data about the urn from which balls are drawn.

Gini (1884-1965); see, *e.g.*, [36]. In view of the discussion recalled just now, Castelnuovo concludes that the study of the hypotheses (causes) in experimental science must be carried out in a different way from that based, like in the Bayes-Laplace theory, on *a posteriori* probabilities. In order to find causes starting from sets of experimental results, he suggests that one should choose appropriate probabilistic structures, considered as candidates to describe the phenomenon which generates the experimental results at issue. Subsequently, on the grounds of suitable criteria, he suggests verification whether the observed results turn out to be more or less likely in conformity with any of the chosen probabilistic structures. Within this framework, Castelnuovo deals with the *theory of* (statistical) *dispersion*[9]. As an application of this way of thinking, he provides a systematic explanation of this theory in Chapter X of CdP, according to the original writings of Wilhelm Lexis (1837-1914), Ladislaus Bortkiewicz (1868-1931), etc. With a view to a better understanding of the Castelnuovo standpoint on inductive logic, it is worth recalling the essentials of the theory of dispersion by resorting to a common simple example of statistical analysis. Take the table of deaths registered in Italy during n successive years. Let m_ν denote the number of people subject to observation in the ν-th year, and x_ν stand for the number of deaths in that group of people, $\nu = 1, 2, \dots, n$. Assuming, for the sake of simplification, that all m_ν are equal, namely $m_\nu = m$ for every ν, one wishes to check the hypothesis $(H) =$"The events *death* associated with the $n \cdot m$ individuals are *stochastically independent* with a *common* unknown probability q", against observations. The theory of dispersion suggests one considers the *Lexis ratio* defined by

$$L := \frac{n}{n-1} \frac{nm-1}{nm} \frac{s^2}{a(1-a/m)}$$

where $a := \sum_{\nu=1}^{m} x_\nu/n$, $s^2 := \sum_{\nu=1}^{m}(x_\nu - a)^2/n$. Assumption (H), combined with the central limit theorem, entails that for large m the value of L turns out to be close to $1 - \sqrt{(2/n)}$. In view of this statement, whose validity essentially rests on (H), the inductive argument followed by Castelnuovo is that (H) ought to be rejected if L turns out to be decidedly smaller or greater than 1, *i.e.*, in other words, if statistical data have *subnormal* or *supernormal* dispersion respectively.

It should be noticed that the argument appears to be incomplete in view of the vagueness of the criterion. Some progress, from the Castelnuovo standpoint, could be made by recalling that: "Under (H) and for large m, the distribution of $(n-1)L$ is approximatively a χ^2_{n-1}-distribution with $(n-1)$ degrees of freedom". Indeed, such a result could be used in this way: Given α in $(0, 1)$, determine ε_1, ε_2 so that $0 < 1 - \varepsilon_1 < 1 < 1 + \varepsilon_2$, and $Prob\{(1 - \varepsilon_1) < \chi^2_{n-1}/(n-1) < 1 + \varepsilon_2\} \geq 1 - \alpha$, and reject (H) whenever $L \notin ((n-1)(1-\varepsilon_1), (n-1)(1-\varepsilon_2))$. It would

[9] Translation of *teoria della dispersione*.

be interesting to investigate the causes of this incompleteness in the Castelnuovo argument: Is it due to lack of knowledge of the above-mentioned asymptotic result, which went back to Friedrich R. Helmert (1843-1917)? Or did Castelnuovo consider such a kind of completion inconsistent with his way of thinking about the entire subject? Apart from this aspect, he carefully considers also the cases of *Poisson trials* (*i.e.*, assuming that $m_\nu = m$ for every $\nu = 1, \dots, n$, the probability of death for the i-th individual in each group is p_i, $i = 1, \dots, m$) and that of *Lexis trials* ($m_\nu = m$ for $\nu = 1, \dots, n$, the probability of death for each individual in the ν-th group is q_ν, $\nu = 1, \dots, n$). In general, in these two cases the use of the normal approximation to find the law of the corresponding Lexis ratio would give rise to more delicate questions. In particular, Cantelli [8] carefully studied the problem to conclude that the aforesaid Gaussian approximation cannot be justified, in the case of Lexis trials, by theoretical probabilistic arguments, because of the lack of stochastic independence between certain random elements.

From this brief account of the explanation of the theory of dispersion, Castelnuovo provides in Chapter X of CdP, one can grasp a few features of his attitude towards statistics: First, the careful consideration given to statistical issues seen both as natural field of application of probability, and as fruitful source of new problems for the theory of probability. Second, distrust of inductive and statistical reasoning based on posterior probability distributions (Bayesian methods), to be replaced with procedures depending on the assumed statistical model only, so that one may reject a model (hypothesis) when subsequent empirical observations turn out to be consistent with some fact to which the model under exam assigns a very low probability. Third, the already emphasized lack of any reference to methods proposed to remedy the vagueness of the formulation of the previous point. Indeed, when CdP was released, remedies consistent with the Castelnuovo standpoint, were already available. This, in particular, alludes to the ideas of test of significance, confidence interval, etc. developed by Ronald A. Fisher (1890-1962) between 1915 and 1925. See, for example, [32, 33]. At least four Fisher's papers appeared in the Gini's journal *Metron* between 1921 and 1925. They are all strictly related to the subject and have played an important role in the development of non-Bayesian statistical inference. See, *e.g.*, [31] and [34].

To the consideration given by Castelnuovo to statistics one can ascribe the careful explanation of the *law of errors of observation* he makes, *e.g.* in Chapters XI and XII of CdP, and to the method of least squares. The presentation of the subject is split into two parts: The former is devoted to the deduction of the Gaussian distribution as law of errors, from two different sets of hypotheses, duly discussed and compared. The latter deals with technical problems of estimation of the parameters of that distribution, but without any hint on the properties of the estimators seen as functions of random samples. As to the deduction of the law of errors, the first statement fits the Carl F. Gauss (1777-1855) approach to the problem, after whom the most famous probability distribution is named. See [35]. Gauss views the unknown value of a given physical quantity as a random number $\tilde{\theta}$, and considers the values from n observations of that quantity as random numbers $\tilde{X}_1, \tilde{X}_2, \ldots$ that, conditionally on $\tilde{\theta}$, are stochastically independent and identically distributed according to a common probability density function $f_{\tilde{\theta}}(x) = g(x - \tilde{\theta})$, $x \in \mathbb{R}$. Then, if q stands for the probability density function of $\tilde{\theta}$,

$$q_{x^{(n)}}(\theta) = \frac{\prod_{i=1}^{n} g(x_i - \theta) q(\theta)}{\int_{\mathbb{R}} \prod_{i=1}^{n} g(x_i - \theta) q(\theta) d\theta} \qquad (\theta \in \mathbb{R})$$

turns out to be the conditional density of $\tilde{\theta}$, at θ, under the hypothesis that $x^{(n)} = (x_1, \ldots, x_n)$ is the observed value of $(\tilde{X}_1, \ldots, \tilde{X}_n)$. Gauss assumes, first, that q is constant (improper density) and, second, that $q_{x^{(n)}}$ is maximum at $\theta = m_n := \sum_{i=1}^{n} x_i/n$, for every $x^{(n)}$ in $\mathbb{R}^n$ and any number n of observations. On the basis of these assumptions, combined with some further regularity conditions, one deduces that the probability density of the error $(\tilde{X}_i - \tilde{\theta})$, under the assumption that $\tilde{\theta}$ is the "true" value, is

$$g(z) = \frac{1}{\sigma\sqrt{(2\pi)}} \exp\Big\{ -\frac{z^2}{2\sigma^2} \Big\} \qquad (z \in \mathbb{R})$$

σ being a strictly positive parameter depending on the accuracy of the measurement process. Castelnuovo, in addition to a generically negative comment on the recourse to an *a posteriori* probability in the previous argument, criticizes the assumption that $q_{x^{(n)}}$ reaches its global maximum at m_n. In fact, even if, on the one side the posterior distribution of $\tilde{\theta}$ tends to concentrate around the point at which $\theta \mapsto \prod_{i=1}^{n} g(x_i - \theta)$ reaches its absolute maximum value, under rather general conditions, on the other side the hypothesis that such a point coincides, in any case, with m_n seems to be too restrictive. Thus, our Author advocates a different argument to explain the form of the law of errors, namely, the argument derived from the meditations of Laplace on the subject starting from 1780 [47] until the publication of *Théorie Analytique* in 1812, [48]. Any error is seen as a random element whose determination (a real number) depends on the effects of a large number of causes, the effect of each cause being "uniformly negligible". An important assumption is that these causes act independently of each other. Thus, any error turns out to be a sum of a large number n of stochastically independent random numbers, whose distributions are constrained to realize the aforesaid condition of uniform negligibility. It is then clear that, to speak of law of errors, one has to guarantee that the distribution of the above sum converges in a proper sense[10], which corresponds to the central li-

[10]Here convergence is meant in the sense of weak convergence of probability measures.

mit problem. Castelnuovo notices that this problem admits a solution under rather general conditions. More specifically, he refers to the hypotheses involved in the various forms of the central limit theorem, mentioned also in the previous section of the present paper, so that the desired limit exists and is given by the Gauss law. There is no doubt that the Laplace approach to the characterization of the law of errors was superior, from a conceptual standpoint, to the Gauss approach, but much more demanding from a mathematical viewpoint. Laplace succeeded in indicating a method (integral transforms of probability laws) to move towards a solution, but he succeeded in overcoming certain technical difficulties only for rather particular cases. Moreover, even in these cases, he was unable to provide a closed form of the limit. As to the explanation given by Castelnuovo, it should be noticed that it represents a progress with respect to Laplace's, even if the forms of central limit theorem he considers are far from being definitive. In fact, the complete solution of the central limit problem was achieved just in the Forties of the last century.

Castelnuovo extends the Laplace characterization to the law of a vector of errors. From a mathematical viewpoint the extension is straightforward, but the description of the applications is so fresh and elegant that Chapter XII of CdP, which deals with the subject, forms, in the opinion of the present author, one of the most accomplished parts of the entire treatise.

4 – Kinetic theory of gases

Castelnuovo was convinced that the raising interest in the theory of probability might derive from some renowned physical applications, especially from the works of James C. Maxwell (1831-1879) on the kinetic theory of gases. Consequently, he devoted the last chapter of every edition of CdP to the formulation of the velocities distribution for Maxwellian molecules, by paying a particular attention to the attempts to prove it and to the role of probability in the proof's strategy. Looking at the chapter devoted to the applications to physics in each of the three editions of CdP, one draws the impression of a deep and restless meditation that lasted a few decades. In fact, whilst in the second edition, the Author looks satisfied with the work made to put the proof of the law of Maxwell in order, in the third[11] he looks at that work with a so strongly critical eye that he completely rewrites the chapter, adhering to a quick, heuristic argument employed, *e.g.*, in *Moderne Physik* (1933) by Max Born (1882-1970). In any case, Castelnuovo starts by recalling the primitive Maxwell's hypothesis according to which the components of velocity of a molecule in three orthogonal directions are seen as independent and identically

[11]The third edition was planned in two volumes, but only the first was published (1947). See [16]. It consists of the entire Vol.I of the second edition (slightly modified) and of a remaking of the chapter dedicated to physics.

distributed random numbers and the phase space is isotropic (the probability density at (u, v, w) of the above three random numbers depends only on the energy $V^2 := u^2 + v^2 + w^2$). These conditions lead immediately to a functional equation whose unique solution is given by the density function

$$\mu(u, v, w) = ae^{-b(u^2+v^2+w^2)}$$

namely, the *Maxwell law.* Castelnuovo observes that Maxwell, unsatisfied of this explanation, proposed to deduce the above expression from the law of the elastic collisions between balls (molecules). A few years after, the new Maxwell proof was reconsidered by Ludwig E. Boltzmann (1844-1906) to prove that the Maxwell law is necessary (not only sufficient) to equilibrium and, then, to deduce that the gas tends to converge to equilibrium after every action made to remove it from that position. The argument used by Boltzmann to state this second part was criticised, especially because it aimed at justifying an irreversible process on the basis of reversible principles, such as the postulates of dynamics. Boltzmann and his disciples tried to maintain the argument based on probability, to prove that convergence to equilibrium has to be understood as a random event with a very high probability. Over time, also Castelnuovo, once again prisoner of his narrow concept of probability, became convinced of a sort of inadequacy of the Boltzmann viewpoint, so that he abandoned the idea of presenting it in the third edition of CdP, on the contrary of what he had made in the second one[12]. He opts for a simplified presentation, still considered in some modern textbooks, which however obscures the main issue under discussion. The sole principle one draws from mechanics is conservation of energy. Then, one indicates the possible values of the energy of a particle, belonging to a gas formed of a (large) number N of particles, by $E_1, \dots, E_n$. A particle is said to be in the state k when it has energy E_k. The state of the gas is characterised by giving the *occupation numbers* $N_1, \dots, N_n$, where N_k stands for the number of particles having the energy E_k $(k = 1, \dots, n)$. By assuming that p_k is the probability of a particle being in the state k, and that the states of the particles are independent, the probability $\pi(N_1, \dots, N_n)$ of the aforesaid state of the gas is given by

$$\pi(N_1, \dots, N_n) = \frac{N!}{N_1! \dots N_n!} p_1^{N_1} \dots p_n^{N_n} \tag{4.1}$$

with $N = N_1 + \cdots + N_n$, that is a Maxwell-Boltzmann statistic. At this stage, one gives grounds for a further assumption, namely, $p_1 = \cdots = p_n = p$ and assumes that

[12]It is worth recalling, at this stage, both a famous paper [40] and a book [41] by Mark Kac (1914-1984), where a *stochastic* model is introduced in the belief it could be a way to understanding the spatially homogeneous Boltzmann equation, especially as far as convergence and rate of convergence to equilibrium of its solutions are concerned. See also [30].

the stationary state of the gas is the one at which (4.1) takes its maximum value, under the condition of the conservation of total energy E, *i.e.*, $\sum_{k=1}^{n} N_k E_k = E$. It turns out that the most probable state $(\overline{N}_1, \ldots, \overline{N}_n)$ satisfies

$$\frac{\overline{N}_k}{N} \sim \alpha e^{-\beta E_k} \qquad (k = 1, \ldots, n)$$

as $N \to +\infty$, α and β being positive constants chosen so that the condition of conservation of energy is fulfilled. Now, since the number of particles, whose coordinates (u, v, w) belong to a cell, increases proportionally with the volume of the cell (provided that the variation is sufficiently small), and the kinetic energy reduces to $u^2 + v^2 + w^2$ therein, then the previous statement leads to: "In the stationary state, the number of molecules of a gas having components of velocity in the rectangle $(u, u+du) \times (v, v+dv) \times (w, w+dw)$ obeys the Maxwell law $ae^{-b(u^2+v^2+w^2)}$, asymptotically, as $N \to +\infty$".

Stated this result, in the aim of illustrating the role played by probability in the deduction of the Maxwell law, Castelnuovo completes the chapter with a number of interesting considerations on problems in mechanical statistics.

5 – Castelnuovo and new generations of probabilists

On the occasion of his last lecture at the Università "La Sapienza" in Rome, after deep expressions of thanks to Leonard J. Savage (1917-1971), de Finetti declared "In tema di ringraziamenti, dovrei naturalmente aggiungere moltissimi altri nomi, anche se, per ovvie ragioni, mi limito a quelli di tre illustri Colleghi che, pur non condividendo la mia posizione (ma appunto perciò ne faccio loro maggior merito) mi hanno aiutato e dato occasione di esporre le mie idee nelle sedi più qualificate: sono Guido Castelnuovo, Maurice Fréchet, Jerzy Neyman"[13]. See [29]. This confirms the Castelnuovo broad mindedness that characterized, *e.g.*, his undertaking to present and recommend for publication a number of notes and memoirs, by different authors, at the *Accademia dei Lincei* and other prestigious mathematical journals, especially between 1916 and 1932. It is interesting to hint at these works, since most of them can be seen as important stages in the development of modern probability. The present discussion focuses on the works of three authors: Cantelli, de Finetti and Andrej N. Kolmogorov (1903-1987). In Section 3 a hint has been given to a paper by Cantelli [8] on the Lexis dispersion. It was presented by Castelnuovo at the

[13] On the subject of acknowledgements, I should of course have to mention many other names; for obvious reasons, however, I shall limit myself to those three Colleagues who, although not sharing my opinions (but their merit is all the greater because of this), have helped me and given me the opportunity of presenting my ideas in the right quarters, that is, Guido Castelnuovo, Maurice Fréchet and Jerzy Neyman.

Lincei in the form of memoir. In his attempt to determine conditions under which one could justify a certain Gaussian approximation, generally assumed as true by previous authors, Cantelli ran up against the integral equation

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2+itx-t^2 f(x)/2} dx = e^{-t^2(1+\overline{f})/2} \qquad (t \in \mathbb{R})$$

where f is an unknown measurable, positive function, and $\overline{f} := \int_{\mathbb{R}} f(x) e^{-x^2/2} dx / \sqrt{2\pi}$. The Gaussian approximation at issue is valid only if there is a nonconstant solution of the integral equation, included in a suitable class of functions determined by the theory of dispersion. Cantelli was able to prove non-existence of such functions, thus disproving one of the mainstays of the Lexis theory of dispersion, but at the same time he drew the attention to the problem of solving the equation without restrictions on f. It is just in a paper of last year [42] that one can finally find the proof that the Cantelli equation admits nonconstant measurable solutions, after a number of attempts by Cantelli himself, Giuseppe Ottaviani, Giovanni Sansone, Paolo Tortorici, Francesco G. Tricomi, etc., between 1918 and 1967. As far as Cantelli is concerned, a merit of Castelnuovo was that of having given Cantelli the opportunity of presenting his basic and original results on the strong law of large numbers between 1916 and 1917. See [5, 6, 7].

Passing to de Finetti, suffice it to notice that Castelnuovo kept a close eye[14] on the draft of certain important de Finetti's papers, before presenting them at the Lincei. Among them: First, the fundamental, extended memoir (1930) on exchangeable events, and three later notes (in 1933) on the extension of the representation theorem from random events to random numbers. See [24] and [28]. Second, the absolutely novel six notes on the definition and basic properties of the law of a process with independent and homogeneous random increments, between 1929 and 1931[15]. See [19-23, 26].

In 1932, Castelnuovo presented two notes by Kolmogorov concerning the representation of the Fourier transform of an infinitely divisible distribution with finite variance. See [45, 46]. This renowned result simply specifies a previous theorem by de Finetti [23, 26] under the condition of σ-additivity of the probabilities involved [a condition that is not necessary to define admissible (coherent) probabilities in the sense of de Finetti]. In 1929, Castelnuovo presented a Kolmogorov note on the law of large numbers [43] and the note [44] on the representation of associative means of a discrete probability distribution, later extended to arbitrary distributions in the paper [27] by de Finetti.

[14]The present author has had the chance of reading a few letters − shown him by Fulvia de Finetti − of the correspondence between Castelnuovo and de Finetti in the years 1929-1931.

[15]In one of the letters mentioned in the last footnote, Castelnuovo makes plain his distrust of probabilities attached to any analytical property of a function.

To conclude these notes on the work of Castelnuovo in supporting the introduction of probability into the Italian official culture, one here summarizes the obituary notice written in *Giornale dell'Istituto Italiano degli Attuari* (1952). The anonymous author (likely, Cantelli as Editor in chief) states that, "Alla iniziativa di G. Castelnuovo si deve la realizzazione della Scuola di Scienze Statistiche e Attuariali presso l'Università di Roma.

Fin dal 1915 la Facoltà di Scienze Matematiche Fisiche e Naturali aveva accolto, per il Suo grande interessamento, i due insegnamenti di Calcolo delle probabilità e di Matematica Attuariale.

Nel 1920 questi due Corsi, sempre per il Suo interessamento, vennero sovvenzionati dalla Fondazione Marco Besso e costituirono il primo germe della Scuola sopra accennata, la cui istituzione ufficiale fu resa possibile dalla riforma Gentile del 1923; così fu istituita, nell'ottobre 1927, la indicata Scuola di cui Egli fu il primo Preside. [...omissis...] Ai corsi di Calcolo delle probabilità e di Matematica attuariale vennero a perfezionarsi, in particolare, giovani matematici di scuole estere, tra i quali alcuni divenuti notissimi nel campo internazionale per i loro contributi apportati al Calcolo delle probabilità e alle sue applicazioni.

Questa Scuola fu poi, nel 1935, abbinata alla Scuola speciale di Statistica"[16] [headed by Corrado Gini] and from this coupling (of the School of Statistics and Actuarial Sciences with the Special School of Statistics) the *Faculty of Statistics* of Rome originated[17].

Acknowledgements

I am grateful to the organizers of the meeting "*Castelnuovo: un ricordo a 150 anni dalla nascita*" (Roma, 5th November 2015) for inviting me to give a lecture. I am deeply indebted to Antonio Lijoi, Giovanna Nappo, Eleonora Perversi for their invaluable comments, and to Maurizio Cornalba for helping me to orient myself in the four volumes of the Castelnuovo *Opere Matematiche*. Finally, I have to

[16] *Thanks to the initiatives of G. Castelnuovo, the School of Statistics and Actuarial Sciences at the University of Rome was established. Starting from 1915 the Faculty of Mathematics, thanks to his great interest, introduced teachings of Probability and Actuarial Mathematics. In the year 1920, the courses on his initiative again, were subsidized by the Besso Foundation and represented the germ of the above-mentioned School, which was officially instituted thanks to the 1923 reform named after the Ministry of Education of the time, Giovanni Gentile. Castelnuovo was the first head of that school when, in 1927, it was made active.* [...omissis...] *Young mathematicians from foreign schools attended the courses of Probability and Actuarial Mathematics in order to improve their knowledge of these subjects. In 1935, the School was linked with the Special School of Statistics.*

[17] It merged, in 2010, with other existing courses to form a new faculty.

thank Fulvia de Finetti for indicating me the letters of Castelnuovo to her father Bruno.

REFERENCES

[1] J. Bernoulli: *Ars Conjectandi*, Thurnisiorum, Basel, 1713.
[2] J. Bertrand: *Calcul des Probabilités*, Gauthier-Villars, Paris, 1889.
[3] E. Borel: *Éléments de la théorie des Probabilités*, Hermann, Paris, 1909.
[4] F. P. Cantelli: *La tendenza ad un limite nel senso del calcolo delle probabilità*, Rendiconti del Circolo Matematico di Palermo, **41** (1916), 191–201.
[5] F. P. Cantelli: *Sulla legge dei grandi numeri*, R. Accad. Naz. Lincei Mf S 5ª, **11** (1916), 329–350.
[6] F. P. Cantelli: *Sulla probabilità come limite della frequenza*, R. Accad. Naz. Lincei Rf S 5ª, **26** (1917), 39–45.
[7] F. P. Cantelli: *Su due applicazioni di un teorema di G. Boole alla statistica matematica*, R. Accad. Naz. Lincei Rf S 5ª, **26** (1917), 295–302.
[8] F. P. Cantelli: *Sullo schema lexiano della dispersione ipernormale*, R. Accad. Naz. Lincei Mf S 5ª, **12** (1918), 395–411.
[9] F. P. Cantelli: *Sulla oscillazione delle frequenze intorno alla probabilità*, Metron, **3** (1923), 167–174.
[10] G. Castelnuovo: *Calcolo delle Probabilità* (I ediz.) Soc. ed. Dante Alighieri, Milano, 1919.
[11] G. Castelnuovo: *Calcolo delle Probabilità* (II ediz.) [Vol. I, 1926; Vol. II, 1928] Zanichelli, Bologna, 1926-1928.
[12] G. Castelnuovo: *Recensione a É. Borel*, 1925; *Principes et Formules*, É. Borel, 1925; *Mécanique statistique*, P. Lévy, 1925; *Calcul des Probabilités* Scientia, **42** (1927), 291–293.
[13] G. Castelnuovo: *Sul problema dei momenti*, Giorn. Ist. Stat. Att., **1** (1930), 137–169.
[14] G. Castelnuovo: *Sur qualques problèmes se rattachant au calcul des probabilités*, Ann. Inst. Henri Poincaré, **2** (1932), 465–490.
[15] G. Castelnuovo: *Recensione a R. v. Mises*, 1931; *Wahrscheinlichkeitsrechnung*, Scientia, **53** (1933), 427–428.
[16] G. Castelnuovo: *Calcolo delle Probabilità*, III ediz., Vol. I Zanichelli, Bologna, 1947.
[17] G. Castelnuovo: *Opere Matematiche*, 4 Voll., Accad. Naz. Lincei, Roma, 2004.
[18] P. L. Chebyshev: *Oeuvres*, publiés par les soins de MM A. Markoff et N. Souin, Commissionaires de l'Acad. Impériale des Sciences, St-Pétersburg, 1899-1907.
[19] B. de Finetti: *Sulle funzioni a incremento aleatorio*, R. Accad. Naz. Lincei Rf S 6ª, **10** (1929), 163–168.
[20] B. de Finetti: *Sulla possibilità di valori eccezionali per una legge ad incrementi aleatori*, R. Accad. Naz. Lincei Rf S 6ª, **10** (1929), 325–329.
[21] B. de Finetti: *Integrazione delle funzioni a incremento aleatorio*, R. Accad. Naz. Lincei Rf S 6ª, **10** (1929), 548–552.
[22] B. de Finetti: *Problemi determinati e indeterminati nel calcolo delle probabilità*, R. Accad. Naz. Lincei Rf S 6ª, **12** (1930), 367–373.

[23] B. de Finetti. (1930) Le funzioni caratteristiche di legge istantanea. *R. Accad. Naz. Lincei* Rf S 6ª **12** 278-282.
[24] B. de Finetti: *Funzione caratteristica di un fenomeno aleatorio*, R. Accad. Naz. Lincei Mf S 6ª, **4** (1930), 86–133.
[25] B. de Finetti: *Sul significato soggettivo della probabilità*, Fund. Math., **17** (1931), 298–329.
[26] B. de Finetti: *Le funzioni caratteristiche di legge istantanea dotate di valori eccezionali*, R. Accad. Naz. Lincei Rf S 6ª, **14** (1931), 259–265.
[27] B. de Finetti: *Sul concetto di media*, Giorn. Istit. Ital. Attuari, **2** (1931), 369–396.
[28] B. de Finetti: *Classi di numeri aleatori equivalenti. La legge dei grandi numeri nel caso dei numeri aleatori equivalenti. Sulla legge di distribuzione dei valori in una successione di numeri aleatori equivalenti (Three articles)*, R. Accad. Naz. Lincei Rf S 6ª, **18** (1933), 107–110; 203–207; 279–284.
[29] B. de Finetti: *La probabilità: guardarsi dalle contraffazioni!*, Scientia, **111** (1976), 283–303. [English translation: *Probability: beware of falsifications!* in Studies in Subjective Probability, H. E. Kyburg, Jr. and H. E. Smokler (eds.), Krieger Publ. C., Huntington), 1980].
[30] E. Dolera – E. Regazzini: *Proof of a McKean conjecture on the rate of convergence of Boltzmann equation solutions*, Probab. Theory Related Fields, **160** (2014), 315–389.
[31] R. A. Fisher: *On the probable error of a coefficient of correlation deduced from a small sample*, Metron, **1** (1921), 3–32.
[32] R. A. Fisher: *On the mathematical formulation of theoretical statistics*, Philosophical Transactions of the Royal Society of London, S.A **222** (1922), 309–368.
[33] R. A. Fisher: *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 1925.
[34] R. A. Fisher: *Applications of "Student's" distribution*, Metron, **5** (1925), 90–104.
[35] C. F. Gauss: *Motus Corporum Celestium*, Friedrich Perthers and I.H. Besser, Hamburg, 1809.
[36] C. Gini: *Concept et mesure de la probabilité*, Dialectica, **3** (1949), 36–54.
[37] H. Hamburger: *Über eine Erweiterung des Stieltjesschen Momenten Problems*, Teil I Math. Annalen, **81** (1920), 235–319.
[38] H. Hamburger: *Über eine Erweiterung des Stieltjesschen Momenten Problems*, Teil II Math. Annalen, **82** (1920), 120–164.
[39] H. Hamburger: *Über eine Erweiterung des Stieltjesschen Momenten Problems*, Teil III Math. Annalen, **82** (1920), 168–187.
[40] M. Kac: *Formulations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955 III 171-197, Univ. of California, Berkeley.
[41] M. Kac: *Probability and Related Topics in Physical Sciences*, Interscience, London-New York, 1959.
[42] V. Kleptsyn – A. Kurtzmann: *A counterexample to the Cantelli conjecture through the Skorokhod embedding problem*, Ann. of Probab., **43** (2015), 2250–2281.
[43] A. N. Kolmogorov: *Sur la loi des grands nombres*, Atti R. Acc. Lincei, Serie VI, Rendiconti, **9** (1929), 470–474.
[44] A. N. Kolmogorov: *Sur la notion de la moyenne*, Atti R. Acc. Lincei, Serie VI, Rendiconti, **12** (1930), 388–391.
[45] A. N. Kolmogorov: *Sulla forma generale di un processo stocastico omogeneo, Un problema di bruno de Finetti*, Atti R. Acc. Lincei, Serie VI, Rendiconti, **15** (1932), 805–808.

[46] A. N. Kolmogorov: *Ancora sulla forma generale di un processo omogeneo*, Atti R. Acc. Lincei, Serie VI, Rendiconti, **15** (1932), 866–869.
[47] P. S. Laplace: *Mémoire sur les probabilités*, Oeuvres, **9** (1780), 383–485.
[48] P. S. Laplace: *Thórie analytique des probabilités*, Courcies, Paris, 1812.
[49] P. Lévy: *Calcul des Probabilités*, Gauthier-Villars, Paris, 1925.
[50] J. W. Lindeberg: *Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung*, Mat. Zeitschr., **15** (1922), 211–225.
[51] J. W. Lindeberg: *Sur la loi de Gauss*, C. R. Acad. Sci. Paris, **174** (1922), 1400–1402.
[52] A. M. Lyapunov: *Nouvelle forme du théorème sue la limite de probabilité*, Notes Acad. Sci. Phys.-Math., Ser 8 **12**, (1901) 1–24.
[53] A. A. Markov: *Démonstration du second théorème du calcul des probabilités*, Bicentenaire de la loi des grandes nombres (1713-1913) Impr. de l'Acad. Impériale des Sciences, St. Pétersburg, 1913.
[54] O. Onicescu: *Le livre de G. Castelnuovo "Calcolo delle probabilità e applicazioni" comme aboutissant de la suite des gronds livres sur les probabilités*, Simposio Intern. di Geometria Algebrica Cremonese, Roma, 1965.
[55] H. Poincaré: *Calcul des Probabilités*, Gauthier-Villars, Paris, 1912.
[56] E. Regazzini: *Teoria e calcolo delle probabilità*, La matematica Italiana tra le due guerre mondiali, Pitagora Ed., Bologna, 1987, 339-386.
[57] E. Regazzini: *Probability theory between the two world wars. A brief historical review*, Metron, **55** (1987), 4–42.

Lavoro pervenuto alla redazione il 4 marzo 2016
ed accettato per la pubblicazione il 29 novembre 2016

Indirizzo dell'Autore:
Eugenio Regazzini – Università degli Studi di Pavia – Dipartimento di Matematica – via Ferrata 1 – 27100 Pavia, Italy
E-mail: eugenio.regazzini@unipv.it