Centre Emile Bernheim
Research Institute in Management Sciences

Solvay
Business School

The Special Status of Mathematical Probability: A Historical Sketch

X. De Scheemaekere and A. Szafarz

The history of the mathematical probability includes two phases: 1) From Pascal and Fermat to Laplace, the theory gained in application fields; 2) In the first half of the 20th Century, two competing axiomatic systems were respectively proposed by von Mises in 1919 and Kolmogorov in 1933. This paper places this historical sketch in the context of the philosophical complexity of the probability concept and explains the resounding success of Kolmogorov's theory through its ability to avoid direct interpretation. Indeed, unlike experimental sciences, and despite its numerous applications, probability theory cannot be tested per se. Rather it relates to practical matters by means of transition hypotheses or bridging principles that match the structure of practical problems with abstract theory. In this respect probability theory has a very special status among scientific disciplines.

JEL Classifications:
Keywords:

CEB Working Paper N° 08/0017
May 2008

Université Libre de Bruxelles – Solvay Business School – Centre Emile Bernheim
ULB CP 145/01 50, avenue F.D. Roosevelt 1050 Brussels – BELGIUM
e-mail: ceb@admin.ulb.ac.be Tel. : +32 (0)2/650.48.64 Fax : +32 (0)2/650.41.88

The Special Status of Mathematical Probability: A Historical Sketch

Xavier De Scheemaekere* and Ariane Szafarz+

Université Libre de Bruxelles

Solvay Business School

Centre Emile Bernheim

May 2008

Abstract

The history of the mathematical probability includes two phases: 1) From Pascal and Fermat to Laplace, the theory gained in application fields; 2) In the first half of the 20th Century, two competing axiomatic systems were respectively proposed by von Mises in 1919 and Kolmogorov in 1933. This paper places this historical sketch in the context of the philosophical complexity of the probability concept and explains the resounding success of Kolmogorov's theory through its ability to avoid direct interpretation. Indeed, unlike experimental sciences, and despite its numerous applications, probability theory cannot be tested per se. Rather it relates to practical matters by means of transition hypotheses or bridging principles that match the structure of practical problems with abstract theory. In this respect probability theory has a very special status among scientific disciplines.

* E-mail: xdeschee@ulb.ac.be

+ Corresponding author. Address: ULB-CEB-CP 145/1, Av. F.D. Roosevelt, 50, 1050 Brussels, Belgium ; Phone: +32.2.650.48.65; Fax: +32.2.650.41.88; E-mail: aszafarz@ulb.ac.be

1 Introduction

The philosophy of probability is a "hot field" (Eagle [2006], p. 773) where the different philosophical arguments are sharply contrasted. In spite of this, there is a remarkable consensus about the mathematical aspects of probability. The philosophical concept of probability has indeed a mathematical hallmark that echoes the success of probability theory. Undeniably, mathematics is the background of most current philosophical debates in this field, but its role as such has not yet been conceptualised. In this paper, we recount the evolution of the theory of probability from Pascal and Fermat to Kolmogorov in order to analyse the evolving status of the mathematical concept of probability. This area is often overlooked by the current interpretations of probability.

Language tells us that probability can describe very different events. Some authors (e.g., van Brakel [1976], pp. 120-1 or Kneale [1949], pp. 1-2) present a profusion of sentences that illustrate the different meanings of the term "probability":

- It will probably rain tomorrow.
- Yesterday, it probably rained in Paris.
- It is probably raining in London.
- The probability is weak that a twenty-year-old man dies within a year.
- The probability that a twenty-year-old man dies within a year is statistically 0,004.
- My game opponent is winning so much that he is probably cheating.
- This statue probably represents Cromwell.
- The probability that the speed of a hydrogen molecule happens to be between v and w is p.
- The probability to win at the lottery is virtually zero.
- The probability that a random chosen integer is even is 1/2.
- The probability to get a double six with two fair dices is 1/36.
- The probability to get a double six with two particular dices is 1/36.

Clearly, this diversity in the use of probabilistic terms is unchallenged in everyday language: Probability relates to the past, the present or the future, to single or collective events, it is objective or subjective, *ex ante* or *ex post*, numerical or not, and it depends on our knowledge, or, on our ignorance or indeed our knowledge of our ignorance. Gillies ([2000], p. 2) observes that most philosophers of probability agree that the various interpretations of probability can be divided into two broad groups, without agreeing on how these groups should be named. Yet, in philosophy, whatever the terminology, probability boils down to a single alternative: It is either epistemic or objective, in a physical non-platonic sense (Howson [1995]). Admittedly, this distinction is not absolute but it remains of crucial importance for understanding the various philosophical theories of probability, which all originate in one or the other aspect of probability (see Gillies [2000]; Galavotti [2005]; Mellor [2005]).

From a mathematical perspective, the dual character of probability – the "Janus-face" of probability, as Hacking ([1975], p. 12) refers to it – is also blatant and it is best characterized as *a priori* versus *a posteriori*. The alternative that has lead to the emergence of the philosophy of probability may not be the same as the one that underpinned the parallel mathematical developments. This nuance implies that the mathematical and philosophical concepts are different. However, the nature of this difference remains to be analysed. In the rest of the paper we examine more precisely, how the mathematical concept of probability tackles the difficult question of the ultimate nature of probability, and the resulting consequences on the relations between the theory, the applications and the philosophy of probability.

In the literature, there are different perspectives on the relation between mathematics and the concept of probability. Occasionally, probability is deemed 'to occupy a peculiar position between the purely mathematical and the obviously scientific', the subject matter of the theory being 'among other things, finite class frequencies, rational degree of belief, limiting relative frequencies, propensities, degrees of logical confirmation and measure on abstract spaces, to name only some of the most important' (Humphreys [1985], p.568). From that angle, the mathematical concept is just another interpretation. More often, the mathematics is implicitly assumed and only referred to as an empty structure waiting for a philosophical interpretation, as it happens paradigmatically (and pragmatically!) in the subjective theory. Some authors also look for the philosophical interpretation that underlies the mathematical theory (e.g., Shafer and Vovk [2006]). However, the mathematical concept of probability is barely considered, and the way it differs from its philosophical counterpart is not explicitly dealt with.

We therefore aim at clarifying the status of the mathematical concept of probability regarding philosophical considerations and emphasize the contrast between, a fruitful and consensual mathematical view, on the one hand, and, its highly controversial epistemological counterpart, on the other.

Section 2 presents the evolution of the mathematical concept of probability from Pascal and Fermat to Laplace. Special emphasis is placed on the various connections between the *a priori* and the *a posteriori* definitions of probability and on the contribution of Bernoulli's theorem. Section 3 displays the divergences between Kolmogorov's and von Mises's axiomatic approaches and the reasons for Kolmogorov's dominance. Section 4 relates this conceptual evolution to the philosophical concept of probability. It suggests that probability theory has the peculiarity of requiring transition hypotheses or bridging principles that match the structure of practical problems and that of the abstract theory. Section 5 provides conclusions.

2 From Pascal to Laplace

The historical origin of probability calculus traces back to the correspondence between Pascal and Fermat in 1654[1]. These two mathematicians analysed some gambling problems according to a general principle: the odds of a game depend on the total number of possible alternatives, which divided into favourable and unfavourable outcomes. In order to count these different outcomes in a practical way, Pascal and Fermat developed what is known today as combinatorial analysis – with Pascal triangle at its centre.

After Pascal and Fermat, some mathematicians like Huygens, Wallis or Schooten further developed the study of gambling games by means of combinatorial calculus, which lead to the advent of Laplace's theory in the 18th century.

[1] Concededly, there were other predecessors such as Cardan, Kepler or Galileo, but most authors agree on the fact that the work of these scientists did not lead to the systematic study and the development of the mathematical theory of probability (see Todhunter, 1865, Hacking, 1975, Szafarz, 1985, Gillies, 2000).

Yet, to understand fully the import of the mathematical concept of probability, another empirically grounded standpoint, must be taken into consideration: the statistical approach, driven by material motives. Statistics about births and deaths, based on population registers, were collected by scientists in mortality tables, so as to measure the damage caused by the plague, death probabilities at each age, life annuities, etc.

This mathematically naïve research provided a first non-idealized field of application for the nascent theory of probability. Indeed, a "demographic" approach to probability necessarily implies the need to count empirically a certain number of events in order to derive statistics. On the contrary, the study of gambling games does not need to refer to any empirical experience for what matters are completely idealized situations.

Interestingly, Pascal's works, *Traité du triangle arithmétique* or *Des combinaisons*, never explicitly mentions the term 'probability', speaking instead of 'chance' or 'hazard'. After Pascal, probabilistic terminology became increasingly rooted in both mathematical and demographical research. There was no explicit preliminary to the various theoretical and empirical results, even if two definitions were implicitly used. On the one hand, the *a priori* probability referred to the ratio (the fraction) of the number of favourable outcomes to the total number of possible outcomes, as long as these are equi-possible. This definition was latent from the time of Pascal and Fermat, but Laplace first expressed it unequivocally. D'Alembert, in an article of the 1754 *Encyclopedia*, 'Croix ou Pile', firmly criticized this formulation, which was already apparent in de Moivre's work in 1718, but only as a property of the probability concept, itself undefined. On the other hand, the *a posteriori* probability of an event referred to the ratio of the number of realizations to the total number of observations.

At first sight, these definitions each have their respective domains of application and they do not share more than a syntactical similitude through the term 'probability'. This separation would have been effective only if theory had been limited to the study of idealized situations (such as gambling games) and only if empirical measures had not been concurrently extrapolated beyond observation. However, this was not the case. In fact quite the reverse situation held. As a result, confusion emerged between both definitions, reinforced by the introduction of probabilistic reasoning in other fields such as astronomy or sociology. Both theorists and practitioners then began to search for a unified concept of probability.

This mission was accomplished in 1713 by the theoretician, Jacques Bernoulli, in a posthumous unfinished work, *Ars Conjectandi*. Bernoulli provided a theorem (the "law of large numbers", as Poisson later called it) that specifies the relation between both approaches. The law of large numbers states the terms in which the *a posteriori* probability can be considered as an approximation of the *a priori* probability[2]. These terms are clear: firstly, the considered event must intrinsically have an *a priori* probability; secondly, the distance between the observed frequency and this *a priori* probability is itself expressed as an *a priori* probability.

The legitimacy of the theorem is precisely delimited. In the case of gambling games, it offers a theoretical justification of the link between *a priori* probabilities and the players' observations of events. Nevertheless, the rapid spreading of the mathematical theory of probability after the publication of *Ars Conjectandi* is to some extent attributable to what is

[2] In modern mathematical language, one would say that the law is about *convergence in probability* of the relative frequency to the *a priori* probability as the number of trials tends to infinity, but a formulation in terms of limit contravenes the author's original intentions.

called the inverse use of Bernoulli's theorem. This inverse use, or misuse, consists of starting from the thesis – proximity between the *a priori* and *a posteriori* probabilities – and deducing the hypothesis – the value of the *a priori* probability – by extrapolation with respect to experimental measurements. This inverse use was highly successful among mathematicians for it made it possible to apply combinatorial calculus in much broader fields than the study of gambling games. This move can also be seen as a appropriation of practice by theory: not only does the theorem state that relative frequency is an approximation of the *a priori* probability, it also allows its inverse use where theoretical results are applied to events where the total number of equi-possible outcomes cannot be evaluated.

Bernoulli's theorem constituted a major historical step. Indeed, it reconciled the different meanings of probability within a unified mathematical structure (see Shafer [1996]). Further, the inverse use of the theorem considerably enlarged the field of applications of probability theory to all repeatable events. This sweeping process started from Pascal and ended up with Laplace.

Accordingly, in the 18th century, many theoretical works – sometimes combining mathematical improvements and metaphysical considerations – developed combinatorial methods in game theory and their applications. Besides Bernoulli, the works of de Moivre, Euler, Lagrange, Condorcet, etc., extended the early contributions of Pascal and Fermat while expanding the application field to the case of infinite equi-possible outcomes.

The *Théorie analytique des probabilities* of Laplace published in 1812 represents the highlight of classical probability theory. For a long time, it embodied the most detailed and complete mathematical and practical exposition. The introduction of the book – later published independently as the *Essai philosophique sur les probabilités* – presents the author's philosophical considerations. Laplace defines probability in terms of equi-possible outcomes and claims that the probabilistic view of the world is only due to our human condition:

> La théorie des hasards consiste à réduire tous les événements d'un même genre à un certain nombre de cas également possibles, c'est-à-dire tels que nous soyons également indécis sur leur existence, et à déterminer le nombre de cas favorables à l'événement dont on cherche la probabilité (Laplace [1840], p. 7).

In this framework, equi-possibility means equi-ignorance and the concept of probability is extended to all imaginable events. Gambling games become a simple example of application to be considered among others. Nevertheless, such games are frequently[3] mentioned in the first part of the book, which is dedicated to the principles. Other practical cases rely on the inverse use of Bernoulli's theorem or even on very different considerations, where it is impossible to determine *a priori* equi-possible outcomes (Laplace, 1840, p. 158). In these situations, Laplace appeals to personal intuition rather than to mathematical rigor. The ambiguity of the term "probability" is critical to Laplace's process for it enabled him to leave out the nature of the probability that is referred to – *a priori*, *a posteriori* or else. This fundamental imprecision was later noted by von Mises ([1938], p. 61).

[3] The game "croix ou pile" appears at pages 12, 16, 18, 19, 23 and 25 of Laplace (1840) and other games (lottery, dices, urns) at pages 7, 9, 13, 15, 19 and 20.

From Pascal to Laplace, probability theory experienced a constant expansion of its fields of application: first, gambling games with Pascal and Fermat, then repeatable events, thanks to the inverse use of Bernoulli's theorem, and finally, according to Laplace, all phenomena that science can study. The following theoreticians not only picked out Laplace's formal errors, but also aspired to strict delimitation of the range of probabilistic events. This ambition was met by the axiomatization process of probability theory.

3 Competing Axiomatic theories

Since the mid-19th century, several mathematicians had been denouncing the pitfalls of Laplace's formalisation. Firstly, the *a priori* nature of his definition confines the legitimate sphere of application to idealized conditions. Secondly, a definition based on the notion of equi-possibility is circular because Laplace appears to use with indifference the terms "equi-possible" and "equi-probable". This criticism was prominently formulated by Reichenbach ([1949], p.353):

> Some authors present the argument in a disguise provided by the concept of equipossibility: cases that satisfy the principle of "no reason to the contrary" are said to be equipossible and therefore equiprobable. This addition certainly does not improve the argument, even if it originates with a mathematician as eminent as Laplace, since it obviously represents a vicious circle. Equipossible is equivalent to equiprobable. […] Even if the degree of probability can be reduced to equiprobability, the problem is only shifted to this concept. All the difficulties of the so-called a priori determination of probability therefore, centre on this issue.

However, equi-possibility was not unanimously rejected. For example, Borel ([1909], p. 16) argues that such circularity is not vicious and that the tautological nature of the term "equi-possibility" does not threaten the definition of probability. Hacking ([1971]) suggests that the confusion originates in the dual nature of probability, which can be objective (physical) or epistemic. The equivocal character of the definition is precisely necessary in the sense that it overlaps with the ambiguous character of probability. Similarly, the axiomatization process of probability theory ended up avoiding the choice of a particular interpretation of probability.

Another objection to Laplace's theory is that the mere formation of equi-possible cases gives rise to problems. According to Laplace, equi-possible outcomes are such that we are in the same state of ignorance concerning their realization. This principle traces back to Pascal, who already remarked that 'hazard is equal' ('*Le hasard est égal*', cited by du Pasquier 1926]). But if probability depends on knowledge, it must be subjective to some extent, as the subjectivists later argued (Ramsey [1931]; de Finetti [1937], and more recently, Jeffrey [2004]).

Additionally, as Kneale ([1949]) repeatedly observes, Laplace's definition requires that we determine with precision the degree of ignorance about an event that we ignore, which is absurd. This conundrum provides the rationale that underlies the various paradoxes of the principle of indifference, and which seriously undermine the logical theory advocated by Keynes ([1921]) and Carnap ([1950]), among others (see Bertrand [1888]; Gillies [2000]).

Facing the drawbacks of classical theory, Hilbert ([1902]) insisted on the necessity of an axiomatization of probability theory. At the same time, a broad Anglo-Saxon empiricist "biometrical" stream was developing, devoted to the elaboration of proper statistical techniques in data analysis. Biometricians were primarily guided by experimental preoccupations (in biology, for example). Even if they did not break the connection between probability theory and statistical applications, a serious shift appeared between their technical considerations and the existing theoretical foundations of probablity.

A new paradigm for probability theory became necessary. Research in that direction ended up in two competing axiomatic systems. One, elaborated by von Mises and enhanced by Wald, was intrinsically descriptive and empirically based, whereas the other, due to Kolmogorov (in the line of Borel's and Frechet's works), was exclusively mathematical by nature.

The theory presented by von Mises ([1919]) leans on the concept of "collective", defined by two axioms. Any series of observations of the same event – or of a mass phenomenon – is a collective if:

I. The relative frequency of that event has a limit value.
II. This limit value remains unchanged if the initial series is replaced by any sub series.

The second axiom enabled von Mises to get rid of deterministic phenomena. It is a definition of randomness, which was later restricted by Wald[4] ([1937]) to a countable number of sub series for consistency.

Starting from the notion of collective, von Mises defines the probability of an event with respect to a collective as the fixed limit of the relative frequency of this event within the collective. This probability is *a posteriori* in the sense that it comes out of an (infinite) number of experiences.

Von Mises' theory was forcefully criticized; in particular during the 1938 colloquium chaired by Fréchet (see Szafarz [1984]). The modification by Wald of the second axiom contributed to a theoretical consolidation as well as to a putting in perspective of the concept of probability (Ville [1939]). As a consequence, a significant part of the empirical roots of von Mises' approach was sacrificed for mathematical rigour.

Furthermore, the first axiom postulates the existence of a fixed limit for the relative frequency, which looks logically embarrassing in some respects. What is the meaning of an infinite sequence of observations regarding the fact that experimental observation is *per se* finite? How should such theoretical results be applied empirically?

In the 17th century, practitioners like Graunt, van Hudden, de Witt or Halley were already using a definition of probability in terms of the observed relative frequency. This approach had the virtue of tackling directly empirical reality. Nevertheless, such a direct and empirical identification destroyed any possibility to establish a proper theory of probability. As a matter of fact, von Mises was well aware of the difficulty to build an axiomatic system on observed

[4] Wald's theory includes as particular cases other axiomatic systems due to Popper, Copeland and Reichenbach.

relative frequencies. Therefore, he proposed an idealization in the form of a limit. Wishing to maintain of sound empirical ground[5], he ended up with a bald mathematical theory.

Aiming at offering to practitioner a tailor-made axiomatic system by using their language ("sequence of observations"), von Mises found himself in a vicious circle. His system prevents dissociating applications from mathematical formalization because any set of observations *de facto* constitutes the first elements of an infinite series (a collective). Paradoxically, in this setting, the empiricist appears intrinsically dependent on the theory. Moreover, at that time, well-developed statistical research provided accurate techniques for descriptive data analysis. As von Mises concedes, these considerations do not need any axiomatic support. So, he claims to interpret the statistician's preliminary results in the light of his own theory while staying true to the idea that statistics is a well-delimited autonomous field (von Mises [1957], p. 167). For all these reasons, practitioners did not take much interest in the frequentist axiomatization.

In addition, since the early 20th century there was a mathematical alternative to von Mises' theory. From Borel and Fréchet to Kolmogorov, a powerful probability theory was developed along the line of the (mathematical) measure theory, able to address a large range of probabilistic issues and containing Laplace's theory as a particular case (Barone and Novikov [1978]; Shafer and Vovk [2006]). Because of its frequentist foundations, von Mises and Wald's theory, on the contrary, could not encompass their classical definition. Indeed, in their setting any *a priori* structure consideration is prohibited, and only an *infinite* sequence of observations could justify that, for instance, when throwing a dice the probability to get a 6 is 1/6.

Beyond the fact that an infinite sequence is pure idealization, frequentist probability exclusively applies to trials repeatable *ad infinitum* in identical conditions. Moreover, if the frequency does not converge – which is possible – one cannot attribute a probability to the event. Several mathematicians mentioned this limitation regarding the applicability of the frequentist theory, as a consequence of the first axiom:

> Nul probabiliste ne se refusera à admettre que les collectifs définis par M. de Mises sont des suites particulièrement intéressantes, qui méritent à ce titre d'attirer très particulièrement l'attention. De même dans la théorie des fonctions, il est bien légitime de s'intéresser surtout aux fonctions dérivables pourvu qu'on ne suppose pas que toute fonction est dérivable (Fréchet [1938], p. 27).

Another drawback of von Mises' system is that it only takes into account ordered observations. But there are always different ways to order one set. Therefore, a given set of observations, ordered in different ways, may lead to different collectives. Kolmogorov's theory does not suffer from any of these problems.

In summary, von Mises' theory was neither descriptive enough for the statisticians, nor powerful enough for the mathematicians. For the practitioners, the combination of mathematical convergence and physical randomness came out onto the impossibility to test experimentally the theoretical results: these cannot be verified *a priori*, because of axiom I, and nor can they be verified *a posteriori*, because of axiom II. From a theoretical perspective, von Mises' axiomatic system, though coherent, had to face a rival which was in the line of

[5] His goal was to elevate probability theory to the level of theoretical physics, as testified by his numerous references to rational mechanics.

Laplace while scrupulously respecting the logico-mathematical standards and which, in addition, came within the scope of a more general formalism.

In 1933, Kolmogorov came up with a fully axiomatic theory in which a probability is simply a normalized measure. This theory reduces, as a particular case, to Laplace's definitions and theorems. It can be seen, according to Fréchet ([1938], p.54) as a 'modernized classical theory'. However, Kolmogorov's axioms do not have the same disadvantages as Laplace's theory concerning, for example, the notion of equi-possible outcomes. Quite the opposite, the analytical nature of Kolmogorov's formalism goes beyond any interpretation.

From a theoretical point of view, Kolmogorov's system was highly satisfactory. It was not only a generalization of the classical theory, but also the source of numerous compelling theorems, which were behind the notions of conditional probability, probability distribution, and stochastic process.

For a long time, statistics remained the leading sphere of application for probability theory. Yet, progressively, Kolmogorov's formalism entered more fields, including quantum theory and economics. Probability theory is now part of many different disciplines, and its status within these has been the subject of important research for the philosophy of probability – consider for example the difference between the concept of probability in quantum physics and in mathematical finance (see De Scheemaekere [2007]). Kolmogorov's mathematical probability has exhibited a remarkable robustness with regard to this enlarged realm of application. This contributed to its success.

However, from a philosophical viewpoint, the relation between this formalism and reality is questionable. Most textbooks on the theory of probability refer to one or the other concept of probability – *a priori* or *a posteriori* – so as to illustrate the reasonableness of the initial axioms. Kolmogorov ([1950], pp. 3-5) himself mentions von Mises' approach as a possible justification for the link between his theory and reality.

Waismann ([1930]) criticizes von Mises for completely misjudging the role of idealization in taking an empirical sequence as a mathematical series and an observed relative frequency as an ideal infinite limit. Unlike von Mises, Kolmogorov does not misjudge the role of idealization, as he establishes a theory, which relies on an axiomatic system that is 'sovereign from a logical perspective and free from any contingency' (Fréchet [1938], p. 45). On the contrary, Kolmogorv's definition is purely mathematical and, hence, independent of any interpretation. In his formalism, empirical verification cannot jeopardize the initial axioms; it can only question the hypotheses (such as independence of the observations) that are assumed by the practitioner so that he can draw on theoretical results.

In brief, Kolmogorov's theory is purely mathematical and it is consequently devoid of immediate descriptive pretensions[6]. As such, it is also beyond the reach of empirical objections and philosophical interpretations. While providing for a large scope of applications, it is fully compatible with modern mathematics, mainly algebra and calculus.

[6] Interestingly, von Mises stigmatized the rival theory as follows: 'Il s'agit d'une sorte de probabilités purement mathématiques, c'est-à-dire des probabilités dont l'objet n'appartient pas au monde physique […] mais à l'arithmétique même'.

Together with the criticisms vis-à-vis von Mises's theory, this adequacy explains the success[7] of Kolmogorov's approach and the subsequent abandoning of the frequentist formalism.

4 Mathematical and philosophical probability

In this section, we analyse the difference between the mathematical and the philosophical concepts of probability in the light of the relation between theory and practice.

In experimental sciences, theoretical results can be tested empirically. This is not the case for probability theory. Indeed, experiments never investigate the same object as the theory. In other terms, probability theory is particular in the sense that applying it does not mean testing it. The practitioner's main goal is to discover bridging principles that enable him to appeal to theoretical results; it is not used to test an abstract axiomatic system. As stated by Shafer ([1996], p. 22):

> We should give back to practical problems their own logic, and we should see the task of application [in probability] as one of bridging, in any of many different ways, the logic of practical problems and the logic of the abstract theory.

Originally, two different concepts of probability coexisted – one *a priori* and the other *a posteriori*. Contemporary probability theory is inspired by the first one and is formulated in a very general mathematical framework, independent of any practical or philosophical interpretation. In order to bridge the gap between theory and practice, one must consequently make "transition" hypotheses with respect to the initial axioms, so as to apply theoretical results. In this perspective, only these transitions hypotheses can be empirically refuted, not the axioms themselves. Moreover, this necessity for filling the gap applies not only to statistics and frequency-related applications but also to any field making use of probability theory, such as financial valuation and quantum physics:

> New applications of probability are stretching the received understandings and making increasingly awkward the inability of these understandings to deal with the quality of application (Shafer [1996], pp. 23-4).

The central issue here is the relation between mathematical probability and its different philosophical counterparts. Actually, Kolmogorov's definition is to be viewed as an empty structure open to philosophical conceptualisation. Potential interpretations are multiple: frequentist, logical, subjective or propensity. As common language does not provide clear delimitation of what probability ultimately means, an explicit semantic bridge is required. Therefore, transition hypotheses are key concepts concerning not only the relation between probability theory and practice but also concerning each philosophical interpretation which, in its own way, relates to the abstract structure of the mathematical theory.

The specificity of a given philosophical theory of probability originates from the bridge linking it to the mathematical definition. In the subjective theory, the correspondence principle with the mathematical theory is provided by the notion of coherence. In a more loose way, the logical theory relates to probability theory via the concept of rationality. The

[7] The success was, however, not immediate. On the reception and acceptance of Kolmogorov's *Grundbegriffe*, see Shafer and Vovk ([2005]).

propension theory (advocated by Popper [1959] and Mellor [2005]) and the frequency theory of probability link with mathematical abstractions through the idea of propension and relative frequency, respectively. These connections are sometimes tight and rigorous, sometimes more intuitive, depending on the philosophy on which they are founded.

Philosophical theories of probability are therefore frequently envisaged as *foundational* with respect to the mathematical theory. Simultaneously, the concept of probability underlying most philosophical analysis obeys implicitly Kolmogorov's mathematical concept. Hence, for most philosophers avoid this circularity probability theory is treated as an issue in the foundations of statistics[8]. Therefore, in this treatment they miss the need for transition hypotheses that relate the mathematical structures to practical and philosophical matters.

5 Conclusions

Two fundamental movements characterize the history of probability. Firstly, its field of application expanded from Pascal's and Fermat's gambling games to Laplace's all-including phenomena. Secondly, and consequently, an axiomatization process became necessary. Two radically different systems were proposed. One, due to von Mises, was an attempt to define randomness from a frequentist point of view. The other, developed by Kolmogorov, stemmed from a pure mathematical perspective: it came up as a particular case of measure theory.

Paradoxically, the absence of any direct link to reality provided Kolmogorov's axiomatization supremacy over its competitor within the mathematical world as well as establishing this approach as the one and only reference in all fields of application of probability theory. By deliberately refusing to grapple with non-formal considerations, Kolmogorov avoided the delicate question of the ultimate signification of probability.

The history of probability theory demonstrates how researchers struggled to get round the semantics of the concept of probability: is it *a priori* or *a posteriori*? Each of them came to grip with the problem in their own way – except, maybe, Pascal and Fermat, who were not explicitly dealing with the notion of probability. Yet, all the definitions can be seriously criticized, with the exception of Kolmogorov's.

In that perspective, probability theory has a special status among scientific disciplines. It benefits from a large spectrum of real-life applications but, unlike other experimental sciences, it cannot be directly tested empirically. It does relate to practical matters by means of transition hypotheses that match the structure of practical problems in relation to the abstract theory. Similarly, philosophical interpretations of probability each relate to the mathematical theory via specific bridging principles, such as rationality, coherence, propension, and so on. As such, the mathematical concept of probability is beyond any practical or philosophical interpretation. The task of applying it practically or interpreting it philosophically by means of bridging principles lies beyond the mathematicians' scope.

Kolmogorov's probability theory can thus be seen as the optimal solution a mathematician can provide to formalize a complex and vaguely delimited object. Whether such phenomenon is observable in other fields of knowledge remains an open question. In any field, mathematicians are driven to more abstraction by the nature of their discipline while

[8] Out of five recent books on the introduction to the philosophy of probability (Gillies [2000]; Hacking [2001]; Jeffrey [2004]; Galavotti [2005]; Mellor [2005]), only the last one provides a global philosophical perspective.

practitioners call for concrete application possibilities. The resulting trade-off can turn in favour of either side, but a situation without any trade-off, as regards probability theory, is certainly rare.

References

Barone, J. and Novikov, A. [1978]: ‘A history of the axiomatic formulation of probability from Borel to Kolmogorov: part I’, *Archive for history of exact sciences*, **18**, pp. 13-190.

Bertrand, J. [1888] : *Calcul des probabilités*, Paris : Gauthiers-Villars.

Borel, E. [1909] : *Eléments de la théorie des probabilités*, Paris : Hermann.

Carnap, R. [1950]: *Logical Foundations of Probability*, Chicago: The University of Chicago Press.

de Finetti, B. [1937] : ‘La prévision : ses lois logiques, ses sources subjectives’, *Annales de l’Institut Henri Poincaré*, **7**, pp. 1-68.

De Scheemaekere, X. [2007]: ‘The epistemology of modern finance’, Working paper, Centre Emile Bernheim, Université Libre de Bruxelles, <ideas.repec.org/p/sol/wpaper/06-017.html>.

du Pasquier, L.G. [1926]: *Le calcul des probabilités, son évolution mathématique et philosophique,* Paris : Hermann.

Eagle, A. [2006]: ‘Probability, A Philosophical Introduction’, Review, *Mind*, **115**, pp. 773-7.

Fréchet, M. [1938] : ‘Exposé et discussions de quelques recherches récentes sur les fondements du calcul des probabilités’, in Wavre, R.(*ed.*), 1938, *Colloque consacré à la théorie des probabilités,* **2**, Paris : Hermann, pp. 23-55.

Galavotti, M. C. [2005]: *Philosophical Introduction to Probability,* Stanford, CA: CLSI Publications.

Gillies, D. [2000]: *Philosophical Theories of Probability*, London: Routledge.

Hacking, I. [1971]: ‘Equipossibility Theories of Probability’, *The British Journal for the Philosophy of Science*, **22**, pp. 339-55.

Hacking, I. [1975]: *The Emergence of Probability, A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference,* Cambridge: Cambridge University Press.

Hacking, I. [2001]: *An Introduction to Probability and Inductive Logic*, Cambridge: Cambridge University Press.

Hilbert, D. [1902] : ‘Sur les problèmes futurs des mathématiques’, in *Comptes rendus du deuxième congrès international des mathématiciens tenu à Paris du 6 au 12 août 1900,* Paris : Gauthier-Villars, pp. 58-114.

Howson, C. [1995]: ‘Theories of Probability’, *The British Journal for the Philosophy of Science*, **46**, pp. 1-32.

Humphreys, P. [1985]: ‘Why Propensities Cannot be Probabilities’, *The Philosophical Review*, **94**, pp. 557-70.

Jeffrey, R. [2004]: *Subjective Probability*: *The Real Thing*, Cambridge: Cambridge University Press.

Keynes, J. M. [1921]: *A Treatise on Probability,* London: Macmillan and Co.

Kneale, W. C. [1949]: *Probability and Induction,* Oxford: Clarendon Press.

Kolmogorov, A. [1950]: *Foundations of the theory of probability*, New York: Chelsea Pub. Co.

Laplace, P. S. [1840]: *Essai philosophique sur les probabilités*, Bruxelles : Hauman et Cie.

Laplace, P. S. [1887]: 'Théorie analytique des probabilités', in Laplace, *Œuvres complètes*, **7**, Paris : Gauthier-Villars.

Mellor, D. H. [2005]: *Probability, A Philosophical Introduction*, New York: Routledge.

Pascal, B. [1963] : *Oeuvres complètes*, Paris : Seuil.

Popper, K. R. [1959]: 'The Propensity Interpretation of Probability', *The British Journal for the Philosophy of Science*, **10**, pp. 25-42.

Ramsey, F. P. [1931]: *The foundations of mathematics and other logical essays*, London: Kegan.

Reichenbach, H. [1949], *The Theory of Probability*, in Reichenbach, M. and Cohen, R. S, 1978, *Hans Reichenbach, Selected Writings, 1909-1953*, **2**, Dordrecht: D. Reidel Publishing Company.

Shafer [1996]: 'The significance of Jacob Bernoulli's Ars Conjectandi for the philosophy of probability today', *Journal of Econometrics*, **75**, pp. 15-32.

Shafer, G. and V. Vovk [2006]: 'The Sources of Kolmogorov's *Grundbegriffe*', *Statistical Science*, **21**, pp. 70-98.

Szafarz, A. [1984] : 'Richard von Mises: L'échec d'une axiomatique', *Dialectica*, **38**, pp. 311-7.

Szafarz, A. [1985] : 'L'évolution du concept de probabilité mathématique de Pascal à Laplace', *Technologia*, **8**, pp. 67-75.

Todhunter, I. [1865]: *A history of the mathematical theory of probability, from the time of Pascal to that of Laplace*, New York: Chelsea Pub. Co.

van Brakel, J. [1976]: 'Some remarks on the prehistory of the concept of statistical probability', *Archive for history of exact sciences*, **16**, 119-36.

Ville, J. [1939] : *Etude critique de la notion de collectif*, Paris : Gauthier-Villars.

von Mises, R. [1919]: 'Grundlagen der Warscheinlichkeitsrechnung', *Mathematische Zeitschrift*, **5**, pp. 52-99.

von Mises, R. [1928]: *Warscheinlichkeit, Statistik und Wahrheit*, Vienne: Springer Verlag. English translation in von Mises, 1957, *Probability, Statistics and Truth*, Landau: Allen and Unwin.

von Mises, R. [1932] : 'Théorie des probabilités, Fondements et Applications', *Annales de l'Institut Henri Poincaré*, **3**, pp. 133-90.

von Mises, R. [1938]: 'Quelques remarques sur les fondements du calcul des probabilités', in Wavre, R.(*ed.*), 1938, *Colloque consacré à la théorie des probabilités*, **2**, Paris : Hermann, pp. 56-66.

Waismann, F. [1930]: ‘Logische Analyse des Wahrscheinlichkeitsbegriffs’, *Erkenntnis*, **1**, pp. 228-48.

Wald, A. [1937]: ‘Die Widerspruchfreiheit des Kollectivbegriffes der Wahrscheinlichkeitsrechnung’, *Ergebnisse eines Mathematischen Kolloquiums*, **8**, pp. 38-72.