

Review for *Annals of Probability*. (Preliminary Version)

By Glenn Shafer, Faculty of Management, Rutgers University.

Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective, by Jan von Plato. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge University Press. 1994. ISBN 0-521-44403-9.

Jan von Plato's *Creating Modern Probability* is a book to celebrate. It will be a source of pleasure for probabilists and a stimulus for historians. Combining a sweeping vision with a sympathetic and thorough marshaling of sources, it brings to life the emergence of measure-theoretic probability in the first third of the twentieth century, thus supplying a satisfying sense of history to a field that has seemed, to too many of its practitioners, to have sprung suddenly into existence with the publication of Andrei Kolmogorov's *Grundbegriffe der Wahrscheinlichkeitsrechnung* in 1933. Von Plato's mathematical precision will satisfy specialists, while his success in organizing his sources with respect to larger questions will provoke and enable a new cycle of historical research and debate.

The historiography of probability has flowered in recent decades. After a long period in which the only history of probability was Isaac Todhunter's 1865 treatise, we now have a whole shelf of scholarly studies, ranging from Ian Hacking's *Emergence of Probability* (1975) to Lorraine Daston's prize-winning *Classical Probability in the Enlightenment* (1988) and Anders Hald's *History of Probability and Statistics and their Applications before 1750* (1990). This flowering has extended to statistics, generally thought to be a distinct branch of knowledge after about 1820, with Theodore Porter's *Rise of Statistical Thinking, 1820-1900* (1986), Stephen M. Stigler's *History of Statistics: The Measurement of Uncertainty before 1900* (1986), and Ian Hacking's *The Taming of Chance* (19--). But it has scarcely extended into the twentieth century. One might imagine that a treaty had been concluded between historians and mathematicians. From Kolmogorov onward, probability theory is mathematics, to be touched by a historian at his or her peril. But historians can say what they want about what J.L. Doob has called the "vague nonmathematical context" of the probability calculations of earlier times.

With grace and finesse, von Plato has conquered the buffer period of 1900-1933 for history. With grace, because his understanding of the mathematics of the period cannot be faulted. With finesse, because he has honored fully the claims of contemporary probabilists. For von Plato, probability after 1933 is *modern*, truly mathematical because it has infinitary concepts at its heart. Before 1900, in contrast, probability is *classical*, part of applied or “mixed” mathematics rather than pure mathematics. Why and how, von Plato asks, did probability become modern? His answer centers on two themes: the demands of statistical physics, which required that probability accommodate infinities with ease in order to handle continuous processes and ergodic theory, and the struggle over the foundations of mathematics, which ended with the triumph of David Hilbert’s formalism, in probability as in other branches of mathematics.

Von Plato’s method is exhaustive and didactic. He summarizes directly almost all the articles and books that constitute his original sources, thus mostly dispensing with the need to cite the secondary literature, which is scant in any case. He does not even cite his own earlier historical articles. The result includes long and sometimes meandering chapters on statistical physics, Kolmogorov, and de Finetti, together with shorter but still substantial chapters on quantum mechanics, the method of arbitrary functions, and von Mises. Fortunately, these large chunks are preceded by two well-structured introductions. The first chapter gives an overview of the book. The second chapter, entitled “Pathways to modern probability,” describes the immediate mathematical and philosophical context of Kolmogorov’s achievement.

Statistical Physics

Von Plato’s long chapter on statistical physics is the heart of the book. It guides us along the trajectory of probability in physics from the work of Clausius (1857) until the advent of quantum mechanics, which has a chapter of its own. We begin with the development of the idea of probability as a time average, especially in the work of Boltzmann, Maxwell, and Gibbs. We hear how these authors came to terms with the reversibility of mechanics; since states of high entropy are vastly more probable than states of low entropy, we are likely to see an increase in entropy whether we look to the future or to the past from a state of low

entropy. We hear about ergodic theory: how Boltzmann, Maxwell, and Gibbs struggled to base it on mechanical laws, how Borel, Poincaré, and the Ehrenfests gave it a more probabilistic turn, how von Mises called for a purely probabilistic account, and how von Neumann and Birkhoff carried it out. We hear how Einstein dealt with these issues, how Einstein and von Smolchowski dealt with Brownian motion, and how the theory of continuous stochastic processes developed from this starting point. The chapter concludes with a discussion of the treatment of radioactivity before quantum mechanics.

This is a broad and busy landscape. Those who want to study it carefully will find von Plato's triangulation immeasurably valuable, but the casual reader will find it difficult to see the forest from the trees. It is here that I most regretted von Plato's reluctance to discuss secondary sources, fancying that by debating the general theses advanced by others he might have provided a more satisfying sense of the grand picture. I would also have been interested to hear more about how statistical physics related to other work in statistics. I was fascinated, for example, to hear about Einstein's treatment of spectral analysis in 1914, but I was left wondering about its relation to the larger history of the statistical analysis of time series. But von Plato aims neither to settle debates about the historical development of statistical physics nor to relate it to other branches of statistics; rather, he aims to see how statistical physics contributed to modern probability.

What did statistical physics contribute to the foundations of probability? Most superficially, perhaps, it raised the problem of the mathematical representation of stochastic processes in continuous time. According to Von Plato, the idea of a probabilistic law as a function of continuous time was first put forward by Boltzmann, in 1872. Von Plato traces the development of the idea in ergodic theory and in the study of Brownian motion, with a glance at Lundberg's work on insurance and Bachelier's efforts to describe the stock exchange. He remains puzzled that the theory of stochastic processes in the modern sense did not finally appear until 1929, in the hands of de Finetti, Hostinsky, and Kolmogorov.

In ergodic theory, von Plato finds a variegated array of influences. Boltzmann introduced the idea of an ensemble of systems, although it has been most associated with the later work of Gibbs. The probability of a state was the average fraction of the time a single system spends in that

state, but talking about a single system makes it difficult to express the idea that macroscopic properties of the system depend on the initial state only through the total energy. If you speak of an ensemble of systems with that energy, you can say that a property holds with high probability. According to von Plato, such talk eventually left ensembles behind, evolving into talk about probability one or measure zero, an idea previously at home only in pure mathematics (as when we compare the rational to the irrational numbers). A landmark in this process was Poincaré's recurrence theorem for the three-body problem, in 1890: from any state, except for exceptional ones of measure zero, a system will eventually return arbitrarily close to that state.

One recurrent theme in the evolution of ergodic theory is that independence from initial conditions follows from slight uncertainties about those conditions, compounded by similar uncertainties about minor external disturbances. According to Von Plato, this was Boltzmann's original argument for the "ergodic hypothesis," the hypothesis that a system's trajectory would eventually run through all states that have the same total energy. This argument did not remain at the center of ergodic theory, but it gained a life of its own, which von Plato charts in a separate chapter, "Classical embeddings of probability and chance." The high point of this story is Poincaré's "method of arbitrary functions," presented in his 1896 *Calcul des probabilités*. Poincaré showed that in many mechanical systems, the probability distribution of very small disturbances in initial conditions has no effect on the probability distribution of the state much later, provided only that the initial probability distribution is continuous. This makes precise a foundation for objective probability that was popular into the 1930s. It was formulated in measure-theoretic terms by Hopf, but it has been of less interest to later probabilists, perhaps simply because Kolmogorov's formulation seems to eliminate the need to worry further about foundations.

Eventually, the mainstream of development of ergodic theory relied not on uncertainties in initial conditions and external disturbances but on the formulation of the ergodic hypothesis that Poincaré suggested in 1894, slightly extending the idea of his recurrence theorem: the trajectories of a system should be dense except for a set of initial conditions of measure zero. By presenting this as a version of Maxwell's thinking, Poincaré opened the way to the views of von Mises, who was led by his own work on the limiting behavior of Markov chains in the 1920s to believe that the

ergodic problem was purely probabilistic, independent of mechanical hypotheses. Von Plato's archival research leads to the conclusion that von Mises influenced von Neumann and Birkhoff, who proved measure-theoretic ergodic theorems in 1931-32. In Birkhoff's theorem, measure theory was dominant in both the hypothesis and the conclusion; the ergodic hypothesis is that no set of states of intermediate measure is invariant under the motion of the system, and the conclusion is that time averages approach unique limits except on a set of trajectories of measure zero. The theory was abstracted even farther from physics by Khintchine and Hopf in 1932. Unfortunately, these purely mathematical theorems represented an abandonment of the original goal of ergodic theory: to demonstrate that statistical behavior follows from mechanical laws.

In summary, we may say that statistical physics gave to the foundations of probability the idea of continuous-time stochastic processes and the conviction that "except for a set of measure zero" is a legitimate trailer for empirically meaningful statements, a way of dispensing with the more intuitive but ultimately confused idea of ensembles. Most importantly, statistical physics created a demand for a probability theory that dealt rigorously with infinities, a demand compounded by the hope that mathematical rigor might compensate for the physicists' own confusion about how macroscopic statistical behavior is related to mechanical laws.

From the Foundations of Mathematics to the Foundations of Probability

The period at the beginning of the twentieth-century in which the measure-theoretic foundations of probability were established was also a period of intense debate over the foundations of mathematics. One of this book's most striking accomplishments is that it places probability clearly in this larger context. For this we can thank von Plato's training as a philosopher of mathematics and his continuing interest in intuitionism; he is a leader in the current effort to found mathematics on theory of types (see von Plato 1995).

The dominant formalist understanding of mathematics, set in place by Hilbert, is now most often contrasted with the views of those, including von Plato, who follow Brouwer in objecting to the law of the excluded middle. But the foundations of probability were more influenced by the earlier

school of French intuitionists, which included Poincaré, Baire, Borel, Lebesgue, and Lusin (see Largeault 1993). As all probabilists know, these scholars were most concerned with the legitimacy and the process of the construction of the infinite from the finite. Von Plato describes at length Borel's attempt to replace geometric (measure-theoretic) probabilistic reasoning with what he hoped would be more meaningful reasoning with "denumerable" probabilities. It was in this context of concern about the legitimacy of the infinite that von Mises formulated the idea of a collective, Hilbert formulated his philosophy of mathematics, and Kolmogorov used Hilbert's philosophy to formulate an interpretation of measure-theoretic probability.

Working before the triumph of quantum mechanics, the French intuitionists did not feel the need for a radical concept of objective chance. They were much more sympathetic with the view that probability arises from the practical impossibility of prediction, as when the future depends too delicately on initial conditions. This meant, however, that the actual random was a real puzzle for them; if the only meaningful infinite objects are those we construct by finitistic means, then what can be meant by an infinite sequence of random choices—or even worse, by a random choice from a continuous distribution? This was the question that von Mises sought to answer, beginning in 1919, with his theory of collectives. A collective is a single random sequence (not an ensemble), say of 0s and 1s, with the property that the proportion of 1s tends to a definite limit, say p , not only in the sequence as a whole but in every infinite subsequence. Today this idea is most often recalled merely to be ridiculed, for it is obvious to us that if p is strictly between zero and one, then there is a subsequence of all 1s and a subsequence of all 0s. As but von Plato points out in his chapter on von Mises, the ridicule is misplaced; it went without saying, in the context of the times, that only subsequences that can be constructed count; the subsequence of all 1s that leaps to our eyes need not exist for a constructivist. Von Mises's enterprise was to legitimize the idea of the random not by challenging constructivism but by shifting its focus from the obviously self-contradictory task of constructing an actual random sequence to the more meaningful task of constructing a description of such a sequence. This enterprise had the respect of von Mises's constructivist contemporaries, including Kolmogorov, who returned to it with his own work on the complexity of sequences in the 1960s.

Hilbert's philosophy of mathematics was based a shift similar to von Mises's but on a grander scale. Hilbert agreed with the intuitionists that only the finitary part of mathematics has meaning, but, as von Plato puts it, he located what was meaningful and finite in metamathematics, where one speaks of properties of formal systems. It is fully legitimate to posit and reason about infinities mathematically, as long as the axiomatic system thus constructed is consistent, for the infinities will introduce no contradictions into the description of any finite reality that may satisfy the axioms. In his first explorations of this idea, Hilbert had thought that consistency would be enough to ensure that infinity would be a conservative extension over the finite: no finitary statements would be provable with infinitary notions that were not already provable with finitary ones. This hope was refuted by Gödel in 1931, but as von Plato notes, the refutation seems to have come too late to have influenced Kolmogorov's understanding of the application of his axioms, which he based on Hilbert's ideas in spite of his own somewhat stronger sense of constructivism.

The key to Kolmogorov's success in axiomatizing probability using measure theory was the fact that Hilbert's philosophy allowed a considerable distance between an axiomatic system and its interpretation. Even when axioms authorize infinities, the connection between the axioms and the reality they describe can be approximate and finitary. Thus measure theory can describe, if from a greater distance, the same randomness that von Mises had sought to describe. Under certain conditions ("which we do not wish to go into here," as Kolmogorov said in the *Grundbegriffe*), one can assume that a real number $P(A)$ is attached to the event A such that (1) one can be practically certain, in a large number of trials, that the relative frequency will differ only little from $P(A)$, and (2) if $P(A)$ is very small, then one can be practically certain that A will not appear in a single trial. Thus was frequentist probability founded on a mathematical framework that made no mention of frequency, using a subjective concept of practical certainty.

Kolmogorov was not the first to try to base probability on measure theory. As von Plato demonstrates, the idea that had been in the air since the turn of the century, long before the project of making measure theory itself fully abstract and independent of geometry was brought to a conclusion, by Frechet in 1914. Hilbert, who had included the axiomatization of probability theory in his famous list of unsolved problems in 1900, even inspired a Göttingen dissertation on the topic in 1907. But it was Kolmogorov who made measure theory work as a foundation for

probability. Three elements were essential to his success: his extension theorem, which showed that continuous-time stochastic processes could be constructed in an essentially finitistic way, his conception of conditional expectation as a function, using the Radon-Nikodym theorem, and the formalist philosophy of infinity that Hilbert developed in the 1920s.

Minor Themes

Statistical physics is not the only science that brought infinitary events and the idea of measure zero into probability. Von Plato's second chapter, "Pathways to modern probability," explores two other pathways for such ideas, both beginning in planetary astronomy. They turn out to be closely related both to each other and to ergodic theory in its eventual form. The first pathway starts with the work of the Swedish astronomer Hugo Glydén, whose understanding of the approximate computation of planetary motions led him, in 1888, to study the probability of convergence of the integers in a continued fraction. The second begins later, in 1909, when another astronomer, Bohl, studied the problem of mean motion in planetary astronomy. The conclusions of Glydén and Bohl, in the hands of Hermann Weyl, became a simple statement about real numbers: if you choose a real number at random, multiply it by 2, 3, 4, ... and take only the decimal parts, you get with probability one a dense set that is uniformly distributed in the unit interval. Von Plato considers Weyl's proof of this statement the first rigorous proof of an ergodic theorem. But at an intuitive level, ergodicity turns out to be a very old idea indeed. In an appendix (or "supplement"), von Plato recounts the efforts by Nicole Oresme (1325-1382) to show that the motions of the heavenly bodies are aperiodic. Oresme's results, as he notes, are easy consequences of Poincaré's recurrence theorem.

Von Plato's chapter on probability in quantum mechanics reviews the old quantum theory, with an emphasis on Einstein's contributions, and then recounts the rapid development of quantum mechanics, from its first invention by Heisenberg in July 1925, through Born's explanation of the statistical interpretation of Schrödinger's wave mechanics in July 1926, to Heisenberg's formulation of the uncertainty relation in March 1927. The theme of the chapter is how quickly most physicists came around to a genuine indeterminism, even though Einstein, who never abandoned the older epistemic interpretation of probability, inspired Born's interpretation. As von Plato insists, the triumph of the quantum revolution preceded

Kolmogorov's formulation of measure-theoretic foundations for probability and thus must be counted as an important constituent of the ambient frequentism within which that formulation was reached. Yet quantum mechanics remained clearly to one side of Kolmogorov's enterprise, for Kolmogorov carefully left room for the traditional epistemic understanding of frequentist probabilities and sought most prominently to legitimize continuous probabilities, whereas the probabilities of quantum mechanics were both radically objective and jarringly finite. It was only much later, with Feynman's path integrals, that quantum theory came into a closer embrace with measure theory.

Von Plato's final chapter is devoted to Bruno de Finetti. One of the best chapters of the book, it summarizes masterfully de Finetti's work on subjectivist probability, providing a clear understanding of his mathematical contributions and his philosophical commitments. We see step by step how de Finetti developed exchangeability, partial exchangeability, his representation theorems, and the law of large numbers for exchangeable events. We get a balanced view of de Finetti's foundations for subjective probability, beginning with his arguments based on coherent bets, but also including his axiomatization of qualitative probability, which he thought sounder. We get a good explanation of de Finetti's problems with Kolmogorov's foundations, especially with the rule of denumerable additivity. We also get a good view of de Finetti's work on continuous stochastic processes in the late 1920s and early 1930s, especially his work on infinitely divisible distributions. The work on stochastic processes has been overshadowed, in recent decades, by interest in de Finetti's subjectivism.

De Finetti's subjective interpretation of probability is not part of the story that culminates in Kolmogorov's *Gundbegriffe*. As von Plato emphasizes, Kolmogorov sought to provide foundations for a frequentist probability. De Finetti's alternative subjectivist conception was fully articulated only later, in the late 1930s, and it was heavily influential only much later, after the work of L.J. Savage. But although we do not need de Finetti in order to understand the rest of the story, the rest of the story does help us towards a fuller understanding de Finetti. It allows von Plato to place de Finetti's representation theorem in the context of the ergodic decomposition theorem, and it allows him to explain the content of de Finetti's insistence on operational meaning, especially insofar as this led de Finetti to reject the meaningfulness of probabilities for infinitary events.

For de Finetti, it only made sense to speak of probabilities for verifiable events. The context of the larger story also gives meaning to a significant lacunae in de Finetti's subjectivist argument: his failure to motivate his subjective interpretation in physics.

Other Perspectives

Any historical study must make choices in emphasis and perspective, and we can expect that future work will go beyond the choices von Plato has made.

The most salient of von Plato's choices is the choice to write from within. As he notes in his introduction, the history of science must be studied both from without and from within; we must place scientific ideas in the wider lives of their authors and their times, and we must also understand the internal logic of the development of these ideas. Von Plato sees his choice as a natural one, because study from within is easier, and perhaps also because it is a prerequisite for an adequate study from without. The study of circumstances of individuals and institutions will have to wait, he believes, "for a general organization of historical studies on probability."

The decision to undertake a study from within does not suffice to mark off the ground to be covered, for there are many layers of scholarly activity, in many different directions, between mathematical probability and larger concerns to which it is relevant, and so it remains to say what is in and what is out. Von Plato's preferences are clear. Quantum mechanics is in; statistics is out. De Finetti is in; Keynes and Jeffreys are out. Even R. A. Fisher's philosophy of probability merits but a few passing words. As von Plato says in his preface, the absence of predecessors left him free to let his own interests act as his guide. But he sometimes justifies his preferences, and he is not always persuasive. No doubt quantum mechanics helped convince mathematicians of the importance of probability, but the accounts of Hacking, Porter, and Stigler suggest that the nineteenth-century successes of social statistics also had a great impact on appreciation of probability, an impact that mathematicians could scarcely have escaped. And social statistics, unlike quantum mechanics, did not direct attention from the continuous back to the discrete. Future reconsiderations of von Plato's topic will need to identify and weigh the evidence about the relative

weight of statistics and physics in the development of measure-theoretic probability.

For my own part, I hope that von Plato's successors will give more sympathy and attention to classical probability, for in his emphasis on the transition to modern probability, he sometimes fails to do justice to the resources and contributions of the older theory. The idea of equiprobability does not exhaust the conceptual or technical content of the older theory, any more than the idea of measure exhausts the content of Kolmogorov's theory. In particular, the older theory had a dynamic aspect; from its very beginning in the work of Pascal and Huygens, it was concerned with changes in the values of expectations. A full understanding of the sources of Kolmogorov's theory would include an understanding of how the ideas of conditional probability and independence evolved out of the dynamic aspects of classical probability.

In the long run, the significance of von Plato's work will hinge most on the fate of the theory he celebrates. Kolmogorov's theory is now two-thirds of a century old. By anchoring it in history, von Plato frees our imagination to turn to the future. Kolmogorov found a way to link infinitary measure-theoretic probability to the finitistic meaning of probability in his culture. Will that linkage remain viable in a twenty-first century in which the applications of probability continue to expand, while measure-theoretic probability plunges ever deeper into pure mathematics? Or will new foundations for these applications? If Kolmogorov was modern, what will be post-modern?

References

Daston, Lorraine (1988). *Classical Probability in the Enlightenment* Princeton University Press.

Hacking, Ian (1975). *The Emergence of Probability*. Cambridge University Press.

Hacking, Ian (19--). *The Taming of Chance*.

Hald, Anders (1990). *A History of Probability and Statistics and their Applications before 1750*. Wiley.

Largeault, Jean (1993). *Intuition et intuitionisme*. Paris: J. Vrin.

von Plato, Jan (1995). The axioms of constructive geometry. *Annals of Pure and Applied Logic*.

Porter, Theodore (1986). *The Rise of Statistical Thinking, 1820-1900*. Princeton University Press.

Stigler, Stephen M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Harvard University Press.