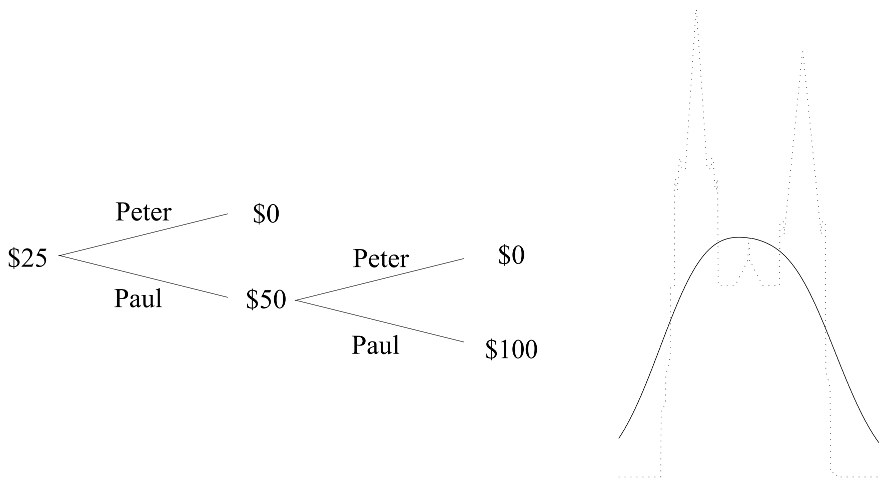
From Cournot's Principle to Market Efficiency

Glenn Shafer



The Game-Theoretic Probability and Finance Project

Working Paper #15

First posted November 26, 2005. Last revised March 10, 2006.

Project web site:
http://www.probabilityandfinance.com

Abstract

The efficient-markets hypothesis was formulated in the early 1960s, when Cournot's principle was no longer widely understood and accepted as a philosophical foundation for probability theory. A revival of Cournot's principle can help us distinguish clearly among different aspects of market efficiency.

Glenn Shafer is Professor in the Rutgers Business School - Newark and New Brunswick, 180 University Avenue, Newark, New Jersey 07102 USA, and in the Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX UK. E-mail gshafer@rutgers.edu; web site www.glennshafer.com.

Contents

1 Introduction

Cournot's principle says that an event of small or zero probability singled out in advance will not happen. From the turn of the twentieth century through the 1950s, many mathematicians, including Chuprov, Borel, Fréchet, Lévy, and Kolmogorov, saw this principle as fundamental to the application and meaning of probability.[1] In their view, a probability model gains empirical content only when it rules out an event by assigning it small or zero probability.

In the 1960s, when probability theory was gaining in importance in economics and especially finance, Cournot's principle was no longer so widely accepted. In fact, the principle had almost disappeared as those who had espoused it in the earlier period passed from the scene. In this article, I argue that its disappearance entailed a loss of clarity in the interpretation of probability, which accounts in part for the high level of confusion in initial formulations of the efficient-markets hypothesis.

The game-theoretic framework for probability (Shafer & Vovk [125]) revives Cournot's principle in a form directly relevant to markets. In this framework, Cournot's principle is equivalent to saying that a strategy for placing bets without risking bankruptcy will not multiply the bettor's capital by a large or infinite factor. It can therefore be applied directly to strategies for exploiting market prices, without assuming the existence of meaningful probability distributions related to these prices.

The claim that an investor cannot make a lot of money using public information is part of the efficient-markets hypothesis as it was formalized in the 1980s [86]. But this efficient-markets hypothesis also says that market prices are discounted expected values with respect to a probability distribution that changes only in accordance with relevant information. This bundling of ideas has enabled scholars to talk past each other. Some (e.g. Malkiel [94]) claim the efficient-markets hypothesis is vindicated when strategies for making money fail. Others (e.g. Shiller [132]) claim it is refuted by any evidence that price changes are not always based on information.

The game-theoretic framework allows us to unbundle the efficient-markets hypothesis in a useful way. This unbundling is encouraged by the framework's success in dealing with classical probability. The framework accommodates both the case where an investor or a bettor may buy any variable at its expected value with respect to a specified probability distribution, as in classical probability theory, and the case where only some variables are priced and offered for sale, as in the incomplete markets in which real investors participate. In the case where all variables are priced, the framework reduces to classical probability, but many classical results extend to the case where only limited prices are given.

As it turns out, the game-theoretic form of Cournot's principle, applied directly to market prices, implies several stylized facts commonly associated with the existence of a whole probability distribution for future value, including

[1] Most of them did not call it "Cournot's principle", but this name, due to Fréchet, was used in the 1950s (see p. 5) and may be the most reasonable and convenient name available today.

the $\sqrt{dt}$ scaling of price changes [150] and CAPM-type relations between the realized average return for a particular security and that of the market [149]. To the extent that they are confirmed by data, these stylized facts can count as demonstrations of the usefulness of that part of the efficient-markets hypothesis represented by Cournot's principle. But this will not, by itself, provide any support for the quite separate hypothesis that price changes are usually or always based on information.

The following sections review the rise and fall of Cournot's principle in classical probability (§2), its new form in game-theoretic probability (§3), and its potential in this new form for probability, economics, and finance theory (§4).

2 The rise and fall of Cournot's principle

This section traces Cournot's principle from its inception in Jacob Bernoulli's discussion of moral certainty in the early 18th century to its disappearance in Joseph Doob's reformulation of mathematical probability theory in the 1950s.

The section is organized on conceptual as well as chronological lines. In §2.1, I trace the relatively uncontroversial concept of moral certainty from the 17th to the 20th century. In §2.2, I trace the development of a more controversial idea—the idea that Cournot's principle is the only bridge from a probability model to the world; this idea first emerged in Cournot's analysis of moral certainty, and it was best articulated by Paul Lévy in the 1920s. In §2.3, I distinguish between the strong form of Cournot's principle, which asserts that a particular event of very small probability will not happen on a particular trial, and the weak form, which asserts merely that events of small probability happen rarely on repeated trials. Then I turn to the history of the opposition; in §2.4, I acknowledge the indifference of British mathematicians and statisticians, and in §2.5, I acknowledge the more explicit opposition of German philosophers. Finally, in §2.6, I explain how Doob's mathematical framework for stochastic processes contributed to the disappearance of Cournot's principle after the second world war.

This section draws heavily on recent papers with Vladimir Vovk on the historical context of Andrei Kolmogorov's contributions to the foundations of probability [127, 128, 151].

2.1 Moral certainty

An event with very small probability is *morally impossible*; it will not happen. Equivalently, an event with very high probability is *morally certain*; it will happen. This principle was first formulated within mathematical probability by Jacob Bernoulli. In his *Ars Conjectandi*, published posthumously in 1713, Bernoulli proved that in a sufficiently long sequence of independent trials of an event, there is a very high probability that the frequency with which the event happens will be close to its probability. Bernoulli explained that we can treat the very high probability as moral certainty and so use the frequency of the event as an estimate of its probability. Beginning with Poisson [23, 109, 133], this conclusion was called the law of large numbers. (Only later, mostly in the last fifty years, was "the law of large numbers" used to designate Bernoulli's theorem itself and its generalizations, which are purely mathematical statements.)

Probabilistic moral certainty was widely discussed in the eighteenth century. In the 1760s, Jean d'Alembert muddled matters by questioning whether the prototypical event of very small probability, a long run of many happenings of an event as likely to fail as happen on each trial, is possible at all. A run of a hundred may be metaphysically possible, he felt, but physically impossible. It has never happened and never will [34–36]. In 1777, George-Louis Buffon argued that the distinction between moral and physical certainty was one of degree. An event with probability 9999/10000 is morally certain; an event with much greater probability, such as the rising of the sun, is physically certain [26, 92].

Augustin Cournot, a mathematician now remembered as an economist and a philosopher of science [97, 98], gave the discussion a nineteenth-century cast in his 1843 treatise on probability [31]. Because he was familiar with geometric probability, Cournot could talk about probabilities that are vanishingly small. He brought physics to the foreground. It may be mathematically possible, he argued, for a heavy cone to stand in equilibrium on its vertex, but it is physically impossible. The event's probability is vanishingly small. Similarly, it is physically impossible for the frequency of an event in a long sequence of trials to differ substantially from the event's probability [31, pp. 57 and 106].

In the second half of the nineteenth century, the principle that an event with a vanishingly small probability will not happen took on a real role in physics, most saliently in Ludwig Boltzmann's statistical understanding of the second law of thermodynamics. As Boltzmann explained in the 1870s, dissipative processes are irreversible because the probability of a state with entropy far from the maximum is vanishingly small [145, p. 80] [123]. Also notable was Henri Poincaré's use of the principle. Poincaré's recurrence theorem, published in 1890 [108, 145], says that an isolated mechanical system confined to a bounded region of its phase space will eventually return arbitrarily close to its initial state, provided only that this initial state is not exceptional. Within any region of finite volume, the states for which the recurrence does not hold are exceptional inasmuch as they are contained in subregions whose total volume is arbitrarily small.

At the turn of the twentieth century, it was a commonplace among statisticians that one must decide what level of probability will count as practical certainty in order to apply probability theory. We find this stated explicitly in 1901, for example, in the articles by Georg Bohlmann and Ladislaus von Bortkiewicz in the section on probability in the *Encyklopädie der mathematischen Wissenschaften* [139, p. 825] [10, p. 861].

Aleksandr Chuprov, professor of statistics in Petersburg, was the champion of Cournot's principle in Russia. He called it Cournot's lemma [29, p. 167] and declared it a basic principle of the logic of the probable [129, pp. 95–96]. Andrei Markov, also in Petersburg, learned about mathematical statistics from

Chuprov [107], and we see an echo of Cournot's principle in Markov's textbook, published in German in 1912 [96, p. 12]:

> The closer the probability of an event is to one, the more reason we have to expect the event to happen and not to expect its opposite to happen.
>
> In practical questions, we are forced to regard as certain events whose probability comes more or less close to one, and to regard as impossible events whose probability is small.
>
> Consequently, one of the most important tasks of probability theory is to identify those events whose probabilities come close to one or zero.

The importance of Cournot's principle was also emphasized in the early twentieth century by Émile Borel. According to Borel, a result of the probability calculus deserves to be called objective when its probability becomes so great as to be practically the same as certainty [11–14]. Borel gave a more refined and demanding scale of practical certainty than Buffon's. A probability of 10^{-6}, he suggested, is negligible at the human scale, a probability of 10^{-15} at the terrestrial scale, and a probability of 10^{-50} at the cosmic scale [15, pp. 6–7].

2.2 Probability's only bridge to the world

Saying that an event of very small or vanishingly small probability will not happen is one thing. Saying that probability theory gains empirical meaning only by ruling out the happening of such events is another. Cournot may have been the first to make this second assertion:

> ... *The physically impossible event is therefore the one that has infinitely small probability*, and only this remark gives substance—objective and phenomenal value—to the theory of mathematical probability [31, p. 78].

Cournot's wording reflects the influence of Immanuel Kant; "objective and phenomenal" refers to Kant's distinction between the noumenon, or thing-in-itself, and the phenomenon, or object of experience [37].

Paul Lévy, a French mathematician who began writing on probability in the 1920s, stands out for the clarity of his articulation of the thesis that Cournot's principle is the only way of connecting a probabilistic theory with the world outside mathematics. In a note published in 1922, Lévy's teacher Jacques Hadamard explained that probability is based on two basic notions: the notion of perfectly equivalent (equally likely) events and the notion of a very unlikely event [66, p. 289]. In his *Calcul des probabilités*, published in 1925, Lévy emphasized the different roles of these two notions. The notion of equally likely events, Lévy explained, suffices as a foundation for the mathematics of probability, but so long as we base our reasoning only on this notion, our probabilities are merely subjective. It is the notion of a very unlikely event that permits the results of the mathematical theory to take on practical significance ([87], pp. 21, 34; see also [88], p. 3). Combining the notion of a very unlikely event with Bernoulli's theorem, we obtain the notion of the objective probability of an event, a physical constant that is measured by relative frequency. Objective probability, in Lévy's view, is entirely analogous to length and weight, other physical constants whose empirical meaning is also defined by methods established for measuring them to a reasonable approximation ([87], pp. 29–30).

Lévy's views were widely shared in France. Starting in the 1940s, Borel called Cournot's principle first "the fundamental law of chance" (la loi fondamentale du hasard) [16] and then "the only law of chance" (la loi unique du hasard) [17, 18]. The latter phrase was taken up by Robert Fortet [85].

Neither Lévy nor Borel used the name "Cournot's principle," which was coined by Maurice Fréchet in 1949. Fréchet's inspiration was Oskar Anderson, who had talked about the Cournotsche Lemma (Cournot's lemma) and the Cournotsche Brücke (Cournot's bridge) [3, 4]. Anderson was following his teacher Chuprov in the use of "lemma." Fréchet felt that "lemma," like "theorem," should be reserved for purely mathematical results and so suggested "principe de Cournot." Fréchet's coinage was used in the 1950s in French, German, and English [42, 120, 121, 140].

2.3 Weak and strong forms of the principle

Fréchet distinguished between strong and weak forms of Cournot's principle [62, 99, p. 6]. The strong form refers to an event of small or zero probability that we single out in advance of a single trial: it says the event will not happen on that trial. The weak form says that an event with very small probability will happen very rarely in repeated trials. Some authors, including Lévy, Borel, and Kolmogorov, adopted the strong principle. Others, including Chuprov and Fréchet himself, preferred the weak principle.

The strong principle combines with Bernoulli's theorem to produce the unequivocal conclusion that an event's probability will be approximated by its frequency in a particular sufficiently long sequence of independent trials. The weak principle combines with Bernoulli's theorem to produce the conclusion that an event's probability will *usually* be approximated by its frequency in a sufficiently long sequence of independent trials, a general principle that has the weak principle as a special case. This was pointed out by Castelnuovo in his 1919 textbook (p. 108). Castelnuovo called the general principle the *empirical law of chance* (la legge empirica del caso):

> In a series of trials repeated a large number of times under identical conditions, each of the possible events happens with a (relative) frequency that gradually equals its probability. The approximation usually improves with the number of trials [28, p. 3].

Although the special case where the probability is close to zero is sufficient to imply the general principle, Castelnuovo thought it pedagogically preferable to begin his introduction to the meaning of probability by enunciating the general principle, which accords with the popular identification of probability with frequency. His approach was influential at the time. Maurice Fréchet and Maurice Halbwachs adopted it in their 1924 textbook [64]. It brought Fréchet to the same understanding of objective probability as Lévy: it is a physical constant that is measured by relative frequency [61, p. 5] [60, pp. 45–46].

The weak point of Castelnuovo and Fréchet's position lies in the modesty of their conclusion: they conclude only that an event's probability is *usually* approximated by its frequency. When we estimate a probability from an observed frequency, we are taking a further step: we are assuming that what usually happens has happened in the particular case. This step requires the strong form of Cournot's principle. According to Kolmogorov [80, p. 240 of the 1965 English edition], it is a reasonable step only if "we have some reason for assuming" that the position of the particular case among other potential ones "is a regular one, that is, that it has no special features." Kolmogorov and his contemporaries considered the absence of special features that would enable one to single out particular trials essential to any application of probability theory to the world. Richard von Mises formalized this absence in terms of rules for selecting subsequences from infinite sequences of trials [142–144], but Kolmogorov did not consider such infinitary principles relevant to applications [127, §A.2]. A finitary principle, one applicable to a single trial, is needed, and this is Cournot's principle.

2.4 British practicality

For Borel, Lévy, and Kolmogorov, probability theory was a mathematical object, and there was a puzzle about how to relate it to the world. Cournot's principle solved this puzzle in a way that minimized the importance of the distinction between subjective and objective meanings of probability. For Borel and Lévy, probabilities begin as subjective but become objective when they are sufficiently close to zero or one and we adopt Cournot's principle. Kolmogorov, faithful to Soviet ideology, avoided any hint of subjectivism but still recognized the role of Cournot's principle in relating the mathematical formalism of probability to the world of frequencies.

The British saw quite a different picture in the late 19th century [114, p. 74ff]. There was little mathematical work on probability in Britain in this period, and in any case the British were not likely to puzzle over how to relate abstractions to the world. They saw probability, to the extent that it was of any use at all, as a way of directly describing something in the world, either belief or frequency. This left them with a quarrel. Many, including Augustus De Morgan, William Stanley Jevons and Francis Edgeworth, said belief [44, 49, 50, 70]. A few, most influentially John Venn [137], said frequency. R. A. Fisher and Harold Jeffreys carried the debate into the 20th century [54, 55, 69]. Neither side had any need for Cournot's principle, and some participants in the debate saw no use even for Bernoulli's theorem [37, 51].

British authors did sometimes discuss the classical puzzles about very unlikely events. Could a pair of dice come up sixes a thousand times running?

Could Shakespeare's plays be produced by drawing random letters from a bag? But they resolved these questions by reminding their readers that rare is not the same as impossible. As Venn put it [137, p. 349], "A common mistake is to assume that a very unlikely thing will not happen at all." I have yet to find, in the period before the Second World War, a British discussion of the French and Russian viewpoint on Cournot's principle.

With the work of Francis Galton in the 1880s and then Karl Pearson in the 1890s [1, 115, 133], the British began to take a leading role in the application and development of statistics, while remaining less interested in the classical theory of probability. One aspect of this development was the emergence of principles of statistical testing. For those on the continent who subscribed to Cournot's principle, no additional principles were needed to justify rejecting a probabilistic hypothesis that gives small probability to an event we single out in advance and then observe to happen [24]. But in the British tradition, the problem of testing "significance" came to be seen as something separate from the meaning of probability itself [55, 101].

2.5 German philosophy

In contrast with Britain, Germany did see a substantial amount of mathematical work in probability during the first decades of the twentieth century, much of it published in German by Scandinavians and eastern Europeans. But the Germans were already pioneering the division of labor to which we are now accustomed, between mathematicians who prove theorems about probability and philosophers (including philosophically minded logicians, statisticians, and scientists) who analyze the meaning of probability. German philosophers did not give Cournot's principle a central role.

The Germans, like the British, argued vigorously at the end of the nineteenth and beginning of the twentieth century about whether probability is subjective or objective. Karl Friedrich Stumpf is remembered as one the most articulate proponents of subjectivism [135], while Johannes von Kries was the most influential objectivist [141].

Von Kries was the most cogent and influential of all the German philosophers who discussed probability in the late nineteenth century. In his *Principien der Wahrscheinlichkeitsrechnung*, which first appeared in 1886, von Kries rejected the philosophy of Laplace and the mathematicians who followed him. As von Kries pointed out, Laplace and his followers started with a subjective concept of probability but then used observations to validate claims about objective probabilities. They seemed to think that objective probabilities exist and can be the subject of reasoning by Bayes's theorem whenever observations are numerous. This nonsensical law of large numbers, von Kries thought, was the result of combining Bernoulli's theorem with d'Alembert's mistaken belief that small probabilities can be neglected.

Von Kries believed that objective probabilities sometimes exist, but only under conditions where equally likely cases can legitimately be identified. Two conditions, he thought, are needed:

- Each case is produced by equally many of the possible arrangements of the circumstances, and this remains true when we look back in time to earlier circumstances that led to the current ones. In this sense, the relative sizes of the cases are *natural*.
- Nothing besides these circumstances affects our expectation about the cases. In this sense, the Spielräume[2] are *insensitive*.

Von Kries's *principle of the Spielräume* was that objective probabilities can be calculated from equally likely cases when these conditions are satisfied. He considered this principle analogous to Kant's principle that everything that exists has a cause. Kant thought that we cannot reason at all without the principle of cause and effect. Von Kries thought that we cannot reason about objective probabilities without the principle of the Spielräume.

Even when an event has an objective probability, von Kries saw no legitimacy in the law of large numbers. Bernoulli's theorem is valid, he thought, but it tells us only that a large deviation of an event's frequency from its probability is just as unlikely as some other unlikely event, say a long run of successes. What will actually happen is another matter. This disagreement between Cournot and von Kries can be seen as a quibble about words. Do we say that an event will not happen (Cournot), or do we say merely that it is as unlikely as some other event we do not expect to happen (von Kries)? Either way, we proceed as if it will not happen. But the quibbling has its reasons. Cournot wanted to make a definite prediction, because this provides a bridge from probability theory to the world of phenomena—the real world, as those who have not studied Kant would say. Von Kries thought he had a different way of connecting probability theory with phenomena.

Von Kries's critique of moral certainty and the law of large numbers was widely accepted in Germany [72]. In an influential textbook on probability, Emmanuel Czuber named Bernoulli, d'Alembert, Buffon, and De Morgan as advocates of moral certainty and declared them all wrong; the concept of moral certainty, he said, violates the fundamental insight that an event of ever so small a probability can still happen [31, p. 15]. This thought was echoed by the philosopher Alexius Meinong [103, p. 591].

This wariness about ruling out the happening of events whose probability is merely very small did not prevent acceptance of the idea that zero probability represents impossibility. Beginning with Wiman's work on continued fractions in 1900, mathematicians writing in German had worked on showing that various sets have measure zero, and everyone understood that the point was to show that these sets are impossible [9, p. 419]. This suggests a great gulf between zero probability and merely small probability. One does not sense such a gulf in the writings of Borel and his French colleagues; for them, the vanishingly small was merely an idealization of the very small.

[2]In German, Spiel means "game" or "play", and Raum (plural Räume) means "room" or "space". In most contexts, Spielraum can be translated as "leeway" or "room for maneuver". For von Kries, the Spielraum for each case was the set of all arrangements of the circumstances that produce it.

Von Kries's principle of the Spielräume did not endure, for no one knew how to use it. But his project of providing a Kantian justification for the uniform distribution of probabilities remained alive in German philosophy in the first decades of the twentieth century [103, 118]. John Maynard Keynes [73] brought it into the English literature, where it continues to echo, to the extent that today's probabilists, when asked about the philosophical grounding of the classical theory of probability, are more likely to think about arguments for a uniform distribution of probabilities than about Cournot's principle.

2.6 The fracture

The destruction wrought in the 1930s and 1940s by Hitler and Stalin and then by the second world war disrupted or destroyed individuals, families, and nations. It also fractured intellectual traditions. In the case of probability theory, mathematical and philosophical traditions that had thrived in Western Europe gave way to new currents of thought, often centered in the Soviet Union and the United States. The mathematical leadership of Paris gave way to Moscow, where philosophical discourse could be dangerous, and to the United States, where it was often despised. The philosophical traditions of mathematicians in continental Europe faded away as English became the dominant language of philosophy of science, now more heavily influenced by German-speaking philosophers who had escaped from central Europe than by mathematicians of any language. Cournot's principle was one victim of this fracture.

In his *Grundbegriffe der Wahrscheinlichkeitsrechnung*, published in 1933, Kolmogorov had articulated a frequentist interpretation of probability that relied on Cournot's principle. He had stated two principles for interpreting probability; Principle A said that probabilities were approximated by frequencies on repeated trials, and Principle B was the strong form of Cournot's principle, which applies to a single trial. The axiomatization of probability in the *Grundbegriffe*, though it added little to earlier formulations by Fréchet and others [128], was widely acknowledged after the second world war as the definitive mathematical foundation for probability. For a short moment, it appeared that Kolmogorov's prestige might carry Cournot's principle into the new age as well. Harald Cramér repeated Kolmogorov's two principles in his influential *Mathematical Methods in Statistics*, written during the war and published in English in 1946. Hans Richter's 1956 probability textbook, from which West Germans learned the new approach to mathematical probability, also recognized the "Cournotsche Prinzip" as the foundation for applications. But such philosophizing fell out of favor among the new generation of mathematicians. Although Kolmogorov's student Yuri Prokhorov kept it alive in the Soviet encyclopedias [117], there was no mention of Cournot's principle in Boris Gnedenko's Курс теории вероятностей, published in 1950, or in Michel Loève's *Probability Theory*, published in 1955, and I have not seen it in any textbook for mathematical probability after Richter's.

In addition to taking the French probabilists seriously, Kolmogorov also showed interest in debates on the foundations of statistics taking place in the

West after the war [78]. But even with his immense mathematical prestige, he took care to make his philosophical comments brief and incidental to the mathematical theory, and his Soviet colleagues usually dared less [5, 79, 83, 91, 129]. The Moscow probabilists became known and admired abroad as formalists, who showed the rest of the world how mathematical probability could dispense with worries about meaning. This formal spirit took hold even in France , with the rise of Bourbaki, a band of mathematical reformers who often looked askance at the generation of mathematicians represented by Fréchet and Lévy and at probability theory itself [21, 25].

In the United States, a pivotal role was played by the mathematician Joseph Doob, who extended Kolmogorov's formalism to accommodate continuous stochastic processes. An experiment with a random outcome is represented in Kolmogorov's formalism by its set E of possible outcomes, together with a set $\mathfrak{F}$ of subsets of E (the *σ-algebra*) and a real-valued function P on $\mathfrak{F}$. For each $A \in \mathfrak{F}$, $\mathrm{P}(A)$ is the probability of A, which is supposed to approximate the frequency with which the outcome falls in A in repeated trials. In a seminal paper on continuous Markov processes, published in 1931 [76], Kolmogorov had used this framework to discuss transition probabilities—the probabilities for a stochastic process's value $y_{t'}$ at time t' conditional on its value y_t, where $t < t'$. This use of the framework lies within the boundaries of Kolmogorov's Principle A (frequency on repeated trials), at least if we have the means of repeatedly starting the process at any particular value y, for Kolmogorov's frequentism required only that the experiment be susceptible of repetition, not that it actually be repeated. But Kolmogorov never showed how to use his framework to describe probabilities for the entire time series or trajectory $y = \{y_t\}_{0 \le t < \infty}$. He did not, that is to say, define a useful σ-algebra $\mathfrak{F}$ for the case where E consists of possible trajectories.[3] This left a mathematical challenge and also a philosophical question, for in many cases it is unrealistic to talk about repeating an entire time series. One thinks, for example, of the daily Dow Jones average from 1896 to the present; we may want to think that this sequence of numbers is random, but the experiment cannot be repeated.[4]

Doob is celebrated for having met the mathematical challenge; he introduced the concept of a *filtration*, a sequence $\{\mathfrak{F}_t\}$ of σ-algebras that grows with time, reflecting the growth of knowledge as values of the time series previously lying in the future are observed, and he generalized Kolmogorov's concept of conditional expectation of one variable given another to the concept of expectation given each of the $\mathfrak{F}_t$. But what of the philosophical problem? If a time series cannot be repeated, then we cannot interpret the probability for a property of the time series as the frequency with which that property occurs. So how do we interpret it? One answer jumps out of the history I have been recounting.

[3]More precisely, he did not do this for the case of continuous time. He did show, in §I.6 of the *Grundbegriffe*, how to construct a probability measure representing a discrete Markov chain.

[4]Of course, hardly any experiment can be repeated exactly, for chances always vary. This point haunted probability theory from its very beginning with Bernoulli [23, §1.3 of the English translation].

Kolmogorov did not really need both Principle A (the frequency interpretation) and Principle B (Cournot's principle), because Principle A can be derived from Principle B when there are repeated trials (Bernoulli's theorem). All he needed was Cournot's principle, and this is available even when there is only a single trial. It tells us that the meaning of the probability measure P lies in the prediction that a property to which P assigns very small probability will not happen. This is, in fact, how we test a hypothesized probability measure for a stochastic process.

Had he been a colleague of Paul Lévy's, living in a Paris unravaged by Hitler, Doob might have settled on this solution. But he was an American, a pragmatist living in a far different world than Lévy or Kolmogorov. Having himself worked as a statistician, Doob believed that the application of mathematical theory could be left to the practitioner. As he told von Mises in a debate at Dartmouth in 1940, a practitioner must use "a judicious mixture of experiments with reason founded on theory and experience" [46, p. 209]. There is no use in a philosopher telling the practitioner how to use the mathematician's formalism.

Doob's attitude did not prevent philosophers from talking about probability. But as I have already mentioned, English-language philosophy of probability after the second world war was dominated by traditions that had developed in the English and German languages. The German scholars Rudolf Carnap, Hans Reichenbach, and Richard von Mises all settled in the United States on the eve of the second world war and published in English treatises on probability that did not mention Cournot's principle [27, 119, 143].

Because of mathematicians' emphasis on the formal character of Kolmogorov's axioms, the one consensus that emerged in English-language philosophy of probability in the postwar years was that the probability calculus has many interpretations. This idea was first articulated with respect to the new measure-theoretic formalism in 1939 by Ernst Nagel, who listed nine interpretations, including multiple versions of the traditional rivals in the English and German traditions, belief and frequency [105, pp. 40–41].

In this environment, where Cournot's principle was fading away, the one person who bothered to articulate a case against the principle was the subjectivist Bruno de Finetti. De Finetti participated in the 1949 Paris conference where Fréchet coined the name, and he may have been the first to use the name in English, when he deplored "the so-called principle of Cournot" [42]. He did not really disagree with the statement that one should act as if an event with a very small probability should not happen. But he took the principle as a tautology, a consequence of the subjective definition of probability, not a principle standing outside probability theory and relating it to the world [43, p. 235] [39].

The one prominent post-war philosopher who might have been expected to champion Cournot's principle was Karl Popper, who taught that all scientific theories make contact with reality by providing opportunities for falsification. Cournot's principle tells us how to find such an opportunity in a probability model: single out an event of very small probability and see if it happens. Popper was sympathetic with Cournot's principle; this is already clear in his celebrated *Logik der Forschung*, published in 1935 [110, §68]. But the picture is muddied by his youthful ambition to axiomatize probability himself [111] and his later effort to say something original about propensity [112]. In *Realism and the Aim of Science* in 1983, he mentions Cournot's principle in passing but suggests somehow replacing it with a theorem (sic) of Doob's on the futility of gambling strategies [113, p. 379]. Recasting Cournot's principle as a principle about the futility of gambling is the very project to which I now turn, but I cannot support this with an appeal to Popper's authority, for he never seems to have appreciated the principle's historical importance and continuing potential.

3 Cournot's principle in game-theoretic form

As I have explained, Shafer and Vovk [125] revive Cournot's principle in a game-theoretic form: a strategy for placing bets without risking bankruptcy will not multiply the bettor's capital by a large or infinite factor. In the case where the bettor can buy or sell any random variable for its expected value, this is equivalent to the classical form of the principle; Jean Ville demonstrated the equivalence in 1939 [138]. But the game-theoretic principle can also be applied to real markets, where only some payoffs are priced.

This section discusses some of the implications of the game-theoretic principle. After reviewing Ville's theorem (§3.1), I sketch the main contribution of [125], which was to show how the game-theoretic principle extends from classical probability games to more general games and to generalize classical limit theorems accordingly (§§3.2 and 3.3). Then I review more recent work, which shows that good forecasts can be obtained by using a quasi-universal test as a foil (§§3.4 and 3.5). This has profound implications for the interpretation of probability, the practice of statistics, and our understanding of markets. I look at some of the implications for markets in §3.6.

3.1 Ville's theorem

Consider a sequence $Y_1, Y_2, \ldots$ of binary random variables with a joint probability distribution P. Suppose, for simplicity, that P assigns every finite sequence $y_1, \ldots, y_n$ of 0s and 1s positive probability, so that its conditional probabilities for Y_n given values of the preceding variables are always unambiguously defined. Following Jean Ville [138], consider a gambler who begins with \$1 and is allowed to bet as he pleases on each round, provided that he does not risk bankruptcy. We can formalize this with the following protocol, where betting on Y_n is represented as buying some number s_n (possibly zero or negative) of tickets that cost $\$\mathrm{P}\{Y_n = 1 | Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1}\}$ and pay $\$Y_n$.

Binary Probability Protocol
Players: Reality, Skeptic
Protocol:
$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:

Skeptic announces $s_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - \mathrm{P}\{Y_n = 1 | Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1}\})$.

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always nonnegative ($\mathcal{K}_n \geq 0$ for all n) no matter how Reality moves.

This is a perfect-information sequential protocol; moves are made in the order listed, and each player sees the other player's moves as they are made. The sequence $\mathcal{K}_0, \mathcal{K}_1, \ldots$ is Skeptic's capital process.

Ville showed that Skeptic's getting rich in this protocol is equivalent to an event of small probability happening, in the following sense:

1. When Skeptic follows a measurable strategy (a rule that gives s_n as a function of $y_1, \ldots, y_{n-1}$),

$$\mathrm{P}\{\sup_n \mathcal{K}_n \geq \frac{1}{\epsilon}\} \leq \epsilon \tag{1}$$

   for every $\epsilon > 0$. (This is because the capital process $\mathcal{K}_0, \mathcal{K}_1, \ldots$ is a nonnegative martingale; Equation (1) is sometimes called *Doob's inequality.*)

2. If A is a measurable subset of $\{0, 1\}^\infty$ with $\mathrm{P}(A) \leq \epsilon$, then Skeptic has a measurable strategy that guarantees

$$\liminf_{n \to \infty} \mathcal{K}_n \geq \frac{1}{\epsilon}$$

   whenever $(y_1, y_2, \ldots) \in A$.

We can summarize these results by saying that Skeptic's being able to multiply his capital by a factor of $1/\epsilon$ or more is equivalent to the happening of an event with probability ϵ or less.

Although Ville spelled out his theory only for the binary case, he made its generality clear. It applies to the following more general protocol, where prices are regular conditional expected values for a known joint probability distribution P for a sequence of random variables $Y_1, Y_2, \ldots$:

PROBABILITY PROTOCOL
Players: Reality, Skeptic
Protocol:
$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $s_n : \mathbb{R} \to \mathbb{R}$ such that
$\mathrm{E}(s_n(Y_n) | Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1})$ exists.
Reality announces $y_n \in \mathbb{R}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n) - \mathrm{E}(s_n(Y_n) | Y_1 = y_1, \ldots, Y_{n-1} = y_{n-1})$.

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always nonnegative ($\mathcal{K}_n \geq 0$ for all n) no matter how Reality moves.

Here Skeptic can buy any measurable function of Y_n on the nth round for its conditional expected value, provided this expected value exists. In this general protocol, as in the binary protocol, there is a probability of ϵ or less that the capital process for a particular strategy will reach $1/\epsilon$ times its initial value, and there is a zero probability that it will diverge to infinity. Conversely, for any event of probability less than ϵ, there is a strategy whose capital process reaches $1/\epsilon$ times its initial value if the event happens, and for any event of probability zero, there is a strategy whose capital process diverges to infinity if the event happens [125, Chapter 8].

In light of these results, we can put both the finitary and infinitary versions of Cournot's principle in game-theoretic terms:

The finitary principle. Instead of saying that an event of small probability singled out in advance will not happen, we say that a strategy chosen by Skeptic, if it avoids risk of bankruptcy, will not multiply his capital by a large factor.

The infinitary principle. Instead of saying that an event of zero probability singled out in advance will not happen, we say that a strategy chosen by Skeptic, if it avoids risk of bankruptcy, will not make him infinitely rich.

As we will see shortly, the game-theoretic principles can be used in more general protocols, where prices are limited and are not necessarily related to a meaningful probability measure for Reality's moves.

Ville's work was motivated by von Mises's notion of a collective [142–144]. Von Mises had argued that a sequence $y_1, y_2, \ldots$ of 0s and 1s should be considered random if no subsequence with a different frequency of 1s can be picked out by a gambler to whom the ys are presented sequentially; this condition, von Mises felt, would keep the gambler from getting rich by deciding when to bet. Ville showed that von Mises's condition is insufficient, inasmuch as it does not rule out the gambler's getting rich by varying the direction and amount to bet.

Ville was the first to use the concept of a martingale as a tool in probability theory. For him, a martingale was a strategy for the player I have been calling Skeptic. From there, he slipped into also using the word for the player's capital process, for once the initial capital is fixed, the strategies and the capital processes are in a one-to-one correspondence. Doob, who borrowed the concept from Ville [45, 47], made it apparently more suitable for general use by stripping away the betting story; for him, a martingale was a merely sequence $\mathcal{K}_1, \mathcal{K}_2, \ldots$ of random variables such that

$$\mathrm{E}(\mathcal{K}_{n+1}|\mathcal{K}_1 = k_1, \ldots, \mathcal{K}_n = k_n) = k_n.$$

In Doob's later formulation [48], which is now standard in the theory of stochastic processes and the theory of finance, we begin with a filtration $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ in a probability space $(P, \Omega, \mathcal{F})$, and we say that random variables $\mathcal{K}_1, \mathcal{K}_2, \ldots$ form a martingale if $\mathcal{K}_n$ is measurable with respect to $\mathcal{F}_n$ and

$$\mathrm{E}(\mathcal{K}_{n+1}|\mathcal{F}_n) = \mathcal{K}_n$$

for all n.

3.2 The game-theoretic framework

The framework of [125] returns to Ville's game-theoretic version of classical probability theory and generalizes it. The generalization has three aspects:

- Instead of beginning with a probability measure and using its conditional probabilities or expected values as prices on each round, we allow another player, Forecaster, to set the prices as play proceeds. This makes the framework "prequential" [38]; there is no need to specify what the price on the nth round would be had Reality moved differently on earlier rounds.

- When convenient, we make explicit additional information, say x_n, that Reality provides to Forecaster and Skeptic before they make their nth moves.

- We allow the story to be multi-dimensional, with Reality making several moves and Forecaster pricing them all.

A convenient level of generality for the present discussion is provided by the following protocol, where $\mathbb{R}^k$ is k-dimensional Euclidean space, $\mathbf{Y}$ is a subset of $\mathbb{R}^k$, and $\mathbf{X}$ is an arbitrary set.

LINEAR FORECASTING PROTOCOL
Players: Reality, Forecaster, Skeptic
Protocol:
$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots, N$:
Reality announces $x_n \in \mathbf{X}$.
Forecaster announces $f_n \in \mathbb{R}^k$.
Skeptic announces $s_n \in \mathbb{R}^k$.
Reality announces $y_n \in \mathbf{Y}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n \cdot (y_n - f_n)$.
Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always nonnegative ($\mathcal{K}_n \geq 0$ for all n) no matter how the other players move.

Here $s_n \cdot (y_n - f_n)$ is the dot product of the k-dimensional vectors s_n and $y_n - f_n$. Notice also that play stops on the Nth round rather than continuing indefinitely. This is a convenient assumption in this section, where we emphasize the finitary picture; we will return to the infinitary picture later.

The linear forecasting protocol covers many prediction problems considered in statistics (where x and y are often called *independent* and *dependent* variables, respectively) and machine learning (where x is called the *object* and y the *label*) [67, 136, 147]. Market games can be included by taking f_n to be a vector of opening prices and y_n the corresponding vector of closing prices for the nth trading period.

A strategy for Skeptic in the linear forecasting protocol is a rule that gives each of his moves s_n as a function of the preceding moves by Reality and Forecaster, $(x_1, f_1, y_1), \ldots, (x_{n-1}, f_{n-1}, y_{n-1}), x_n, f_n$. A strategy for Forecaster is a rule that gives each of his moves f_n as a function of the preceding moves by Reality and Skeptic, $(x_1, s_1, y_1), \ldots, (x_{n-1}, s_{n-1}, y_{n-1}), x_n$. One way of prescribing a strategy for Forecaster is to choose a probability distribution for $(x_1, y_1), (x_2, y_2), \ldots$ and set f_n equal to the conditional expected value of y_n given $(x_1, y_1), \ldots, (x_{n-1}, y_{n-1}), x_n$. We will look at other interesting strategies for Forecaster in §3.5.

How can one express confidence in Forecaster? The natural way is to assert Cournot's principle: say that a legal strategy for Skeptic (one that avoids $\mathcal{K}_n < 0$ no matter how the other players move) will not multiply Skeptic's initial capital by a large factor.

Once we adopt Cournot's principle in this form, it is natural to scale the implications of our confidence in Forecaster the same way we do in classical probability. This means treating an event that happens only when a specified legal strategy multiplies the capital by $1/\epsilon$ as no more likely than an event with probability ϵ.

To formalize this, consider a possible sequence of moves by Reality and Forecaster,

$$(x_1, f_1, y_1), \ldots, (x_N, f_N, y_N). \tag{2}$$

The space of all such sequences, $\{\mathbf{X} \times \mathbb{R}^k \times \mathbf{Y}\}^N$, is the *sample space* for our protocol, and a subset of it is an *event*. The *upper probability* of an event E is

$$\overline{\mathrm{P}}E := \inf\{\epsilon \mid \text{ Skeptic has a strategy that guarantees } \mathcal{K}_N \geq 1/\epsilon \text{ if Forecaster and Reality satisfy } E \text{ and } \mathcal{K}_N \geq 0 \text{ otherwise}\}.$$

Roughly, $\overline{\mathrm{P}}E$ is the smallest ϵ such that Skeptic can multiply his capital by $1/\epsilon$ if E happens without risking bankruptcy if E fails. When $\overline{\mathrm{P}}E$ is small, we say that E is *morally impossible*.

The *lower probability* of an event E is

$$\underline{\mathrm{P}}E = 1 - \overline{\mathrm{P}}E^{\mathrm{c}},$$

where E^{c} is E's complement with respect to the sample space. When $\underline{\mathrm{P}}(E)$ is close to one, we say that E is *morally certain*.

As in classical probability, we can combine Cournot's principle with a form of Bernoulli's theorem to obtain a statement about relative frequency in a long sequence of events. In a sufficiently long sequence of events with upper probability 0.1 or less, for example, it is morally certain that no more than about 10% of the events will happen [126, §5.3]. This is a martingale-type result; rather than insist that the events be independent in some sense, we assume that the upper probability for each event is calculated at the point in the game where the previous event is settled.

3.3 Extending the classical limit theorems

One of the main contributions of Shafer & Vovk [125] was to show that game theory can replace measure theory as a foundation for classical probability.

We showed in particular that classical limit theorems, especially the strong law of large numbers and the law of the iterated logarithm, can be proven constructively within a purely game-theoretic framework. From Ville's work, we know that for any event with probability zero, there is a strategy for Skeptic that avoids bankruptcy for sure and makes him infinitely rich if the event fails. But constructing the strategy is another matter. In the case of the events of probability zero associated with the classical theorems, we did construct the requisite strategies; they are computable and continuous.

We provided similar constructions for classical results that do not require an infinite number of rounds of play to be meaningful: the weak law of large numbers, finitary versions of the law of the iterated logarithm, and the central limit theorem. The game-theoretic central limit theorem gives conditions under which upper and lower probabilities for the value of an average of many outcomes will approximately coincide and equal the usual probabilities computed from the normal distribution.

The game-theoretic results are more powerful than the measure-theoretic ones in the respects I listed at the beginning of §3.2: the prices can be provided by Forecaster (an actual forecaster or a market) rather than by a probability distribution known in advance, and Forecaster and Skeptic can use information x that is not itself priced or probabilized. In addition, new results emerge when betting is restricted in some way. A new one-sided central limit theorem arises if Forecaster makes only one-sided betting offers [125, Chapter 5], and laws of large numbers can be established for market prices [125, Chapter 15].

3.4 Is there a universal test?

Within the measure-theoretic formalization of probability, it is axiomatic that the union of a countable number of events of probability zero itself has probability zero. In the early twentieth century, there was considerable hesitation about this axiom. Even Émile Borel, who introduced it into probability, was uncomfortable with it. Maurice Fréchet and Bruno de Finetti debated it [40, 57]. It was finally accepted by most mathematicians because it is useful (in proving the limit theorems, for example) and apparently harmless. Because only a finite number of events can be observed, an assumption about infinite collections of events, being untestable, should not get us into trouble [77, p. 14].

Once the countable additivity of probability was more or less universally accepted, it became natural to discuss universal statistical tests. If we imagine that we can observe an infinite sequence of random variables $Y_1, Y_2, \ldots$, then any subset of $\mathbb{R}^\infty$ that is assigned probability zero by a probability measure P on $\mathbb{R}^\infty$ defines a test of the hypothesis that P is the joint probability distribution for $Y_1, Y_2, \ldots$. Given that we have only a finite number of mathematical and logical symbols and can combine them in at most a countable number of ways, we can define at most a countable number of subsets of $\mathbb{R}^\infty$ that have measure zero. Their union, say E, seems to define a universal test: reject P if and only if the observed sequence $y_1, y_2, \ldots$ falls in E. This idea was advanced by Abraham Wald in the 1930s, in defense of von Mises's and Ville's idea of using batteries of tests to define randomness for sequences [154–156] [41, pp. 15–16]. It was given definitive form within measure-theoretic probability in 1966 by Per Martin-Löf, who demonstrated the existence of what he called universal tests of stochasticity [102]. More recently, this concept has been elaborated in the theory of algorithmic complexity [89, §2.5].

The thesis that one can find a universal test appears less plausible when we abandon the infinitary picture for the more realistic finitary picture, where we test using events of small probability rather than events of zero probability. When we consider two events E_1 and E_2 whose probabilities are so small as to make them morally impossible, we will surely say that their disjunction $E_1 \cup E_2$ is also morally impossible, for the sum of two small numbers is also small, even if not quite as small.[5] But we cannot count on the sum of many small numbers being small, and so we cannot say that the union of many morally impossible events is always morally impossible.

The picture is similar in the game-theoretic framework. In this framework, we are testing not a probability distribution but the hypothesis that Forecaster is a good forecaster. A test for Skeptic is not an event with small probability but a strategy for Skeptic that does not risk bankruptcy. We reject the hypothesis when the strategy makes Skeptic sufficiently rich (infinitely rich or many times richer than he was initially). We combine tests not by taking unions of events but by averaging strategies.

Suppose, as usual, that Skeptic starts with 1, and suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are strategies for Skeptic that do not risk bankruptcy. Then $(\mathcal{S}_1 + \mathcal{S}_2)/2$ does not risk bankruptcy, and we can say the following:

Infinitary case. If $\mathcal{S}_1$ or $\mathcal{S}_2$ make Skeptic infinitely rich, then $(\mathcal{S}_1 + \mathcal{S}_2)/2$ will also make him infinitely rich.

Finitary case. If $\mathcal{S}_1$ or $\mathcal{S}_2$ increases Skeptic's capital from 1 to some large number C, then $(\mathcal{S}_1 + \mathcal{S}_2)/2$ will also increase it to a large number, at least $C/2$.

This is parallel to the way testing using $E_1 \cup E_2$ is related to testing using E_1 or E_2. In the infinitary case, we can combine two tests perfectly, obtaining a test that rejects Forecaster's if either of the separate tests rejects. In the finitary case, the combination is not so perfect; the combined test does reject if either of the separate tests rejects, but perhaps not so strongly.

When we look closer, however, the game-theoretic approach provides some new insights.

Consider first the infinitary case. In the measure-theoretic approach pioneered by Borel, Fréchet, and Kolmogorov, we use countable additivity to prove limit theorems such as the law of large numbers. In the game-theoretic approach, we have no such arbitrary general axiom, but there is an obvious way

[5] Kolmogorov made the same point by saying that if two events are both practically certain, then their simultaneous happening is also practically certain, but not quite as certain [77, pp. 4–5]. After the decline of Cournot's priniciple, this came to be seen as paradoxical, as in Kyburg's "lottery paradox" [84].

to try to combine a countable number of strategies $\mathcal{S}_1, \mathcal{S}_2, \ldots$. We try to form a strategy $\mathcal{S}$ by taking the linear combination of these strategies using positive coefficients $\alpha_1, \alpha_2, \ldots$ that add to one. This means that $\mathcal{S}$'s move in the situation $(x_1, f_1, y_1), \ldots, (x_{n-1}, f_{n-1}, y_{n-1}), x_n, f_n$ should be

$$\mathcal{S}((x_1, f_1, y_1), \ldots, (x_{n-1}, f_{n-1}, y_{n-1}), x_n, f_n)$$
$$= \sum_{j=1}^{\infty} \alpha_j \mathcal{S}_j((x_1, f_1, y_1), \ldots, (x_{n-1}, f_{n-1}, y_{n-1}), x_n, f_n). \quad (3)$$

As it turns out, this works for the strategies we need to combine to prove the classical limit theorems (see, e.g., [125, p. 67]), but it does not work for arbitrary strategies $\mathcal{S}_1, \mathcal{S}_2, \ldots$ in arbitrary instances of the linear forecasting protocol, because there is no guarantee that the right hand side of (3) will converge. This vindicates, in some respects, the critics of countable additivity. The general axiom turns out not to be necessary after all.

The new insights provided by the game-theoretic approach in the finitary case are more complicated to explain but also more important. As it turns out, the aspects of disagreement between forecasts f_n and outcomes y_n that we really want to test are relatively limited. To see this, consider binary probability forecasting again:

Binary Probability Protocol with Forecaster and Objects
Players: Reality, Forecaster, Skeptic
Protocol:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:
Reality announces $x_n \in \mathbf{X}$.
Forecaster announces $p_n \in [0, 1]$.
Skeptic announces $s_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - p_n)$.

Restriction on Skeptic: Skeptic must choose the s_n so that his capital is always nonnegative ($\mathcal{K}_n \geq 0$ for all n) no matter how the other players move.

In this protocol, where Forecaster gives a probability p_n on each round, taking into account the previous outcomes $y_1, \ldots, y_{n-1}$ and auxiliary information $x_1, \ldots, x_n$, we are mainly interested in two aspects of the agreement between the probabilities p_n and the outcomes y_n:

Calibration. Whenever there are a large number of rounds on which p_n is close to some fixed probability p^*, we want the frequency with which $y_n = 1$ on those rounds to be approximately equal to p^*.

Resolution. We want this approximate equality between frequency and p^* to remain true when we consider only rounds where p_n is close to p^* and also x_n is close to some fixed value x^* in the object space $\mathbf{X}$.

As it turns out [153], we can often average strategies that reject Forecaster's performance over a grid of values of (x^*, p^*) that are sufficiently dense to capture all deviations of practical interest. This average strategy, which is testing for calibration and resolution, will not necessarily test for more subtle deviations by $y_1, y_2, \ldots$ from the forecasts $p_1, p_2, \ldots$, such as those associated with the law of the iterated logarithm or Ville's refutation of von Mises's theory, but these more subtle deviations may hold little interest. So the average strategy can be regarded, for practical purposes, as a universal test. To avoid confusion, I call it a *quasi-universal strategy.*

3.5 Defensive forecasting

In cases where we have a quasi-universal strategy, a new opportunity opens up for Forecaster. Forecaster will do well enough if he can avoid rejection by that strategy. Formally, he needs a winning strategy in a version of the game where Skeptic is required to follow the quasi-universal strategy but Reality is free to move as she pleases. Does Forecaster have such a winning strategy? The surprising answer is yes.

This is easiest to see in the case where the quasi-universal strategy gives a move for the nth round that is continuous in the forecast p_n. As it happens, this is not an unreasonable requirement. We can construct quasi-universal strategies for calibration and resolution that are continuous in this respect, and there is even a philosophical argument for ruling out any discontinuous strategy for Skeptic: discontinuous functions are not really computable [22, 100].

As it turns out, it is easy to show that for any forecast-continuous strategy for Skeptic there exists a strategy for Forecaster that does not allow Skeptic's capital to grow, regardless of what Reality does. Let me repeat the simple proof given in [148, 153]. It begins by simplifying so that Forecaster's job seems to be even a little harder. Instead of requiring that the entire forecast-continuous strategy for Skeptic be announced at the beginning of the game, we ask only that Skeptic announce his strategy for each round before Forecaster's move on that round. And we drop the restriction that Skeptic avoid risk of bankruptcy. This produces the following protocol:

BINARY FORECASTING AGAINST CONTINUOUS TESTS
Players: Reality, Forecaster, Skeptic
Protocol:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:
Reality announces $x_n \in \mathbf{X}$.
Skeptic announces continuous $S_n : [0, 1] \to \mathbb{R}$.
Forecaster announces $p_n \in [0, 1]$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + S_n(p_n)(y_n - p_n)$.

Here S_n is Skeptic's strategy for the nth round; it gives his move as a function of Forecaster's not-yet-announced move p_n.

Theorem 1 *Forecaster has a strategy that ensures* $\mathcal{K}_0 \geq \mathcal{K}_1 \geq \mathcal{K}_2 \geq \cdots$.

Proof Because S_n is continuous, Forecaster can use the following strategy:

- if the function $S_n(p)$ takes the value 0, choose p_n so that $S_n(p_n) = 0$;
- if S_n is always positive, take $p_n := 1$;
- if S_n is always negative, take $p_n := 0$.

This guarantees that $S_n(p_n)(y_n - p_n) \leq 0$, so that $\mathcal{K}_n \leq \mathcal{K}_{n-1}$. ∎

Some readers may question the philosophical rationale for requiring that S_n be continuous. As it turns out, dropping this requirement does not cost us much; Forecaster can still win if we allow him to randomize [152]. This means that instead of telling Reality his probability p_n, Forecaster may give Reality only a probability distribution P_n for p_n, with the value p_n to be drawn from P_n out of sight of Reality or perhaps after Reality has selected y_n.

A strategy for Forecaster is what one usually calls a probability model; given the previous outcomes $y_1, \ldots, y_{n-1}$ and auxiliary information $x_1, \ldots, x_n$, it gives a probability p_n for $y_n = 1$. Such probabilities can be used in any repetitive decision problem [146]. So Theorem 1's guarantee that they are valid, in the sense that they pass any reasonable test of calibration and resolution, has immense practical significance.

When he follows the strategy described by Theorem 1, is Forecaster using experience of the past to predict the future? He is certainly taking the past into consideration. The moves for Skeptic recommended by the quasi-universal strategy signal emerging discrepancies that Skeptic would like to take advantage of, and the strategy for Forecaster chooses his p_n to avoid extending these discrepancies. But because they succeed regardless of the y_n, it is awkward to call the p_n predictions. The are really only descriptions of the past, not predictions of the future.

The fact that we can always make good probability forecasts undermines some popular ideas about stochasticity. Indeed, to the extent that everything is stochastic, stochasticity has no content. We can still point to quantum mechanics as an extraordinarily successful stochastic theory, whose probabilities appear to withstand all tests, not merely tests of calibration and resolution. Less extreme but also remarkable, there are cases where relatively simple probability models—exchangeable models or Markov models, for example—are successful. In these cases, which go beyond merely being able to give sequential probabilities that beat tests of calibration and resolution, it is reasonable to claim predictive insight; perhaps it is even reasonable to claim that we have caught a glimpse of causal regularities [124]. But bare stochasticity, it seems, is no regularity at all.

3.6 Implications for market prices

Organized exchanges, in which a buyer or seller can always find a ready price for a particular commodity or security, are forecasting games. So we can ask whether Cournot's principle holds in such exchanges, and we can consider the implications of its holding. It is often said that in an efficient market, an investor cannot make a lot of money without taking undue risk. Cournot's principle makes this precise by saying that he will not make a lot of money without risking bankruptcy; he starts with a certain initial capital, and on each round of trading he risks at most a portion of his current capital. In the next section, I will say more about how this formulation relates to established formulations of the efficient-markets hypothesis. Here, in preparation, I explain how Cournot's principle alone can explain certain stylized facts about prices that are often explained using stochasticity.

3.6.1 The $\sqrt{dt}$ effect

Consider first the stylized fact that changes in market prices over an interval of time of length dt scale as $\sqrt{dt}$. In a securities market where shares are traded 252 days a year, for example, the typical change in price of a share from one year to the next is $\sqrt{252}$, or about 16, times as large as the typical change from one day to the next. There is a standard way of explaining this. We begin by assuming that price changes are stochastic, and we argue that successive changes must be uncorrelated; otherwise someone who knew the correlation (or learned it by observation) could devise a trading strategy with positive expected value. Uncorrelatedness of 252 successive daily price changes implies that their sum, the annual price change, has variance 252 times as large and hence standard deviation, or typical value, $\sqrt{252}$ times as large. This is a simple argument, but stochastic ideas intervene in two places, first when price changes are assumed to be stochastic, and then when market efficiency is interpreted as the absence of a trading strategy with positive expected value. As I now explain, we can replace this stochastic argument with a purely game-theoretic argument, in which Cournot's principle expresses the assumption of market efficiency.

For simplicity, consider the following protocol, which describes a market in shares of a corporation. Investor plays the role of Skeptic; he tries to make money, and Cournot's principle says he cannot get very rich following the rules, which do not permit him to risk bankruptcy. Market plays the roles of Forecaster (by giving opening prices) and Reality (by giving closing prices). For simplicity, we suppose that today's opening price is yesterday's closing price, so that Market gives only one price each day, at the end of the day. When Investor holds s_n shares during day n, he makes $s_n(y_n - y_{n-1})$, where y_n is the price at the end of day n.

THE MARKET PROTOCOL
Players: Investor, Market
Protocol:

$\mathcal{K}_0 := 1$.
Market announces $y_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots, N$:
  Investor announces $s_n \in \mathbb{R}$.
  Market announces $y_n \in \mathbb{R}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - y_{n-1})$.

Restriction on Investor: Investor must choose the s_n so that his capital is always nonnegative ($\mathcal{K}_n \geq 0$ for all n) no matter how Market moves.

For simplicity, we ignore the fact that the price y_n of a share cannot be negative.

Since there is no stochastic assumption here, we cannot appeal to the idea of the variance of a probability distribution for price changes to explain what $\sqrt{dt}$ scaling means. But we can use

$$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_n - y_{n-1})^2} \tag{4}$$

as the typical daily change, and we can compare it to the magnitude of the change we see over the whole game, say

$$\max_{0<n\leq N} |y_n - y_0| \tag{5}$$

The quantity (5) should have the same order of magnitude as $\sqrt{N}$ times the quantity (4). Equivalently, we should have

$$\sum_{n=1}^{N}(y_n - y_{n-1})^2 \sim \max_{0<n\leq N}(y_n - y_0)^2, \tag{6}$$

where $\sim$ is understood to mean that the two quantities are of the same order of magnitude.

Does Cournot's principle give us any reason to think that (6) should hold? Indeed it does. As it turns out, Investor has a legal strategy (one avoiding bankruptcy) that makes a lot of money if (6) is violated. Market (who here represents all the other investors and speculators) wants to set prices so that Investor will not make a lot money, and we just saw, in §3.5 that he can more or less do so. So we may expect (6) to hold.

The strategy that makes money if (6) is violated is an average of two strategies, one a momentum strategy (holding more shares after the price goes up), the other a contrarian strategy (holding more shares after the price goes down).

1. The momentum strategy is based on the assumption that Investor can count on $\sum(y_n - y_{n-1})^2 \leq E$ and $\max(y_n - y_0)^2 \geq D$, where D and E are known constants. On this assumption, the strategy is legal and turns \$1 into \$$D/E$ or more for sure.

2. The contrarian strategy is based on the assumption that Investor can count on $\sum(y_n - y_{n-1})^2 \geq E$ and $\max(y_n - y_0)^2 \leq D$, where D and E are known constants. On this assumption, the strategy is legal and turns \$1 into \$$E/D$ or more for sure.

If the assumptions about $\sum(y_n - y_{n-1})^2$ and $\max(y_n - y_0)^2$ fail, then the strategy fails to make money, but Investor can still avoid bankruptcy. For details, see [150].

3.6.2 The game-theoretic CAPM

The Capital Asset Pricing Model (CAPM), popular in finance theory for almost forty years, assumes that a firm whose shares are traded in a securities market has a stable level of risk relative to the market as a whole. The risk for a security s is defined in terms of a probability model for the returns of all the securities in the market; it is the theoretical regression coefficient

$$\beta_s = \frac{\mathrm{Cov}(R_s, R_m)}{\mathrm{Var}(R_m)}, \tag{7}$$

where R_s is a random variable whose realizations are s's returns, and R_m is a random variable whose realizations are a market index's returns.[6] The CAPM says that

$$\mathrm{E}(R_s) = r + \beta_s(\mathrm{E}(R_m) - r), \tag{8}$$

where r is rate of interest on government debt, assumed to be constant [30, p. 197]. Because $\mathrm{E}(R_m) - r$ is usually positive, this equation suggests that securities with higher β have higher average returns. The equation has found only weak empirical confirmation, but it continues to be popular because it suggests plausible ways of analyzing decision problems faced by financial managers.

As it turns out, a purely game-theoretic argument based on Cournot's principle leads to an analogous equation involving only observed returns, with no reference to a probability distribution. The game-theoretic equation is

$$\overline{r}_s \sim r' + b_s(\overline{r}_m - r'), \tag{9}$$

where

$$\overline{r}_s := \frac{1}{N}\sum_{n=1}^{N} s_n, \qquad \overline{r}_m := \frac{1}{N}\sum_{n=1}^{N} m_n,$$

and

$$b_s := \frac{\sum_{n=1}^{N} s_n m_n}{\sum_{n=1}^{N} m_n^2}, \qquad r' := \overline{r}_m - \frac{1}{N}\sum_{n=1}^{N} m_n^2,$$

s_n and m_n being the actual returns of s and the market index, respectively, over period n. This is analogous to (8), inasmuch as r' measures the performance of

[6] Here "return" means simple return; $R = (p_{n+1} - p_n)/p_n$, where p_n is the the price of the share (or the level of the market index) at time n. All expected values, variances, and covariances are with respect to probabilities conditional on information known at time n.

the market as a whole, and the other quantities are empirical analogues of the theoretical quantities in (8).

The interpretation of (9) is similar to the interpretation of the game-theoretic version of $\sqrt{dt}$ scaling, equation (6); a speculator can make money to the extent it is violated. Given the approximations in the derivation of (9), as well as the existence of transaction costs and other market imperfections, we can expect the relation to hold only loosely, but we can ask whether it is any looser in practice than the empirical relations implied by CAPM. If not, then the very approximate confirmation of CAPM that has been discerned in data might be attributed to (9), leaving nothing that can be interpreted as empirical justification for the stochastic assumptions in CAPM. For details, see [149].

4 The return of Cournot's principle

In this concluding section, I discuss what probability, economics, and finance can gain from a revival of Cournot's principle.

4.1 Why probability needs Cournot's principle

Until the middle of the twentieth century, specialists in mathematical probability generally assumed that any probability can be known, either a priori or by observation. Those who understood probability as a measure of belief did not question the presumption that one can know one's beliefs. Those who understood probability as relative frequency assumed that one can observe frequencies. Those who interpreted probability using Cournot's principle did so on the assumption that they would know the probabilities they wanted to test; you would not check whether an event of small probability happened unless you had conjectured it had small probability.

The observations necessary for estimating a numerical probability may be hard to come by. But at worst, Cournot suggested, they could be made by a superior intelligence who represents the limits of what humans can observe [97, pp. 146–150]. Here Cournot was drawing an analogy with the classical understanding of determinism. Classical determinism required more than the future being determined in some theological sense; it required that the future be predictable by means of laws that can be used by a human, or at least by a superior intelligence whose powers of calculation and observation are human-like.

The presumption that probabilities be knowable leads to the apprehension that some events may not have probabilities. Perhaps there are three categories of events:

1. Those we can predict with certainty.
2. Those we can predict only probabilistically.
3. Those that we can predict neither with certainty nor probabilistically.

Most probabilists did think that there are events in the third category. Kolmogorov said so explicitly, and he did not speak of them as events whose probabilities cannot be known; he spoke of them as events that do not have probabilities [81, p. 1]. John Maynard Keynes and R. A. Fisher, each in his own way, also insisted that not every event has a numerical probability [56, 73, 74].

Doob's success in formalizing the concept of a probability measure for an arbitrary stochastic process destabilized this consensus. As I have already emphasized, there are many cases where we cannot repeat an entire stochastic process—cases where there is only one realization, one time series. In these cases, the probability measure assigns probabilities to many events that are not repeated. Having no direct frequency interpretation, these probabilities cannot be verified in any direct way. Because Doob did not appeal to Cournots principle or provide any other guidance about their meaning, his followers looked in other directions for understanding. Many looked towards mechanisms, such as well-balanced dice, that produce or at least simulate randomness. As they saw it, phenomena must be produced in some way. Deterministic phenomena are produced by deterministic mechanisms, indeterministic phenomena by chance mechanisms. The probabilities, even if unverifiable and perhaps unknowable, are meaningful because they have this generative task.

The growing importance of this way of seeing the world is evidenced by a pivotal article published by Jerzy Neyman in 1960 [106]. According to Neyman, science was moving into a period of dynamic indeterminism,

> ...characterized by the search for evolutionary chance mechanisms capable of explaining the various frequencies observed in the development of phenomena studied. The chance mechanism of carcinogenesis and the chance mechanism behind the varying properties of the comets in the Solar System exemplify the subjects of dynamic indeterministic studies. One might hazard the assertion that every serious contemporary study is a study of the chance mechanism behind some phenomena. The statistical and probabilistic tool in such studies is the theory of stochastic processes...

As this quotation confirms, Neyman was a frequentist. But his rhetoric suggests that the initial meaning of probabilities lies in their relation to how phenomena are generated rather than in their relation to frequencies. He wants to explain frequencies, but he does not ask that every probability have a frequency interpretation. Perhaps it is enough that successive probability predictions be well calibrated and have good resolution in the sense explained in §3.4.

What is most striking about Neyman's vision is that stochastic processes appear as the only alternative to deterministic models. The third category of phenomena, those we can predict neither with certainty nor probabilistically, has disappeared. This way of thinking has become ever more dominant since 1960. In many branches of science, we now hear casual references to "true," "physical," or "objective" probabilities, without any hesitation about their existence. An indeterministic process is assumed to be a stochastic process, regardless of whether we do or even can know the probabilities. The naïveté derided by von Kries 120 years ago is once again orthodoxy.

The game-theoretic results reported in §3 provide a framework for regaining the philosophical sophistication of von Kries, Keynes, Fisher, and Kolmogorov, without abandoning the successes achieved by the theory of stochastic processes. Whenever we test a stochastic process empirically, we are applying Cournot's principle to known (hypothetical) probabilities. When we have less than a stochastic process, a model giving only limited prices or probabilities, we can still test it via Cournot's principle, without regarding it as part of some unknowable yet somehow still meaningful full stochastic process.

The results on defensive forecasting reviewed in §3.5 also provide new insights. They show that in a certain limited sense, our third category is indeed empty. Any quantity or event that can be placed in a series (in a time series, not necessarily a series of independent repetitions) can be predicted probabilistically, at least with respect to that series. This suggests that talk about chance mechanisms is also empty. Defensive forecasting works for any time series, regardless of how it is generated. The idea of a chance mechanism adds nothing.

4.2 Why economics needs Cournot's principle

The suggestion that market prices are expected values goes back at least to Cournot [31, Chapter V]. But whose expected values are they? Does the market have a mind? And how do learn more details about the probabilities that presumably accompany these expected values? These questions scarcely troubled Cournot, who still lived in a Laplacean world where one does not fret too much about mankind's ability to discern the proper measure of its ignorance. But as critiques such as that of von Kries accumulated, this tranquility became less and less tenable. By the beginning of the twentieth century, it seemed to require either the philosophical naïveté of a Bachelier [6] or the obscurity of an Edgeworth [104].

Insofar as probability is concerned, today's economists are descendants not of Edgeworth and Keynes, but of Doob, Neyman, and de Finetti. It was Paul Samuelson, perhaps, who first brought Doob's measure-theoretic version of martingales into economics, in the 1965 article now recognized as a first step in the formulation of the concept of informational efficiency [122]. Doob's mathematics was not enough to bring Samuelson back to Cournot's tranquility, as we see in the conclusion of the article:

> I have not here discussed where the basic probability distributions are supposed to come from. In whose minds are they *ex ante*? Is there any *ex post* validation of them? Are they supposed to belong to the market as a whole? And what does that mean? Are they supposed to belong to the "representative individual," and who is he? Are they some defensible or necessitous compromise of divergent expectation patterns? Do price quotations somehow produce a Pareto-optimal configuration of *ex ante* subjective probabilities?

This paper has not attempted to pronounce on these interesting questions.

These questions point towards a deconstruction of probability, but Samuelson's profession has moved in the opposite direction, resolutely embracing, under the slogan of "rational expectations," the hypothesis that the world of trade has true probabilities in the sense of Neyman, happily congruent with its inhabitants' subjective probabilities in the sense of de Finetti [93]. The affect may be bullheadedness rather than tranquility, but economics has reimposed a unity on objective and subjective probability.

The game-theoretic framework allows us to give up this willful reunification, renouncing Cournot's tranquility but retaining Cournot's principle. Once we reclaim Cournot's principle as a way of interpreting probabilities, we can interpret prices directly in the same way; neither a representative individual nor a mind for the market is needed. There is no need to posit probabilities, objective or subjective, that go beyond the prices. These prices are no one's opinions. They are created by competition [68]. They form a martingale not in Doob's sense but in Ville's sense, and because they are the prices at which investors can trade, they generate other martingales in Ville's sense, even though no probabilities intervene. Moreover, because the finitary version of the game-theoretic treatment produces error bounds that tighten as the number of instances grows, we can regain the Knightian distinction between risks faced by an insurance company, which competes in evaluating average outcomes in a long series of trails it observes along with competitors, and the uncertainties faced by an entrepreneur or venture capitalist, who encounters a smaller number of opportunities in a series more privately observed [75].

4.3 Why finance needs Cournot's principle

We can always invent a probability distribution with respect to which prices form a martingale. As Eugene Fama explained in his celebrated 1970 article [52], the central problem in understanding market efficiency is to explain what it means for this probability distribution to reflect fully all available information. Some have questioned whether he ever provided an explanation, but his 1976 finance text [53] added a gloss that has endured: the probabilities should correspond to rational expectations [86]. Somehow, the probabilities should be right. They should connect properly with what really happens in the future.

This brings us back again to Cournot's principle, and to the thesis of this article, that we can test market prices directly using the game-theoretic version of Cournot's principle, without positing full probability distributions.

The results on defensive forecasting that I reported in §3.5 tell us that it is possible to set prices that will foil nearly any trading strategy that does not risk bankruptcy. In markets with reasonable liquidity, there are many speculators trying to do this doable task, and so it is reasonable to hypothesize that they have succeeded. In other words, the game-theoretic version of Cournot's principle is a very plausible hypothesis. But it predicts only a sort of equilibrium in speculation, not an equilibrium among rational investors with well-founded probabilities concerning the risks of different investments and well-defined preferences over those risks [20].

Does it tell us that prices fully reflect all available information? In one sense, yes. It tells us that there is no information available that will allow a speculator to beat the prices. But it does not say that the available information determines prices. It does not rule out there being prices just as consistent with all available information that differ from the actual prices by a factor of 2, or a factor of 3, or a factor of 10. It does not rule out there being prices just as consistent with all the available information that would result in vastly different allocations of capital among competing projects [134]. It does not rule out long slow swings in prices based on no information at all or variability far beyond that justified by the flow of new information [116, 131].

Future empirical work on the hypothesis of informational efficiency in capital markets should, I believe, try to unbundle the game-theoretic hypothesis expressed by Cournot's principle from the quite distinct hypothesis that price changes are largely due to new information. Much empirical work that has already been done, related to the anomalies I have just mentioned, may in the future be seen as first steps in this direction.

Acknowledgements

This paper draws on collaboration with Vladimir Vovk and has benefited from correspondence and conversation with innumerable colleagues, many of whom are thanked in earlier books and papers cited here.

The author has also benefited from the discussion of the paper at the 8th International Conference of the Cournot Centre for Economic Research in Paris, December 1, 2005, especially discussion by André Orléan, Bernard Walliser, and John Vickers. Comments by Alain de Fontenay and Marietta Peytcheva were also helpful.

References

[1] John Aldrich. The language of the English biometric school. *International Statistical Review*, 71(1):109–129, 2003.

[2] Aleksandr D. Aleksandrov, Andrei N. Kolmogorov, and Mikhail A. Lavrent'ev, editors. Математика, ее содержание, методы и значение. Nauka, Moscow, 1956. The Russian edition had three volumes. The English translation, *Mathematics, Its Content, Methods, and Meaning*, was first published in 1962 and 1963 in six volumes by the American Mathematical Society, Providence, RI, and then republished in 1965 in three volumes by the MIT Press, Cambridge, MA. Reprinted by Dover, New York, 1999.

[3] Oskar Nikolaevich Anderson. *Einführung in die mathematische Statistik*. Springer, Vienna, 1935.

[4] Oskar Nikolaevich Anderson. Die Begründung des Gesetzes der grossen Zahlen und die Umkehrung des Theorems von Bernoulli. *Dialectica*, 3(9/10):65–77, 1949.

[5] Anonymous. Account in Russian of a conference on statistics in Moscow in 1954. Вестник статистики *(Bulletin of Statistics)*, 5:39–95, 1954.

[6] Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure, 3e série*, 17:21–86, 1900. This was Bachelier's doctoral dissertation. Reprinted in facsimile in 1995 by Éditions Jacques Gabay, Paris. An English translation, by A. James Boness, appears on pp. 17–78 of *The Random Character of Stock Market Prices*, edited by Paul H. Cootner, MIT Press, Cambridge, MA, 1964.

[7] Raymond Bayer, editor. *Congrès International de Philosophie des Sciences, Paris, 1949; IV: Calcul des Probabilités*. Number 1146 in *Actualités Scientifiques et Industrielles*. Hermann, Paris, 1951.

[8] Jacob Bernoulli. *Ars Conjectandi*. Thurnisius, Basel, 1713. Edith Sylla's English translation, *The Art of Conjecturing, together with Letter to a Friend on Sets in Court Tennis*, was published by Johns Hopkins University Press in 2005. Oscar Sheynin's English translation of Part IV, dated 2005, is at www.sheynin.de.

[9] Felix Bernstein. Über eine Anwendung der Mengenlehre auf ein aus der Theorie der säkularen Störungen herrührendes Problem. *Mathematische Annalen*, 71:417–439, 1912.

[10] Georg Bohlmann. Lebensversicherungs-Mathematik. In *Encyklopädie der mathematischen Wissenschaften, Bd. I, Teil 2*, pages 852–917. Teubner, Leipzig, 1901.

[11] Émile Borel. La valeur pratique du calcul des probabilités. *Revue du mois*, 1:424–437, 1906. Reprinted in [19], Volume 2, pp. 991–1004.

[12] Émile Borel. *Éléments de la théorie des probabilités*. Gauthier-Villars, Paris, 1909. Third edition 1924. The 1950 edition was translated into English by John E. Freund and published as *Elements of the Theory of Probability* by Prentice-Hall in 1965.

[13] Émile Borel. *Le Hasard*. Alcan, Paris, 1914. The first and second editions both appeared in 1914, with later editions in 1920, 1928, 1932, 1938, and 1948.

[14] Émile Borel. Sur les probabilités universellement négligeables. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 190:537–540, 1930. Reprinted as Note IV of [15].

[15] Émile Borel. *Valeur pratique et philosophie des probabilités*. Gauthier-Villars, Paris, 1939. Reprinted in 1991 by Éditions Jacques Gabay.

[16] Émile Borel. *Le jeu, la chance et les théories scientifiques modernes*. Gallimard, Paris, 1941.

[17] Émile Borel. *Les probabilités et la vie*. Presses Universitaires de France, Paris, 1943. Later editions in 1946, 1950, 1958, and 1967. The 1958 edition was translated into English by Maurice Baudin and published as *Probabilities and Life* by Dover, New York, in 1962.

[18] Émile Borel. *Probabilité et certitude*. Presses Universitaires de France, Paris, 1950. An English translation, *Probability and Certainty*, was published in 1963 by Walker, New York.

[19] Émile Borel. *Œuvres de Émile Borel*. Centre National de la Recherche Scientifique, Paris, 1972. Four volumes.

[20] David Bourghelle, Olivier Brandouy, Roland Gillet, and André Orléan, editors. *Croyances, Représentations Collectives et Conventions en France*. Economica, Paris, 2005.

[21] Marcel Brissaud, editor. *Écrits sur les processus aléatoires: Mélanges en hommage à Robert Fortet*. Lavoisier, Paris, 2002.

[22] L. E. J. Brouwer. Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritte. Erster Teil. Allgemeine Mengelehre. *Koninklijke Nederlandse Akademie van Wetenschschappen Verhandelingen*, 5:1–43, 1918.

[23] Bernard Bru. Poisson, le calcul des probabilités et l'instruction publique. In Michel Métivier, Pierre Costabel, and Pierre Dugac, editors, *Siméon Denis Poisson et la science de son temps*, pages 51–94. Éditions de l'École Polytechnique, Palaiseau, 1981. Translated as Poisson, the probability calculus, and public education, in the *Electronic Journal for History of Probability and Statistics*, www.jehps.net, Vol. 1, No. 2, November 2005.

[24] Bernard Bru. Borel, Lévy, Neyman, Pearson et les autres. *MATAPLI*, (60):51–60, octobre 1999.

[25] Bernard Bru. Présentation de l'œuvre de Robert Fortet. In Brissaud [21], pages 19–48.

[26] George-Louis Buffon. Essai d'arithmétique morale. In *Supplément à l'Histoire naturelle*, volume 4, pages 46–148. Imprimerie Royale, Paris, 1777.

[27] Rudolf Carnap. *Logical Foundations of Probabililty*. University of Chicago Press, Chicago, 1950.

[28] Guido Castelnuovo. *Calcolo delle probabilitá*. Albrighi e Segati, Milan, Rome, and Naples, 1919. Second edition in two volumes, 1926 and 1928. Third edition 1948.

[29] Aleksandr Aleksandrovich Chuprov. Очерки по теории статистики *(Essays on the theory of statistics)*. Sabashnikov, Saint Petersburg, second edition, 1910. The first edition appeared in 1909. The second edition was reprinted by the State Publishing House, Moscow, in 1959.

[30] Thomas E. Copeland and J. Fred Weston. *Financial Theory and Corporate Policy*. Addison-Wesley, Reading, MA, third edition, 1988.

[31] Antoine-Augustin Cournot. *Exposition de la théorie des chances et des probabilités*. Hachette, Paris, 1843. Reprinted in 1984 as Volume I (B. Bru, editor) of [32].

[32] Antoine-Augustin Cournot. *Œuvres complètes*. Vrin, Paris, 1973–1984. 10 volumes, with an eleventh to appear.

[33] Harald Cramér. *Mathematical Methods in Statistics*. Princeton University Press, Princeton, NJ, 1946.

[34] Jean d'Alembert. Réflexions sur le calcul des probabilités. In *Opuscule mathématiques*, volume 2, pages 1–25. 1761.

[35] Jean d'Alembert. Doutes et questions sur le calcul des probabilités. In *Mélanges de littérature, d'histoire, et de philosophie*, volume 5, pages 275–304. 1767.

[36] Lorraine Daston. D'Alembert's critique of probability theory. *Historia Mathematica*, 6:259–279, 1979.

[37] Lorraine Daston. How probabilities came to be objective and subjective. *Historia Mathematica*, 21:330–344, 1994.

[38] A. Philip Dawid. Statistical theory: The prequential approach (with discussion). *Journal of the Royal Statistical Society. Series A*, 147:278–292, 1984.

[39] A. Philip Dawid. Probability, causality and the empirical world: A Bayes–de Finetti–Popper–Borel synthesis. *Statistical Science*, 19:44–57, 2004.

[40] Bruno de Finetti. A proposito dell'estensione del teorema delle probabilità totali alle classi numerabili. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:901–905, 1063–1069, 1930.

[41] Bruno de Finetti. *Compte rendu critique du colloque de Genève sur la théorie des probabilités*. Hermann, Paris, 1939. This is the eighth fasicule of [157].

[42] Bruno de Finetti. Recent suggestions for the reconciliation of theories of probability. In Jerzy Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 217–225. University of California Press, Berkeley and Los Angeles, 1951.

[43] Bruno de Finetti. Notes de M. B. de Finetti sur le "Rapport général". In *Les mathématiques et le concret* [63], pages 232–241.

[44] Augustus De Morgan. *An Essay on Probabilities, and on their application to Life Contingencies and Insurance Offices*. Longman, Orme, Brown, Green & Longmans, London, 1838. Reprinted by Arne Press, New York, 1981.

[45] Joseph L. Doob. Regularity properties of certain families of chance variables. *Transactions of the American Mathematical Society*, 47:455–486, 1940.

[46] Joseph L. Doob. Probability as measure. *Annals of Mathematical Statistics*, 12:206–214, 1941. This article stems from a discussion with Richard von Mises at a meeting of the Institute of Mathematical Statistics in Hanover, New Hampshire, in September 1940.

[47] Joseph L. Doob. Application of the theory of martingales. In *Le Calcul des Probabilités et ses Applications*, Colloques Internationaux, pages 23–27. Centre National de la Recherche Scientifique, Paris, 1949.

[48] Joseph L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[49] Francis Y. Edgeworth. *Metretike, or the Method of Measuring Probabilities and Utility*. Temple, London, 1887.

[50] Francis Y. Edgeworth. *Writings in Probability, Statistics, and Economics*. Edward Elgar, Cheltenham, United Kingdom, 1996. Three volumes. Edited by Charles R. McCann, Jr.

[51] Robert Leslie Ellis. On the foundations of the theory of probabilities. *Transactions of the Cambridge Philosophical Society*, 8(1):1–6, 1849. The paper was read on February 14, 1842. Part 1 of Volume 8 was published in 1843 or 1844, but Volume 8 was not completed until 1849.

[52] Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417, 1970.

[53] Eugene F. Fama. *Foundations of Finance*. Basic Books, New York, 1976.

[54] Ronald A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, A*, 222:309–368, 1922.

[55] Ronald A. Fisher. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh, first edition, 1925. Thirteenth edition 1958.

[56] Ronald A. Fisher. *Statistical Methods and Scientific Inference.* Oliver and Boyd, Edinburgh, 1956.

[57] Maurice Fréchet. Sur l'extension du théorème des probabilités totales au cas d'une suite infinie d'événements. *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere*, 63:899–900, 1059–1062, 1930.

[58] Maurice Fréchet. *Généralités sur les probabilités. Variables aléatoires.* Gauthier-Villars, Paris, 1937. This is Book 1 of Fréchet (1937–1938). The second edition (1950) has a slightly different title: *Généralités sur les probabilités. Éléments aléatoires.*

[59] Maurice Fréchet. *Recherches théoriques modernes sur la théorie des probabilités.* Gauthier-Villars, Paris, 1937–1938. This work consists of two books, [58] and [61]. The two books together constitute fasicule 3 of Volume 1 of Émile Borel's *Traité du calcul des probabilités et ses applications.*

[60] Maurice Fréchet. Exposé et discussion de quelques recherches récentes sur les fondements du calcul des probabilités. In Rolin Wavre, editor, *Les fondements du calcul des probabilités*, pages 23–55. 1938. This is the second fascicule of [157].

[61] Maurice Fréchet. *Méthode des fonctions arbitraires. Théorie des événements en chaîne dans le cas d'un nombre fini d'états possibles.* Gauthier-Villars, Paris, 1938. This is Book 2 of Fréchet (1937–1938). Second edition 1952.

[62] Maurice Fréchet. Rapport général sur les travaux du Colloque de Calcul des Probabilités. In Bayer [7], pages 3–21.

[63] Maurice Fréchet. *Les mathématiques et le concret.* Presses Universitaires de France, Paris, 1955.

[64] Maurice Fréchet and Maurice Halbwachs. *Le calcul des probabilités à la portée de tous.* Dunod, Paris, 1924.

[65] Boris V. Gnedenko. Курс теории вероятностей *(Theory of Probability).* Nauka, Moscow, first edition, 1950. Sixth edition 1988.

[66] Jacques Hadamard. Les principes du calcul des probabilités. *Revue de métaphysique et de morale*, 39:289–293, 1922. A slightly longer version of this note, with the title "Les axiomes du calcul des probabilités", was included in *Oeuvres de Jacques Hadamard*, Tome IV, pp. 2161–2162. Centre National de la Recherche Scientifique, Paris, 1968.

[67] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning Theory: Data Mining, Inference, and Prediction.* Springer, New York, 2001.

[68] Frederich A. Hayek. Competition as a discovery procedure. In *New Studies in Philosophy, Politics, Economics and History of Ideas*, pages 179–190. University of Chicago Press, Chicago, 1978.

[69] David Howie. *Interpreting Probability: Controversies and Developments in the Early Twentieth Century*. Cambridge University Press, Cambridge, 2002.

[70] William Stanley Jevons. *The principles of science: a treatise on logic and scientific method*. Macmillan, London, New York, 1874. Two volumes. Second edition 1887. Reprinted by Dover, 1958.

[71] Norman L. Johnson and Samuel Kotz. *Leading Personalities in Statistical Sciences*. Wiley, New York, 1997.

[72] Andreas Kamlah. Probability as a quasi-theoretical concept; J. V. Kries' sophisticated account after a century. *Erkenntnis*, 19:239–251, 1983.

[73] John Maynard Keynes. *A Treatise on Probability*. Macmillan, London, 1921.

[74] John Maynard Keynes. The general theory of employment. *Quarterly Journal of Economics*, 51:209–233, 1937.

[75] Frank H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston, 1921.

[76] Andrei N. Kolmogorov. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104:415–458, 1931. Translated into English on pp. 62–108 of [82].

[77] Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933. An English translation by Nathan Morrison appeared under the title *Foundations of the Theory of Probability* (Chelsea, New York) in 1950, with a second edition in 1956.

[78] Andrei N. Kolmogorov. The main problems of theoretical statistics (abstract). In Второе всесоюзное совещание по математической статистике *(Second National Conference on Mathematical Statistics)*, pages 216–220, Tashkent, 1948. English translation in [130], pp. 219–224.

[79] Andrei N. Kolmogorov. Summary, in Russian, of his address to a conference on statistics in Moscow in 1954. In Anonymous [5], pages 46–47. English translation in [130], pp. 225–226.

[80] Andrei N. Kolmogorov. Теория вероятностей (Probability theory). In Математика, ее содержание, методы и значение, pages 33–71 of Part 4 in the 1963 English edition; 252–284 of Volume 2 in the 1965 English edition. 1956. Chapter XI of [2].

[81] Andrei N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. *Russian Mathematical Surveys*, 38(4):29–40, 1983.

[82] Andrei N. Kolmogorov. *Selected Works of A. N. Kolmogorov. Volume II: Probability Theory and Mathematical Statistics*. Kluwer, Dordrecht, 1992.

[83] Samuel Kotz. Statistics in the USSR. *Survey*, 57:132–141, 1965.

[84] Henry E. Kyburg Jr. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, Connecticut, 1961.

[85] François Le Lionnais, editor. *Les grands courants de la pensée mathématique*. Cahiers du Sud, Paris, 1948. This volume was reprinted by Blanchard in 1962 and by Hermann in 1998. An English translation, *Great Currents of Mathematical Thought*, was published by Dover, New York, in 1971.

[86] Stephen F. LeRoy. Efficient capital markets and martingales. *Journal of Economic Literature*, 27:1583–1621, 1989.

[87] Paul Lévy. *Calcul des probabilités*. Gauthier-Villars, Paris, 1925.

[88] Paul Lévy. *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris, 1937. Second edition 1954.

[89] Ming Li and Paul Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, New York, second edition, 1997.

[90] Michel Loève. *Probability theory: foundations, random sequences*. Van Nostrand, Princeton, NJ, 1955. Second edition 1960.

[91] George G. Lorentz. Mathematics and politics in the Soviet Union from 1928 to 1953. *Journal of Approximation Theory*, 116:169–223, 2002.

[92] Jeff Loveland. Buffon, the certainty of sunrise, and the probabilistic reductio ad absurdum. *Archive for History of Exact Sciences*, 55:465–477, 2001.

[93] Robert E. Lucas Jr. and Edward C. Prescott. Investment under uncertainty. *Econometrica*, 39:659–681, 1971.

[94] Burton G. Malkiel. The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17:59–82, 2003.

[95] Andrei A. Markov. Исчисление вероятностей *(Calculus of Probability)*. Типография Императорской Академии Наук, Saint Petersburg, 1900. Second edition 1908, fourth 1924.

[96] Andrei A. Markov. *Wahrscheinlichkeitsrechnung*. Teubner, Leipzig, 1912. Translation of second edition of [95]. Free on-line at http://historical.-library.cornell.edu.

[97] Thierry Martin. *Probabilités et critique philosophique selon Cournot*. Vrin, Paris, 1996.

[98] Thierry Martin. *Bibliographie cournotienne*. Annales littéraires de l'Université Franche-Comté, Besançon, 1998.

[99] Thierry Martin. Probabilité et certitude. In Thierry Martin, editor, *Probabilités subjectives et rationalité de l'action*, pages 119–134. CNRS Éditions, Paris, 2003.

[100] Per Martin-Löf. *Notes on Constructive Mathematics*. Almqvist & Wiksell, Stockholm, 1970.

[101] Per Martin-Löf. Response to Sverdrup's polemical article, Tests without power. *Scandinavian Journal of Statistics*, 2:161–165, 1975.

[102] Per Martin-Löf. Statistics from the point of view of statistical mechanics. Notes by Ole Jørsboe. Matematisk Institut. Aarhus University. 36 pp, December 1966–February 1967.

[103] Alexius Meinong. *Über Möglichkeit und Wahrscheinlichkeit: Beiträge zur Gegenstandstheorie und Erkenntnistheorie*. Barth, Leipzig, 1915.

[104] Philip Mirowski, editor. *Edgeworth on Chance, Economic Hazard, and Statistics*. Rowman & Littlefield, Lanham, MD, 1994.

[105] Ernst Nagel. *Principles of the Theory of Probability*. University of Chicago Press, Chicago, 1939. Volume I, Number 6 of the International Encyclopedia of Unified Science, edited by Otto Neurath.

[106] Jerzy Neyman. Indeterminism in science and new demands on statisticians. *Journal of the American Statistical Association*, 55:625–639, 1960.

[107] Kh. O. Ondar, editor. *The Correspondence Between A. A. Markov and A. A. Chuprov on the Theory of Probability and Mathematical Statistics*. Springer, New York, 1981. Translated from the Russian by Charles M. and Margaret Stein.

[108] Henri Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica*, 13:1–271, 1890.

[109] Siméon-Denis Poisson. *Recherches sur la probabilité des judgments en matière criminelle et en matière civile, précédés des règles générale du calcul des probabilités*. Bachelier, Paris, 1837.

[110] Karl R. Popper. *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*. Springer, Vienna, 1935. An English translation, *The Logic of Scientific Discovery*, with extensive new appendices, was published by Hutchinson, London, in 1959.

[111] Karl R. Popper. A set of independent axioms for probability. *Mind*, 47:275–277, 1938.

[112] Karl R. Popper. The propensity interpretation of probability. *British Journal for hte Philosophy of Science*, 37:25–42, 1959.

[113] Karl R. Popper. *Realism and the Aim of Science.* Hutchinson, London, 1983.

[114] Theodore M. Porter. *The Rise of Statistical Thinking, 1820–1900.* Princeton University Press, Princeton, NJ, 1986.

[115] Theodore M. Porter. *Karl Pearson: The Scientific Life in a Statistical Age*. Princeton University Press, Princeton, 2004.

[116] James M. Poterba and Lawrence Summers. Mean reversion in stock prices: evidence and implications. *Journal of Financial Economics*, 22:27–59, 1988.

[117] Yuri V. Prokhorov and B. A. Sevast'yanov. Probability theory. In M. Hazewinkel, editor, *Encyclopaedia of Mathematics*, volume 7, pages 307–313. Kluwer, 1987. An updated and annotated translation of the *Soviet Mathematical Encyclopaedia*.

[118] Hans Reichenbach. *Der Begriff der Wahrscheinlichkeit für die mathematische Darstellung der Wirklichkeit.* Barth, Leipzig, 1916.

[119] Hans Reichenbach. *The theory of probability, an inquiry into the logical and mathematical foundations of the calculus of probability.* University of California Press, Berkeley, 1949. English translation, by Ernest H. Hutten and Maria Reichenbach, of *Wahrscheinlichkeitslehre; eine Untersuchung über die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung*, published in Leiden by A. W. Sijthoff, 1935.

[120] Hans Richter. Zur Begründung der Wahrscheinlichkeitsrechnung. *Dialectica*, 8:48–77, 1954.

[121] Hans Richter. *Wahrscheinlichkeitstheorie.* Springer, Berlin, 1956.

[122] Paul A. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6:41–50, Spring 1965.

[123] Eugene Seneta. Boltzmann, Ludwig Edward. In *Leading Personalities in Statistical Sciences* [71], pages 353–354.

[124] Glenn Shafer. *The Art of Causal Conjecture.* The MIT Press, Cambridge, MA, 1996.

[125] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game*. Wiley, New York, 2001.

[126] Glenn Shafer and Vladimir Vovk. On a review by V. N. Tutubalin. http://www.probabilityandfinance.com/reviews/tut_response1.pdf, 2002.

[127] Glenn Shafer and Vladimir Vovk. The origins and legacy of Kolmogorov's *Grundbegriffe*. Working Paper No. 4, http://www.probabilityandfinance.com, 2005.

[128] Glenn Shafer and Vladimir Vovk. The sources of Kolmogorov's *Grundbegriffe*. *Statistical Science*, 2006. To appear.

[129] Oscar Sheynin. *Aleksandr A. Chuprov: Life, Work, Correspondence. The making of mathematical statistics.* Vandenhoeck & Ruprecht, Göttingen, 1996.

[130] Oscar Sheynin, editor. *From Davidov to Romanovsky. More Russian papers on probability and statistics. S. N. Bernstein, B. V. Gnedenko, A. N. Kolmogorov, A. M. Liapunov, P. A. Nekrasov, A. A. Markov, Kh.O. Ondar, T. A. Sarymsakov, N. V. Smirnov.* Hänsel-Hohenhausen, Egelsbach, Germany, 1998b. Translations from Russian into English by the editor. Deutsche Hochschulschriften No. 2579. In microfiche.

[131] Robert J. Shiller. *Market Volatility.* The MIT Press, Cambridge, Massachusetts, 1989.

[132] Robert J. Shiller. From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*, 17:83–104, 2003.

[133] Stephen M. Stigler. *The History of Statistics: The Measurement of Uncertainty before 1900.* Harvard University Press, Cambridge, MA, 1986.

[134] Joseph E. Stiglitz. The allocation role of the stock market: Pareto optimality and competition. *Journal of Finance*, 36:235–251, 1981.

[135] Karl Friedrich Stumpf. Ueber den Begriff der mathematischen Wahrscheinlichkeit. *Sitzungsberichte der philosophisch-philologischen und der historischen Classe der k. b. Adademie der Wissenschaften zu München*, pages 37–120, 1892.

[136] Vladimir Vapnik. *The Nature of Statistical Learning Theory.* Springer, New York, 1996.

[137] John Venn. *The Logic of Chance.* Macmillan, London and New York, third edition, 1888.

[138] Jean-André Ville. *Étude critique de la notion de collectif.* Gauthier-Villars, Paris, 1939. This differs from Ville's dissertation, which was defended in March 1939, only in that a 17-page introductory chapter replaces the dissertation's one-page introduction.

[139] Ladislaus von Bortkiewicz. Anwendungen der Wahrscheinlichkeitsrechnung auf Statistik. In *Encyklopädie der mathematischen Wissenschaften, Bd. I, Teil 2*, pages 821–851. Teubner, Leipzig, 1901.

[140] Guy von Hirsch. Sur un aspect paradoxal de la théorie des probabilités. *Dialetica*, 8:125–144, 1954.

[141] Johannes von Kries. *Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung*. Mohr, Freiburg, 1886. The second edition, which appeared in 1927, reproduced the first without change and added a new 12-page foreward.

[142] Richard von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 5:52–99, 1919.

[143] Richard von Mises. *Wahrscheinlichkeitsrechnung, Statistik und Wahrheit*. Springer, Vienna, 1928. Second edition 1936, third 1951. A posthumous fourth edition, edited by his widow Hilda Geiringer, appeared in 1972. English editions, under the title *Probability, Statistics and Truth*, appeared in 1939 and 1957.

[144] Richard von Mises. *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Deuticke, Leipzig and Vienna, 1931.

[145] Jan von Plato. *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*. Cambridge University Press, Cambridge, 1994.

[146] Vladimir Vovk. Competitive on-line learning with a convex loss function. Working Paper No. 14, http://www.probabilityandfinance.com, 2005.

[147] Vladimir Vovk, Akex Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World*. Springer, New York, 2005.

[148] Vladimir Vovk, Ilia Nouretdinov, Akimichi Takemura, and Glenn Shafer. Defensive forecasting for linear protocols. In *AISTATS 2005: Proceedings of the 10th International Workshop on Artificial Intelligence and Statistics, Barbados, January 6-8, 2005*, pages 365–372, 2005. http://www.gatsby.ucl.ac.uk/aistats/proceedings.htm. Extended version: Working Paper No. 10, http://www.probabilityandfinance.com.

[149] Vladimir Vovk and Glenn Shafer. The game-theoretic capital asset pricing model. Working Paper No. 1, http://www.probabilityandfinance.com, 2002.

[150] Vladimir Vovk and Glenn Shafer. A game-theoretic explanation of the $\sqrt{dt}$ effect. Working Paper No. 5, http://www.probabilityandfinance.com, 2003.

[151] Vladimir Vovk and Glenn Shafer. Kolmogorov's contributions to the foundations of probability. *Problems of Information Transmission*, 39(1):21–31, 2003. Extended version: Working Paper No. 6 at http://www.probabilityandfinance.com.

[152] Vladimir Vovk and Glenn Shafer. Good randomized sequential probability forecasting is always possible. *Journal of the Royal Statistical Society, Series B*, 67:747–764, 2005. Extended version: Working Paper No. 7 at http://www.probabilityandfinance.com.

[153] Vladimir Vovk, Akimichi Takemura, and Glenn Shafer. Defensive forecasting. In *Algorithmic Learning Theory: Proceedings of the 16th International Conference, ALT 2005, Singapore, October 8-11, 2005*, page 459. Springer-Verlag, 2005. Lecture Notes in Computer Science, Volume 3734. Extended version: Working Paper No. 8, http://www.probabilityandfinance.com.

[154] Abraham Wald. Sur la notion de collectif dans le calcul des probabilités. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 202:180–183, 1936.

[155] Abraham Wald. Die Widerspruchfreiheit des Kollectivbegriffes der Wahrscheinlichkeitsrechnung. *Ergebnisse eines Mathematischen Kolloquiums*, 8:38–72, 1937.

[156] Abraham Wald. Die Widerspruchfreiheit des Kollectivbegriffes. In Rolin Wavre, editor, *Les fondements du calcul des probabilités*, pages 79–99. 1938. This is the second fascicule of [157].

[157] Rolin Wavre. *Colloque consacré à la théorie des probabilités.* Hermann, Paris, 1938–1939. The colloquium was held in October 1937 at the University of Geneva. The proceedings were published by Hermann in eight fasicules in their series *Actualités Scientifiques et Industrielles.* The first seven fasicules appeared in 1938 as numbers 734 through 740; the eighth, de Finetti's summary of the colloquium, appeared in 1939 as number 766.