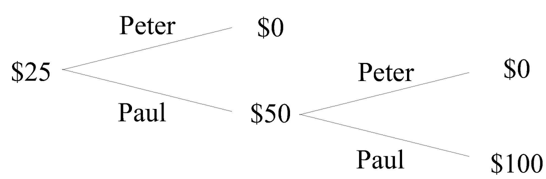
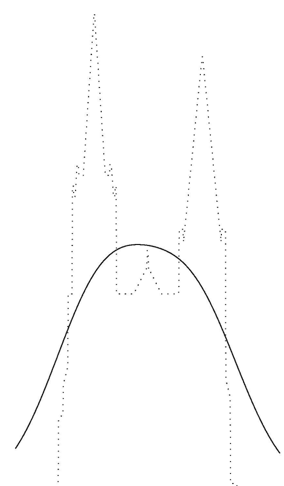
How the game-theoretic foundation for probability resolves the Bayesian vs. frequentist standoff

Glenn Shafer, Rutgers University





The Game-Theoretic Probability and Finance Project

Working Paper #56

First posted August 23, 2020. Last revised August 25, 2020.

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

Abstract

The game-theoretic foundation for probability, which begins with a betting game instead of a mere assignment of probabilities to events, can serve as the basis for all the probability mathematics used by mathematical statistics. It can also generalize frequentist inference so that it stands beside Bayesian inference as a way of using betting. The generalization is vast.

When you bet on an event, the event happens or not; you win or lose. So the frequency of the event is central. But there are more complicated bets. You might put \$10 on the table on the understanding that you will get back either \$0 or \$5 or \$30. This is the first step beyond frequentism. The next step is to recognize that we can test any probability forecaster by betting against their forecasts. It does not matter whether the forecasts are made by a statistical model, a neural net, or a physical model.

This generalization can dissolve the illusion of competing subjective and objective interpretations of probability. A betting game involves two players, one who offers bets and one who selects from the offers. There may be objective and subjective elements on both sides.

1 Introduction

The conventional picture of the Bayesian vs. frequentist divide conflates two distinctions. One concerns the meaning of probability — whether it is subjective or objective. The other concerns a choice between two methods of inference — one that relies on the type of reasoning introduced by Thomas Bayes (thus using Bayes's theorem and conditional probability calculations), and one that relies on the type of reasoning introduced earlier by Jacob Bernoulli (thus using significance tests, confidence intervals, etc.). Bayesians are said to believe that probabilities are subjective and that Bayes's theorem usually suffices for inference. Frequentists are said to believe that probabilities are frequencies and that this justifies Bernoullian methods.

This chapter explains how the game-theoretic foundation for probability, which begins with a betting game instead of a mere assignment of probabilities to events, can dissolve the conflation and transform the entire notion of frequentism, generalizing Bernoullian inference so that it stands beside Bayesian inference as a way of using betting.

The project of using betting games as the starting point for mathematical probability dates back to Jean Ville, who used it in 1939 to generalize and criticize Richard von Mises's earlier project of axiomatizing probability as a theory about frequencies. Ville's idea was further developed by Per Martin-Löf, Claus-Peter Schnorr, Phil Dawid, and others, including Vladimir Vovk and myself. But it is only now being deployed in mathematical statistics.[1]

The simplest kind of bet is an all-or-nothing bet. It has only two outcomes: you win or you lose. You put your money on the table, and you get back a larger amount known in advance or you get back nothing. But there are more complicated bets — bets with many possible outcomes. You might put $10 on the table on the understanding that you will get back either $0 or $5 or $30. When we make an all-or-nothing bet on an event, it is natural focus our minds on the frequency of the event in repeated trials, even if these repeated trials are entirely imaginary. But when we make bets that are not all-or-nothing, we are taking a first step beyond frequentism. This step, which I discuss in §2 and in [22], already permits a substantial generalization of the Bernoullian methodology of significance tests and confidence intervals. And it already begins to dissolve the illusion of competing subjective and objective interpretations of probability.

In §3, I take the next step by pointing out the great generality of testing by betting. We can test any probability forecaster by betting against their forecasts. It does not matter whether the forecaster is a statistical model, a machine-learning method, a physical model,, or a soothsayer. Neither do the topics of the forecasts matter. The forecaster might shift capriciously from one topic to another. So the betting generalization of Bernoullian methodology is

[1]For an overview of Ville's influence, see [4]. Key milestones include [27, 17, 18, 20, 7]. For my work with Vovk, see our recent book, *Game-Theoretic Foundations for Probability and Finance* [25], and the working papers posted at www.probabilityandfinance.com. For the application to mathematical statistics, see especially [22].

vast.

In §4, I define a formal game for testing a probability forecaster and explain how this game can be used as a starting point for a generalization of probability theory, thus providing for the betting method the same mathematical grounding that conventional probability theory provides for the familiar Bernoullian methods.

In §5, I discuss how the same formal game can represent Bayesian inference. The Bayesian picture translates into ways of obtaining and using strategies for the probability forecaster. Other ways of finding such strategies are also available.

I conclude with two sections of a more historical nature. In §6, I document some precedents for the way the adjective "Bernoullian" is used in this chapter. In §7, I discuss why testing by betting, in spite of seeming so natural, has previously been so little used in mathematical statistics.

2 Beyond all-or-nothing bets

For both testing and prediction, as Ville understood, betting generalizes rather than displaces the frequency theory of probability. In the frequentist tradition, we test a probabilistic hypothesis by singling out an event to which it gives small probability; the hypothesis is discredited at least to some extent if the event then happens.[2]

We can make the singling out an event of small probability p more vivid by pretending that the statistician is betting against the event at odds $p : (1-p)$, so that the discredit is associated with winning $(1-p)/p$ times as much as risked. Once we think about testing in this way, it is natural to generalize by allowing the statistician to buy any nonnegative payoff S for the expected value $\mathbf{E}(S)$ the hypothesis assigns it. The discredit is then measured by the factor $S/\mathbf{E}(S)$ by which the bet multiplies the money it risks. Prediction is similarly generalized; once we have provisionally accepted a probabilistic hypothesis, betting against it is associated with predicting outcomes for which the bet will not multiply by a large factor the money risked.

Are the probabilities and expected values in this picture subjective or objective? An answer to this question must begin with the fact that a bet involves at least two parties. Usually one party proposes or offers the bet; the other decides whether to take it. Even when our story about betting is only imagined by a single scientist or statistician, the story involves two parties, the one that offers to sell any nonnegative payoff S for $\mathbf{E}(S)$, and the one who decides which payoff S to buy. There may be elements of subjectivity and objectivity on both sides. The expected values $\mathbf{E}(S)$ may represent the opinions of a single person,

[2] Readers unfamiliar with the frequentist tradition may find this historical fact puzzling or illogical. Perhaps it is best explained as a result of wanting to use Bernoulli's theorem as a justification for equating probability with frequency. That theorem says that the frequency of an event in repeated trials will be as close as desired to its probability *with high probability*, and so equating probability with frequency requires ruling out events of small probability. See [24] for further discussion.

or they may represent the predictions of a well established theory. The test S may be chosen on a whim, or it may be a conventional choice corresponding to a serious alternative. In the end, neither the subjectivity or the objectivity can be eliminated entirely from the story. Betting requires an agent, but the idea that probabilities say something true about the world is inherent in the project of testing them.

When statisticians test with bets that are not all-or-nothing, they have some new and very useful tools. As explained in [23], the factor $S/\mathbf{E}(S)$, considered as measure of the strength of the statistical evidence, is more easily communicated to laypeople than a p-value. We also acquire a logic for sequential testing, because a sequence of bets form a single test when each is allowed to use only the capital remaining from the previous one, the final measure of the evidence against the hypothesis then being the ratio of the final to the initial capital.[3] In this context, moreover, the choice of a test $\mathcal{S}$ implies an alternative hypothesis, thus tightening the logic, linking test outcome to prior expected outcome under an alternative in a way that the standard concepts of p-value and power are not linked.

3 Improvisation: Bob tests Alice

For at least twenty years, since Leo Breiman famously contrasted the culture of modeling with the culture of prediction [5], statisticians have struggled to reconcile the free-wheeling methods of machine learning with the probabilistic vision that still defines mathematical statistics. Even when the predictions produced by dynamic neural networks, for example, take the form of probabilities, these probabilities live outside R. A. Fisher's statistical models and Andrei Kolmogorov's global probability measures [11, 16]. This is also true of the probabilistic predictions made by the physical models that dominate weather prediction. How can we test probabilistic predictions when they are not derived from a statistical model or a probability measure?

This question is easily answered once we accept testing by betting. We can test any probability forecaster by betting against each of its successive predictions, no matter how these predictions are generated and no matter how improvised they are. No model is required.

Suppose Alice announces odds for sports events. One week she looks at the roster for a tennis tournament and assigns each player a probability of winning. The next week she announces probabilities for the outcome of a game between Real Madrid and Barcelona—probabilities for Real Madrid winning, for Barcelona winning, and for a tie. Then she announces a probability distribution for the winning point spread between the Nets and the 76ers. And so on. Bob can challenge Alice's prowess as a probability forecaster by betting at the odds Alice announces. If Bob begins with \$1, bets each time using only what remains

[3] The condition that each bet only risk the capital from the previous one is essential. If you can draw on an unlimited line of credit, then you can always expect to multiply the capital you actually use; see [6].

of this initial capital together with subsequent winnings, and walks away with \$100 after a year of betting, he will have put a big dent in Alice's reputation as a forecaster. Alice may plead that she was merely unlucky, but she cannot claim success as a forecaster.

We can glean some general insights from this example.

1. Bob can challenge and discredit Alice without giving alternative probabilities. He does not even need to believe that there are meaningful or reliable probabilities for the events in question.

2. Bob does not need to risk real money. He can bet with play money. His goal is to make a point, not to get rich. When he uses play money, he does not need a counterparty to his bets. So Alice is not risking real money either; she is risking only her reputation as a forecaster.

3. Even if real money is risked by both parties, we can posit that the amounts are too small to matter. At question is the factor by which Bob manages to multiply the money he risks. If Bob risked only his initial \$1 and ended up with \$100, he could have achieved the same factor of 100 by making bets only 1% as great, risking 1 cent and ending up with \$1.

4. Alice may know more about the sports and the competitors than Bob. If Alice has a good reputation for using knowledge available to her to forecast sports outcomes, and yet Bob succeeds in making money on her forecasts, then we may conjecture that Alice's additional information is not very relevant.

5. Bob may know more about the sports and the competitors than Alice. If Alice has a good reputation as a sports forecaster, and yet Bob makes money on her forecasts, then we may conjecture that things known to Bob but not to Alice are relevant.

6. If Bob does not make money betting against Alice's probabilities—if he begins with \$1 and ends up with only 10 cents or perhaps \$1.10—then we have no evidence against Alice's probabilities. If we know Bob to be very clever and very knowledgeable about the events in question, then this result may be taken as evidence that Alice is doing her job well.

Bob can also test Alice when she gives repeatedly updated probabilities for the same outcome. Suppose, for example, that every week during the football season in the United States she gives probabilities for which team will emerge as the champion of the National Football League. In this case, Bob can test her by buying and selling weekly. Every time Alice assigns a probability to each team being the final champion, Bob buys a payoff that depends on which is the final champion, then sells it back to her the next week at the prices she gives then. If he uses only an initial stake and subsequent winnings for his betting but multiplies that initial stake substantially, then Alice will be discredited. What other way could she be discredited? What other way could any empirical meaning be assigned to her probabilities?

When a forecaster repeatedly changes their probabilities concerning a future event, it may be possible to test their insight even if the event itself is aborted. In early 2020, www.fivethirtyeight.com was regularly giving new probabilities for the season's championship of the National Basketball Association. The betting by testing method might have been used to discredit (or support) this forecaster even though the championship was never settled. The remainder of the season was canceled on March 12 because of the COVID-19 virus.

Financial markets and the prediction markets that imitate them also provide a setting for testing by betting. In the case of financial markets, it is the efficiency of a market rather than the sagacity of a forecaster that is at stake. This way of testing market efficiency has not yet been widely implemented, but see [25, Chapter 16] and [30]. Both financial and prediction markets are constantly being tested by the participants themselves, but the transaction costs and the other restrictions they impose (limitations on short selling in financial markets and limitations on the amounts bet in prediction markets) leave room for thought experiments in the form of imagined testing by betting.

4 Strategies for Bob: Bernoullian inference via game theory

The usual theory of significance testing and confidence intervals is justified by results in mathematical probability — the law of large numbers, the central limit theorem, and all their variations and generalizations. Is there analogous justification for the expansive use of testing by betting advocated here? Yes.

Suppose Alice's announcements always take the form of a probability distribution P on some finite nonempty set $\mathcal{Y}$ and that Bob bets by choosing and buying a payoff $f(y)$, where f is a real-valued function on $\mathcal{Y}$ and $y \in \mathcal{Y}$ is the outcome of the sporting event, not yet known. He pays P's expected value for this payoff, $\mathbf{E}_P(f)$. Alice and Bob alternate moves, say for N rounds. Write $\mathcal{K}_0$ for Bob's initial capital, and $\mathcal{K}_n$ for his capital at the end of the nth round. Write $\mathcal{P}(\mathcal{Y})$ for the set of all probability distributions on $\mathcal{Y}$. Write $\mathbb{N}$ for the natural numbers. With this notation, we can lay out the rules for play in Alice's and Bob's game as follows.

Protocol 1 (Testing Alice's probabilistic forecasts)**.**
Parameter: $N \in \mathbb{N}$
  $\mathcal{K}_0 = 1$.
  FOR $n = 1, 2, \dots, N$*:*
    Alice announces a finite nonempty set $\mathcal{Y}_n$ *and* $P_n \in \mathcal{P}(\mathcal{Y}_n)$.
    Bob announces $f_n : \mathcal{Y}_n \to [0, \infty)$ *such that* $\mathbf{E}_{P_n}(f_n) = \mathcal{K}_{n-1}$.
    Reality announces $y_n \in \mathcal{Y}_n$.
    $\mathcal{K}_n := f_n(y_n)$.

Here, and in the other protocols discussed in this chapter and in [25], each player sees the other players' moves as they are made; the game is one of *perfect* *information*. This does not rule out the possibility that players acquire private information—information not available to the other players—at the outset or as play proceeds.

The requirement that Bob's move f_n always be non-negative and have expected value equal to his current capital $\mathcal{K}_{n-1}$ enforces the condition that he risk no more than his initial capital $\mathcal{K}_0$. He is not be allowed to borrow or otherwise bring more money into the game when play goes badly for him. Because $\mathcal{K}_0 = 1$, his final capital $\mathcal{K}_N$ is the amount by which he multiplies this initial capital.

Within this setup, the mathematical results of probability theory are theorems of game theory — theorems about whether one or more of the players have strategies that guarantee certain goals. Consider, for example, the goal $\mathcal{K}_N \geq 30$ for Bob. He does not have a strategy guaranteed to achieve this goal, because Reality can keep him from ever increasing his capital. But there are many goals he can achieve. For example, if $E \subseteq \mathcal{Y}$, $\epsilon > 0$, and X is the number of the outcomes $y_1, \ldots, y_N$ in E, then Bob has a strategy that will achieve $\mathcal{K}_N \geq N\epsilon^2$ unless the moves by Alice and Reality satisfy

$$\left| X - \frac{\sum_{n=1}^{N} P_n(E)}{N} \right| < \epsilon; \tag{1}$$

see [25, Exercise 2.8]. The agreement between Alice and Reality represented by (1) is one aspect of the game-theoretic law of large numbers. Probability's other classical limit theorems, the law of the iterated logarithm and the central limit theorem in their myriad guises, also have game-theoretic formulations. We can put these game-theoretic results in more familiar form if we set

$$\overline{\mathbb{P}}(A) := \inf \left\{ \alpha \mid \text{Bob can guarantee that } \mathcal{K}_N \geq \frac{1}{\alpha} \text{ when } A \text{ happens} \right\} \tag{2}$$

for every set A of possibilities for the moves $\mathcal{Y}_1, P_1, y_1, \ldots, \mathcal{Y}_N, P_N, y_N$ by Alice and Reality. The existence of a strategy for Bob guaranteeing $\mathcal{K}_N \geq N\epsilon^2$ unless Alice and Reality make (1) happen can then be written as

$$\overline{\mathbb{P}} \left(\left| X - \frac{\sum_{n=1}^{N} P_n(E)}{N} \right| \geq \epsilon \right) \leq \frac{1}{N\epsilon^2}.$$

In [25], Vovk and I vary and generalize Protocol 1 in a great variety of ways. We imagine that play continues indefinitely instead of ending after N rounds. We suppose that N is a stopping time—a rule that stops play when and if moves so far satisfy some condition. We limit the moves one of the players can make or even replace the player with a fixed strategy. We have Reality announce other new information on each round. We introduce other players. We allow Alice to offer fewer bets, in the spirit of imprecise probability [1]. We allow the $\mathcal{Y}_n$ to be infinite. We even consider continuous time. The definition (2) can be adapted to all these generalizations, thus providing a rigorous generalization of probability theory and mathematical statistics.

5 Strategies for Alice, Bayesian and non-Bayesian

Let us simplify Protocol 1 by fixing a finite nonempty set $\mathcal{Y}$ and requiring Alice to always choose $\mathcal{Y}_n$ equal to $\mathcal{Y}$. In this case, a probability distribution P for Reality's moves $y_1, \ldots, y_N$ in Protocol 1 can serve as a strategy for Alice. Alice simply uses P's conditional probability distribution $P(y_n|y_1, \ldots, y_{n-1})$ as her nth move P_n. This is a very special kind of strategy for Alice; when choosing P_n, it takes into account only Reality's previous moves, ignoring Bob's moves and any other information Alice might have or acquire.

Strictly speaking, of course, a probability distribution P for $y_1, \ldots, y_N$ may not define all the needed conditional probabilities; $P(y_n|y_1, \ldots, y_{n-1})$ is not defined when P gives $y_1, \ldots, y_{n-1}$ probability zero. So a more accurate formulation is that a system of conditional probabilities—a family of probability distributions—can serve as a strategy for Alice.

Write $\mathbb{S}$ for the set of all sequences of elements of $\mathcal{Y}$ of length $N-1$ or less; this includes the "empty sequence", denoted by $\Box$. With this notation, we can say that the strategy for Alice is a family $(P_s)_{s \in \mathbb{S}}$ of probability distributions on $\mathcal{Y}$.

Suppose we announce such a strategy for Alice to all the players at the outset and require Alice to play it. This leaves Alice with no role to play. Removing her, we obtain the following protocol.

Protocol 2 (Testing a probability distribution).
Parameters: $N \in \mathbb{N}$, *finite nonempty set* $\mathcal{Y}$, *family* $(P_s)_{s \in \mathbb{S}}$

$\mathcal{K}_0 = 1$.
FOR $n = 1, 2, \ldots, N$*:*
Bob announces $f_n : \mathcal{Y} \to [0, \infty)$ *such that* $\mathbf{E}_{P_{y_1,\ldots,y_{n-1}}}(f_n) = \mathcal{K}_{n-1}$.
Reality announces $y_n \in \mathcal{Y}$.
$\mathcal{K}_n := f_n(y_n)$.

Given the family $(P_s)_{s \in \mathbb{S}}$, we can define a global probability distribution P for $y_1, \ldots, y_N$ using the usual the formula for a joint probability as a product of conditional probabilities:

$$P(y_1, \ldots, y_N) = P(y_1)P(y_2|y_1) \ldots P(y_N|y_1, \ldots, y_{N-1}).$$

Alternatively, we can use the rule of iterated expectation to define the global expectation operator:

$$\mathbf{E}(f(y_1, \ldots, y_N)) := \mathbf{E}_{P_\Box}(\mathbf{E}_{P_{y_1}}(\cdots \mathbf{E}_{P_{y_1,\ldots,y_{N-1}}}(f(y_1, \ldots, y_N)) \cdots)). \quad (3)$$

This manner of defining P corresponds to the reasoning about the value of expectations that we find in the work of Blaise Pascal and other early probabilists (c.f. [25, p. 36]).

Changing your predictions for future events by conditioning a probability distribution on what has happened so far is sometimes considered the essence of

Bayesian inference. Our formulation respects, moreover, the conditions of coherence on which Bayesians insist. When Alice prices all payoffs $f_n : \mathcal{Y} \to [0, \infty)$, allowing Bob to buy or sell f_n at the same price, the prices must cohere by all being expected values with respect to a probability distribution. The analogous weaker condition of coherence used in the theory of imprecise probability is required when Alice gives separate buying and selling prices [25, Chapter 6].

Another aspect of Bayesian inference is the idea of averaging probability distributions. Suppose Alice is uncertain about what probability distribution on $\mathcal{Y}^N$ will resist Bob's efforts but feels confident that one of the distributions in a certain class $(P_\theta)_{\theta\in\Theta}$ would do the job. In this case, the Bayesian method is to average these distributions with respect to some prior distribution on Θ and use the result as P. Under certain conditions, we know, the P thus obtained will do asymptotically as well any P_θ. But this is not, of course, the only way Alice might proceed. Another interesting possibility is to average possible strategies for Bob and play against the average; see [25, Chapter 12].

6 Who said "Bernoullian and Bayesian"?

I have been using the adjective *Bernoullian* rather than *frequentist* to designate statistical methods that follow Jacob Bernoulli's example rather than that of Thomas Bayes. This recognizes Bernoulli as the first to state a theory of direct statistical estimation, just as Bayes was the first to state Bayes's formula. It also allows us to contrast Bernoullian and Bayesian methods without asserting anything about how probabilities are to be interpreted.

I have borrowed this use of "Bernoullian" from some prominent predecessors. Here are some quotations documenting their use of it.

- Francis Edgeworth used *Bernoullian* in this way in 1918, contrasting "the *direct* problem associated with the name of Bernoulli" with "the *inverse* problem associated with the name of Bayes" [10].

- Richard von Mises made a similar remark in German in 1919 ([28], page 5): "Man kann die beiden großen Problemgruppen ... als den Bernoullischen und den Bayesschen Ideenkreis charakterisieren." In English: "We can call the two large groups of problems the Bernoullian and Bayesian circles of ideas."

- A. P. Dempster explicitly advocated the use of "Bernoullian" and "Bayesian" in 1966 [8]. In 1968 [9], in a review of three volumes of collected papers by Jerzy Neyman and E. S. Pearson, Dempster wrote

  > Neyman and Pearson rode roughshod over the elaborate but shaky logical structure of Fisher, and started a movement which pushed the Bernoullian approach to a high-water mark from which, I believe, it is now returning to a more normal equilibrium with the Bayesian view.

- Ian Hacking, probably inspired by Dempster, used *Bernoullian* repeatedly in his 1990 book, *The Taming of Chance* [14]. For example, when discussing Poisson's interest in changes in the chance of conviction by a jury (page 97), he wrote:

  > Laplace had two ways in which to address such questions. One is Bernoullian, and attends to relative frequencies; the other is Bayesian, and is usually now interpreted in terms of degrees of belief. Laplace almost invited his readers not to notice the difference.

The use of "Bernoullian" by Edgeworth and von Mises predated the introduction of the term "frequentist" by Ernest Nagel in 1936 [19]. It is also notable that von Mises, generally recognized in the mid-20th century as the leading proponent of "the frequency theory of probability", always contended that Bayes's formula provides the correct method of statistical inference [29].

7 Why does testing by betting seem novel?

Testing by betting is part of our culture. It is commonplace to challenge a strongly expressed opinion by offering a contrary bet or demanding that the opinionated party offer odds. We know, moreover, that mathematical probability grew out of a calculus for betting. So how can testing by betting seem novel in mathematical statistics? Surely this can only be because mathematicians have deliberately put it out of mind. Suppressed it because it is dangerous. Sometimes too immoral. More often, simply too subjective. We want science, and its handmaiden statistics, to be objective.

The celebrated letters between Blaise Pascal and Pierre Fermat in 1654 was about betting, and Christian Huygens's pamphlet, published a few years later, was about betting — about how stakes should be set for bets in games of chance. Pascal's and Huygens's arguments, moreover, were game-theoretic [21]. Hans Freudenthal and Stephen Stigler have emphasized Huygens's argument for equal chances at a and b being worth $(a+b)/2$ [13, 26]. To have equal chances at a and b, Huygens explains, you may play a game in which you and an opponent both stake $(a+b)/2$, with the winner taking the whole $a+b$ but giving the loser back b. Here equal chances might mean simply that the players are treated equally. Huygens was talking about games of chance, but his reasoning could apply just as well to a game of skill, provided that the two players have agreed to play on even terms.

The concept of probability and even the word "probability" did not appear in Pascal's, Fermat's, and Huygens's writings about their calculus for betting. Many authors did write about probability in the 17th century and earlier, but their probabilities were not numerical and were not modeled on games of chance [12, 15]. This changed only with Bernoulli's *Ars conjectandi*, published in 1713 [2, 3]. In order to make Pascal's and Huygens's calculus into a theory of probability, Bernoulli replaced Huygens's game theory with calculations based on equally possible cases. Why? One obvious motivation was to make probability numbers appear more objective. Betting has an ineradicable subjective element.

Today, when the objectivity of science is under attack, we may feel more tempted than ever to avoid the notion of betting in discussions of statistical testing. But the raging controversies about the meaning of significance testing suggest that this strategy is failing. As I argue in [23], talking about betting may help us communicate the results of statistical tests in a way that better enables the public to understand both their value and their limitations.

References

[1] Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors. *Introduction to Imprecise Probabilities.* Wiley, 2014. 6

[2] Jacob Bernoulli. *Ars Conjectandi.* Thurnisius, Basel, 1713. 9, 10

[3] Jacob Bernoulli. *The Art of Conjecturing, together with Letter to a Friend on Sets in Court Tennis.* Johns Hopkins University Press, Baltimore, 2006. Translation of [2] and commentary by Edith Sylla. 9

[4] Laurent Bienvenu, Glenn Shafer, and Alexander Shen. On the history of martingales in the study of randomness. *Electronic Journal for History of Probability and Statistics*, 5(1), 2009. 1

[5] Leo Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 2001. 3

[6] Harry Crane and Glenn Shafer. Risk is random: The magic of the d'Alembert, 2020. Working Paper 56, www.probabilityandfinance.com. 3

[7] A. Philip Dawid. Statistical theory: The prequential approach (with discussion). *Journal of the Royal Statistical Society. Series A*, 147(2):278–292, 1984. 1

[8] Arthur P. Dempster. Further examples of inconsistencies in the fiducial argument. *Annals of Mathematical Statistics*, 34(3):884–891, 1966. 8

[9] Arthur P. Dempster. Crosscurrents in statistics; Review of *The Selected Papers*, by E. S. Pearson, *Joint Statistical Papers*, by Jerzy Neyman and E. S. Pearson, and *A Selection of Early Statistical Papers*, by J. Neyman. *Science*, 160:661–663, 1968. 8

[10] Francis Edgeworth. Mathematical representation of statistics: A reply. *Journal of the Royal Statistical Society*, 81(2):322–333, 1918. 8

[11] Ronald A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London (A)*, 222:309–368, 1922. 3

[12] James Franklin. *The Science of Conjecture: Evidence and Probability before Pascal*. Johns Hopkins, Baltimore, 2001. Second edition 2015. 9

[13] Hans Freudenthal. Huygens' foundations of probability. *Historia Mathematica*, 7:113–117, 1980. 9

[14] Ian Hacking. *The Taming of Chance*. Cambridge University Press, New York, 1990. 9

[15] Sven K. Knebel. *Wille, Würfel und Wahrscheinlichkeit: Das System der moralischen Notwendigkeit in der Jesuitenscholastik 1550–1700*. Meiner, Hamburg, 2000. 9

[16] 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. 3

[17] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6):602–619, 1966. 1

[18] Per Martin-Löf. The literature on von Mises' Kollectivs revisited. *Theoria*, 35:12–37, 1969. 1

[19] Ernest Nagel. The meaning of probability. *Journal of the American Statistical Association*, 31(193):10–30, 1936. 9

[20] Claus-Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*. Springer, 1971. 1

[21] Glenn Shafer. Pascal's and Huygens's game-theoretic foundations for probability. *Sartoniana*, 32:117–145, 2019. 9

[22] Glenn Shafer. Martingales at the casino. In Laurent Mazliak and Glenn Shafer, editors, *The Splendors and Miseries of Martingales: Their History from the Casino to Mathematics*. Birkhäuser, 2021. 1

[23] Glenn Shafer. Testing by betting: A strategy for statistical and scientific communication (with discussion). *Journal of the Royal Statistical Society, Series A*, to appear, 2021. 3, 10

[24] Glenn Shafer and Vladimir Vovk. The sources of Kolmogorov's *Grundbegriffe*. *Statistical Science*, 21(1):70–98, 2006. 2

[25] Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance*. Wiley, Hoboken, New Jersey, 2019. 1, 5, 6, 7, 8

[26] Stephen M. Stigler. Chance is 350 years old. *Chance*, 20(4):33–36, 2007. 9

[27] Jean Ville. *Étude critique de la notion de collectif*. Gauthier-Villars, Paris, 1939. 1

[28] Richard von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 5:52–99, 1919. 8

[29] Richard von Mises. On the correct use of Bayes' formula. *Annals of Mathematical Statistics*, 13(2):156–165, 1942. 9

[30] Wei Wu and Glenn Shafer. Testing lead-lag effects under game-theoretic efficient market hypotheses, 2007. Working Paper 23, www.probabilityandfinance.com. 5