Using Game Theory to Reunify Subjective and Objective Probability

Glenn Shafer, Rutgers University

December 13, 2019

Seminar in Logic, Games and Language

CUNY Graduate Center

Belief and frequency co-existed in the calculus of games of chance as it was taught in Europe beginning in the 13th century. They still co-existed in the theory of mathematical probability that Jacob Bernoulli based on that calculus. But they came apart in the middle of the 19th century. This talk shows how game theory can bring them back together.

My new book with Volodya Vovk (*Game-Theoretic Foundations for Probability and Finance*, Wiley, May 2019) bases mathematical probability on a game with three players: Forecaster (who offers bets), Skeptic (who decides which offers to take), and Reality (who decides the outcomes). Forecaster is the Bayesian. Skeptic is the frequentist. See the working papers at www.probabilityandfinance.com.

1. Testing a sports forecaster

2. Formalizing the game

3. Strategies for Bob
   (Unifying subjective and objective probability)

4. Strategies for Alice
   (Recovering probability theory)

Testing a sports forecaster

Alice announces probabilities for sports events.

- Week 1: Alice announces a probability of winning for each of the 128 players in the Wimbledon men's singles.
- Week 2 (after the Wimbledon is settled): Alice announces probabilities for a soccer game—P(Real Madrid wins), P(Barcelona wins), P(tie).
- Week 3: Alice announces a probability distribution for the point spread in a game between the Nets and the 76ers.
- And so on.
- Each competition is settled before the next probabilities are announced.

Fragment from Nate Silver's https://fivethirtyeight.com/

UPDATED DEC. 11, 2019, AT 7:22 AM

2019-20 NBA Predictions

TEAM	CONFERENCE	MAKE FINALS	WIN FINALS
Bucks21-3	East	47%	26%
Clippers18-7	West	37%	20%
Rockets15-8	West	29%	15%
76ers18-7	East	26%	12%

- Week 1: Probabilities for each of the 128 Wimbledon men.
- Week 2: P(Real Madrid), P(Barcelona), P(tie).
- Week 3: Probabilities for point spread between Nets and 76ers.

How can you test Alice?

One way is to try to make money at the odds she offers.

Can you think of any other way?

Bob tests Alice by betting.

- Bob bets at the odds Alice announces.

- In other words, Bob buys random variables for their expected values (as given by Alice).

- Example:
  --According to Alice's probabilities, the expected age of the Wimbledon winner is 28.
  --So Bob can buy the age of the winner, in dollars, for $28.

Bob tests Alice by betting.

- When Bob starts with $1,000 and walks away with $100,000, he puts a big dent in Alice's reputation.

- Maybe she was just unlucky, but she cannot claim success as a forecaster.

The only fair measure of Bob's success is **how much he multiplies the money he risks**.

Bob's gaining $99,000 is not impressive if he risked $1 Billion to do it.

But when he multiplies his money by 100, this is something like a "p-value" of 1/100.

- Week 1: Probabilities for each of the 128 Wimbledon men.
- Week 2: P(Real Madrid)=3%, P(Barcelona)=90%, P(tie)=7%.
- Week 3: Probabilities for point spread between Nets and 76ers.
- Week 4: Number of touchdowns by Broncos.

Suppose Bob starts with $100 and does not risk more than that.

- Bob buys age of Wimbledon winner for $28.
  Winner turns out to be 25.
  Now Bob has $97.
- Bob pays $97 for ($0 if Madrid, $100 if Barcelona or tie).
  Madrid wins.
  Now Bob has $0.

Now Bob has to stop betting, because he is out money.
Bob is not allowed to risk more than his original $100.*

*Important detail: Bob is not allowed to borrow more money from someone with infinitely deep pockets; otherwise he could "go martingaling".

Common sense about testing by betting

- Bob need not have alternative probabilities.
  --Maybe he doesn't think there are meaningful probabilities for the events.
- Bob need not risk real money.
  --He can bet with play money. His goal is to make a point, not to get rich.
- Alice need not risk real money either.
  --She is risking only her reputation.
- Alice may know more than Bob.
  --If she has a good reputation, and yet Bob multiplies his money, then maybe her additional information is not very relevant.
- Bob may know more than Alice.
  --If Alice has a good reputation, and yet Bob multiplies his money, maybe his additional information is relevant.
  --If Bob has a good reputation and yet does not multiply his money, maybe Alice is doing a good job.

Bob is more “frequentist” than Bayesian.

Bayesian inference involves bets on hypotheses
—bets that are never settled.

Bob’s bets are settled.

Formalizing the game

Notation

- Finite nonempty set $\mathcal{Y}$
- Probability distribution P on $\mathcal{Y}$
- Real-valued function f on $\mathcal{Y}$
- Bob bets buys the payoff $f(y)$, where $y \in \mathcal{Y}$ is the outcome.
- Bob pays f's expected value under P.
- $\mathcal{P}(\mathcal{Y})$ is the set of all probability distributions on $\mathcal{Y}$.
- $\mathbb{N}$ is the natural numbers, $1, 2, \ldots$.

If f takes only the values 0 and $C > 0$, then buying f is an *all-or-nothing* bet, a bet on the event

$$\{y | f(y) = C\}.$$

Other bets can be compounded from all-or-nothing bets.

Make P do triple duty:

- $P(y)$ is the probability P assigns to $y \in \mathcal{Y}$.
- $P(A)$ is the probability P assigns to $A \subseteq \mathcal{Y}$.
- $P(f)$ is the expected value P assigns to f:

$$P(f) = \sum_{y \in \mathcal{Y}} f(y) P(y).$$

Sometimes we write $P(f(y))$ for $P(f)$.

Rules of play

$\mathcal{K}_0 = 1.$

FOR $n = 1, 2, \ldots, N$:

  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 $P_n(f_n) = \mathcal{K}_{n-1}$.

  Reality announces $y_n \in \mathcal{Y}_n$.

  $\mathcal{K}_n := f_n(y_n)$.

- Alice and Bob alternate moves for N rounds.
- $\mathcal{K}_0$ is Bob's initial capital.
- $\mathcal{K}_n$ for Bob's capital at end of nth round.

$\mathcal{K}_0 = 1$.
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 $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}_n$.
  $\mathcal{K}_n := f_n(y_n)$.

Perfect-information protocol:

Each player sees the other players' moves as they are made.

Players may also acquire private information—information not available to the other players—at the outset or as play proceeds.

Bob is not allowed to risk more than his initial capital $\mathcal{K}_0 = 1$. So $\mathcal{K}_N$ is the factor by which he multiplies the money he risks.

$\mathcal{K}_0 = 1$.
FOR $n = 1, 2, \ldots, N$:
  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 $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}_n$.
  $\mathcal{K}_n := f_n(y_n)$.

We can make the protocol into a game by specifying goals.
For example, we might give Bob the goal $\mathcal{K}_N \geq 30$.

Bob does not have a strategy that guarantees $\mathcal{K}_N \geq 30$, because Reality can keep him from ever increasing his capital.

But Bob can achieve other goals.

Strategies for Bob
(Unifying subjective and objective probability)

Something Bob can achieve

$\mathcal{K}_0 = 1$.
FOR $n = 1, 2, \ldots, N$:
Alice announces a finite nonempty set $\mathcal{Y}_n$ and $P_n \in \mathcal{P}(\mathcal{Y}_n)$.
Glenn announces a subset A_n of $\mathcal{Y}_n$.
Bob announces $f_n : \mathcal{Y}_n \to [0, \infty)$ such that $P_n(f_n) = \mathcal{K}_{n-1}$.
Reality announces $y_n \in \mathcal{Y}_n$.
$\mathcal{K}_n := f_n(y_n)$.

Let X be the fraction of the A_n that happen.

Bob is a frequentist.

He thinks X should approximate the average probability Alice gives to the A_n.

Theorem (Game-Theoretic Law of Large Numbers).
Bob has a strategy that guarantees that either

$$X \approx \frac{\sum_{n=1}^{N} P_n(A_n)}{N}$$

or $\mathcal{K}_N$ is very large.

$\mathcal{K}_0 = 1.$
FOR $n = 1, 2, \ldots, N$:
  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 $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}_n$.
  $\mathcal{K}_n := f_n(y_n)$.

Are Alice's probabilities subjective or objective?

- Because we have more than one player, we can reconcile the two.
- Alice may consider her probabilities subjective.
- But Bob is testing their objectivity.
- We might locate Alice's subjectivity in her willingness to offer bets.
- But digging deeper, we can locate it
  --in her belief that the offers will not allow Bob to multiply his capital and
  --in her consequent belief that things will average out as predicted.

Strategies for Alice
(Recovering probability theory)

$\mathcal{K}_0 = 1$.
FOR $n = 1, 2, \ldots, N$:
  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 $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}_n$.
  $\mathcal{K}_n := f_n(y_n)$.

A probability distribution for Reality's moves $y_1, \ldots, y_N$ is a strategy for Alice: Use the conditional probabilities for y_n given $y_1, \ldots, y_{n-1}$ as P_n.

This is a special kind of strategy for Alice.

- It takes account only of Reality's previous moves, ignoring Bob's moves and any private information Alice might have or acquire.
- It also assumes that $\mathcal{Y}_1, \ldots, \mathcal{Y}_N$ are fixed at the outset.

Simplify by assuming that a finite nonempty set $\mathcal{Y}$ is fixed at the outset, and Alice always chooses $\mathcal{Y}$.

$\mathcal{K}_0 = 1$.
FOR $n = 1, 2, \ldots, N$:
  Alice announces $P_n \in \mathcal{P}(\mathcal{Y})$.
  Bob announces $f_n : \mathcal{Y} \to [0, \infty)$ such that $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}$.
  $\mathcal{K}_n := f_n(y_n)$.

$\mathcal{K}_0 = 1.$
FOR $n = 1, 2, \dots, N$:
  Alice announces $P_n \in \mathcal{P}(\mathcal{Y})$.
  Bob announces $f_n : \mathcal{Y} \to [0, \infty)$ such that $P_n(f_n) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}$.
  $\mathcal{K}_n := f_n(y_n)$.

$P(y_n | y_1, \dots, y_{n-1})$ is not defined when P gives $y_1, \dots, y_{n-1}$ probability zero.

So a more accurate formulation is that a system of conditional probabilities—a family of probability distributions—is a strategy for Alice.

More notation

- $\mathbb{S}$ is the set of all sequences of elements of $\mathcal{Y}$ of length $N-1$ or less.
- This includes the "empty sequence", denoted by $\Box$.
- We call the elements of $\mathbb{S}$ *situations*.
- We sometimes write $y_1 \dots y_n$ for a typical situation, with the understanding that $y_1 \dots y_n = \Box$ when $n = 0$.

The strategy is a family $(P_s)_{s \in \mathbb{S}}$ of probability distributions on $\mathcal{Y}$.

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

$$\mathbb{P}(y_1,\ldots,y_N) = P_{\Box}(y_1)P_{y_1}(y_2)\cdots P_{y_1,\ldots,y_{N-1}}(y_N).$$

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

$$\mathbb{P}(f(y_1,\ldots,y_N)) := P_{\Box}(P_{y_1}(\cdots P(_{y_1,\ldots,y_{N-2}}(P_{y_1,\ldots,y_{N-1}}(y_N))\cdots)).$$

Suppose we announce the strategy $(P_s)_{s\in\mathbb{S}}$ 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.

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 $P_{y_1,\ldots,y_{n-1}}(f_n) = \mathcal{K}_{n-1}$.

Reality announces $y_n \in \mathcal{Y}$.

$\mathcal{K}_n := f_n(y_n)$.

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 $P_{y_1,\ldots,y_{n-1}}(f_n) = \mathcal{K}_{n-1}$.

Reality announces $y_n \in \mathcal{Y}$.

$\mathcal{K}_n := f_n(y_n)$.

Although Alice is no longer in the game, we can imagine her watching what Bob and Reality do.

She has a probability distribution $\mathbb{P}$ for Reality's moves $y_1, \ldots, y_N$.

What does she think Bob can accomplish?

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 $P_{y_1,\ldots,y_{n-1}}(f_n) = \mathcal{K}_{n-1}$.

Reality announces $y_n \in \mathcal{Y}$.

$\mathcal{K}_n := f_n(y_n)$.

Suppose Bob follows a strategy that depends only on Reality's moves $y_1, \ldots, y_N$.

Then $\mathcal{K}_N$ is a function of $y_1, \ldots, y_N$
— a random variable with respect to $\mathbb{P}$.

By the rule of iterated expectation, $\mathbb{P}(\mathcal{K}_N(y_1, \ldots, y_N)) = 1$.

By Markov's inequality, $\mathbb{P}(\mathcal{K}_N \geq C) \leq \dfrac{1}{C}$.

Parameters: $N \in \mathbb{N}$, finite nonempty set $\mathcal{Y}$, family $(P_s)_{s \in \mathbb{S}}$

$\mathcal{K}_0 = 1$.

FOR $n = 1, 2, \dots, N$:

Bob announces $f_n : \mathcal{Y} \to [0, \infty)$ such that $P_{y_1,\dots,y_{n-1}}(f_n) = \mathcal{K}_{n-1}$.

Reality announces $y_n \in \mathcal{Y}$.

$\mathcal{K}_n := f_n(y_n)$.

$$\mathbb{P}\left(\mathcal{K}_N \geq C\right) \leq \frac{1}{C}$$

This relates testing by betting to traditional testing.

- Traditional testing
  —rejects when an event of small probability happens,
  —says a small p-value is negative evidence.
- But when C is large, $\mathcal{K}_N \geq C$ is an event of small probability.
- So testing by betting is a generalization of statistical testing.

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 $P_{y_1, \ldots, y_{n-1}}(f_n) = \mathcal{K}_{n-1}$.

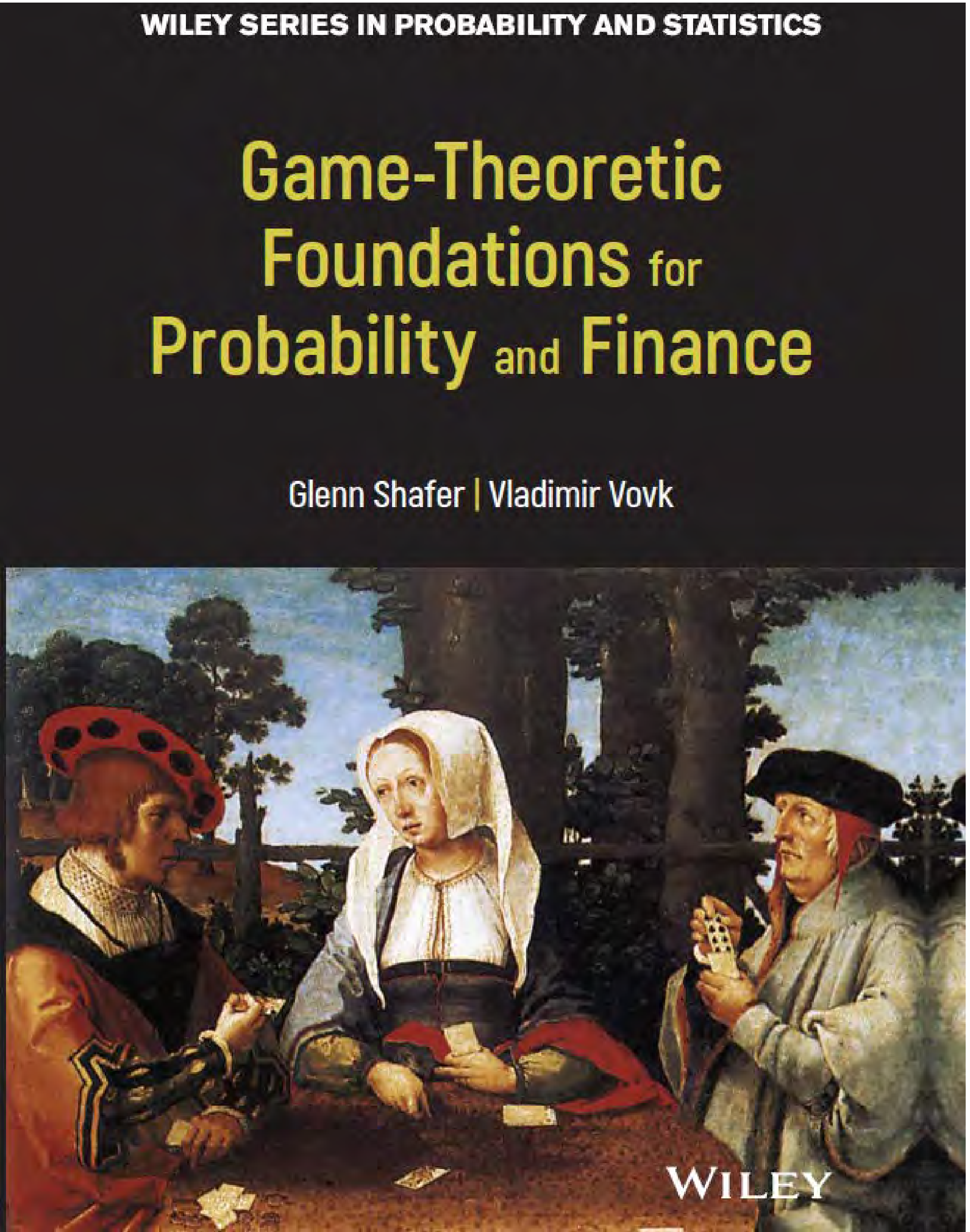
Reality announces $y_n \in \mathcal{Y}$.

$\mathcal{K}_n := f_n(y_n)$.

$$\mathbb{P}(\mathcal{K}_N \geq C) \leq \frac{1}{C}$$

Testing by betting is more general than traditional testing.

- Neither Alice nor Bob have to follow a strategy.
- Alice might offer fewer bets on each round —(imprecise probabilities).
- Alice can change $\mathcal{Y}$ unpredictably.



See also 75 working papers at www.probabilityandfinance.com.

Working Paper 54, “The language of betting as a strategy for statistical and scientific communication”, shows that testing by betting can be used in standard statistical problems, where the betting can be thought of as buying likelihood ratios. This also leads to more flexible methods of meta-analysis.

Working Paper 53, “Pascal’s and Huygens’ game-theoretic foundations for probability”, explains that game-theoretic probability is as old as measure-theoretic probability.

May 2019