Betting interpretations of probability

Glenn Shafer

June 21, 2010

Third Workshop on Game-Theoretic Probability and Related Topics

Royal Holloway, University of London

Outline

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

1. **Probability began with betting.**
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

Beyond any doubt or argument, numerical probability originated in betting.

Henry IV, Part II:

We all that are engaged to this loss
Knew that we ventured on such dangerous seas
That if we wrought out life 'twas ten to one.
And yet we ventured, for the gain proposed
Choked the respect of likely peril feared.

In *The Science of Conjecture: Evidence and Probability before Pascal*, James Franklin reports that "ten to one" appears six times in Shakespeare.

1. Probability began with betting.
2. **To interpret probability, relate a story about betting to the world.**
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

Two ways to interpret probability

1. Consensus view since World War II

- Mathematical probability is a set of axioms.
- **Interpret the AXIOMS.** Relate the axioms and the numbers in some way to the world.

2. My 1990 proposal
   (Can the various meanings of probability be reconciled?)

- Mathematical probability is a story about betting.
- **Interpret the STORY.** Relate the players and their moves in some way to the world.

The consensus view is explained in "Interpretations of Probability" in the online Stanford Encyclopedia of Philosophy.

There are five interpretations of Kolmogorov's axioms:

1. Classical

2. Logical

3. Frequency

4. Propensity

5. Subjective

Betting is relevant only to the subjective interpretation: "Subjective probabilities are traditionally analyzed in terms of betting behavior."

Interpretations of Kolmogorov

1. Classical
2. Logical
3. Frequency
4. Propensity
5. Subjective

What do you think?

- Which ones are "betting interpretations"?
- Is the classical interpretation related to betting? (Pascal thought so.)
- Does the frequency interpretation relate a betting story to something in the world? (Von Mises thought so.)

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. **There is more than one way of doing this.**
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

My 1990 proposal

- Mathematical probability is a story about betting.
- **Interpret the STORY instead of interpreting KOLMOGOROV'S AXIOMS.** Relate the players and their moves in some way to the world.

There is no single *correct* interpretation of probability.

We can improve all interpretations by making the mathematical object being interpreted (related to the world) richer.

My 1990 proposal

- Mathematical probability is a story about betting.
- **Interpret the STORY.** Relate the players and their moves in some way to the world.

The game-theoretic framework sharpens the "story" to a precise game.

The simplest version has three players, who move in order and see each others' moves.

On each round,

Forecaster offers bets.
Skeptic chooses a bet.
Reality decides the outcome.

On each round,

Forecaster offers bets.
Skeptic chooses a bet.
Reality decides the outcome.

In what ways can we relate the players and their moves to the world?

Possible interpretations of Forecaster:

-- a theory (e.g. quantum mechanics)
-- a model or strategy for prediction
-- a market
-- a television weather forecaster

Possible interpretations of Skeptic:

-- a scientist testing a theory or model or weather forecaster
-- a strategy for testing
-- an investor

Possible interpretations of the bets:

-- Someone is willing to offer the bets.
-- Someone is disposed to accept the bets. **[What does "disposed" mean?]**
-- It is safe to offer the bets. **[What does "safe" mean?]**
-- Someone conjectures it is safe to offer the bets.

On each round,

Forecaster offers bets.
Skeptic chooses a bet.
Reality decides the outcome.

This game sharpens the picture advocated by de Finetti beginning in 1930s.

But de Finetti considered only a subjective interpretation.

De Finetti's interpretion is a behavioral theory about YOU (Forecaster).

- Skeptic's only role is to enforce Forecaster's coherence.
- Reality is hardly mentioned.
- There is no clear specification of the players' information or goals.

Modern game theory was just taking shape (von Neumann 1927, Borel 1938).

De Finetti's betting game

Bruno de Finetti, 1937

Three players: Forecaster, Skeptic, and Reality

For each event ***A***, Forecaster announces P(***A***).
For each real ***s***, Forecaster offers to pay ***s*** P(***A***) in order to get ***s*** if ***A*** happens.

Skeptic chooses from these bets.

Reality decides which events happen.

De Finetti is interested in only one interpretation:

- Forecaster (You) is a person in the world.
- Forecaster's offers are his beliefs.
- What properties must these beliefs have to avoid sure loss?

Robert Fortet's betting game (1951)

Two players: Forecaster and Reality

For each $x \leq 1$, Forecaster chooses between

- paying x and getting back 1 if ***A*** happens
- paying 1-x and getting back 1 if ***A*** doesn't happen

or says he cannot choose.

Reality decides whether ***A*** happens or not.

Fortet used this game to define Forecaster's upper and lower probabilities.

Lower probability = upper bound of x for which he prefers the bet on ***A***.

Upper probability = lower bound of x for which he prefers the bet against ***A***.

Smith's betting game

Cedric Smith, 1961

Three players: Bob, Charles, Umpire

Umpire names a positive number ***w***.

Bob says whether he will bet on ***B*** against ***C*** at these odds.

Umpire names a positive stake ***x***, not too large.

Charles pays Bob ***x*** if ***B*** happens.
Bob pays Charles ***wx*** if ***C*** happens.

- Interpretation is similar de Finetti's, inasmuch as we identify Bob with a real person and study his opinions.
- Umpire is our Skeptic, missing from de Finetti.
- Charles is a passive counterparty.
- Reality is implicit.

Walley's betting game [loosening of de Finetti's]

Peter Walley, 1991.

Three players: Forecaster, Skeptic, and Reality

For each uncertain payoff ***X*** and each number ***p***, Forecaster announces whether he is disposed to pay ***p*** for ***X***.

Skeptic decides which dispositions to take advantage of.

Reality decides the values of uncertain payoffs.

Interpretation the same as de Finetti's. But you may get only upper and lower probabilities.

Phil Dawid (2004):

I myself have been enormously influenced by the uncompromisingly personalist approach to Probability set out by de Finetti. At the same time, I do take seriously the criticism that, if this theory is to be more than a branch of Psychology, some further link with external reality is needed.

My way of saying it:

1. An intepretation of probability is a way of linking a betting story to the world.

2. There is no single correct interpretation.

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. **Game theoretic probability teaches us how to use betting strategies as statistical tests.**
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

Measure-theoretic probability begins with a probability space:
• Classical: elementary events with probabilities adding to one.
• Modern: space with filtration and probability measure.

Probability of A = total of probabilities for elementary events that favor A.

Game-theoretic probability begins with a game:
• One player offers prices for uncertain payoffs.
• Another player decides what to buy.

Probability of A = initial stake you need in order to get 1 if A happens.

Theorems in game-theoretic framework

We prove a theorem (e.g., law of large numbers) by constructing a strategy that multiplies the capital risked by a large factor if the theorem fails.

Statistics in game theoretic framework

A statistical test is a strategy for trying to multiply the capital risked.

Cournot's principle

Measure-theoretic: Event of small probability will not happen. Reject the theory if it does.

Game-theoretic: Strategy for multiplying capital by large factor will not succeed. Reject the theory if it does.

Phil Dawid (2004):

We regard a probabilistic theory as falsified if it assigns probability unity to some prespecified theoretical event ***A***, and observation shows that the physical counterpart of the event ***A*** is in fact false.

My way of saying it:

1. We reject a probabilistic theory if it assigns lower probability close to one for a prespecified event that happens.

2. Equivalently, we reject if a strategy for gambling at prices offered by the theory succeeds in multiplying the capital risked by a large factor.

Different from de Finetti:

Instead of relating the betting game to the world by identifying Forecaster with a real person, we relate it to the world by identifying Reality's moves with something that happens in the world and asserting that Forecaster's offers are "safe" in the sense of Cournot's principle.

de Finetti's response, 1955 (my translation from the French):

"When an event has extremely small probability, it is appropriate to act as if it will not happen." In the subjective theory, this does not stop being true…, but it becomes a tautology… The definition of subjective probability is based on the behavior of the person who evaluates it: it is the measure of the sacrifices he thinks appropriate to accept to escape from the damage that would happen along if the event happens (premium for insurance, betting, etc.).

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. **Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.**
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

As an empirical theory, game-theoretic probability makes predictions: A will not happen if there is a strategy that multiplies your capital without risking bankruptcy when A happens.

Defensive forecasting:

Amazingly, predictions that pass all statistical tests are possible (defensive forecasting).

Defensive forecasting

Under repetition, good probability forecasting is possible.

- We call it defensive because it defends against a quasi-universal test.

- Your probability forecasts will pass this test even if reality plays against you.

Why Hilary Putnam thought good probability prediction is impossible. . .

> FOR $n = 1, 2, \ldots$
> Forecaster announces $p_n \in [0, 1]$.
> Skeptic announces $s_n \in \mathbb{R}$.
> Reality announces $y_n \in \{0, 1\}$.
> Skeptic's profit $:= s_n(y_n - p_n)$.

Reality can make Forecaster uncalibrated by setting

$$y_n := \begin{cases} 1 & \text{if } p_n < 0.5 \\ 0 & \text{if } p_n \geq 0.5 \end{cases}$$

Skeptic can then make steady money with

$$s_n := \begin{cases} 1 & \text{if } p < 0.5 \\ -1 & \text{if } p \geq 0.5 \end{cases}$$

But Skeptic's move

$$s_n = \begin{cases} 1 & \text{if } p < 0.5 \\ -1 & \text{if } p \geq 0.5 \end{cases}$$

is discontinuous in p. This infinitely abrupt shift—an artificial idealization—is crucial to the counterexample.

Forecaster can defeat any strategy for Skeptic if

- the strategy for Skeptic is continuous in p, or
- Forecaster is allowed to randomize, announcing a probability distribution for p rather than a sharp value for p.

See Working Papers 7 & 8 at www.probabilityandfinance.com.

Nuances:

1. When a probabilistic theory successfully predicts a long sequence of future events (as quantum mechanics does), it tells us something about phenomena.

2. When a probabilistic theory predicts only one step at a time (basing each successive prediction on what happened previously), it has practical value but tells us nothing about phenomena. Defensive forecasting pass statistical tests *regardless of how events come out*.

3. When we talk about the probability of an isolated event, which different people can place in different sequences, we are weighing arguments. This is the place of evidence theory.

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. **Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.**
6. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

Phil Dawid (2004):

…causal modeling and inference do not differ in any qualitative way from regular probabilistic modeling and inference. Rather, they differ quantitatively, because they have a different scope and ambition: namely, to identify and utilize relationships that are *stable* over a shifting range of environments ("*regimes*").

My way of saying it:

1. Successful prediction with feedback has little causal content, because it depends on the sequence of situations, however chosen.

2. Only a prediction strategy that is not adaptive (does not depend on feedback) can claim causal meaning.

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. **Assessment of evidence requires judgement outside any framework for repetition.**
8. **Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).**
9. Non-additive belief functions also rely on constitutive judgments of irrelevance.

De Moivre's argument for $\mathrm{P}(A\&B) = \mathrm{P}(A)\mathrm{P}(B|A)$

Assumptions

1. $\mathrm{P}(A) =$ price of a ticket that pays 1 if A happens.
2. $\mathrm{P}(A)x =$ price of a ticket that pays x if A happens.
   (Here x can be any real number.)
3. After A happens (we learn A and nothing else),
   $\mathrm{P}(B|A)x =$ price of a ticket that pays x if B happens.

Argument

1. Pay $\mathrm{P}(A)\mathrm{P}(B|A)$ to get $\mathrm{P}(B|A)$ if A happens. If A does happen, pay $\mathrm{P}(B|A)$ to get 1 if B happens.
2. So $\mathrm{P}(A)\mathrm{P}(B|A)$ is the cost of getting 1 if $A\&B$ happens.

De Finetti's adopted De Moivre's argument for $\mathrm{P}(A\&B) = \mathrm{P}(A)\mathrm{P}(B|A)$, changing "price" to "an individual's price".

Assumptions

1. $\mathrm{P}(A)x$ = price at which I will sell a ticket that pays x if A happens.
2. After A happens (we learn A and nothing else), $\mathrm{P}(B|A)x$ = price at which I will sell a ticket that pays x if B happens.

Argument

1. You pay me $\mathrm{P}(A)\mathrm{P}(B|A)$ to get $\mathrm{P}(B|A)$ if A happens. If A does happen, you pay me $\mathrm{P}(B|A)$ to get 1 if B also happens.
2. So $\mathrm{P}(A)\mathrm{P}(B|A)$ is what you need to pay me to get 1 if $A\&B$ happens.

The game-theoretic argument for $\mathsf{P}(B|A) = \frac{\mathsf{P}(A\&B)}{\mathsf{P}(A)}$

Context Winning against given prices means multiplying your capital by a large factor buying or selling the tickets priced (and others like them in the long run).

Hypothesis You will not win against $\mathsf{P}(A)$ and $\mathsf{P}(A\&B)$.

Conclusion You still will not win if after A (and nothing else) is known, $\mathsf{P}(A\&B)/\mathsf{P}(A)$ is added as a new probability for B.

How to prove it Show that if $\mathcal{S}$ is a strategy against all three probabilities, then there exists a strategy $\mathcal{S}'$ against $\mathsf{P}(A)$ and $\mathsf{P}(A\&B)$ alone that risks the same risks and payoffs.

Proof: Let M be the amount of B tickets $\mathcal{S}$ buys after learning A. To construct $\mathcal{S}'$ from $\mathcal{S}$, delete these B tickets and add

$$M \text{ tickets on } A\&B \quad \text{and} \quad -M\frac{\mathsf{P}(A\&B)}{\mathsf{P}(A)} \text{ tickets on } A$$

to $\mathcal{S}$'s purchases in the initial situation.

- The tickets added have zero total initial cost:

$$M\mathsf{P}(A\&B) - M\frac{\mathsf{P}(A\&B)}{\mathsf{P}(A)}\mathsf{P}(A) = 0.$$

- The tickets added and the tickets deleted have the same net payoffs:

$$\begin{array}{ll} 0 & \text{if } A \text{ does not happen;} \\ -M\dfrac{\mathsf{P}(A\&B)}{\mathsf{P}(A)} & \text{if } A \text{ happens but not } B; \\ M\left(1-\dfrac{\mathsf{P}(A\&B)}{\mathsf{P}(A)}\right) & \text{if } A \text{ and } B \text{ both happen.} \end{array}$$

Comments

1. Game-theoretic advantage over de Finetti: the condition that we learn only A and nothing else (relevant) has a meaning without a prior protocol (see my 1985 article on conditional probability).

2. Winning against probabilities by multiplying the capital risked over the long run: To understand this fully, learn about game-theoretic probability.

1. Probability began with betting.
2. To interpret probability, relate a story about betting to the world.
3. There is more than one way of doing this.
4. Game theoretic probability teaches us how to use betting strategies as statistical tests.
5. Under repetition with feedback, good probability forecasting is possible. Non-additivity is not needed.
6. Probabilities merit being called *causal* if they make good long-run predictions (pass statistical tests) without feedback. Causal probabilities may be additive or non-additive.
7. Assessment of evidence requires judgement outside any framework for repetition.
8. Conditional probability (Bayes) relies on constitutive judgements of irrelevance (independence).
9. **Non-additive belief functions also rely on constitutive judgments of irrelevance.**

A betting interpretation for probabilities and Dempster-Shafer degrees of belief

- On-line in *International Journal of Approximate Reasoning*.
- At www.probabilityandfinance.com as working paper 31.
- Or arXiv:1001.1653v1