

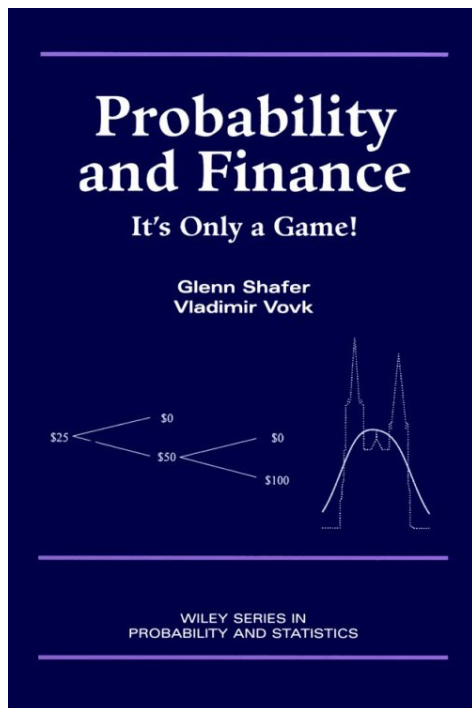
Basics and Interpretation of Game-Theoretic Probability

Glenn Shafer

GTP 2012 Tokyo

November 12, 2012

1. Game theoretic probability
2. Game theoretic upper and lower probability
3. Defensive forecasting
4. Probability judgment



Glenn Shafer
Vladimir Vovk

Wiley
2001

2nd edition in preparation

Shafer and Vovk (2001): base probability on game theory instead of measure theory.

- **To test a probabilistic theory:** Bet at the odds it gives.
- **To prove that something happens with probability one:** Devise a strategy that multiplies the capital it risks by a large factor if the theorem fails.

These ideas can be applied to forecasting, causality, and inference.

1. Game theoretic probability
2. Game theoretic upper and lower probabilities
3. Defensive forecasting
4. Probability judgment

- **Probabilities derive from betting offers.**
Not from the measure of sets.
- **Test probabilities by betting.**
Refute alleged probabilities by making money.
- **Prove “probability one” by betting strategy.**
 - Do not say that the property will fail on a set of measure zero.
 - Say that its failure will mean the success of a betting strategy.

- **Probabilities derive from betting offers.**
Not from the measure of sets.
- **Test probabilities by betting strategy.**
Refute alleged probabilities by making money.

Qualification:

Not all probabilities are testable assertions.

We test probabilities (or betting offers or other prices) when they are offered as theories about the world. Their meaning in this case is that no one will profit unusually by taking advantage of these prices.

But I can offer you prices without claiming that you will not make a lot of money at those prices.

1. Game theoretic probability
2. Game theoretic upper and lower probabilities
3. Defensive forecasting
4. Probability judgment

Probabilities derive from betting offers.

The offers may determine less than a probability distribution.

1. The stock market gives a price but not a probability distribution for tomorrow's value of a share of Google.

2. A forecaster who gives a probability for rain tomorrow every day for a year does not give a joint probability distribution for the 365 outcomes.

In such cases, we get only upper and lower probabilities.

1. Game theoretic probability
2. Game theoretic upper and lower probability
3. **Defensive forecasting**
4. Probability judgment

- In the game-theoretic framework, it can be shown that **good probability forecasting is possible**.
- Once a sequence of events is fixed, **you can give probabilities that pass statistical tests**.
- The forecasting **defends** against the tests.

Defensive forecasting continued

- **Once a sequence of events is fixed**, you can give probabilities that pass statistical tests.
- **The only role of the observer** is to place the event in a sequence.
- Advance knowledge is not needed.
- The sequence need not be “iid”; this concept is not even defined.

1. Game theoretic probability
2. Game theoretic upper and lower probability
3. Defensive forecasting
4. **Probability judgment**

Jeyzy Neyman's **inductive behavior**

A statistician who makes predictions with 95% confidence has two goals:
be informative
be right 95% of the time

Question: Why isn't this good enough for **probability judgment**?

Answer: Because two statisticians who are right 95% of the time may tell the court different and even contradictory things.

They are placing the current event in different sequences.

Probability judgment continued

- Good probability forecasting requires a sequence.
- It does not require repetition of the “same” event.
- Each event remains unique.

Probability judgment: Assessment of the relevance or irrelevance of experience from different sequences for which we have good probability forecasters.

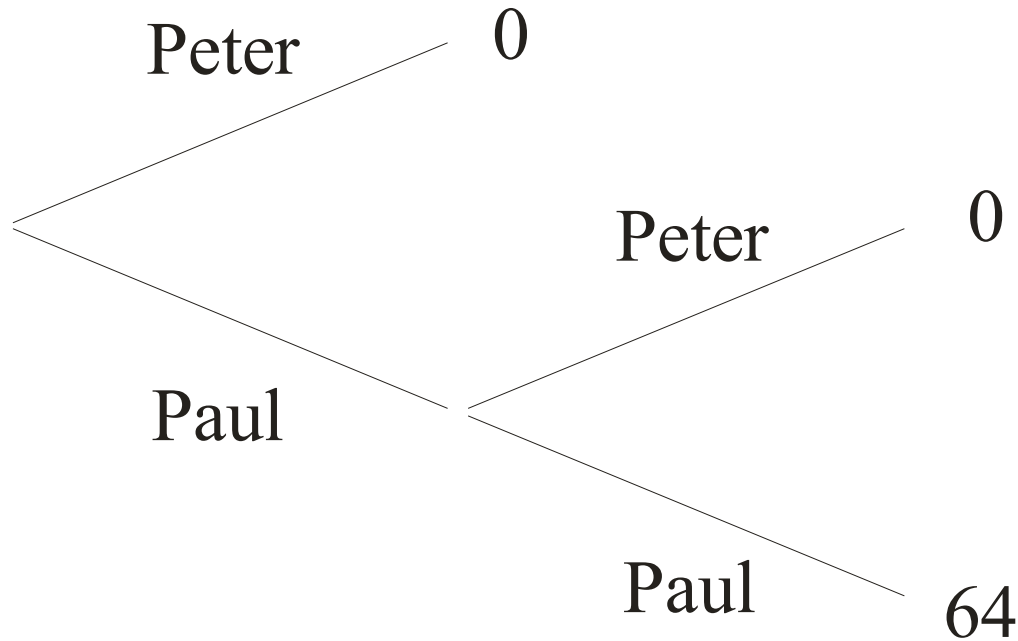
1. Game theoretic probability

The contrast between measure-theoretic & game-theoretic probability began in 1654.

Pascal = game theory

Fermat = measure theory

Pascal's question to Fermat in 1654



Paul needs 2 points to win.
Peter needs only one.

If the game must be broken off,
how much of the stake should
Paul get?



Blaise Pascal (1623–1662).

Fermat's answer (measure theory)

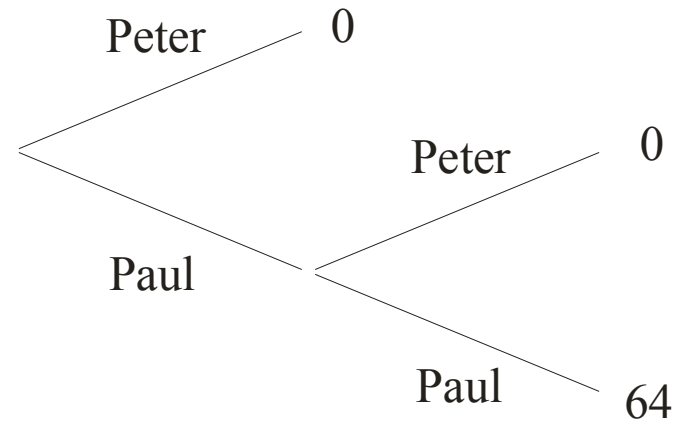
Count the possible outcomes.

Suppose they play two rounds. There are 4 possible outcomes:

1. Peter wins first, Peter wins second
2. Peter wins first, Paul wins second
3. Paul wins first, Peter wins second
4. Paul wins first, Paul wins second

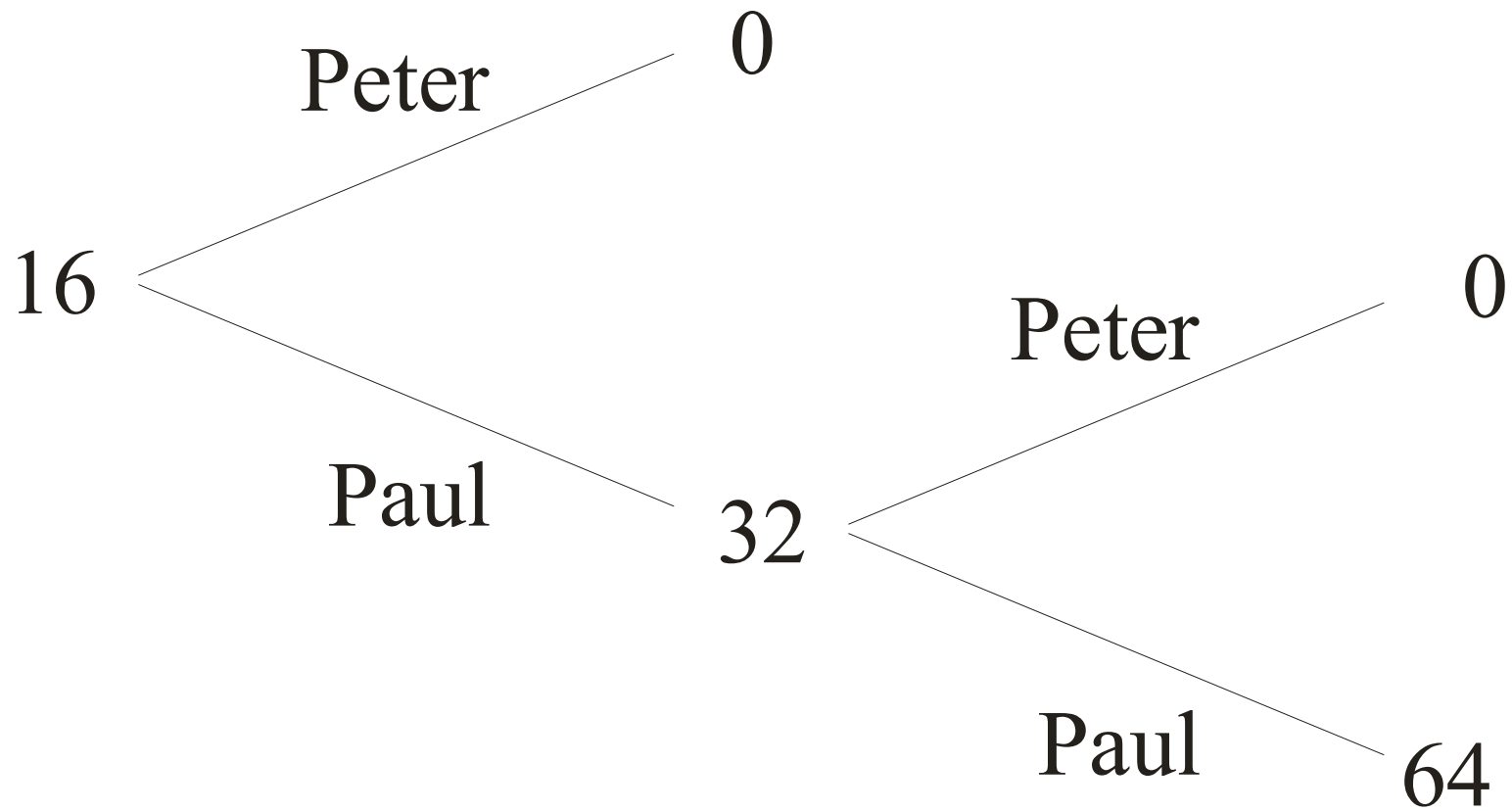
Paul wins only in outcome 4.
So his share should be $\frac{1}{4}$, or
16 pistoles.

Pascal didn't like the
argument.



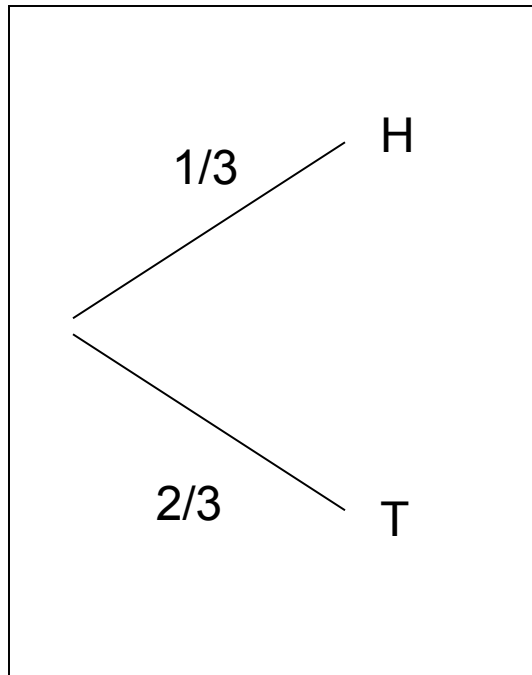
Pierre Fermat, 1601-1665

Pascal's answer (game theory)

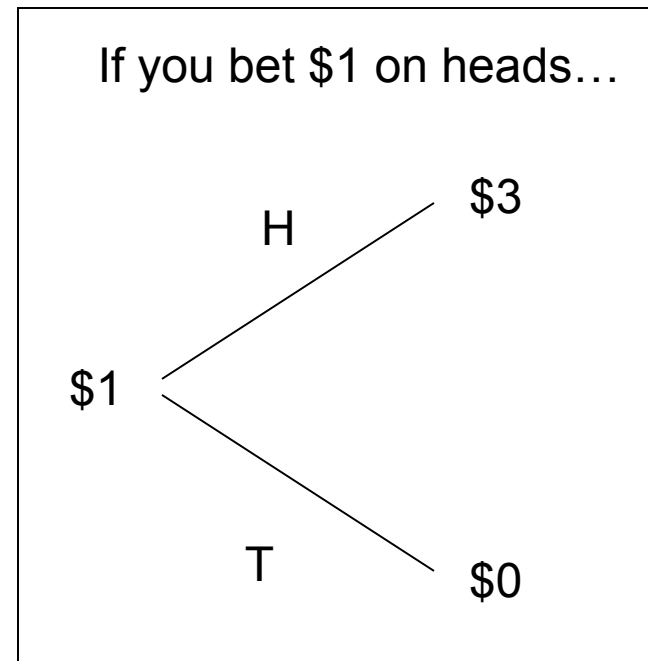


Another probability problem: HH before TT

PROBABILITIES



PAYOFFS

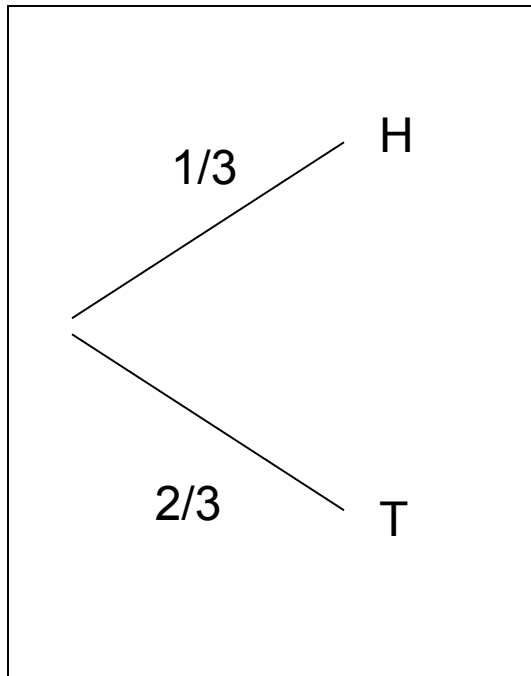


Toss the biased coin repeatedly.

**What is the probability of
HH before TT?**

Measure-theoretic question

PROBABILITIES



**What is the probability
of HH before TT?**

Measure-theoretic solution

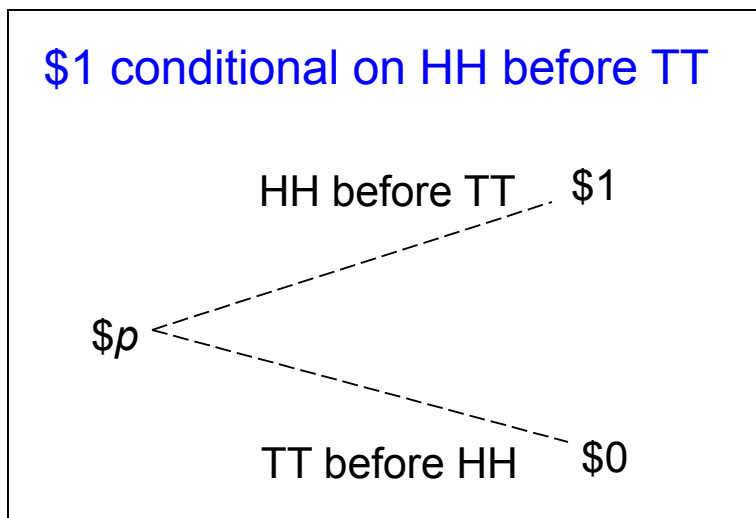
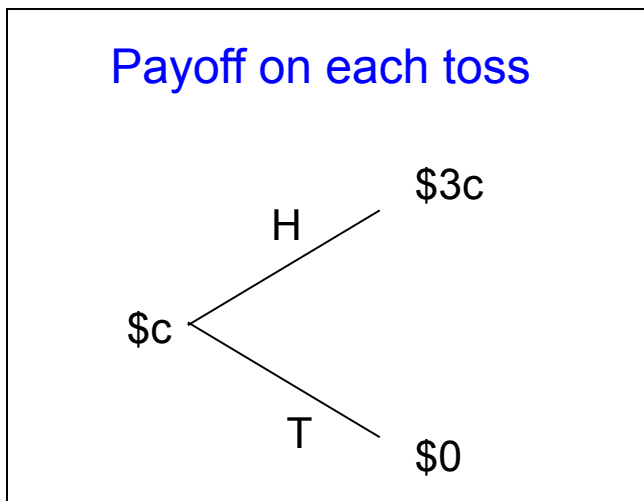
Probability	Outcome
$(\frac{1}{3})^2$	HH
$(\frac{1}{3})^2 \frac{2}{3}$	TTHH
$(\frac{1}{3})^3 \frac{2}{3}$	HTHH
$(\frac{1}{3})^3 (\frac{2}{3})^2$	THTHH
$(\frac{1}{3})^4 (\frac{2}{3})^2$	HTHTHH
etc.	

Summing the series, we find

$$\text{probability} = \frac{5}{21}$$

Game-theoretic question

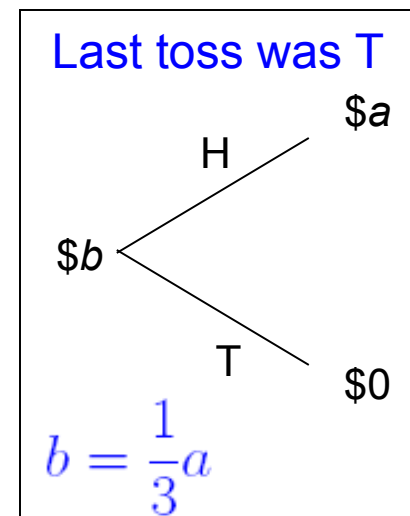
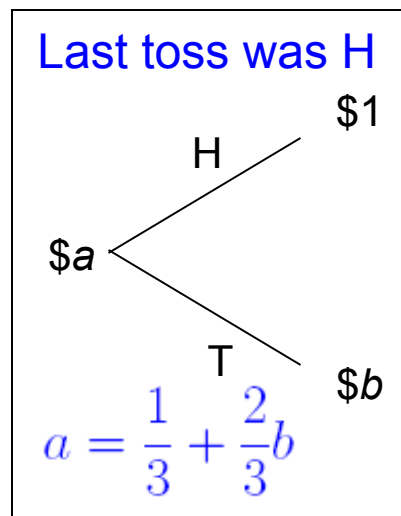
What is $\$p$, the price of a $\$1$ payoff conditional on HH before TT?



Game-theoretic solution

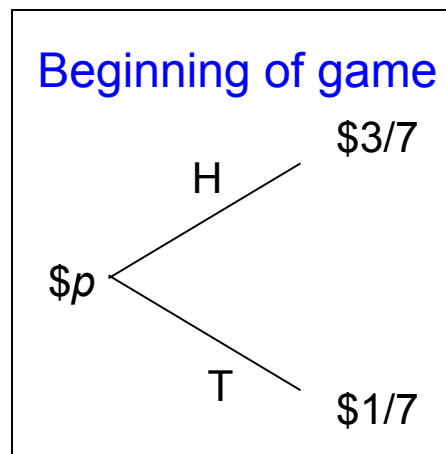
a = price when last toss was H

b = price when last toss was T



$$a = \frac{3}{7}$$

$$b = \frac{1}{7}$$



$$p = \frac{1}{3} \times \frac{3}{7} + \frac{2}{3} \times \frac{1}{7}$$

$$= \frac{5}{21}$$

Caveat

In Pascal's problem, the prices are not necessarily assertions about the world. Perhaps the players have different levels of skill but have nevertheless agreed to play at even odds.

This anticipates the modern theory of option pricing, where so-called "risk-neutral" probabilities are merely prices derived from other prices, not assertions about whether what will make money in the future.

Fermat's combinatorial or measure-theoretic approach leads to a metaphysics of possibility. By the 1690s, Jacob Bernoulli was writing about "equally possible cases" that happen equally often.

Pascal's betting approach is more flexible. Betting prices have many meanings. If we choose to use giving betting prices as a theory about the world, the meaning of this theory lies in the way it is tested, not in any metaphysics about possibilities that precede realities.

Game theoretic upper and lower probability

Measure-theoretic probability:

- Classical: elementary events with probabilities adding to one.
- Modern: space with filtration and probability measure.

Probability of A = total of probabilities for elementary events favoring A

Game-theoretic probability:

- One player offers prices for uncertain payoffs.
- Another player decides what to buy.

Probability of A = initial stake needed to obtain the payoff

[1 if A happens and 0 otherwise]

If no strategy delivers exactly the 0/1 payoff:

Upper probability of A = initial stake needed to obtain at least the payoff

[1 if A happens, 0 otherwise]

Objective interpretation of probability

Cournot's principle

Commonly used by mathematicians before WWII

An event of very small probability will not happen.

To avoid lottery paradox, consider only events with simplest descriptions.
(Wald, Schnorr, Kolmogorov, Levin)

Ville's principle

Equivalent to Cournot's principle when upper probabilities are probabilities

You will not multiply the capital you risk by a large factor.

Mathematical definition of probability:

$P(A)$ = stake needed to obtain \$1 if A happens, \$0 otherwise

Objective interpretation of game-theoretic probability:

You will not multiply the capital you risk by a large factor.

Subjective interpretation of game-theoretic probability:

I don't think you will multiply the capital you risk by a large factor.

Unlike de Finetti, we do not need behavioral assumptions (e.g., people want to bet or can be forced to do so).

To make Pascal's theory part of modern game theory, we must **define the game precisely.**

- Rules of play
- Each player's information
- Rule for winning

A game between Forecaster and Reality

Forecaster gives probabilities for a sequence x_1, x_2, \dots of 1s and 0s.

Before Reality announces x_n , Forecaster announces probability p_n for $x_n = 1$.

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

Forecaster announces $p_n \in [0, 1]$.

Reality announces $x_n \in \{0, 1\}$.

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

Forecaster announces $p_n \in [0, 1]$.

Reality announces $x_n \in \{0, 1\}$.

Clarifications:

1. The phenomena need not be binary. We assume $x_n \in \{0, 1\}$ only for simplicity.
2. Reality's move space may change from round to round.
3. Perfect information: All players hear announcements as they are made.
4. In addition to x_1, \dots, x_{n-1} , Forecaster may have other newly acquired information.

Forecaster is tested by a third player, Skeptic, who tries to get rich from Forecaster's betting offers.

Players: Forecaster, Reality, Skeptic

Protocol:

$\mathcal{K}_0 := 1$.

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

Forecaster announces $p_n \in [0, 1]$.

Skeptic announces $M_n \in \mathbb{R}$.

Reality announces $x_n \in \{0, 1\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(x_n - p_n)$.

Winner: Skeptic wins if $\mathcal{K}_n \geq 0$ for all n and $\lim_{n \rightarrow \infty} \mathcal{K}_n = \infty$.
Otherwise Forecaster and Reality win.

Example of a game-theoretic probability theorem.

$\mathcal{K}_0 := 1.$

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

Forecaster announces $p_n \in [0, 1].$

Skeptic announces $s_n \in \mathbb{R}.$

Reality announces $y_n \in \{0, 1\}.$

$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - p_n).$

Skeptic wins if

(1) \mathcal{K}_n is never negative **and**

(2) either $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (y_i - p_i) = 0$

or $\lim_{n \rightarrow \infty} \mathcal{K}_n = \infty.$

Theorem Skeptic has a winning strategy.

Ville's strategy

$$\mathcal{K}_0 = 1.$$

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

Skeptic announces $s_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$$\mathcal{K}_n := \mathcal{K}_{n-1} + s_n(y_n - \frac{1}{2}).$$

Ville suggested the strategy

$$s_n(y_1, \dots, y_{n-1}) = \frac{4}{n+1} \mathcal{K}_{n-1} \left(r_{n-1} - \frac{n-1}{2} \right), \text{ where } r_{n-1} := \sum_{i=1}^{n-1} y_i.$$

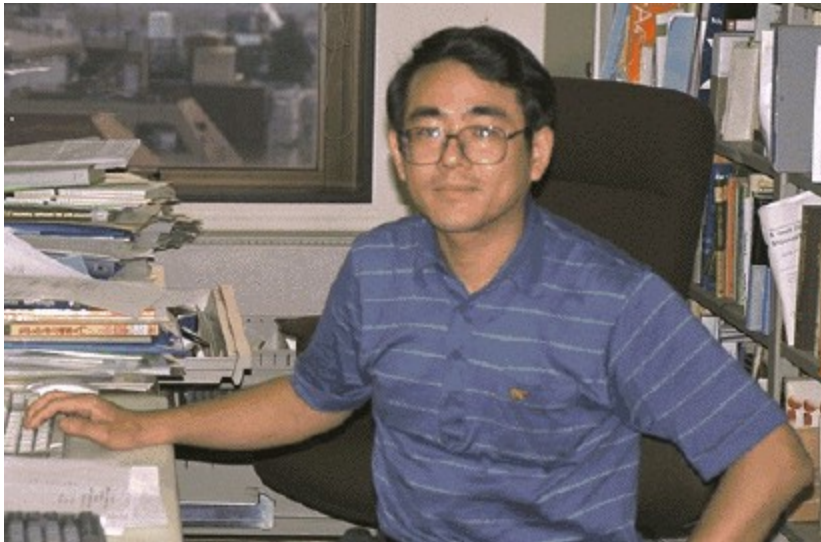
It produces the capital

$$\mathcal{K}_n = 2^n \frac{r_n!(n-r_n)!}{(n+1)!}.$$

From the assumption that this remains bounded by some constant C , you can easily derive the strong law of large numbers using Stirling's formula.

Defensive forecasting

The name was introduced in Working Paper 8 at www.probabilityandfinance.com, by Vovk, Takemura, and Shafer (September 2004). See also Working Papers 7, 9, 10, 11, 13, 14, 16, 17, 18, 20, 21, 22, and 30.



Akimichi Takemura



Volodya Vovk

Crucial idea: all the tests (betting strategies for Skeptic) Forecaster needs to pass can be merged into a single **portmanteau test** for Forecaster to pass.

1.If you have two strategies for multiplying capital risked, divide your capital between them.

2.Formally: average the strategies.

3.You can average countably many strategies.

4.As a practical matter, there are only countably many tests (Abraham Wald, 1937).

5.I will explain how Forecaster can beat any single test (including the portmanteau test).

A. How Forecaster beats any single test

B. How to construct a portmanteau test for binary probability forecasting

- Use law of large numbers to test calibration for each probability p .**
- Merge the tests for different p .**

How Forecaster can beat any single test \mathcal{S}

Skeptic adopts a continuous strategy \mathcal{S} .

FOR $n = 1, 2, \dots$

Reality announces $x_n \in \mathbf{X}$.

Forecaster announces $p_n \in [0, 1]$.

Skeptic makes the move s_n specified by \mathcal{S} .

Reality announces $y_n \in \{0, 1\}$.

Skeptic's profit $:= s_n(y_n - p_n)$.

Theorem Forecaster can guarantee that Skeptic never makes money.

We actually prove a stronger theorem. Instead of making Skeptic announce his entire strategy in advance, only make him reveal his strategy for each round in advance of Forecaster's move.

FOR $n = 1, 2, \dots$

Reality announces $x_n \in \mathbf{X}$.

Skeptic announces continuous $S_n : [0, 1] \rightarrow \mathbb{R}$.

Forecaster announces $p_n \in [0, 1]$.

Reality announces $y_n \in \{0, 1\}$.

Skeptic's profit $:= S_n(p_n)(y_n - p_n)$.

Theorem. Forecaster can guarantee that Skeptic never makes money.

FOR $n = 1, 2, \dots$

Reality announces $x_n \in \mathbf{X}$.

Skeptic announces continuous $S_n : [0, 1] \rightarrow \mathbb{R}$.

Forecaster announces $p_n \in [0, 1]$.

Reality announces $y_n \in \{0, 1\}$.

Skeptic's profit $:= S_n(p_n)(y_n - p_n)$.

Theorem Forecaster can guarantee that Skeptic never makes money.

Proof:

- If $S_n(p) > 0$ for all p , take $p_n := 1$.
- If $S_n(p) < 0$ for all p , take $p_n := 0$.
- Otherwise, choose p_n so that $S_n(p_n) = 0$.

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

FOR $n = 1, 2, \dots$

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.

Probability judgment

To conclude, I will discuss how Bayesian conditioning can be understood as an example of probability judgment.

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

Assumptions

1. $P(A)$ = price of a ticket that pays 1 if A happens.
2. $P(A)x$ = 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),
 $P(B|A)x$ = price of a ticket that pays x if B happens.

Argument

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

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

Assumptions

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

Argument

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

De Finetti interpreted De Moivre's prices in a particular way.

There are other ways.

In game-theoretic probability (Shafer and Vovk 2001) we interpret the prices as a prediction.

The prediction: You will not multiply by a large factor the capital you risk at these prices.

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

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 $P(A)$ and $P(A\&B)$.

Conclusion You still will not win if after A (and nothing else) is known, $P(A\&B)/P(A)$ 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 $P(A)$ and $P(A\&B)$ alone that risks the same risks and payoffs.

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

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

to S 's purchases in the initial situation.

- The tickets added have zero total initial cost:

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

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

0	if A does not happen;
$-M \frac{P(A\&B)}{P(A)}$	if A happens but not B ;
$M \left(1 - \frac{P(A\&B)}{P(A)}\right)$	if A and B both happen.

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.

Cournotian understanding of Dempster-Shafer

- Fundamental idea: transferring belief
- Conditioning
- Independence
- Dempster's rule

Fundamental idea: transferring belief

- Variable ω with set of possible values Ω .
- Random variable \mathbf{X} with set of possible values \mathcal{X} .
- We learn a mapping $\Gamma : \mathcal{X} \rightarrow 2^\Omega$ with this meaning:

If $\mathbf{X} = x$, then $\omega \in \Gamma(x)$.

- For $A \subseteq \Omega$, our belief that $\omega \in A$ is now

$$\mathbb{B}(A) = \mathbb{P}\{x | \Gamma(x) \subseteq A\}.$$

Cournotian judgement of independence: Learning the relationship between \mathbf{X} and ω does not affect our inability to beat the probabilities for \mathbf{X} .

Example: The sometimes reliable witness

- Joe is reliable with probability 30%. When he is reliable, what he says is true. Otherwise, it may or may not be true.

$$\mathcal{X} = \{\text{reliable, not reliable}\} \quad \mathbb{P}(\text{reliable}) = 0.3 \quad \mathbb{P}(\text{not reliable}) = 0.7$$

- Did Glenn pay his dues for coffee? $\Omega = \{\text{paid, not paid}\}$
- Joe says “Glenn paid.”

$$\Gamma(\text{reliable}) = \{\text{paid}\} \quad \Gamma(\text{not reliable}) = \{\text{paid, not paid}\}$$

- New beliefs:

$$\mathbb{B}(\text{paid}) = 0.3 \quad \mathbb{B}(\text{not paid}) = 0$$

Cournotian judgement of independence: Hearing what Joe said does not affect our inability to beat the probabilities concerning his reliability.

Example: The more or less precise witness

- Bill is absolutely precise with probability 70%, approximate with probability 20%, and unreliable with probability 10%.

$\mathcal{X} = \{\text{precise, approximate, not reliable}\}$

$$\mathbb{P}(\text{precise}) = 0.7 \quad \mathbb{P}(\text{approximate}) = 0.2 \quad \mathbb{P}(\text{not reliable}) = 0.1$$

- What did Glenn pay? $\Omega = \{0, \$1, \$5\}$

- Bill says "Glenn paid \$ 5."

$$\Gamma(\text{precise}) = \{\$5\} \quad \Gamma(\text{approximate}) = \{\$1, \$5\} \quad \Gamma(\text{not reliable}) = \{0, \$1, \$5\}$$

- New beliefs:

$$\mathbb{B}\{0\} = 0 \quad \mathbb{B}\{\$1\} = 0 \quad \mathbb{B}\{\$5\} = 0.7 \quad \mathbb{B}\{\$1, \$5\} = 0.9$$

Cournotian judgement of independence: Hearing what Bill said does not affect our inability to beat the probabilities concerning his precision.

Conditioning

- Variable ω with set of possible values Ω .
- Random variable X with set of possible values \mathcal{X} .
- We learn a mapping $\Gamma : \mathcal{X} \rightarrow 2^\Omega$ with this meaning:

If $X = x$, then $\omega \in \Gamma(x)$.

•

$\Gamma(x) = \emptyset$ for some $x \in \mathcal{X}$.

- For $A \subseteq \Omega$, our belief that $\omega \in A$ is now

$$\mathbb{B}(A) = \frac{\mathbb{P}\{x | \Gamma(x) \subseteq A \ \& \ \Gamma(x) \neq \emptyset\}}{\mathbb{P}\{x | \Gamma(x) \neq \emptyset\}}.$$

Cournotian judgement of independence: Aside from the impossibility of the x for which $\Gamma(x) = \emptyset$, learning Γ does not affect our inability to beat the probabilities for X .

Example: The witness caught out

- Tom is absolutely precise with probability 70%, approximate with probability 20%, and unreliable with probability 10%.

$\mathcal{X} = \{\text{precise, approximate, not reliable}\}$

$$\mathbb{P}(\text{precise}) = 0.7 \quad \mathbb{P}(\text{approximate}) = 0.2 \quad \mathbb{P}(\text{not reliable}) = 0.1$$

- What did Glenn pay? $\Omega = \{0, \$1, \$5\}$

- Tom says "Glenn paid \$ 10."

$$\Gamma(\text{precise}) = \emptyset \quad \Gamma(\text{approximate}) = \{\$5\} \quad \Gamma(\text{not reliable}) = \{0, \$1, \$5\}$$

- New beliefs:

$$\mathbb{B}\{0\} = 0 \quad \mathbb{B}\{\$1\} = 0 \quad \mathbb{B}\{\$5\} = 2/3 \quad \mathbb{B}\{\$1, \$5\} = 2/3$$

Cournotian judgement of independence: Aside ruling out his being absolutely precise, what Tom said does not help us beat the probabilities for his precision.

Independence

$\mathcal{X}_{\text{Bill}} = \{\text{Bill precise, Bill approximate, Bill not reliable}\}$

$$\mathbb{P}(\text{precise}) = 0.7 \quad \mathbb{P}(\text{approximate}) = 0.2 \quad \mathbb{P}(\text{not reliable}) = 0.1$$

$\mathcal{X}_{\text{Tom}} = \{\text{Tom precise, Tom approximate, Tom not reliable}\}$

$$\mathbb{P}(\text{precise}) = 0.7 \quad \mathbb{P}(\text{approximate}) = 0.2 \quad \mathbb{P}(\text{not reliable}) = 0.1$$

Product measure:

	$\mathcal{X}_{\text{Bill} \ \& \ \text{Tom}} =$	$\mathcal{X}_{\text{Bill}} \times \mathcal{X}_{\text{Tom}}$
$\mathbb{P}(\text{Bill precise, Tom precise}) =$		$0.7 \times 0.7 = 0.49$
$\mathbb{P}(\text{Bill precise, Tom approximate}) =$		$0.7 \times 0.2 = 0.14$
	etc.	

Cournotian judgements of independence: Learning about the precision of one of the witnesses will not help us beat the probabilities for the other.

Nothing novel here. Dempsterian independence = Cournotian independence.

Example: Independent contradictory witnesses

- Joe and Bill are both reliable with probability 70%.

- Did Glenn pay his dues? $\Omega = \{\text{paid, not paid}\}$

- Joe says, "Glenn paid." Bill says, "Glenn did not pay."

$$\Gamma_1(\text{Joe reliable}) = \{\text{paid}\} \quad \Gamma_1(\text{Joe not reliable}) = \{\text{paid, not paid}\}$$

$$\Gamma_2(\text{Bill reliable}) = \{\text{not paid}\} \quad \Gamma_2(\text{Bill not reliable}) = \{\text{paid, not paid}\}$$

- The pair (Joe reliable, Bill reliable), which had probability 0.49, is ruled out.

$$\mathbb{B}(\text{paid}) = \frac{0.21}{0.51} = 0.41 \quad \mathbb{B}(\text{not paid}) = \frac{0.21}{0.51} = 0.41$$

Cournotian judgement of independence: Aside from learning that they are not both reliable, what Joe and Bill said does not help us beat the probabilities concerning their reliability.

Dempster's rule (independence + conditioning)

- Variable ω with set of possible values Ω .
- Random variables X_1 and X_2 with sets of possible values \mathcal{X}_1 and \mathcal{X}_2 .
- Form the product measure on $\mathcal{X}_1 \times \mathcal{X}_2$.
- We learn mappings $\Gamma_1 : \mathcal{X}_1 \rightarrow 2^\Omega$ and $\Gamma_2 : \mathcal{X}_2 \rightarrow 2^\Omega$:

If $X_1 = x_1$, then $\omega \in \Gamma_1(x_1)$.

If $X_2 = x_2$, then $\omega \in \Gamma_2(x_2)$.

- So if $(X_1, X_2) = (x_1, x_2)$, then $\omega \in \Gamma_1(x_1) \cap \Gamma_2(x_2)$.
- Conditioning on what is not ruled out,

$$\mathbb{B}(A) = \frac{\mathbb{P}\{(x_1, x_2) | \emptyset \neq \Gamma_1(x_1) \cap \Gamma_2(x_2) \subseteq A\}}{\mathbb{P}\{(x_1, x_2) | \emptyset \neq \Gamma_1(x_1) \cap \Gamma_2(x_2)\}}$$

Cournotian judgement of independence: Aside from ruling out some (x_1, x_2) , learning the Γ_i does not help us beat the probabilities for X_1 and X_2 .

You can suppress the Γ s and describe Dempster's rule in terms of the belief functions

Joe: $\mathbb{B}_1\{\text{paid}\} = 0.7$ $\mathbb{B}_1\{\text{not paid}\} = 0$
 Bill: $\mathbb{B}_2\{\text{not paid}\} = 0.7$ $\mathbb{B}_2\{\text{paid}\} = 0$

		Bill	
		0.7 not paid	0.3 ??
Joe	0.7 paid	0.49	Paid
	0.3 ??	Not paid	0.09

$$\mathbb{B}(\text{paid}) = \frac{0.21}{0.51} = 0.41$$

$$\mathbb{B}(\text{not paid}) = \frac{0.21}{0.51} = 0.41$$

Dempster's rule is unnecessary. It is merely a composition of Cournot operations: formation of product measures, conditioning, transferring belief.

But Dempster's rule is a unifying idea. Each Cournot operation is an example of Dempster combination.

- Forming product measure is Dempster combination.
- Conditioning on A is Dempster combination with a belief function that gives belief one to A .
- Transferring belief is Dempster combination of (1) a belief function on $\mathcal{X} \times \Omega$ that gives probabilities to cylinder sets $\{x\} \times \Omega$ with (2) a belief function that gives probability one to $\{(x, \omega) | \omega \in \Gamma(x)\}$.