SIPTA Online School 2020, University of Liverpool Institute for Risk and Uncertainty

Game-theoretic foundations for statistical testing and imprecise probabilities

Remote lectures by **Glenn Shafer**. December 9th and 10th, 2020.

Lecture 4. Axiomatic game-theoretic probability.

Reading: Chapers 6-9 of *Game-Theoretic Foundations for Probability and Finance*, by Glenn Shafer and Vladimir Vovk, Wiley, 2019

Revised March 25, 2021

1. Glance at a broader landscape
2. De Finetti as Forecaster; Shafer/Vovk as Skeptic
3. Local and global
4. Axioms
5. Global upper expected value
6. The supermartingale of upper expected values
7. Lévy’s zero-one law
8. Classical, measure-theoretic, game-theoretic
9. Independence and causality

1. Glance at a broader landscape

We have a perpetual regression defining probabilities in terms of probabilities in terms of probabilities…

R. A. Fisher, 1958

Previous lectures studied game-theoretic probability using examples.

This lecture gives an axiomatic account.

Axiomatic account clarifies relation to:

- Bruno de Finetti’s subjective probability,
- the imprecise-probability generalization of de Finetti,
- measure-theoretic frequentism.

Some mathematicians who defined probability in terms of bets considered fair.

Blaise Pascal
1623 - 1662

Christiaan Huygens
1629 - 1695

Abraham De Moivre
1667-1754

No known portrait

Thomas Bayes
1702-1761

Some mathematicians who defined probability in terms of a person's willingness to bet.

Joseph Bertrand
1822-1900
France

Emile Borel
1871-1956
France

Bruno de Finetti
1906-1985
Italy

Three ways of understanding probability in terms of betting:

1. Fair betting rates. (Pascal, Huygens, De Moivre, Bayes)

2. A person's betting rates. (Bertrand, Borel, de Finetti)

3. Betting rates that cannot be beat. (Shafer/Vovk)

Shafer/Vovk closest to frequentism.

From my RSS paper:

Are the probabilities tested subjective or objective? The probabilities may represent someone's opinion, but the hypothesis that they say something true about the world is inherent in the project of testing them.

In measure-theoretic probability, we can prove that

frequency = probability in the limit

with measure-theoretic probability one.

In game-theoretic probability, can prove that

frequency = probability in the limit

with game-theoretic probability one.

The frequentist thinks this justifies interpreting probability as frequency.

The frequentist’s claim is just as legitimate (or illegitimate) with game-theoretic probability as with measure-theoretic probability.

2. De Finetti + imprecise probabilists as Forecaster; Shafer/Vovk as Skeptic

Lower and upper previsions represent commitments to act/behave in certain ways.

Gert de Cooman, 2003

Probability forecasting

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

Forecaster makes betting offers.

Skeptic decides which, if any, to take.

Bruno de Finetti’s subjective theory took Forecaster’s viewpoint. He replaced the traditional Italian term for expected value, *speranza matematica* (mathematical hope), with *previsione* (forecast).

Founders of “imprecise probabilities” followed de Finetti’s example.

- Peter Williams 1975
- Peter Walley 1981, 1991
- Gert de Cooman (Launched *Imprecise Probabilities Project* with Walley in 1996.)

Society for Imprecise Probability, Theory and Applications
SIPTA http://www.sipta.org/

Peter Walley
Born 1953
Australia
Ph.D., Cornell, 1979

Gert de Cooman
Born 1964
Ghent, Belgium

Prof. De Cooman will talk on "Randomness and Imprecision" this Monday (March 29) at 1pm EST in the Rutgers Foundations of Probability seminar.

If you are interested, contact me for the zoom link.

This talk uses the martingale-theoretic approach of game-theoretic probability to incorporate imprecision into the study of randomness. We associate (weak Martin-Löf, computable, Schnorr) randomness with interval, rather than precise, forecasting systems. The richer mathematical structure this uncovers, allows us to, amongst other things, better understand existing results for the precise limit.

Bruno de Finetti, 1931

… to measure numerically the degree of belief that a certain subject *O* has towards an event *E* … assume … that he might be forced to keep a betting shop …

… it is the decision of subject *O* … to define the price *p* of one ticket, giving the right to cash one lira in the eventuality that …. *E* occurs; having done so, he commits himself to sell or to buy at such a price as many tickets as the public will want. …

… Any competitor ….

From de Finetti's "Sul significato soggettivo della probabilità", 1931; translation by Mara Kahle.

To make this rigorously game-theoretic:

- name the competitor,
- specify the players' information.

Peter Williams

From 9th ISIPTA, page 20.

In de Finetti's picture, as Williams noted in 1975,

> ...an "opponent" is free to choose the stake and consequently whether the bet is on or against the event in question.

Williams proposed

> ...to relax the requirement that the individual take either side of an acceptable bet.

So the class of bets the agent considers acceptable would be only a cone rather than a linear space.

Williams gave axioms for *upper and lower conditional previsions* representing the relaxed model.

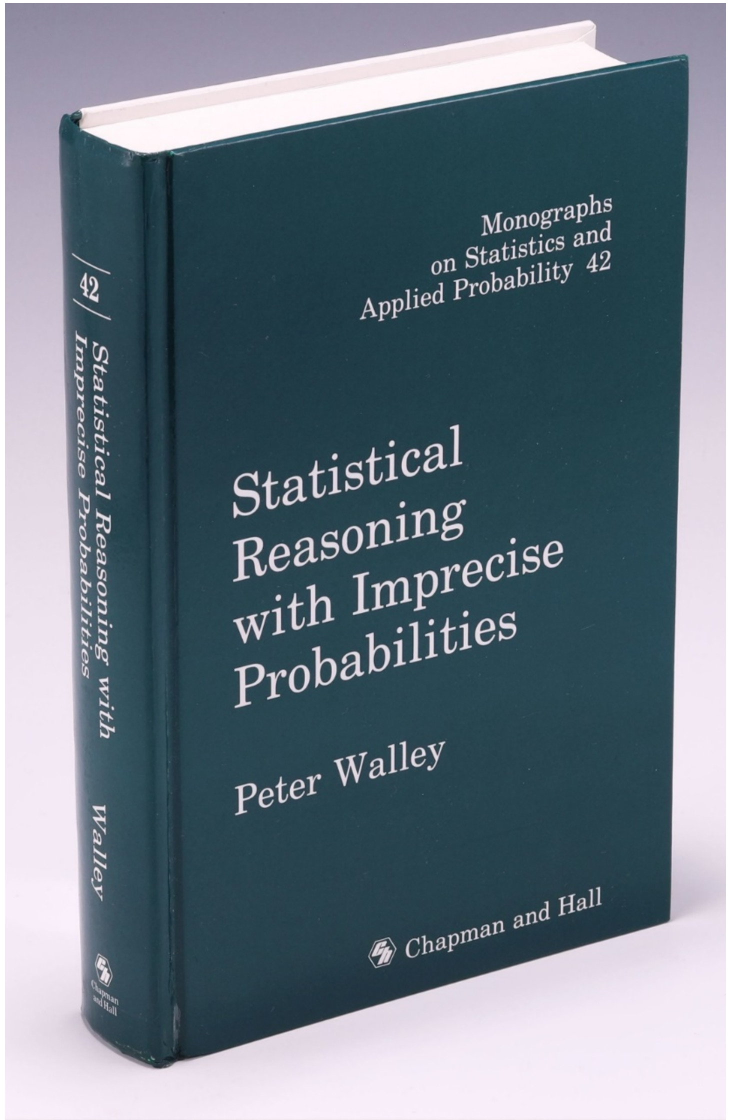
Williams, Peter M. (1975a). "Coherence, strict coherence and zero probabilities." *Fifth International Congress of Logic, Methodology and Philosophy of Science*. Vol. VI, pp. 29–33.

– (1975b). *Notes on conditional previsions*. Tech. rep. Published as (Williams 2007). School of Mathematical & Physical Sciences, University of Sussex.

– (1976). "Indeterminate probabilities." *Formal Methods in the Methodology of Empirical Sciences*. Conference for Formal Methods in the Methodology of Empirical Sciences. (Warsaw, Poland, June 17–21, 1974). Ed. by Marian Przełęcki, Klemens Szaniawski, Ryszard Wójcicki, & Grzegorz Malinowski. Vol. 103. Synthese Library. D. Reidel Publishing Company & Ossolineum Publishing company, pp. 229–246. DOI: 10.1007/978-94-010-1135-8_16.

– (2007). "Notes on conditional previsions." *International Journal of Approximate Reasoning* 44. Published version of (Williams 1975b), pp. 366–383. DOI: 10.1016/j.ijar.2006.07.019.

Peter Walley's very rare 1991 book



- Like de Finetti and Williams, Walley took Forecaster's viewpoint.
- Like Williams, he used the terms *lower prevision* and *upper prevision*.
- For Walley, a lower prevision $\underline{P}$ "represents a disposition to accept gambles of the form $X - \mu$ whenever $\mu < \underline{P}(X)$." (page 101)

Forecaster's viewpoint is default in SIPTA.

For an objective understanding of forecasting, we need Skeptic's viewpoint.

Game-theoretic probability emphasizes <u>upper</u> expectation.
Imprecise probability theory emphasizes <u>lower</u> prevision.

Why?

Upper if you are Skeptic. Lower if you are Forecaster.

Game-Theoretic Foundations, page 131:

- Because most people buy more often than they sell, ordinary language is more developed for buying.
- Skeptic's buying prices are given by the lower functional.
- Forecaster's buying prices are given by the upper functional.

3. Local and global

Think global, act local.

Attributed to Patrick Geddes, 1915

Testing a forecaster

Forecaster announces a probability distribution P on $\mathcal{Y}$.
Skeptic announces $S : \mathcal{Y} \to [0, \infty)$ such that $\mathbf{E}_P(S) = 1$.
Reality announces $y \in \mathcal{Y}$.
$\mathcal{K} := S(y)$.

The variable S on the <u>local</u> sample space $\mathcal{Y}$ has expected value $\mathbf{E}_P(S)$.

Testing a forecaster over time

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots, N$:
Forecaster announces a probability distribution P_n on $\mathcal{Y}$.
Skeptic announces $S_n : \mathcal{Y} \to [0, \infty)$ such that $\mathbf{E}_{P_n}(S_n) = \mathcal{K}_{n-1}$.
Reality announces $y_n \in \mathcal{Y}$.
$\mathcal{K}_n := S_n(y_n)$.

Local variables S_n have expected values $\mathbf{E}_{P_n}(S_n)$.

We also have <u>global</u> variables. In general, they have global upper and lower expected values.

Testing a forecaster over time

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

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

  Forecaster announces a probability distribution P_n on $\mathcal{Y}$.

  Skeptic announces $S_n : \mathcal{Y} \to [0, \infty)$ such that $\mathbf{E}_{P_n}(S_n) = \mathcal{K}_{n-1}$.

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

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

Local variables S_n have expected values $\mathbf{E}_{P_n}(S_n)$.

We also have a global sample space Ω:

variables with upper expected values,

events with upper probabilities.

Ω consists of paths of the form $P_1 y_1 \ldots P_N y_N$.

We can prove

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

See Lecture 3, slide 39.

From page 98 of *Game-Theoretic Foundations*

Abstract	Local	Global
outcome space $\mathcal{Y}$	Reality's move space $\mathcal{Y}$	sample space Ω
outcome $y \in \mathcal{Y}$	Reality's move $y \in \mathcal{Y}$	path $\omega \in \Omega$
event $E \subseteq \mathcal{Y}$	local event $E \subseteq \mathcal{Y}$	global event $E \subseteq \Omega$
upper probability $\overline{\mathbf{P}}(E)$	upper probability $\overline{\mathbf{P}}(E)$	upper probability $\overline{\mathbb{P}}(E)$
extended variable $f : \mathcal{Y} \to \overline{\mathbb{R}}$	local variable $f : \mathcal{Y} \to \overline{\mathbb{R}}$	global variable $X : \Omega \to \overline{\mathbb{R}}$
upper expected value $\overline{\mathbf{E}}(f)$	upper expected value $\overline{\mathbf{E}}(f)$	upper expected value $\overline{\mathbb{E}}(X)$
upper expectation $\overline{\mathbf{E}}$	local upper expectation $\overline{\mathbf{E}}$	global upper expectation $\overline{\mathbb{E}}$

Here $\overline{\mathbb{R}} = [-\infty, \infty]$. For abstract theory, it is best to include variables that take infinite values. We use the convention $-\infty + \infty = \infty$.

4. Axioms

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.

Albert Einstein, 1921

Axioms for upper expectations from *Game-Theoretic Foundations*, p. 113:

Given a nonempty set $\mathcal{Y}$, we call a functional $\overline{\mathbf{E}} : \mathbb{R}^{\mathcal{Y}} \to \mathbb{R}$ an *upper expectation on* $\mathcal{Y}$ if it satisfies these five axioms:

Axiom E1. If $f_1, f_2 \in \mathbb{R}^{\mathcal{Y}}$, then $\overline{\mathbf{E}}(f_1 + f_2) \leq \overline{\mathbf{E}}(f_1) + \overline{\mathbf{E}}(f_2)$.

Axiom E2. If $f \in \mathbb{R}^{\mathcal{Y}}$ and $c \in (0, \infty)$, then $\overline{\mathbf{E}}(cf) = c\overline{\mathbf{E}}(f)$.

Suboptimality allowed. **Axiom E3.** If $f_1, f_2 \in \mathbb{R}^{\mathcal{Y}}$ and $f_1 \leq f_2$, then $\overline{\mathbf{E}}(f_1) \leq \overline{\mathbf{E}}(f_2)$.

Coherence. **Axiom E4.** For each $c \in \mathbb{R}$, $\overline{\mathbf{E}}(c) = c$.

Axiom E5. If $f_1 \leq f_2 \leq \cdots \in [0, \infty]^{\mathcal{Y}}$, then $\overline{\mathbf{E}}\left(\lim_{k\to\infty} f_k\right) = \lim_{k\to\infty} \overline{\mathbf{E}}(f_k)$.

We call Axiom E5 the *continuity axiom*. We call $\overline{\mathbf{E}}(f)$ *f's upper expected value*.

Continuity axiom <ins>**not**</ins> needed for major results; simplifies some proofs.

Some direct consequences of the axioms:

If $f \in \overline{\mathbb{R}}^{\mathcal{Y}}$ *and* $c \in \mathbb{R}$, *then* $\overline{\mathbf{E}}(f + c) = \overline{\mathbf{E}}(f) + c$.

If $f \in \overline{\mathbb{R}}^{\mathcal{Y}}$, *then* $\inf f \leq \overline{\mathbf{E}}(f) \leq \sup f$.

If $f_1, f_2, \cdots \in [0, \infty]^{\mathcal{Y}}$, *then* $\overline{\mathbf{E}}\left(\sum_{k=1}^{\infty} f_k\right) \leq \sum_{k=1}^{\infty} \overline{\mathbf{E}}\left(f_k\right)$.

Require continuity axiom.

If $f \in \overline{\mathbb{R}}^{\mathcal{Y}}$, *then* $f > 0 \Longrightarrow \overline{\mathbf{E}}(f) > 0$.

Kolmogorov (1933) on the continuity axiom:

> …Since the new axiom is essential only for infinite fields of probability, it is hardly possible to explain its empirical meaning.… In describing any actual observable random process, we can obtain only finite fields of probability. Infinite fields of probability occur only as idealized models of real random processes. *This understood, we limit ourselves arbitrarily to models that satisfy Axiom VI.* So far this limitation has been found expedient in the most diverse investigations.

Axioms for pricing non-negative payoffs (*Game-Theoretic Foundations*, p. 104)

Axiom E1$^{[0,\infty]}$**.** *If* $f_1, f_2 \in [0,\infty]^{\mathcal{Y}}$, *then* $\overline{\mathbf{E}}(f_1 + f_2) \leq \overline{\mathbf{E}}(f_1) + \overline{\mathbf{E}}(f_2)$.

Axiom E2$^{[0,\infty]}$**.** *If* $f \in [0,\infty]^{\mathcal{Y}}$ *and* $c \in (0,\infty)$, *then* $\overline{\mathbf{E}}(cf) = c\overline{\mathbf{E}}(f)$.

Axiom E3$^{[0,\infty]}$**.** *If* $f_1, f_2 \in [0,\infty]^{\mathcal{Y}}$ *and* $f_1 \leq f_2$, *then* $\overline{\mathbf{E}}(f_1) \leq \overline{\mathbf{E}}(f_2)$.

Axiom E4$^{[0,\infty]}$**.** *For each* $c \in [0,\infty)$, $\overline{\mathbf{E}}(c) = c$.

Axiom E5$^{[0,\infty]}$**.** *If* $f_1 \leq f_2 \leq \ldots$ *are in* $[0,\infty]^{\mathcal{Y}}$, *then*

$$\overline{\mathbf{E}}\left(\lim_{k\to\infty} f_k\right) = \lim_{k\to\infty} \overline{\mathbf{E}}(f_k).$$

Extra axiom **Axiom E6**$^{[0,\infty]}$**.** *If* $f \in [0,\infty]^{\mathcal{Y}}$ *and* $c \in (0,\infty)$, *then* $\overline{\mathbf{E}}(f + c) = \overline{\mathbf{E}}(f) + c$.

Functional satisfying these axioms can be extended upper expectation on all payoffs.

Beginning with Williams, axioms for imprecise probabilities have been stated in two equivalent forms:

- Axioms for lower/upper previsions
- Axioms for the set of offered gambles

The essential axioms for offered gambles say that

1. offered gambles can be combined,
2. any multiple or fraction of an offered gamble is offered,
3. sure payoff is not offered.

Natural axioms for offered gambles:

Axiom G1. If $g_1, g_2 \in \mathbf{G}$, then $g_1 + g_2 \in \mathbf{G}$.

Axiom G2. If $c \in [0, \infty)$ and $g \in \mathbf{G}$, then $cg \in \mathbf{G}$.

Axiom G3. If $g_1 \in \overline{\mathbb{R}}^{\mathcal{Y}}$, $g_2 \in \mathbf{G}$, and $g_1 \leq g_2$, then $g_1 \in \mathbf{G}$.

Axiom G4. If $g \in \mathbf{G}$, then $\inf g \leq 0$.

To make equivalent to axioms for upper expectations, add:

Axiom G0. If $g \in \overline{\mathbb{R}}^{\mathcal{Y}}$ and $g - \epsilon \in \mathbf{G}$ for every $\epsilon \in (0, \infty)$, then $g \in \mathbf{G}$.

Continuity axiom:

Axiom G5. If $g_1 \leq g_2 \leq \ldots$ are all in $\mathbf{G}$, and g_1 is bounded below, then $\lim_{k \to \infty} g_k \in \mathbf{G}$.

$$\overline{\mathbf{E}}_{\mathbf{G}}(f) := \inf\{\alpha \in \mathbb{R} \mid f - \alpha \in \mathbf{G}\}$$

$$\mathbf{G}_{\overline{\mathbf{E}}} := \{g \in \overline{\mathbb{R}}^{\mathcal{Y}} \mid \overline{\mathbf{E}}(g) \leq 0\}$$

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \ldots, N$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Testing the hypothesis that $p \leq \frac{1}{2}$

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M \geq 0$.
Reality announces $y \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M(y - \frac{1}{2})$.

Testing a probability forecaster

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Forecaster announces a probability distribution P_n on $\mathcal{Y}$.
Skeptic announces $S_n : \mathcal{Y} \to \overline{\mathbb{R}}$ such that $\mathbf{E}_{P_n}(S_n)$ exists.
Reality announces $y_n \in \mathcal{Y}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + S_n(y_n)$.

Local offers satisfy the axioms if Skeptic can

1. have negative capital and
2. waste money.

In each protocol, the offers satisfy the axioms for offers.

Major result in *Game-Theoretic Foundations* (2019)
(first posted as a working paper in 2009)

If a protocol’s local upper expectations satisfy the axioms, then the global one does too.

Generalizes Kolmogorov’s extension theorem.

Closely related to

- Doob’s martingale convergence theorem,
- Lévy’s zero-one law.

Continuity axiom not required.

5. Global upper expected value

It is well known that in the Middle Ages all scholastic philosophers advocated Aristotle's "infinitum actu non datur" as an irrefutable principle.

Georg Cantor, 1866

Abstract protocols

PARAMETER: Nonempty set $\mathcal{Y}$

Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.

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

Forecaster announces an upper expectation $\overline{\mathbf{E}}_n$ on $\mathcal{Y}$.

Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{Y}}$ such that $\overline{\mathbf{E}}_n(f_n) \leq \mathcal{K}_{n-1}$.

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

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

PARAMETERS: Nonempty set $\mathcal{Y}$; upper expectation $\overline{\mathbf{E}}$ on $\mathcal{Y}$

Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.

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

Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{Y}}$ such that $\overline{\mathbf{E}}(f_n) \leq \mathcal{K}_{n-1}$.

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

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

For abstract theory, the second protocol is adequate. World combines Forecaster and Reality.

PARAMETERS: Nonempty set $\mathcal{Y}$; upper expectation $\overline{\mathbf{E}}$ on $\mathcal{Y}$

Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.

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

Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{Y}}$ such that $\overline{\mathbf{E}}(f_n) \leq \mathcal{K}_{n-1}$.

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

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

Set	Elements of the set
T	supermartingales
L	supermartingales that converge in $\overline{\mathbb{R}}$
M	martingales

Definition of global upper expected value

$$\overline{\mathbb{E}}(X) := \inf\left\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T}, \inf \mathcal{T} > -\infty, \liminf_{n\to\infty} \mathcal{T}_n \geq X\right\}$$

Recall simpler definition for the finite-horizon protocol for testing *p*.

$$\overline{\mathbb{E}}(X) := \inf\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T} \text{ and } \mathcal{T}_N \geq X\}$$

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.

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

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

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

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

Definition of global upper expected value

$$\overline{\mathbb{E}}(X) := \inf\left\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T}, \inf \mathcal{T} > -\infty, \liminf_{n\to\infty} \mathcal{T}_n \geq X\right\}$$

We can also use the same definition in a non-initial situation:

$$\overline{\mathbb{E}}_s(X) := \inf\left\{\mathcal{T}(s) \mid \mathcal{T} \in \mathbf{T}, \inf \mathcal{T} > -\infty, \forall \omega \in \Omega_s : \liminf_{n\to\infty} \mathcal{T}_n(\omega) \geq X(\omega)\right\}$$

$$\overline{\mathbb{E}}(X) = \overline{\mathbb{E}}_\square(X)$$

Definition of "almost sure"

An event E is *almost sure* if there exists a nonnegative supermartingale that tends to ∞ on all paths outside E.

Definition of global upper expected value

$$\overline{\mathbb{E}}(X) := \inf\left\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T}, \inf \mathcal{T} > -\infty, \liminf_{n\to\infty} \mathcal{T}_n \geq X\right\}$$

Lévy's zero-one law will imply that $\overline{\mathbb{E}}$ is an upper expectation. (page 160)

Once we know that $\overline{\mathbb{E}}$ is an upper expectation, it is easy to show that E is almost sure if and only if $\overline{\mathbb{P}}(E^c) = 0$. (page 161)

6. The supermartingale of upper expected values

… [Doob's] theorem had ancestors in two different frameworks: in Lebesgue's theory of integration there was his proof (1903) of the theorem about differentiation almost everywhere; in Borel's theory of denumerable probabilities there was his statement and proof by probabilistic arguments (1909) of the almost sure convergence of frequencies in the game of heads or tails—the first version of the strong law of law numbers.

Bernard Bru, 2009

Doob’s martingale convergence theorem

A nonnegative supermartingale that begins with a finite value converges in $\mathbb{R}$ almost surely.

This can be proven in our abstract protocol using Doob’s upcrossing argument.

Definition of global upper expected value

$$\overline{\mathbb{E}}_s(X) := \inf\left\{\mathcal{T}(s) \mid \mathcal{T} \in \mathbf{T}, \inf \mathcal{T} > -\infty, \forall \omega \in \Omega_s : \liminf_{n\to\infty} \mathcal{T}_n(\omega) \geq X(\omega)\right\}$$

Proposition 7.7. *If X is a bounded below global variable and $s \in \mathbb{S}$, then*

$$\overline{\mathbb{E}}_s(X) = \inf\left\{\mathcal{T}(s) \mid \mathcal{T} \in \mathbf{T} \text{ and } \forall \omega \in \Omega_s : \limsup_{n\to\infty} \mathcal{T}_n(\omega) \geq X(\omega)\right\}$$

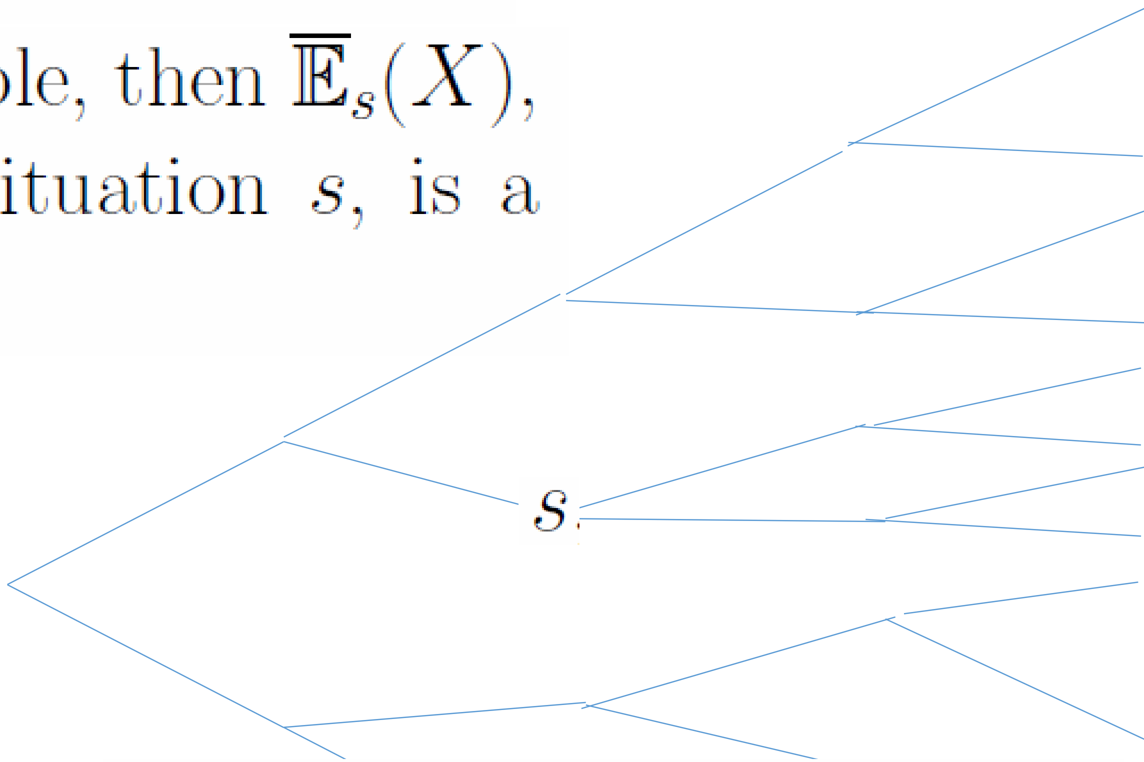
$$= \inf\left\{\mathcal{L}(s) \mid \mathcal{L} \in \mathbf{L} \text{ and } \forall \omega \in \Omega_s : \lim_{n\to\infty} \mathcal{L}_n(\omega) \geq X(\omega)\right\}$$

The Supermartingale of Upper Expected Values

Theorem. If X is a global variable, then $\overline{\mathbb{E}}_s(X)$, considered as a function of the situation s, is a supermartingale.

We can also write the supermartingale as $\mathbb{E}_0(X), \mathbb{E}_1(X), \ldots$.

(Any process can be written both as a sequence of random variables and as a function on the situations.)



Here s is a situation at time 2. $\mathbb{E}_s(X)$ for this particular s is one of the 4 possible values of $\mathbb{E}_2(X)$.

7. Lévy's zero-one law

The first martingale convergence theorem is the celebrated Paul Lévy 0-1 law. It is perhaps one of the most beautiful results of probability theory.

Michel Loève, 1973

Lévy's zero-one law

Theorem 8.1. *If X is a bounded-below global variable in Protocol 7.1, then*

$$\liminf_{n\to\infty} \overline{\mathbb{E}}_n(X) \geq X \quad a.s.$$

We prove it with a more complicated version of Doob's upcrossing argument.

Corollary 8.9. *Suppose X is a bounded global variable and has an expected value. Then*

$$\lim_{n\to\infty} \mathbb{E}_n(X) = X \quad a.s.$$

What Lévy said

Corollary 8.10. *Suppose the event E has a probability. Then*

$$\lim_{n\to\infty} \mathbb{P}_n(E) = \mathbf{1}_E \quad a.s.$$

Why we call it a Zero-one law

8. Classical, measure-theoretic, game-theoretic

The true basis of the probability calculus is the principle of compound probability, which allows us to replace two experiences with a single experience.

Paul Lévy, 1954

A century ago, probability theory was often based on two axioms:

Total probability. For mutually exclusive A and B,

$$\mathbf{Prob}(A \text{ or } B) = \mathbf{Prob}(A) + \mathbf{Prob}(B).$$

Compound probability. For successive A and B,

$$\mathbf{Prob}(A \text{ and } B) = \mathbf{Prob}(A) \times \mathbf{Prob}(B \text{ after } A \text{ happens}).$$

The rule of compound probability allows us to construct a distribution for a sequence of variables $Y_1, Y_2, \ldots$ step-by-step.

- First specify a distribution for Y_1.
- Then a distribution for Y_2 for each value y_1 of Y_1.
- Then a distribution for Y_3 for each pair y_1, y_2 of possible values of Y_1, Y_2.
- Etc.

We still do it this way in math stat.

Cassius Ionescu Tulcea
1923-2021

In 1949, as measure theory was ascendant, Ionescu Tulcea gave measurability conditions under which the classical construction produces a probability measure.

Now we pretend that the measure comes first, thus imposing the awkward notion of “conditional expectation” on ourselves.

Alexander Philip Dawid
Born 1946

In 1985, Dawid gave new theoretical life to the classical construction by calling its ingredients a *probability forecasting system*.

In game-theoretic probability, it is a *strategy for forecaster*.

Protocol 9.1.

PARAMETER: Measurable space $(\mathcal{Y}, \mathcal{F})$

Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.

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

Forecaster announces $P_n \in \mathcal{P}(\mathcal{Y})$.

Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{Y}}$ such that $P_n(f_n) \leq \mathcal{K}_{n-1}$.

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

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

- Let ϕ be a strategy for Forecaster that uses only information in the game.
- Call ϕ a *probability forecasting system.*
- Let P_ϕ be the probability measure on $\mathcal{Y}^\infty$ constructing by using ϕ's moves as conditional probabilities.

Ionescu Tulcea

When Forecaster uses a probability measure's conditional probabilities as a strategy, the global expected values agree with the probability measure.

Protocol 9.2.

PARAMETERS: Measurable space $(\mathcal{Y}, \mathcal{F})$, probability forecasting system ϕ

Skeptic announces $\mathcal{K}_0 \in \overline{\mathbb{R}}$.

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

Skeptic announces $f_n \in \overline{\mathbb{R}}^{\mathcal{Y}}$ such that $\phi_n(y_1, y_2, \ldots)(f_n) \leq \mathcal{K}_{n-1}$.

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

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

Ville's Theorem. Every bounded measurable function X on $\mathcal{Y}^\infty$ has a game-theoretic expected value in Protocol 9.2, and $\overline{\mathbb{E}}(X) = \mathbf{E}_{P_\phi}(X)$.

Lévy’s zero-one law also implies a duality between game-theoretic and measure-theoretic probability:

Theorem 9.7 in Game-Theoretic Foundations
The upper expected values we obtain as infima over game-theoretic supermartingales are also suprema over probability measures.

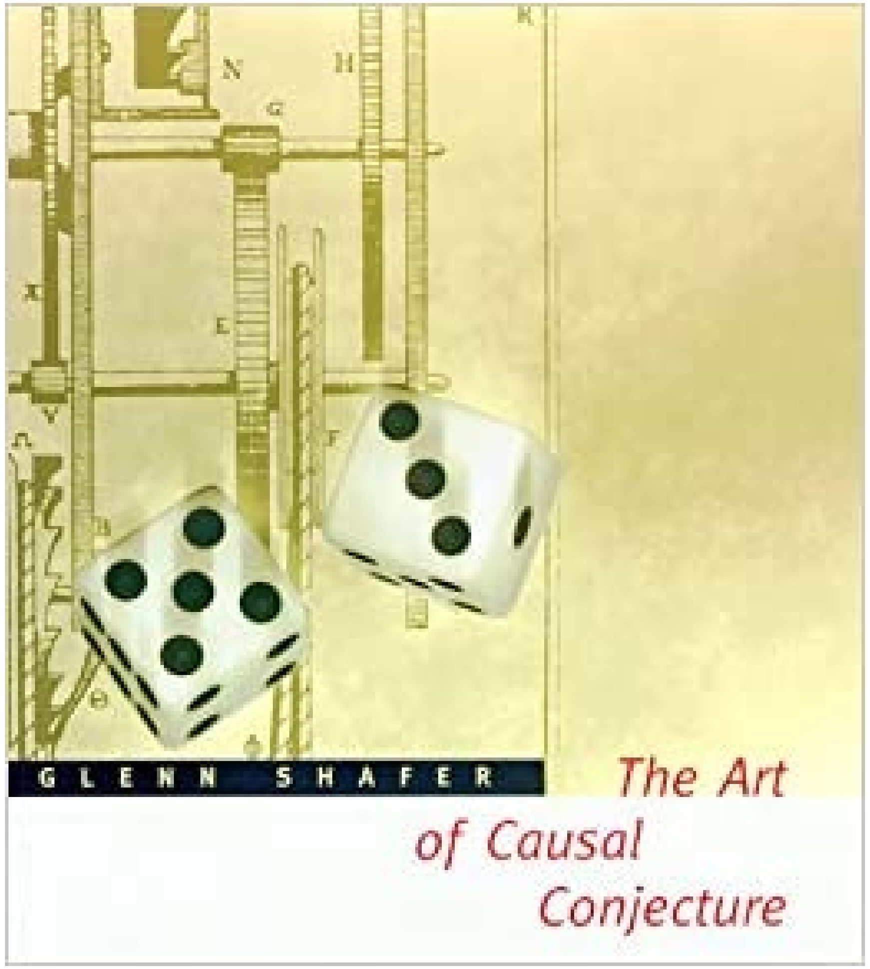
Much stronger regularity conditions required here.

9. Independence and causality

Were counterfactuals to have objective meaning, we might take them to be basic, and define probability and causality in terms of them.

Glenn Shafer, 1996

Exercise 7.11 (causal independence). Show that if X and Y are nonnegative global variables in a testing protocol such that $X(\omega)$ depends only on ω_j and $Y(\omega)$ only on ω_k, where $j < k$, then $\overline{\mathbb{E}}(XY) = \overline{\mathbb{E}}(X)\overline{\mathbb{E}}(Y)$. $\square$



I would love to see the results of my 1996 book on probabilistic causality (probability trees) restated in game-theoretic terms.