When are probabilities predictions?

Glenn Shafer, Rutgers University

Philog (Seminar in Philosophy, Logic and Games)
January 31, 2024

These slides are posted at www.glennshafer.com, talk # 210.

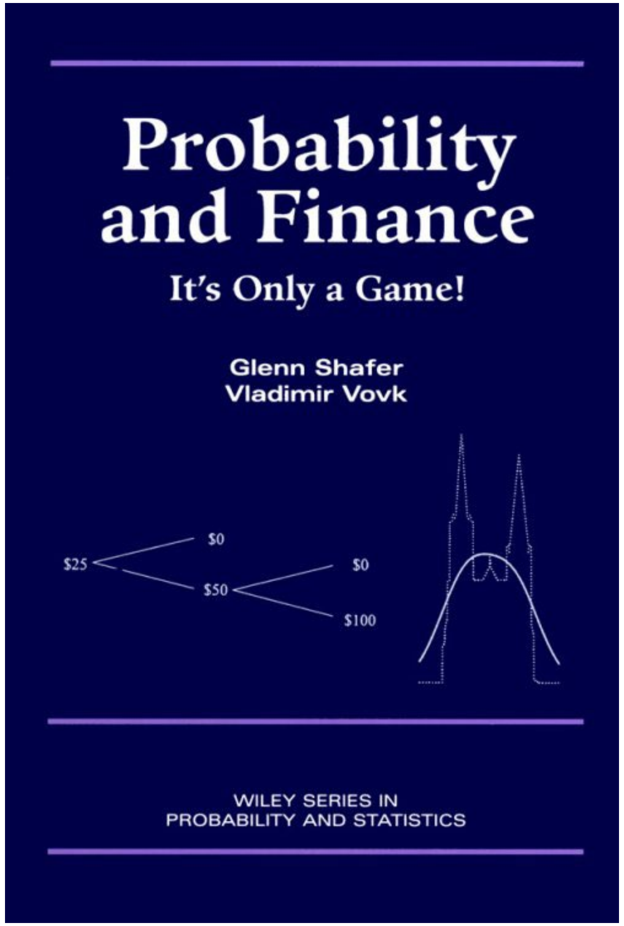
Abstract

In everyday English, a forecast is something less than a prediction. It is more like an estimate. When an economist forecasts 3.5% inflation in the United States next year, or my weather app forecasts 0.55 inches of rain, these are not exactly predictions. When the forecaster gives rain a 30% probability, this too is not a prediction. A prediction is more definite about what is predicted and about predicting it.

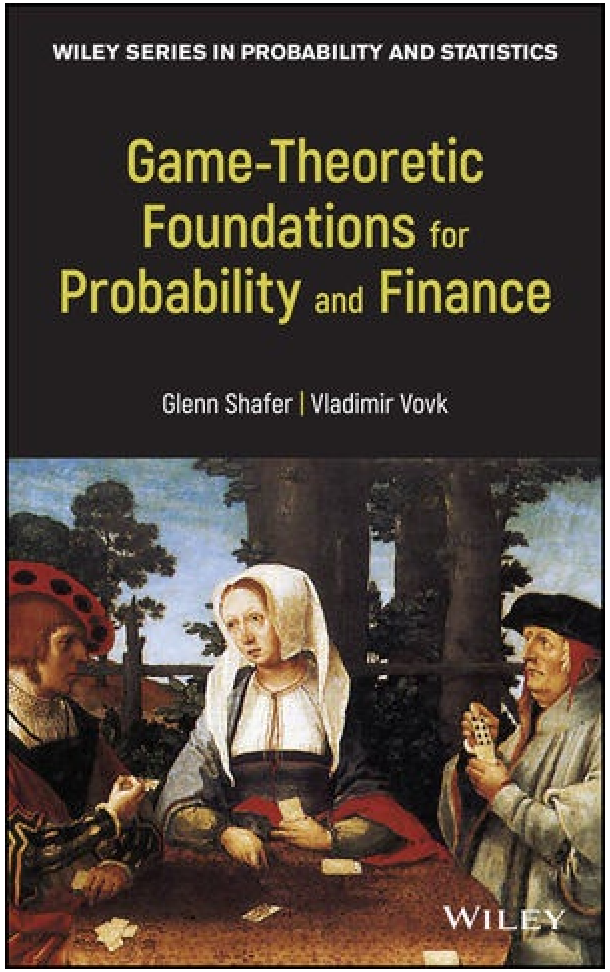
We might say that a probability is a prediction when it is very close to one. But this formulation has a difficulty: there are too many high probabilities. There is a high probability against every ticket in a lottery, but we cannot predict that no ticket will win.

Game-theoretic statistics resolves this problem by showing how some high probabilities are simpler than others. The simpler ones qualify as predictions.

This story has roles for Cournot's principle, Kolmogorov's algorithmic complexity, and de Finetti's *previsione*. See www.probabilityandfinance.com and my two books on the topic with Vladimir Vovk.



2001



2019

Showed by example that the classical limit theorems can be proven in game theory.

- Each proof is a betting strategy.
- So more constructive than measure theory.

- Puts game-theoretic probability on a par with measure-theoretic probability as abstract theory.
- New applications (forecasting, decision, CAPM, equity premium, stochastic calculus, calibration, etc.)

Outline

1. In everyday English, a "forecast" is less definite and categorical than a "prediction". Do you agree?

2. Bruno de Finetti advocated calling probabilities and expected values "forecasts" and saw no respectable role at all for the word "prediction".

3. Contrary to de Finetti, I contend that we can use some high probabilities as predictions: The simple ones.

4. Game-theoretic probability explains the notion of a simple probability: it is one proven by a simple betting strategy.

5. This is the game-theoretic version of Cournot's principle.

Part 1

Do you agree that “forecast” is less definite and categorical than “prediction”?

Or do you think that “forecast” and “prediction” are simply synonyms?

Glenn’s argument for “forecast” being less categorical than “predict”.

When I forecast an inch of rain tomorrow, no one imagines that I expect exactly an inch.

When I predict that my team will win tomorrow’s game, I may be trying to convince you that I know for sure.

History

Crop forecasting was important in the mid 19th century.
How much cotton will be produced?
Jamie Pietruska, *Looking Forward*, Chicago 2017.

In 1923, the president of the American Statistical Association was selling the Harvard Business Forecasts.
Walter A. Friedman, *Fortune Tellers: The Story of America's First Economic Forecasters*, Princeton 2014.

My dictionary

Here is how the unabridged dictionary on my shelf, the 2011 edition of *The American Heritage Dictionary of the English Language*, begins its definition of "forecast":

> To estimate or predict in advance, especially to predict (weather conditions) by analysis of meteorological data.

And here is how it begins its definition of "predict":

> To state, tell about, or make known in advance, especially on the basis of special knowledge. . .

The suggestion that the subject of the verb might be providing only an estimate appears in the leading definition of "forecast" but not in the leading definition of "predict".

Financial forecasting in 2024

Financial professionals usually use “forecast” rather than “prediction” to refer to an estimate of a future number.

- Government Finance Officers Association: A financial forecast is a fiscal management tool that presents estimated information based on past, current, and projected financial conditions.
- Harvard Business School: Financial forecasting is important because it informs business decision-making regarding hiring, budgeting, predicting revenue, and strategic planning.
- Investopedia: Earnings forecasts are based on analysts' expectations of company growth and profitability.

But when you are hyping your forecast to a mass audience, you call it a "prediction".

Some google hits

"sports prediction"	938,000
"sports forecasting"	55,500
"sports forecast"	91,800
"election prediction"	517,000
"election forecasting"	89,100
"election forecast"	798,000

Part 2

Bruno de Finetti advocated calling probabilities and expected values “forecasts”.

He saw no respectable role at all for the word “prediction”.

Pairs like *forecast / prediction*:

French	*prévision / prédiction*
Italian	*previsione / predizione*

Can you tell me about other languages?

In 1970, in *Teoria Delle Probabilità, Sintesi introduttiva con appendice critica,* Bruno de Finetti noted the difference between *previsione* and *predizione* and proposed that *previsione* (forecast) should replace the traditional *speranza matematica* (mathematical expectation).

An English translation, *Theory of Probability: A critical introductory treatment,* appeared in 1974/1975. It translated *previsione* by *prevision* rather than by the understandable *forecast.*

The first two paragraphs of Section 1.2 of Chapter III of Bruno de Finetti's *Teoria Delle Probabilità. Sintesi introduttiva con appendice critica*, Giulio Einaudi, 1970:

1.2. *Previsione, non predizione*. Per usare queta parola, «previsione», biognerà insistere e ricordare quale sia il senso ben previso che ad essa (e derivati) si deve dare e daremo costantemente e scrupolosamente nel seguito, distinguendolo ed anzi contrapponendolo a un altro che nel linguaggio corrente le vience forse più comunemente attribuito, e per il quale riserviamo l'altro termine, «*predizione*».

Fare una *predizione* significherebbe (usando il termine nel senso che proponiamo) avventurarsi a cercar di «indovinare», fra le alternative possibili, quella che avverrà, cosí come pretendono spesso non solo sedicenti maghi e profeti ma anche esperti ed altre persone incline a precorrere il futuro nella fucina della loro fantasia. Pertanto, fare una «predizione» significherebbe non già uscire dall'ambito della logica del certo ma semplicemente intrudervi insieme alla verità accertate e ai dati rilevati altre affermazioni e altri date che si pretende indovinare. Né basta attenuare il carattere «profetico» di siffatte enunciazioni cautelandosi con i riempitivi («credo», «forse», ecc.) già menzionati, ché essi o rimangono aggiunte posticce sprovviste di autentico significato o richiedono d'essere effettivamente tradotti in termini probabilistici, sostituendo la predizione con un previsione.

Translation from the Italian, with the help of ChapGTP 3.5, Google Translator, and other dictionaries.

1.2. *Forecast, not prediction*. To use this word, "forecast," it will be necessary to insist on and remember the well-defined sense that must be given to it (and its derivatives), and that we will consistently and scrupulously give to it in the sequel, distinguishing and even contrasting this sense with another sense that is perhaps more commonly attributed to it in everyday language, and for which we reserve the other term, "*prediction*."

Making a *prediction* would mean (using the term in the sense we propose) venturing to try to "guess," among the possible alternatives, the one that will occur, as often done not only by self-proclaimed magicians and prophets but also by experts and other individuals inclined to foresee the future in the forge of their imagination. Therefore, making a "prediction" would mean not leaving the realm of the logic of certainty but simply injecting into it, along with the ascertained truths and the collected data, other statements and other dates that one claims to divine. Nor is it enough to attenuate the "prophetic" character of such statements by taking precautions with the fillers already mentioned ("I believe," "maybe," etc.), because they either remain artificial additions devoid of authentic meaning or need to be effectively translated into probabilistic terms, replacing the prediction with a forecast.

Beginning of Section 1.3 of Chapter III of Bruno de Finetti's *Teoria Delle Probabilità. Sintesi introduttiva con appendice critica*, Giulio Einaudi, 1970:

1.3. La *previsione*, nel senso in cui abbiamo detto di voler usare questa parola, no si propone di indoviare nulla: non afferma --- come la predizione --- un qualcosa che potrà risultare o vero or falso trasformando velleitariamente l'incertezza in pretesa ma fasulla certezza. Riconosce (come sembrerebbe dover essere ovvio) che l'incerto è incerto, che in fatto di affermazioni tutto quel che si può dire oltre ciò che è detto dalla logica del certo è illegittimo...

My translation:

Forecasting, in the sense in which we have said we want to use this word, does not aim to divine anything: it does not assert—like prediction—something that may turn out to be true or false, whimsically transforming uncertainty into a false claim of certainty. It recognizes (as it would seem obvious) that the uncertain is uncertain, that when it comes to assertions, anything beyond what is dictated by the logic of certainty is illegitimate...

Part 3

Contrary to de Finetti, I think we can use some high probabilities as predictions.

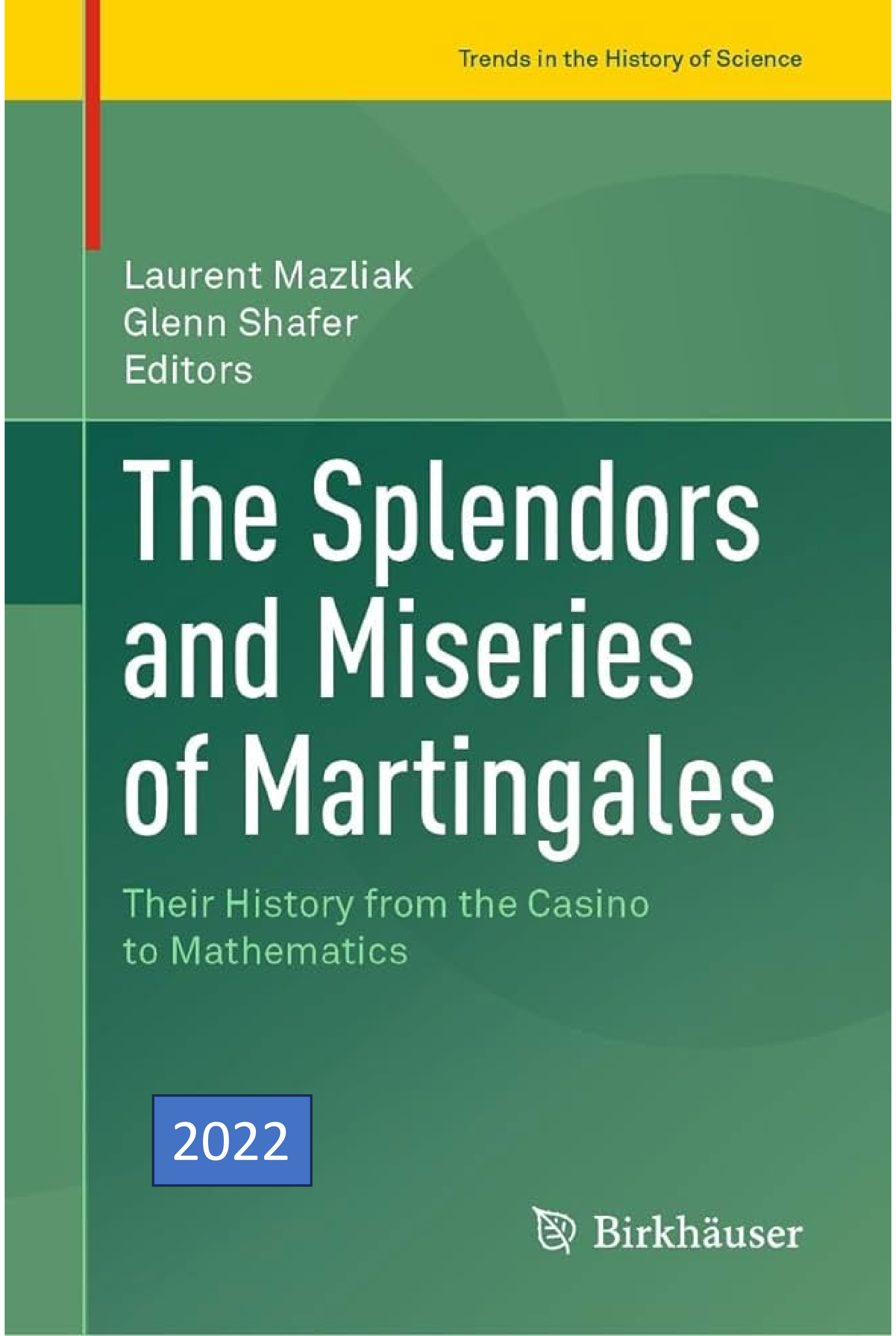
Because of the lottery paradox, we cannot use all high probabilities as predictions.

But we can use simple high probabilities as predictions.

Part 4

Game-theoretic probability explains the notion of a simple probability.

It is one proven by a simple betting strategy.



2022

Pierre Crépel interviewed Jean Ville in 1984, taking notes in French.

I turned his notes into a narrative in English, published on pages 375-391 of this book.

Jean Ville explained (p. 383):

> "The more complicated a probability law, the longer it takes to describe the martingale that would make it happen. See Kolmogorov."

Jean Ville explained:

The more complicated a probability law, the longer it takes to describe the martingale that would make it happen.

See Kolmogorov.

What did he mean?

Examples of probability laws:

law of large numbers

law of the integrated logarithm

These probability laws give high probabilities to certain events.

Jean Ville explained:

The more complicated a probability law, the longer it takes to describe the martingale that would make it happen. See Kolmogorov.

What did he mean?

The capital process of a gambling strategy is called a *martingale*.

Simple strategy = simple martingale.

$P(E)=1$ & $P(E^c)=0$

$\Leftrightarrow$

There is a betting strategy that multiplies its money infinitely unless E happens.

$P(E)=0.95$ & $P(E^c)=0.05$

$\Leftrightarrow$

There is a betting strategy that multiplies its money by 20 unless E happens.

Borel's strong law of large numbers (1909)

Consider an infinite sequence of independent trials of an event with probability p. Write

- Let r_n be the number of times the event happens in the first n trials.
- Let $\overline{y}_n$ be the frequency: $\overline{y}_n := r_n/n$.

Borel proved that $\mathbb{P}\left(\lim_{n\to\infty}\overline{y}_n = p\right) = 1$.

Ville gave a game-theoretic proof. He showed that the martingale

$$\mathcal{T}_n := \frac{r_n!(n-r_n)!}{(n+1)!}p^{-r_n}(1-p)^{-(n-r_n)}$$

goes to ∞ unless $\overline{y}_n \to p$.

Chebyshev's law of large numbers

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

$$\mathbb{P}\left(|\bar{y}_N - p| \geq \epsilon\right) \leq \frac{p(1-p)}{\epsilon^2 N}$$

Set $w_i := y_i - p$,

$$\mathcal{T}_n := \left(\sum_{i=1}^{n} w_i\right)^2 - \sum_{i=1}^{n} w_i^2 + (1-2p)\sum_{i=1}^{n} w_i,$$

and

$$\mathcal{U}_n := \frac{\mathcal{T}_n + Np(1-p)}{\epsilon^2 N^2}.$$

The process $\mathcal{U}$ is a nonnegative martingale that multiplies its money by $\epsilon^2 N/p(1-p)$ if $|\overline{y}_N - p| \geq \epsilon$.

Game-theoretic probability generalizes to the case where the forecaster offers fewer bets.

FOR $n = 1, \ldots, N$:
  Forecaster announces $\mu_n \in [-1, 1]$.
  Skeptic announces $z_n \in \mathbb{R}$.
  Reality announces $y_n \in [-1, 1]$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + z_n(y_n - \mu_n)$.

$$\overline{\mathbb{P}}\left(|\bar{y}_N - \overline{\mu}_N| \geq \epsilon\right) \leq \frac{4}{\epsilon^2 N}$$

The nonnegative martingale:

$$\mathcal{T}_n := \frac{r_n!(n-r_n)!}{(n+1)!} p^{-r_n}(1-p)^{-(n-r_n)}$$

The betting strategy:

- Risk $(r_n+1)/(n+1)$ on the event happening.
- Risk $(n-r_n+1)/(n+1)$ on the event not happening.

The proof that $\mathcal{T}_n \to \infty$ if $\overline{y}_n \not\to p$:

Apply Stirling's formula to $\mathcal{T}_n$.

Part 5

Cournot's principle:

- We connect probabilities with phenomena by predicting that events with high probability will happen.
- We discredit probabilities by observing the happening of an event with high probability.

Game-theoretic version:

- We predict with (upper) probabilities that are high and also simple.
- We discredit forecaster by betting against him.

Jacob Bernoulli (1654-1705) said that high numerical probability is a prediction (i.e., the thing is morally certain).

Condorcet (1743-1794) said that the principle “high probability = prediction” is outside the mathematics of probability.

Cournot (1801-1877) said it is the only way to connect numerical probability with phenomena.

How Cournot said it

A probability of 1000 to 1 is almost considered equivalent to certainty, and one can hardly make the same judgement about a probability of 12 to 1.

. . . *The physically impossible event is therefore the one that has infinitely small probability*, and only this remark gives substance — objective and phenomenal value — to the theory of mathematical probability

In practice, moreover, and in the world of realities, what mathematicians call an infinitely small probability is and can only be an exceedingly small probability. The tip of this very sharp needle is not a mathematical point like the apex of the cone in question. Viewed through a magnifying glass, it becomes a *blunt* tip.

Traditionally, high probabilities were interpreted as predictions.

See my working paper:

"That's what all the old guys said"

(www.probabilityandfinance.com/articles/60.pdf)

I quote almost 100 scholars, from John of Salisbury (1115-1180) to A. Philip Dawid (born 1946).

Most adopted Cournot's principle in one form or another.

Four languages:

- Latin (*probabilitas*)
- French (*probabilité*)
- English (*probability*)
- German (*Wahrscheinlichkeit*)

Status of Cournot's principle

De Finetti (Trieste, 1951): There are many variations of these fallacious opinions:

(i) the mere misinterpretation of the correct Neyman formulation,

(ii) the recourse to the so called "principle of Cournot" (rejecting the possibility of events with "very small probability"),

(iii) the direct adoption of a frequency definition of probability or of an assumption connecting frequency and probability ("empirical law of randomness").

Ray Briggs (Stanford philosophy), in the *Stanford Encyclopedia of Philosophy*:

Standard probability theory rejects *Cournot's Principle*, which says events with low or zero probability will not happen. But see Shafer (2005) for a defense of Cournot's Principle.

Alan Hajek (Australian National University, philosophy)

The principle still has some currency, having been recently rehabilitated and defended by Shafer.

In *Ten Great Ideas About Chance* (2018), Persi Diaconis (Stanford statistics) and Bryan Skyrms (Irvine philosophy) dismiss Jacob Bernoulli's use of his law of large numbers as "Bernoulli's swindle". Cournot's principle, they say,

> . . . is a remarkably persistent fallacy, easy to swallow in the absence of rigorous thinking. We find it in the French mathematician and philosopher Cournot (1843), who holds that small-probability events should be taken to be physically impossible. He also held that this principle . . . connects probabilistic theories to the real world. . .
>
> This mantra was repeated in the twentieth century by very distinguished probability theorists, including Emile Borel, Paul Levy, Andrey Markov, and Andrey Kolmogorov. We cannot help but wonder whether this was to some extent a strategy for brushing off philosophical interpretational problems, rather than a serious attempt to confront them.

John of Salisbury, c. 1115–1180

Sola enim probabilitas dialectico sufficit

Thomas Aquinas, 1225–1274

Et ideo sufficit probabilis certitudo...

Jean Gerson, 1363–1429

non enim consurgit certitudo moralis ex evidentia demonstrationis,
sed ex probabilibus conjecturis

Thomas Granger, 1578–1627

Many probabilities concurring prevail much.

René Descartes, 1596–1650

deux sortes de certitudes. La premiere est apelée morale,
c'est à dire suffisante pour regler nos mœurs…

Antoine Arnauld, 1612–1694, and Pierre Nicole, 1625–1695

nous nous devons contenter d'une certitude morale dans les choses qui ne
sont pas susceptibles d'une certitude metaphysique

John Locke, 1632–1704

some of them border so near upon certainty, that we make no doubt at all about them

Jacob Bernoulli, 1655–1705

Something is *morally certain* if its probability comes so close to complete certainty that the difference cannot be perceived

John Arbuthnot, 1667–1735

Georges-Louis Buffon, 1707–1788

David Hume, 1711–1776

Denis Diderot, 1713–1784

Jean Le Rond d'Alembert, 1717–1783

Nicolas de Condorcet, 1743–1794

Pierre Simon Laplace, 1749–1827

Joseph Fourier, 1768–1830

André-Marie Ampère, 1775–1836

Siméon Denis Poisson, 1781–1840

Thomas Galloway, 1796–1851

Antoine Augustin Cournot, 1801–1877

Augustus De Morgan, 1806–1871

Jules Gavarret, 1809–1890

John Venn, 1834–1923
Wilhelm Lexis, 1837–1914
Hermann Laurent, 1841–1908
Ludwig Boltzmann, 1844–1906
Paul Mansion, 1844–1919
Francis Edgeworth, 1845–1926
Henri Poincaré, 1854–1912
Andrei Markov, 1856–1922
Karl Pearson, 1857–1936
Guido Castelnuovo, 1865–1952
Jacques Hadamard, 1865–1963
Ladislaus von Bortkiewicz, 1868–1931
Georg Bohlmann, 1869–1928
Arthur Lyon Bowley, 1869–1957
Emile Borel, 1871–1956
George Udny Yule, 1871–1951
Aleksandr Chuprov, 1874–1926

Maurice Fréchet, 1878–1973
Evgeny Slutsky, 1880–1948
Richard von Mises, 1883–1953
James V. Uspensky, 1883–1947
Hermann Weyl, 1885–1955
Paul Lévy, 1886–1971
Oskar Anderson, 1887–1960
Charlie Dunbar Broad, 1887–1971
R. A. Fisher, 1890–1962
Harold Jeffreys, 1891–1989
Thornton Fry, 1892–1991
Harald Cramér, 1893–1985
Jerzy Neyman, 1894–1981
David van Dantzig, 1900–1959
Karl Popper, 1902–1994
Abraham Wald, 1902–1950
Marshall Stone, 1903–1989
Andrei Kolmogorov, 1903–1987
Carl Hempel, 1905–1997
Hans Freudenthal, 1905–1990
William Feller, 1906–1970

Joseph Doob, 1910–2004
Jean Ville, 1910–1989
Trygve Haavelmo, 1911–1999
Hans Richter, 1912–1978
Charles Stein, 1920–2016
Yuri Prokhorov, 1929–2013, and Boris Sevast'yanov, 1923—2013
David R. Cox, 1924–2022, and David V. Hinkley, 1944–2019
John Stewart Bell, 1928–1990
Henry Kyburg, Jr., 1928–2007
Hugh Everett III, 1930–1982
Terrence Fine, 1939–2021
Per Martin-Löf, born 1942
Donald Gillies, born 1944
A. Philip Dawid, born 1946

Two ways game-theoretic probability can improve data analysis
https://arxiv.org/abs/2308.14959

"That's what all the old guys said": The many faces of Cournot's principle
http://probabilityandfinance.com/articles/60.pdf