Modernizing Cournot's Principle

Glenn Shafer, Rutgers University

Department of Statistics and Data Science
Carnegie Mellon University

March 25, 2024

These slides were prepared on March 17, 2024 and are available at www.glennshafer.com.

Abstract

To understand multiple testing, we need to revisit Cournot's principle, which says that events of high probability are practically certain. This principle was considered fundamental by scores of authorities, including Aquinas, Bernoulli, Condorcet, Borel, Levy, Kolmogorov, Ville, and Doob.

Critics of Cournot's principle often evoke the lottery paradox: an event with small probability always happens. Why was this paradox overlooked before the 1960s? Possible answers: (1) Earlier authors thought about "certainty" differently. (2) They did not begin with a probability measure.

To modernize Cournot's principle, let's (1) replace certainty with prediction and (2) use only simple high-probability forecasts as predictions. This works best with game-theoretic probability (testing by betting).

Condorcet taught that Cournot is outside probability's mathematics. What does this say about multiple testing and multiple prediction?

Cournot says that events of high probability are practically certain. This principle was considered fundamental by Aquinas, Bernoulli, Condorcet, Borel, Levy, Kolmogorov, Ville, Doob, and many others.

Critics often evoke the lottery paradox: a small probability event always happens.

Why was the lottery paradox overlooked before the 1960s?

1. Earlier authors thought about "certainty" differently.
2. Earlier authors did not begin with a probability measure.

To modernize Cournot's principle,

1. **replace certainty with prediction** and
2. **use only simple high-probability forecasts as predictions**.

This works best with game-theoretic probability.

Condorcet taught that Cournot is outside probability's mathematics. What does this tell us about multiple testing and multiple prediction?

Part 1. Cournot’s principle in its classical form

Part 2. The lottery paradox

Part 3. Replace practical certainty by prediction

Part 4. Cournot’s principle in game-theoretic form

Part 5. What Condorcet teaches about multiple testing

Part 1. Cournot's principle in classical form

My working paper:

"That's what all the old guys said"

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

Quotes nearly 100 scholars over nearly 1000 years.

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

Before probability was numerical

Thomas Aquinas, 1225–1274

Et ideo sufficit probabilis certitudo...

And therefore probable certainty is sufficient...

Jean Gerson, 1363–1429

Denique certitudo quae moralis dici potest vel civilis tangitur ab Aristotele . . . non enim consurgit certitudo moralis ex evidentia demonstrationis, sed ex probabilibus conjecturis, grossis et figuralibus, magis ad unam partem quam ad alteram.

... the certainty that can be called moral or civil is touched on by Aristotle ... moral certainty arises not from the evidence of demonstration, but from probable conjectures, broad and figurative, more on one side than on the other.

	Singlular	Plural
Nominative	mos	mores
Genitive	moris	morum
Dative	mori	moribus
Accusative	morem	mores
Ablative	more	moribus
Vocative	mos	mores

1. Custom, usage, wont
2. Regular practice, rule, law
3. In plural: ways, conduct, character, morals

Cicero, we are told, coined the noun *probabilitas* and the adjective *moralis*.

2.17 Jacob Bernoulli, 1655–1705

Bernoulli's celebrated book on probability, *Ars Conjectandi*, was published posthumously in 1713 [14]. Here are two brief quotations, translated by Edith Sylla [15]:

- From Chapter I of Part IV: Something is *morally certain* if its probability comes so close to complete certainty that the difference cannot be perceived. ...

- From Chapter II of Part IV: Because ...it is rarely possible to obtain certainty that is complete in every respect, necessity and use ordain that what is only morally certain be taken as absolutely certain. It would be useful, accordingly, if definite limits for moral certainty were established by the authority of the magistracy. for instance, it might be determined whether 99/100 of certainty suffices or whether 999/1000 is required. ...

2.23 Nicolas de Condorcet, 1743–1794

In his famous and lengthy eulogy of Buffon, delivered to the Academy of sciences and published in 1790, we find the following passage [43, pp. 36–37]:

> Mr. de Buffon proposed that we assign a precise value to the very large probability that we can consider moral certainty, and beyond this to ignore the small possibility of a contrary event. This principle is true when we only want to make ordinary use of a calculation; and in this sense all men have adopted it in practice and all philosophers have followed it in their reasoning. But it ceases to be correct if we introduce it into the calculus itself, and especially if we want to use to establish theories, to explain paradoxes, and to prove or refute general rules. Besides, this probability, which may be called moral certainty, must be greater or smaller according to the nature of the objects considered and the principles that should guide our conduct; and it would have been necessary to fix the degree of probability at which it begins to be reasonable to believe and allowed to act for each type of truth and action.

2.27 Antoine Augustin Cournot, 1801–1877

. . . *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

… 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 …

* Cournot contrasted physical certainty/impossibility with metaphysical certainty/impossibility.

2.58 Paul Lévy, 1886–1971

We can only discuss the objective value of the notion of probability when we know the theory's verifiable consequences. They all flow from this principle: a sufficiently small probability can be neglected. In other words: *a sufficiently unlikely event can in practice be considered impossible.*

2.76 Jean Ville, 1910-1989

… *events having probabilities very close to* 1 *are practically certain (and therefore those whose probabilities are very small are practically impossible).*

In this way, we deal with two kinds of probabilities in the axiomatic theory: those that are close to 0 or to 1, which have a subjective meaning, quasi-impossibility or quasi-certainty, and those that are close neither to zero nor to 1, which have no subjective meaning when taken in isolation.

Lévy said <u>objective value</u>.

Ville, 24 years younger, said <u>subjective</u> meaning.

Why the difference?

Two sides of certainty:

The subject is certain about the object.

2.75 Joseph Doob, 1910–2004

If one starts with mathematical probability theory the obvious general operational translation principle is that one should ignore real events that have small probabilities. How small is "small" depends on the context, for example, the demands of a client on a statistician. Somewhat more precisely, one first makes a judgment on the possibility of the application of probability in a given context; if so, one then sets up a model and comes to operational decisions based on the principle that hypotheses must be reexamined if they ascribe small probability to a key event that actually happens. (This is, of course a great oversimplification.) ...

Part 2. The lottery paradox

Too many high probabilities.

Why not a problem before ≈ 1960?

Why was the lottery paradox not a problem before ≈ 1960?

One explanation:

People didn't think all probabilities are frequencies.

Draw 5 numbers between 1 and 90 without replacement.

- 5,273,912,160 possible outcomes
- 43,949,268 if you do not specify the order of the 5 numbers

The French government ignored the possibility that anyone would correctly guess the 5 numbers.

Casanova's Lottery, Stephen M. Stigler, 2022

Condorcet ignored it too.

His "lottery paradox" was that Bayes's rule fails when a less than perfectly reliable witness tells you which numbers were drawn.

Why was the lottery paradox not a problem before ≈ 1960?

Because probability was constructive.
A probability measure was not the starting point.

Georg Bohlmann's probability axioms, German encyclopedia of mathematics, 1901

1. If E is certain, then $p = 1$. If E is impossible, then $p = 0$.
2. Let p_1 be the probability that E_1 happens, p_2 the probability that E_2 happens, and p the probability that E_1 or E_2 happens. If E_1 and E_2 are mutually exclusive, then $p = p_1 + p_2$.
3. Let p_1 be the probability that E_1 happens, p_2' the probability that E_2 happens when one knows that E_1 has happened, and p the probability that E_1 and E_2 both happen. Then $p = p_1 p_2'$.

Today many (most?) mathematicians and philosophers consider the lottery paradox a decisive objection to Cournot’s principle.

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.

Persi Diaconis & Brian Skyrms

They call Cournot's principle

> . . . 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.

Part 3. Replace practical certainty by prediction

Cournot said…

- Model says events with high probability are practically certain.
- Test model by checking whether they happen.

Glenn says…

- Predict some events with high simple probabilities.
- Test by betting.

Cournot et al. said…

- Model says events with high probability are practically certain.
- Test model by checking whether they happen.

Glenn says…

- Predict some events with high simple probabilities.
- Test by betting.

- Prediction is obviously both subjective and objective.
- Prediction must be made in advance.
- Time and computational complexity limit us to simple predictions.
- We cannot make 5 billion predictions.

Distinguish between forecasts and predictions.

All probabilities (and all expected values) are forecasts.

We single out some high probabilities as predictions.

“Forecast” is 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.

Bruno de Finetti advocated calling probabilities and expected values “forecasts” (*previsione*).

He saw no respectable role at all for the word “prediction” (*predizione*).

Contrary to de Finetti, I think it is reasonable to use some probabilities as predicions – some simple ones.

Part 4. Cournot's principle in game-theoretic form

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

A simple probability is a probability proven by a simple betting strategy.

Cournot's principle in game-theoretic form

1. Test forecasters (including probability models) by betting (fixed strategy not required).

2. Predict using events with simple probabilities close to 0 or 1.

$$P(E) = \inf\left\{\alpha \,\middle|\, \begin{array}{c} \exists \text{ nonnegative supermartingale } T \text{ such} \\ \text{that } T_0 = \alpha \text{ and } T_N \geq 1 \text{ if } E \text{ happens} \end{array}\right\}$$

A probability is *simple* if it is proven by a simple supermartingale (= simple betting strategy).

Part 5. What Condorcet teaches us about multiple testing

To evaluate the success of testing a forecaster by betting, consider the reputation of the bettor, the total capital used, the rationale for the betting, etc.

When making multiple predictions, consider the simplicity of the proofs as well as the total capital they use.

EXTRA SLIDES

Simple supermartingales for the law of large numbers

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 $\overline{y}_n \to p$.

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 $\mathcal{T}_n$.

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

Chebyshev's law of large numbers

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

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

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

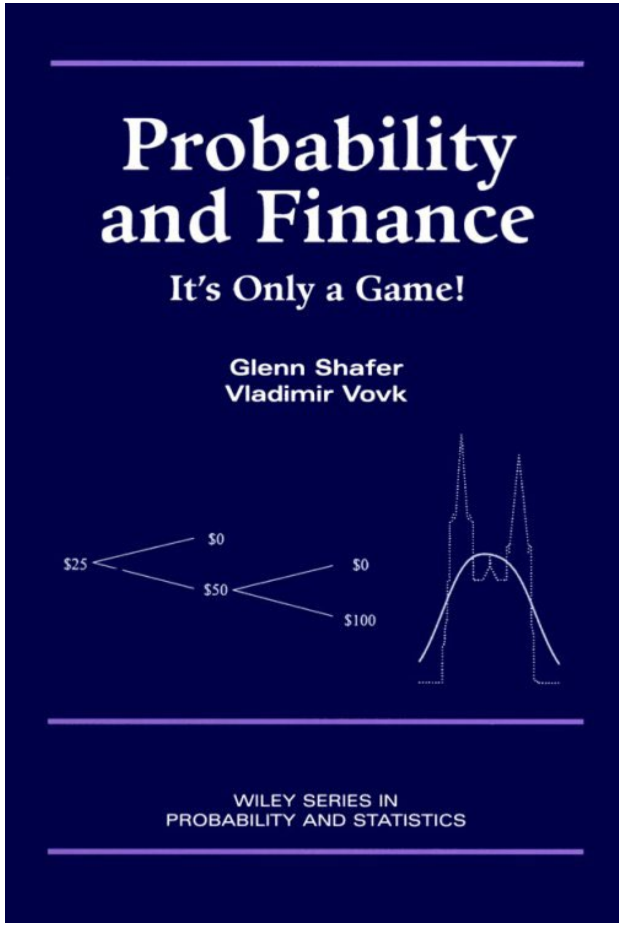
The process $\mathcal{U}$ is a nonnegative martingale that multiplies its money by $\epsilon^2 N/p(1-p)$ if $|\overline{y}_N - p| \geq \epsilon$.

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

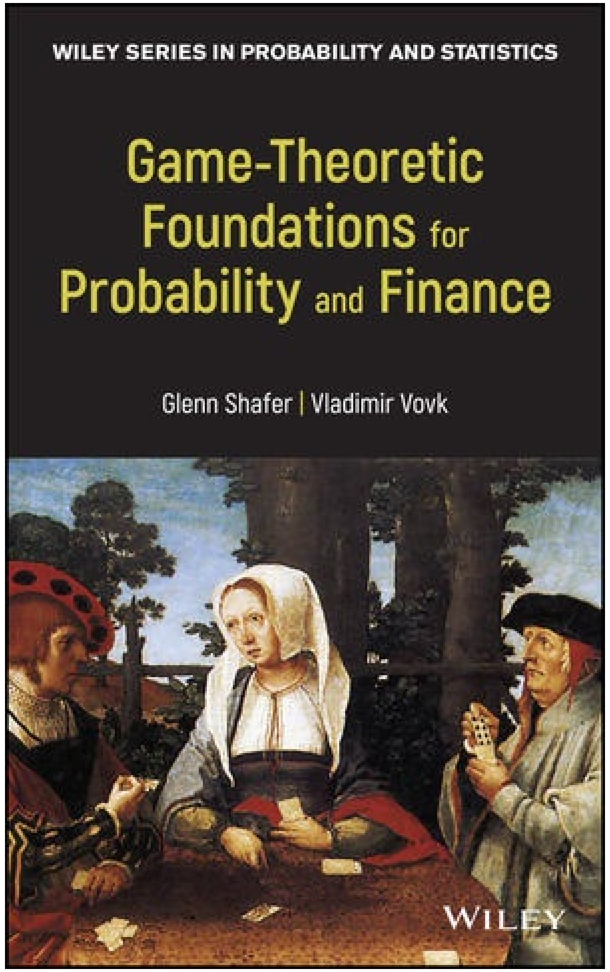
FOR $n = 1, \ldots, N$:
  Forecaster announces $\mu_n \in [-1, 1]$.
  Skeptic announces $z_n \in \mathbb{R}$.
  Reality announces $y_n \in [-1, 1]$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + z_n(y_n - \mu_n)$.

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

References



2001

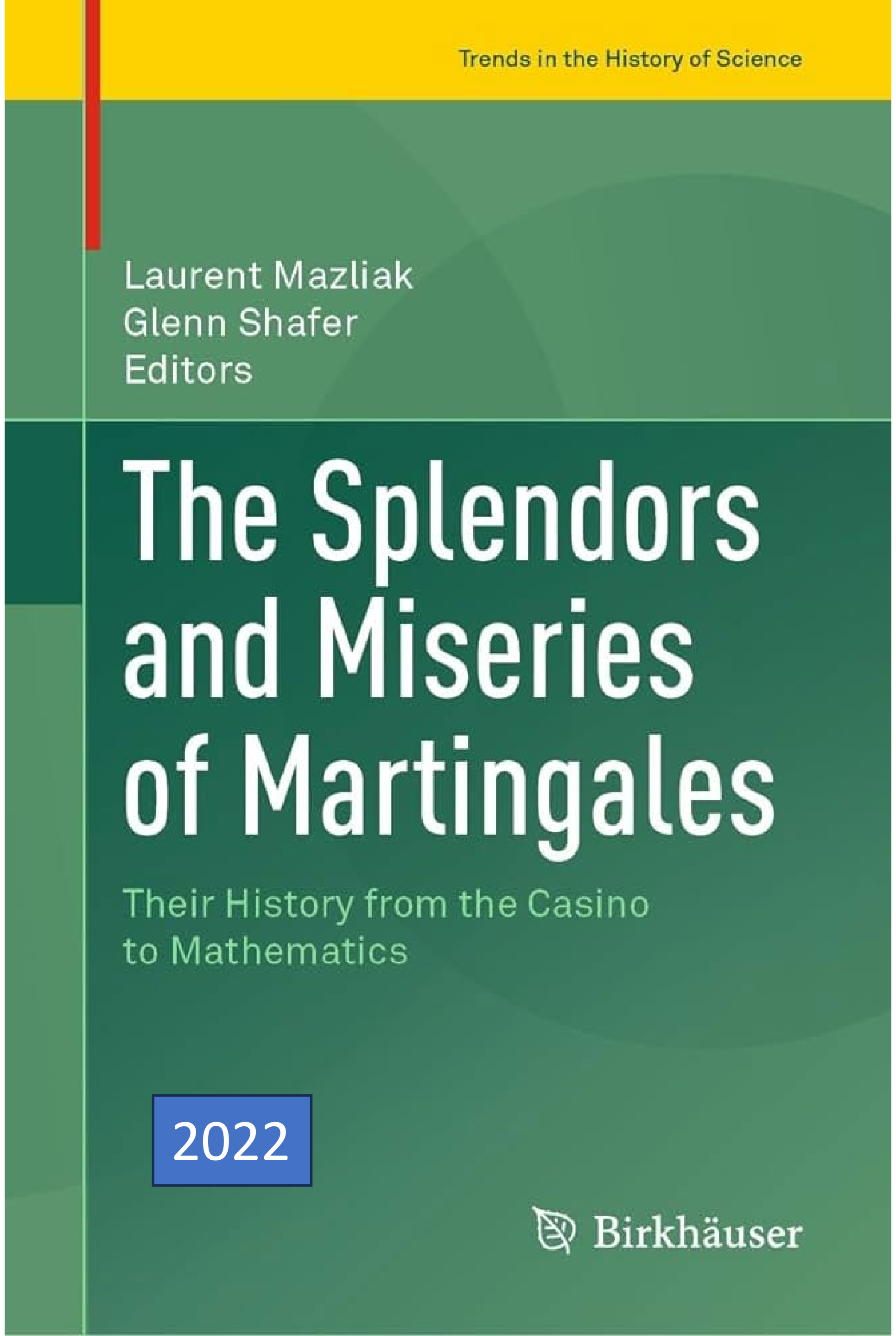


2019

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.)



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 betting strategy is called a *martingale*.
Simple strategy = simple martingale.

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

$\Leftrightarrow$

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

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

More old guys

2.15 John Locke, 1632–1704

Locke published his *An Essay Concerning Human Understanding* in 1689 [124]. Chapter XV of Book IV, entitled "Of probability", includes this passage:

> …most of the propositions we think, reason, discourse—nay, act upon, are such as we cannot have undoubted knowledge of their truth: yet some of them border so near upon certainty, that we make no doubt at all about them; but assent to them as firmly, and act, according to that assent, as resolutely as if they were infallibly demonstrated, and that our knowledge of them was perfect and certain.

2.26 André-Marie Ampère, 1775–1836

Whereas Bernoulli, d'Alembert, and Buffon had proposed selecting some number less than one that would suffice for moral certainty, Ampère realized that he could develop a theory of gambler's ruin with a more demanding concept of moral certainty. In his 1802 book *Considerations sur la théorie mathématique du jeu* [3], he defined this concept on p. 3:

> If we represent absolute certainty, the certainty resulting from mathematical demonstration for example, by unity, as is usually done, then we can consider moral certainty to be any variable fraction that never becomes equal to unity but can get close enough to it as to exceed any particular fraction.

2.28 Augustus De Morgan, 1806–1871

From page 396 of De Morgan's entry "Theory of Probabilities", on pages 393–490 of Volume II of Encyclopædia Metropolitana, Griffin, London, 1849

> *Mathematical certainty* (a thing perhaps impossible in the strictest sense) is the terminus or limit towards which our impressions approach as our knowledge becomes greater and greater, and is never attained as long as any doubt whatsoever remains. *Practical certainty* is that high degree of probability on which the mind acts at once, without thinking the counter-probabilities sufficiently large to be taken into account; and it depends upon the character of the individual.

2.42 Andrei Markov, 1856–1922

The closer the probability of an event is to one, the more reason we have to expect the event to happen and not to expect its opposite to happen.

In practical questions, we are forced to regard as certain events whose probability comes more or less close to one, and to regard as impossible events whose probability is small.

Consequently, one of the most important tasks of probability theory is to identify those events whose probabilities come close to one or zero.

2.55 Richard von Mises, 1883–1953

In classical physical statistics one starts by making certain plausible assumptions, according to the methods of probability calculus, about initial probabilities as well as transition probabilities, and derives from them statements about the course of events to be expected with very high probability. The value of this "high" probability is so near to 1 that the statements are practically indistinguishable from those which are called "deterministic". In all cases that can be checked the agreement between observation and calculation proves to be excellent.

2.62 Harold Jeffreys, 1891–1989

...repeated verifications of consequences of a fact will make it practically certain that the next consequence of it will be verified. This accounts for the confidence that we actually have in inductive inferences.

2.68 Abraham Wald, 1902–1950

… "The event E has a probability near to one" is translated into "it is practically certain that the event E will occur in a single trial."

2.73 Bruno de Finetti, 1906–1985

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”).

De Finetti on forecasts

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...

“Forecast” in English

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

Old 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 <u>what is predicted</u> and about <u>predicting it</u>.

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.