The Eternal Debate between Bernoulli and Leibniz

Glenn Shafer
EMS2013, Budapest
July 2013

1. What Bernoulli and Leibniz said
2. The 300 year argument
   1713 - 1763 - 1813 - 1863 - 1913 - 1963 - 2013
3. The role of judgment?
4. Defensive forecasting

1. What Bernoulli and Leibniz said
2. The 300 year argument

   1713 - 1763 - 1813 - 1863 - 1913 - 1963 - 2013
3. The role of judgment?
4. Defensive forecasting

What Bernoulli and Leibniz said

Bernoulli/Leibniz correspondence in 1703

Bernoulli: Use past frequencies to give probabilities for future.

Leibniz: Things change. Nature has its habits, but only mostly.

Bernoulli

If I perceive … that it happens 1000 times that the young person outlives the old person and the reverse happens only 500 times,

then I may safely enough conclude that it is twice as probable that a young person will outlive an old one as the reverse.

Leibniz

Who is to say that the following result will not diverge somewhat from the law of all the preceding ones – because of the mutability of things?

New diseases attack mankind.

Bernoulli later (in *Ars conjectandi*, 1713):

1. May need new observations.
2. One argument, to be combined with others.

Leibniz later (1714):

Yes. What is done more or less

is more or less doable

in the present state of things…

Edith Sylla, 2006: Translation of *Ars Conjectandi* with introduction and notes.

1. What Bernoulli and Leibniz said
2. The 300 year argument
   1713 - 1763 - 1813 - 1863 - 1913 - 1963 - 2013
3. The role of judgment?
4. Defensive forecasting

The 300 year argument

The 300 year argument

Bernoulli’s theorem always has place in theory.

Scope of application always contested.

Dipping in every 50 years…

1763
Life insurance

1813
Laplace triumphant

1863
Numbers everywhere
Bernoulli/Laplace debunked

1913
Bernoulli resurgent

1963
Stochastic processes
Inductive behavior

2013
Bernoulli resurgent
Bayes everywhere

1763

Anders Hald: "The usefulness of probability theory was convincingly demonstrated by application to problems of life insurance."

- Abraham de Moivre: *Annuities upon Lives* 1725, 1743
- Thomas Simpson: *The Doctrine of Annuities and Reversions* 1742

Astronomy? No Bernoulli in Tycho Brahe, Galileo, or Kepler.

Anders Hald, 1990: *A History of Probability and Statistics and Their Applications before 1750*

1813

In 1812, Laplace publishes his monumental *Théorie analytique des probabilités*.

- Bernoulli and Bayes in harmony.

- Probability applied freely to science and human events.

1863

Numbers and statistics everywhere

- Error theory used
- Statistical mechanics emerging
- Insurance still sold

But mathematical probability stagnant
Laplace and Bernoulli debunked

- John Venn, 1866, rejects Bernoulli.
- Joseph Bertrand, 1887, ridicules Laplace.
  (www.jehps.net, Volume 8)
- Many deplore Laplace's application of probability to judgment.

By 1863, Bernoulli's theorem has become a puzzle.

"I may safely enough conclude that it is twice as probable that a young person will outlive an old one as the reverse."

Probability appears twice, first as confidence, then as frequency.

How to make sense of this in world of data?

Bernoulli: “safely enough” = moral certainty.

Jakob Bernoulli

“Something is *morally certain* if its probability is so close to certainty that the shortfall is imperceptible.”

“Something is *morally impossible* if its probability is no more than the amount by which moral certainty falls short of complete certainty.”

Cournot tried to make 19th century sense of this idea.

Antoine Cournot
1801–1877

Maurice Fréchet, 1878–1973, proposed the name *Cournot's principle*.

Cournot discussed both *moral impossibility* (very small probability) and *physical impossibility* (infinitely small probability).

A physically impossible event is one whose probability is infinitely small. This remark alone gives substance—an objective and phenomenological value—to the mathematical theory of probability.

Antoine-Augustin Cournot. *Exposition de la théorie des chances et des probabilités*. Hachette, Paris, 1843

1913

Bernoulli's theorem resurgent again

- Chebyshev's school flourishing
- Pearson and company bring mathematical statistics to Britain
- New energy across western Europe

Bernoulli flourishes every 100 years: 1713, 1813, 1913.

Two less often remembered Bernoulli acolytes c. 1913

- Alexander Alexandrovich Chuprov
  1874-1926

- Ladislaus von Bortkiewicz
  1868-1931

Aleksandr Chuprov
1874–1926

Only Chuprov came close to repeating Cournot's claim that the the principle of moral certainty is the meaning of probability.

Petersburg Polytechnical Institute

In his *Essays on the Theory of Statistics* (in Russian 1909 and 1910), Chuprov called the principle that an event of small probability will not happen Cournot's lemma, because we use it to get from Bernoulli's theorem to the law of large numbers.

It was, he said, the basic principle of the logic of probable.

Chuprov's analysis of Bernoulli

Bernoulli's theorem: Frequency will be close to the true probability with very high probability.

Cournot's principle: Event with very high probability specified in advance (or simple) will happen.

Law of large numbers: Frequency will be close to the true probability.

Alexander Alexandrovich Chuprov 1874-1926

Essays in the theory of Statistics. Moscow 1909.

Reviving Bernoulli's "stochastic"

Bernoulli: *Ars Conjectandi sive Stochastice*

(the art of conjecturing or guessing)

Revived by Ladislaus von Bortkiewicz (1868-1931) in 1917.

von Bortkiewicz:

1. The probability calculus is mathematics.
2. Stochastik is the science of applying it to the real world.

1963

Jerzy Neyman

How did Neyman think about Bernoulli's theorem?

Neyman agreed with Leibniz!

- iid unrealistic as scientific model
- Science = stochastic processes

Neyman in 1960

In order that the applied statistician be in a position to cooperate effectively with the modern experimental scientist, the theoretical equipment of the statistician must include familiarity and capability of dealing with stochastic processes.

Neyman saw four periods in the history of indeterminism in science.

1. Marginal indeterminism: 19th century. Science deterministic except in the domain of errors of measurement.

2. Static indeterminism: End the 19th, beginning of 20th century. Populations were main subject of scientific study. Idea of independent draws from populations was dominant.

3. Static indeterministic experimentation: 1920 to 1940. R. A. Fisher dominant. Statistical testing and estimation developed.

4. Dynamic indeterminism: Already in full swing in 1960. Every serious study in science is a study of some evolutionary chance mechanism.

Neyman's **inductive behavior**

Bernoulli's theorem is fundamental even though science's stochastic processes are never iid.

The statistician intervenes and chooses a sequence of 95% predictions to make.

A statistician who makes predictions with 95% confidence has two goals only:

- be informative
- be right 95% of the time

Neyman's **inductive behavior**

Your 95% predictions only need to be right 95% of the time.

You can do this without knowing full stochastic model.

2013

How has Neyman's philosophy fared?

Perhaps followed in practice but philosophically unpopular

- Without iid repetition, model seems insufficiently validated.
- Probability judgment should not depend arbitrary choice of past sequence.

Two statisticians who are right 95% of the time may tell the court contradictory things.

They are placing the litigated event in different sequences.

2013

Has Leibniz won the argument?

Yes but.

- We teach Bernoulli's theorem.
- We use Bernoulli's theorem.
- Leibniz's winning only means we don't believe in what we are doing.

OK. Profiling is wrong. But we have to do it.

1. What Bernoulli and Leibniz said
2. The 300 year argument

   1713 - 1763 - 1813 - 1863 - 1913 - 1963 - 2013

3. The role of judgment?
4. Defensive forecasting

The role of judgment?

Bernoulli’s ambitions

1. Explain how probabilities can be learned from experience.

2. Explain how probabilities can be used (and combined) as arguments.

Judgments essential to use of Bernoulli's theorem

- Moral certainty: the event of high probability we have defined is simple enough it should have happened.

  We will not lose our bet.

- Relevance: This makes the sequence relevant.

- Irrelevance: Nothing else we have learned changes our confidence in our bet.

1. What Bernoulli and Leibniz said
2. The 300 year argument
   1713 - 1763 - 1813 - 1863 - 1913 - 1963 - 2013
3. The role of judgment?
4. Defensive forecasting

Defensive forecasting

Game-theoretic probability

- **Probabilities derive from betting offers.**

  Not from the measure of sets.

- **Probabilities are tested by betting strategies.**

- **Probability theorems are proven by betting strategies.**

  -- Do not say that the property fails on a set of measure zero.

  -- Say that its failure implies the success of a betting strategy.

Defensive forecasting

- 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.