Richard von Mises

The forgotten Bayesian

Glenn Shafer, Rutgers University

GERMAN PROBABILITY AND STATISTICS DAYS

TU Dresden

14 March 2025

1. The person

2. His *Kollektiv*

3. The Bayesian

4. Frequentism

The person

- 1883   Born in Lviv (Lemberg)

  Vienna was the family home.

  Studied at Technischen Hochschule Wien
- 1908   Habilitation in Brno (Brünn)
- 1909   Universität Straßburg
- WWI   Austro-Hungarian air force
- 1918   Frankfurt
- 1919   Technischen Hochschule Dresden

Georg-Schumann-Straße 7, Dresden-Südvorstadt

1920-1933   Universität Berlin

Institut für Angewandte Mathematik

Zeitschrift für Angewandte Mathematik und Mechanik

(ZAMM)

1933—1939 Istanbul

1939—1953 Harvard

1944   Gordon McKay Professor

1953   Death

Garrett Birkhoff, ZAMM, 1983

Mission: transplant applied math from Europe.

Applied Math at Brown University in mid-1941

Influence on Harvard Math Department

Most stimulated during war by fellow philosophers
(≈ transplanted Vienna circle)

Conflicting testimony

Garrett Birkhoff

In dossier for Harvard appointment, von Mises says he was "most interested in the theory of probability" for 15 years.

But Harvard invited him for fluid dynamics.

He taught hydro- and aerodynamics almost exclusively.

Hilda Geiringer

In US, became deeply interested in "British-American" stat.

Repeatedly lectured on stat to undergraduate/graduate students.

His *Kollektiv*

The word 'limit' is in fact used in mathematics only as a concise way of making certain statements concerning small deviations.

Collective = sequence of objects with labels
   Infinite idealization, like Euclid's infinite lines.

Probability of label = limiting frequency
   First the collective, then the probability.

Randomness or irregularity axiom
   Limiting frequency unchanged by selection.
   This can be dropped in applications.

A real *Kollektiv* (Bruns, Fechner) is a collection of objects, not necessarily ordered.

An idealized *Kollektiv* (von Mises) is infinite and is ordered so that "limiting frequency" makes sense.

In response to Doob, von Mises wrote, "There is no doubt collectives exist". He meant finite ones.

Each probability or probability distribution has its own collective.

Underlying and justifying a calculation of probabilities from probabilities is a construction of collectives from collectives.

Once you understand this foundation, you calculate with probabilities in the usual way.

Understanding the von Mises picture

Real finite collectives

Motivation

Idealized infinite collectives

Justification

Probability rules / calculations

von Mises

Real finite collectives

⇓ Motivation

Idealized infinite collectives

⇓ Justification

Probability calculations

Doob

What is desired is a mathematical theory which runs parallel to the physical facts, when properly idealized, but which has its own independent justification.

Review of 2nd ed. of *Wahrscheinlichkeit, Statistik und Wahrheit*, *Bull. Amer. Math. Soc.*, 43(5):316-317, 1937, p. 316.

The four operations on collectives

1. Selection/Auswahl

   Make new collective by selecting every third entry.

2. Mixing /Mischung

   Simplify collective with labels 1,2,3,4,5,6 to collective with labels even,odd.

3. Partition/Teilung

   Make new collective by selecting entries with labels 2,4,6.
   So probability of 2 becomes 1/3.
   Conditional probability. Von Mises correctly attributed it to Bayes.

4. Combination/Verbindung

   Combine two collectives to get a joint collective.
   Rule of compound probability / stochastic process

The Bayesian

Von Mises was an old-fashioned Bayesian.

After learning about "British-American" statistics, he was more Bayesian than ever.

Probability, Statistics and Truth

1st English edition, 1939, p. 64:

> The partition rule has been first derived by the Englishman, Bayes, in the middle of the 18th century.

2nd English edition,1957, p. 159:

> We can only hope that statisticians will return to the use of the simple, lucid reasoning of Bayes's conceptions...

THE ANNALS
OF MATHEMATICAL STATISTICS

1938-1943

Von Mises the only Bayesian in the group.

20

Von Mises on Bayes in practice

"On the correct use of Bayes' formula"
Annals of Mathematical Statistics, 13 (1942), pp. 156-165

- Prior probabilities (perhaps unknown) are frequencies in a real collective.

- Incomplete info re prior $\Rightarrow$ incomplete info re posterior

- Bernstein-von Mises: Posterior independent of prior when many observations

- Confidence intervals answer the wrong question.

Von Mises unread

Colin Howson and Peter Urbach:

> In Bayesian terms, von Mises (though he was not, at any rate consciously, a Bayesian) was in effect arguing...

Scientific Reasoning, The Bayesian Approach, 3rd edition, 2006, p. 78.

What is frequentism?

In the 1930s, European mathematicians debated two proposed mathematical foundations for probability:
    von Mises vs (measure theory and functional analysis)

Meanwhile, American philosophers studied two German philosophers who had a frequency interpretation of probability:
    von Mises & Reichenbach

Von Mises & Reichenbach:

Häufigkeitstheorie der Wahrscheinlichkeit

Frequentist

- Coined by philosopher Ernest Nagel 1936,
- Used by statistician Maurice Kendall 1949.

Frequentism

- First used by philosopher Donald Williams in 1945.

Ernest Nagel, 1936

The central idea of the frequency interpretation of probability is that by the probability of a proposition or an "event" is meant the relative frequency of the "event" in an indefinite class of events.

Neyman on von Mises

1957

This clear cut analysis of ... probability ... is due to the recently deceased German philosopher and probabilist Richard von Mises...

... often called the "frequentist" view.

Neyman on von Mises
1961

...the exchange of opinions and the studies surrounding the definition of probability given by Richard von Mises, clarified the thinking considerably.

- On the one hand, this dispute brought out the superiority of Kolmogoroff's axiomatization of the theory.

- On the other hand, the same dispute established firmly von Mises's philosophical outlook on "frequentist" probability as a useful tool in indeterministic studies of phenomena.

In 1977, Neyman advocated both

- a frequentist theory of probability (von Mises) and

- a frequentist theory of statistics (point estimation, tests, p-values, confidence intervals).

Neyman's names

frequentist interpretation of probability

frequentist methodology of statistics

This double use of "frequentist" is a source of confusion.

My names

Interpretation of probability
    frequency vs belief

Statistical methodology
    Bernoullian vs Bayesian

Names for the two classical statistical methodologies

Laplace	a priori	a posteriori
Cournot 1843		règle de Bayes
German authors, 19th century		Bayessche Regel
English authors, 19th century	direct probability	Inverse probability
Von Mises	Bernoullische	Bayessche
Neyman	frequentist	Bayesian
Dempster	Bernoullian	Bayesian

My names

Interpretation of probability:
        frequency vs belief

But I add a third interpretation: forecast.

The forecast may be a belief, but it aims to be pass statistical tests.

Shafer & Vovk, *Game-Theoretic Foundations for Probability and Finance*, 2019