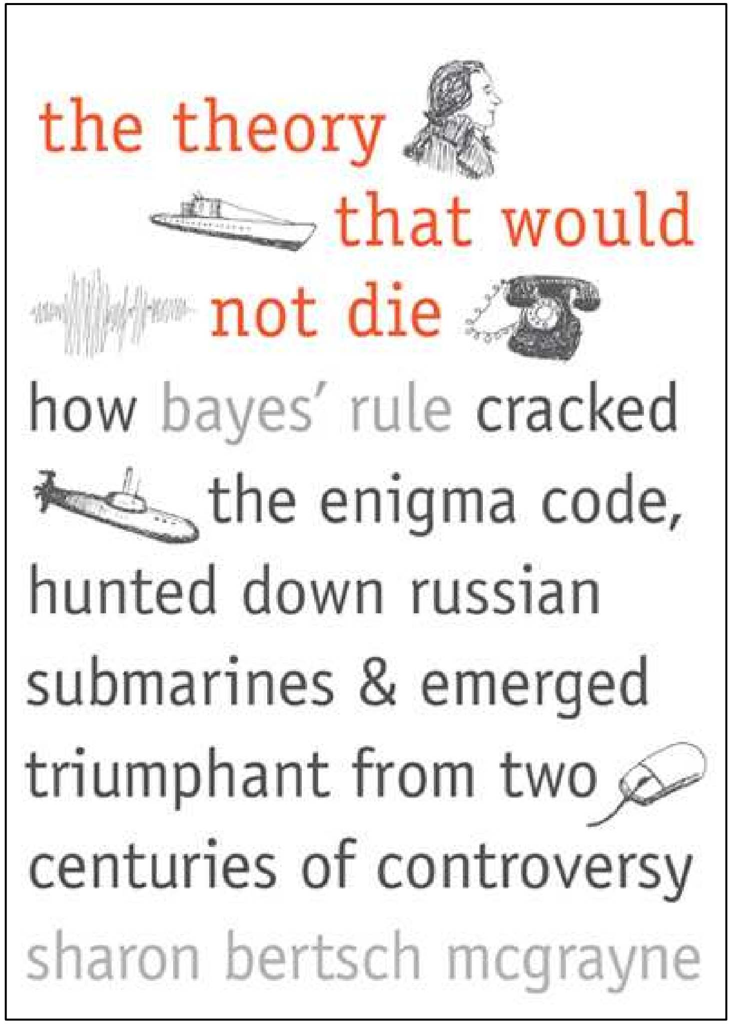
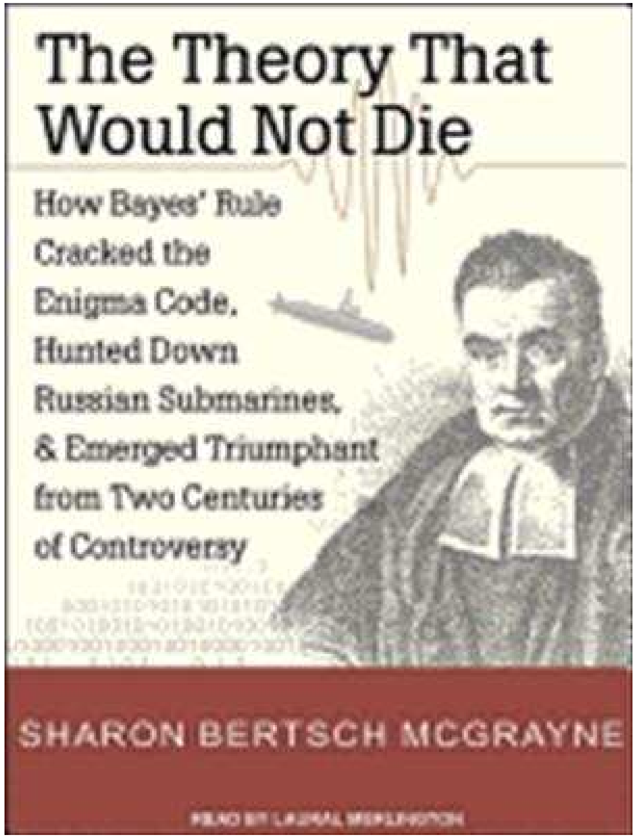
# III Subjective probabilities

# 1. Bayesianism

# III.1.1 – The Emperor's New Cloth or The Solution to Everything?

# According to a recent book Bayesianism is indeed the "Solution to Everything"



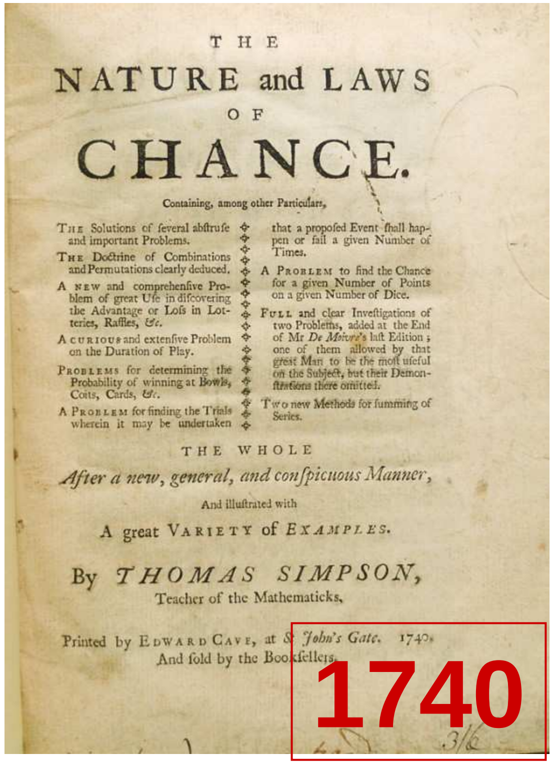



*Bayesianism has been of good use to*

- Prove that smoking causes lung cancer
- Predict winner in the US presidential elections
- Show that cholesterol causes heart attacks
- Find the author of a 18th century document
- Improve the US insurance industry
- Crack the German Enigma code during WW II
- Track down a lost American H-bomb
- Hunt down Soviet submarines
- Investigate nuclear power safety
- Predict the shuttle Challenger accident
- . . . and much more!

# III.1.2 The historical background to Bayesianism

Thomas
Simpson
1710-61
Mathema-
tician



THE

NATURE and LAWS

OF

CHANCE.

Containing, among other Particulars,

THE Solutions of several abstruse and important Problems.

THE Doctrine of Combinations and Permutations clearly deduced.

A NEW and comprehensive Problem of great Use in discovering the Advantage or Loss in Lotteries, Raffles, &c.

A CURIOUS and extensive Problem on the Duration of Play.

PROBLEMS for determining the Probability of winning at Bowls, Coits, Cards, &c.

A PROBLEM for finding the Trials wherein it may be undertaken that a proposed Event shall happen or fail a given Number of Times.

A PROBLEM to find the Chance for a given Number of Points on a given Number of Dice.

FULL and clear Investigations of two Problems, added at the End of Mr *De Moivre's* last Edition; one of them allowed by that great Man to be the most useful on the Subject, but their Demonstrations there omitted.

TWO new Methods for summing of Series.

THE WHOLE

*After a new, general, and conspicuous Manner,*

And illustrated with

A great VARIETY of *EXAMPLES.*

By *THOMAS SIMPSON,*

Teacher of the Mathematicks.

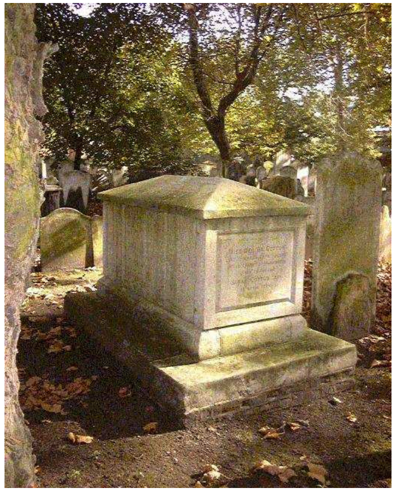
Printed by EDWARD CAVE, at *St John's Gate.* 1740.

And sold by the Booksellers.

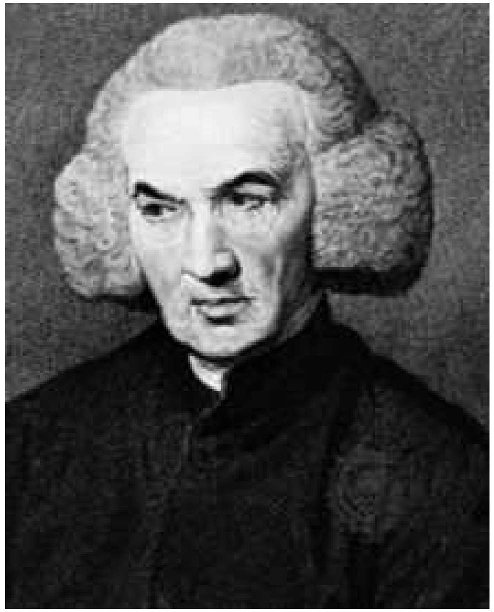


David Hume 1711-76

*"There is no First Cause"*



Thomas Bayes
1701-61
Dissident clergyman



Richard Price
1723-91
Radical priest

Friend of the
American and
French revolutions

Supported women's
rights

LII. *An Eſſay towards ſolving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

**1763**

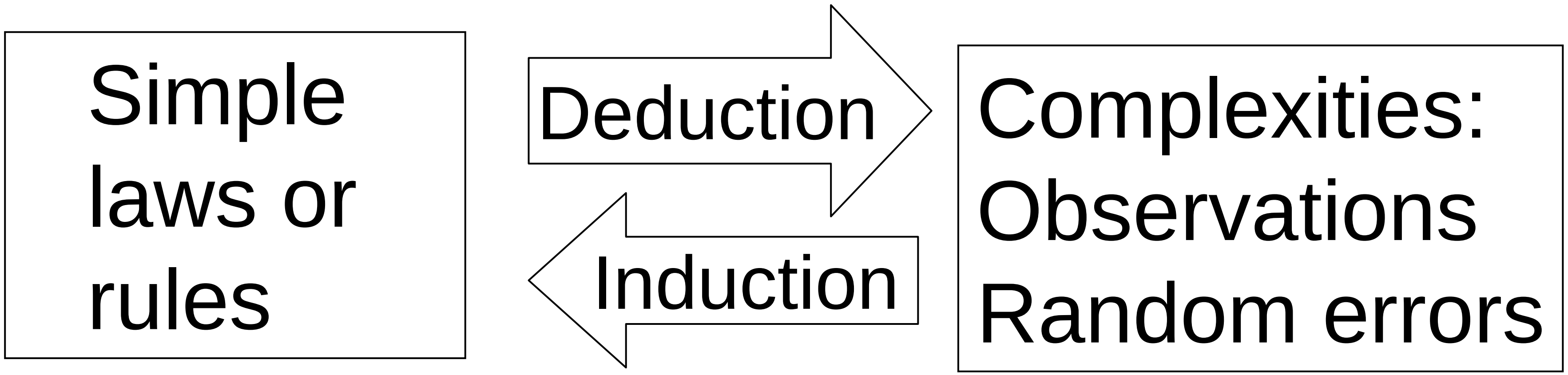
Dear Sir,

Read Dec. 23, 1763.

I Now ſend you an eſſay which I have found among the papers of our deceaſed friend Mr. Bayes, and which, in my opinion, has great merit, and well deſerves to be preſerved. Experimental philoſophy, you will find, is nearly intereſted in the ſubject of it; and on this account there ſeems to be particular reaſon for thinking that a communication of it to the Royal Society cannot be improper.

# The possible roots to Thomas Bayes' 1757 theorem

- How to treat randomness
- Is God the Ultimate Cause?
- How to draw conclusions from observations



The first Bayesian was Simone de Laplace 1749-1827

Bruno de Finetti (1906-85)

# III.1.3 What is Bayesianism?

# Frequentists and Bayesians differ in their definition of "probability"

Classical definition: from tossing coins or dice *Both camps agree*

Frequentist definition: the long-run frequency of a "repeatable event" *Agreement*

Bayesian definition: a person's degree of belief in an event, given the information available *Sharp disagreements*

Bayesianism consists of two parts, one non-controversial, one *highly controversial*

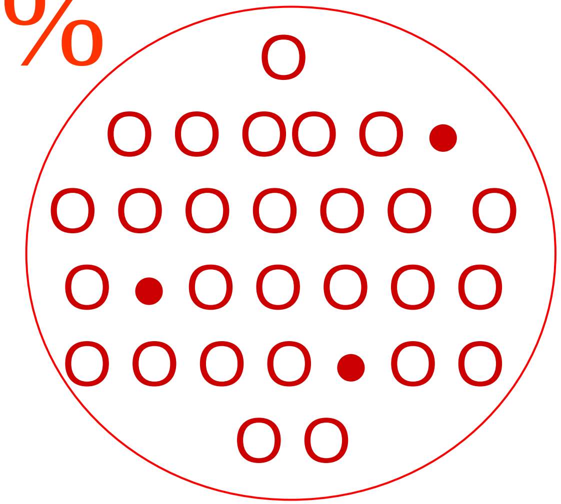
**NON-CONTROVERSIAL: The equation for conditional probabilities**

$$\text{prob}(A|B) = \frac{\text{prob}(A)\cdot\text{prob}(B|A)}{\text{prob}(B)}$$

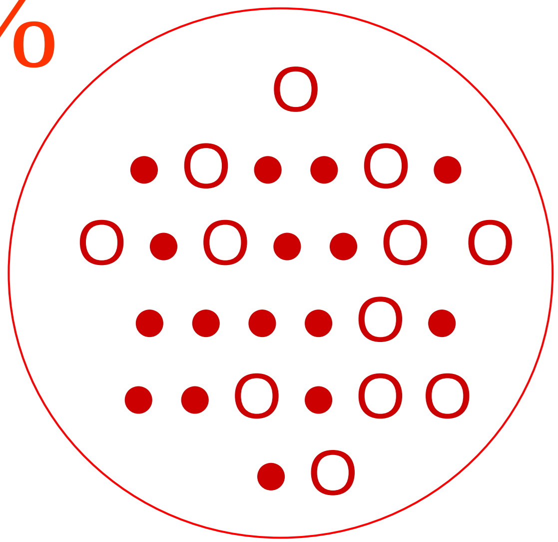
# III.1.4 The non-controversial part – conditional probabilities

# Classical probability

10%



60%



● ● ● 0% of cases

● ● o 3 % - “ -

● o o 24% - “ -

o o o 73% - “ –

● ● ● 22 % of cases

● ● o 43 % - “ -

● o o 29 % - “ -

o o o 6 % - “ –

*Bayesian statistics address the problem:*
*- What is the unknown proportion of red and white balls?*


Draw three balls at random

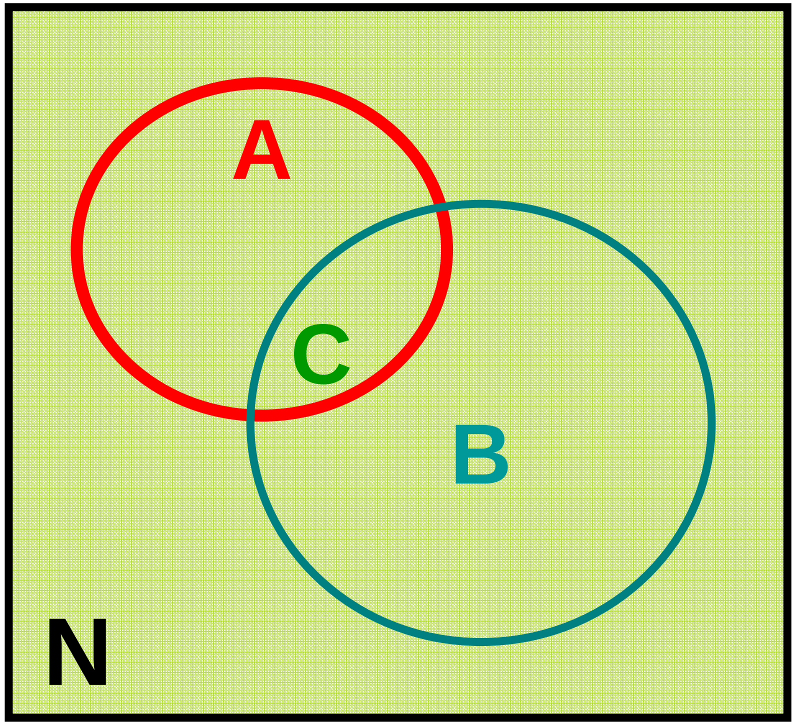
What conclusions can we draw about the "true" probability?

*Bayesian statistics deals with* ***inverse probabilities***

This can be solved by using the non-controversial "Bayes Rule" (to be derived)

$$\text{prob}(A|B) = \frac{\text{prob}(A)\cdot\text{prob}(B|A)}{\text{prob}(B)}$$

# What does A|B mean?

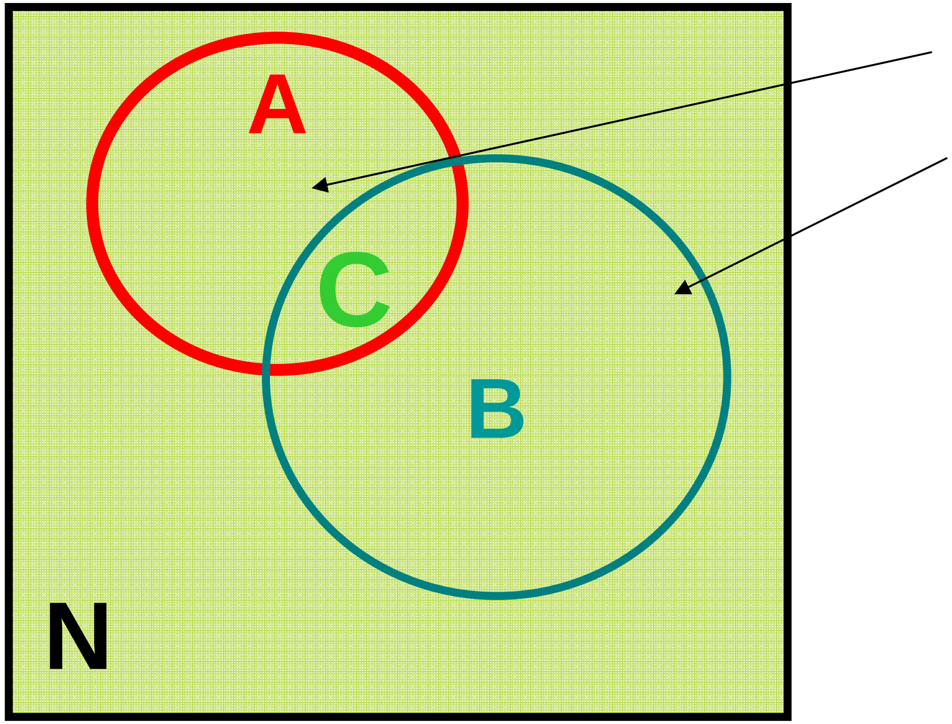



p(B|A) means probability of B *given* A (within the A area) = C/A

p(A|B) means probability of A *given* B (within the B area) = C/B

# III.1.5 Deriving Bayes's equation for single outcomes

# Deriving Bayes' theorem




$p(A) = A/N$

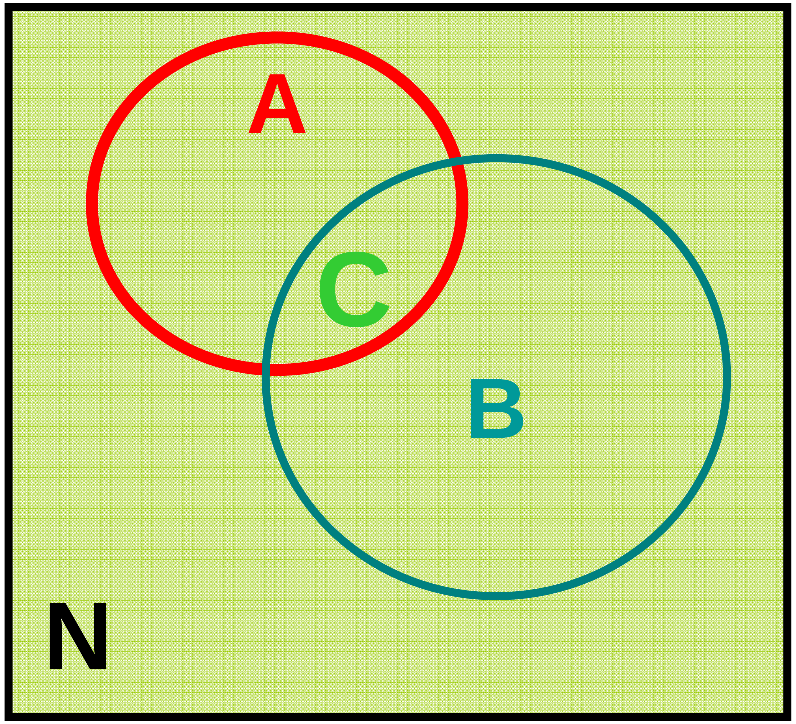
$p(B) = B/N$

$$\frac{p(A)}{p(B)} = \frac{A}{B}$$

$p(B|A) = C/A$

$p(A|B) = C/B$

# Deriving Bayes' theorem




$$A \cdot p(B|A) = C$$

$$B \cdot p(A|B) = C$$

$$p(B) \cdot p(A|B) = p(A) \cdot p(B|A)$$

$$p(A|B) = \frac{p(A) \cdot p(B|A)}{p(B)}$$

# III.1.5 The controversial part - subjective probabilities

Bayesianism consists of two parts, one non-controversial, one *highly controversial*

**NON-CONTROVERSIAL: The derived equation for conditional probabilities**

***HIGHLY CONTROVERSIAL: The use of subjective probabilities (degrees of belief) and the use of this term – the "prior"***

$$\text{prob}(A|B) = \frac{\text{prob}(A) \cdot \text{prob}(B|A)}{\text{prob}(B)}$$

# The controversy is not about mathematics but about philosophy!

The controversial bit, **"the prior"** which can even be **a pure guess**

$$p(A|B) = \frac{p(A) \cdot p(B|A)}{p(B)}$$

The "prior" p(A) + new information p(B|A)/p(B) updates the **probability of A**, i.e. p(A|B)

# Objections to subjective probabilities:

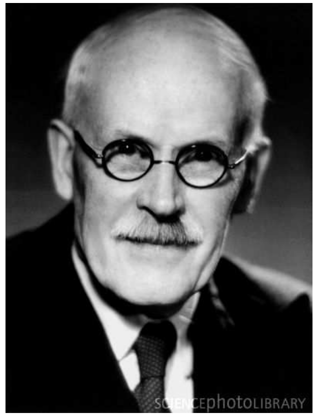
*"Science and mathematics are supposed to be objective and not subjective"*

# The statistical community has since long been divided into Frequentists and Bayesians

Karl Pearson
1857-1936

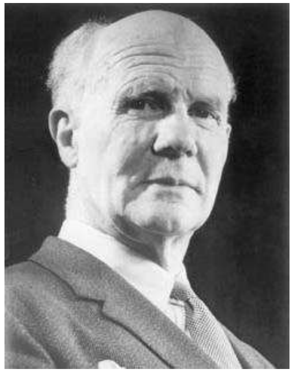
Ronald S Fisher
1890-1962

Robert O. Schleifer
1914-94



Harold Jeffreys
1891-1989

Jerzy Neyman
1894-1981



Egon Pearson
1895-1980

Leonard J. Savage
1917-71

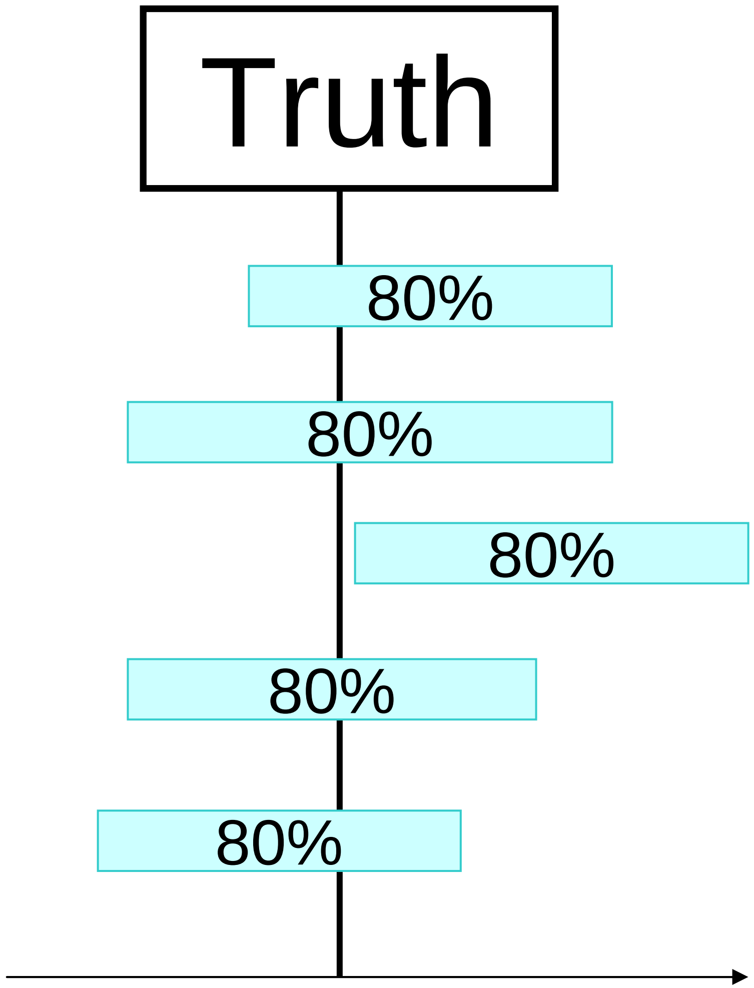
Howard Raiffa 1924 -

# Frequentist approach (the long-established majority):

**Probability** is defined as the long-run frequency of a "repeatable event".

It developed a notion of confidence interval, with a probability that it is covering the true value.

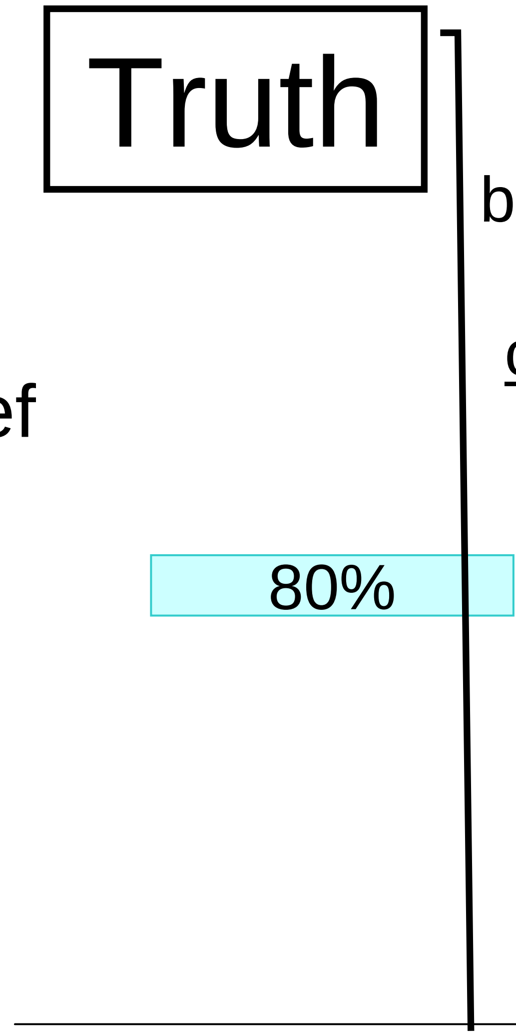
**Confidence** intervals



# Bayesian approach
(a growing minority):

**Probability** is (also) defined as a person’s degree of belief in an event, given the information available.

It developed a notion of credible interval with a probability that the true value will be within the interval.




Whereas confidence intervals are based only on the data **(clean),** credible intervals incorporate **(dirty)** problem-specific contextual information from the prior distribution


The Frequentists have for long regarded the Bayesians as despiteful heretics

**… but it works!**

# *As a consequence three types of Bayesians have developed*

1. **Open Bayesians** (openly using and promoting Bayes theorem)

2. **Covert Bayesians** (using Bayes' theorem without mentioning it)

3. **Latent Bayesians** (using Bayes theorem without being aware of it)

You had to "come out" as a Bayesian

# III.1.7 Why do we need Bayesianism?

a) To allow subjective probabilities

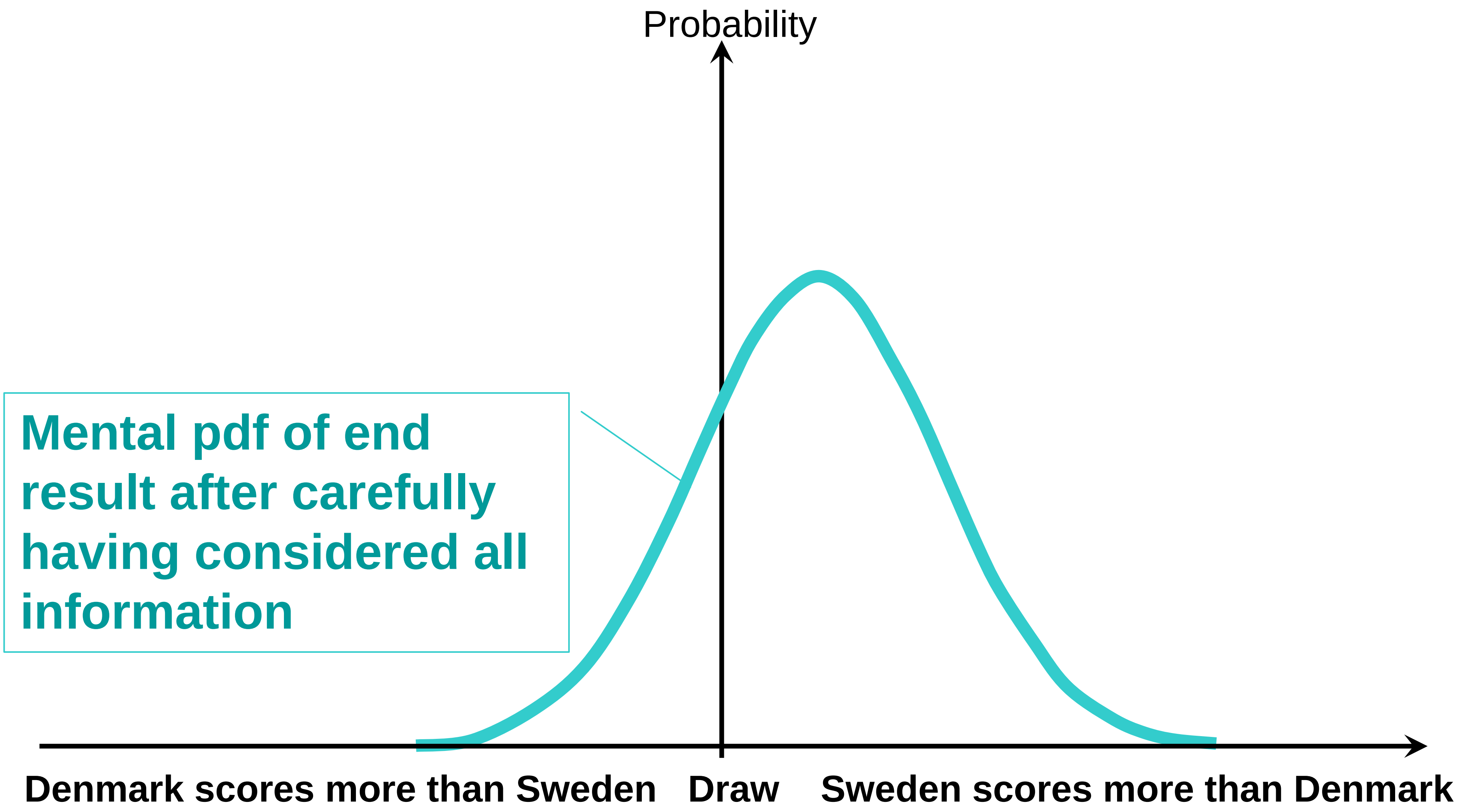
b) To avoid over-confident probabilities

c) To allow updating of preliminary probabilities

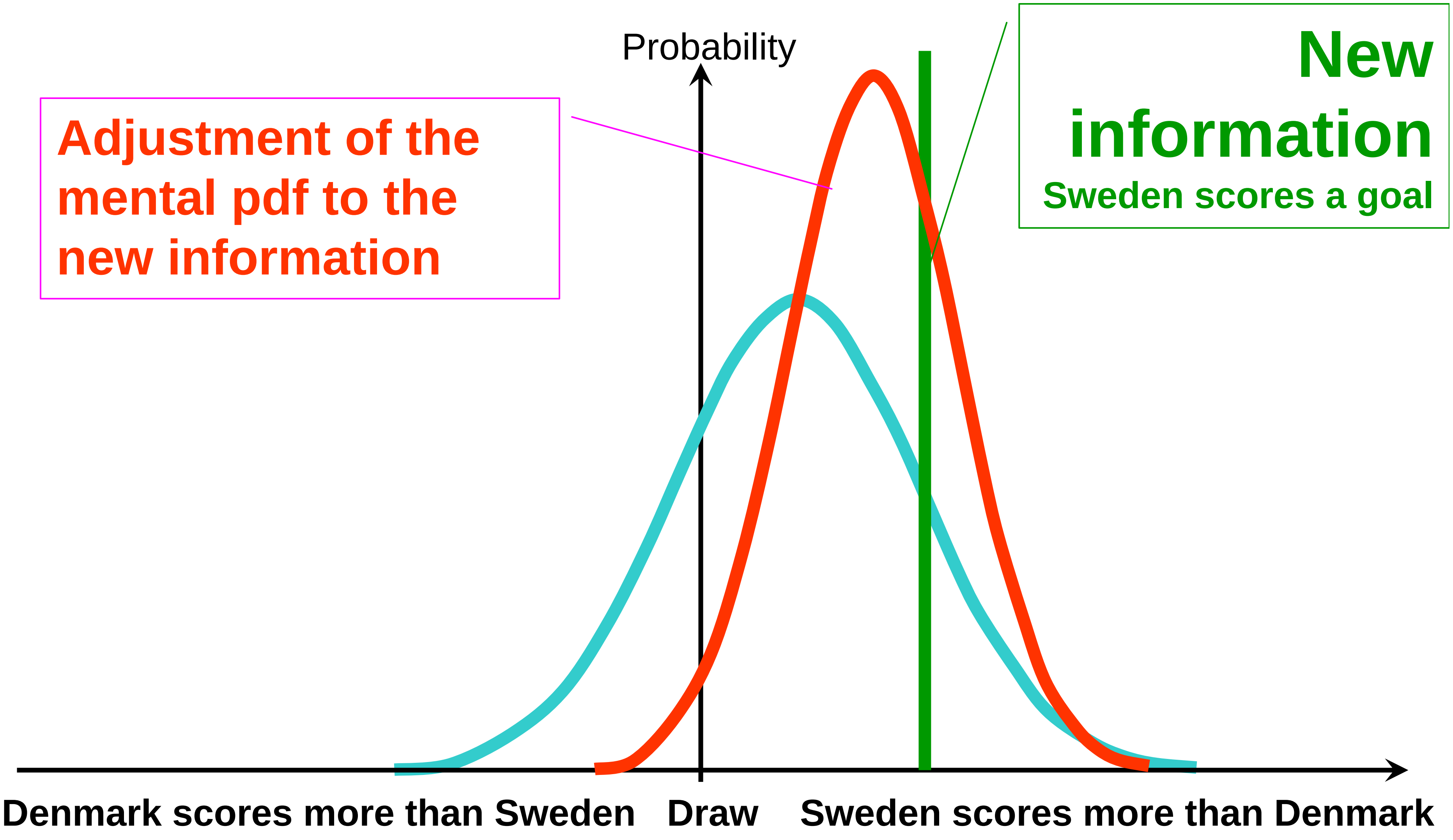
"Far better having an **approximate** answer to the **right** question, which is often <u>vague</u>, than an **exact** answer to the **wrong** question, which can always be made <u>precise</u>". *John W. Tukey (1915-2000)*

# a1) Updating of subjective probabilities

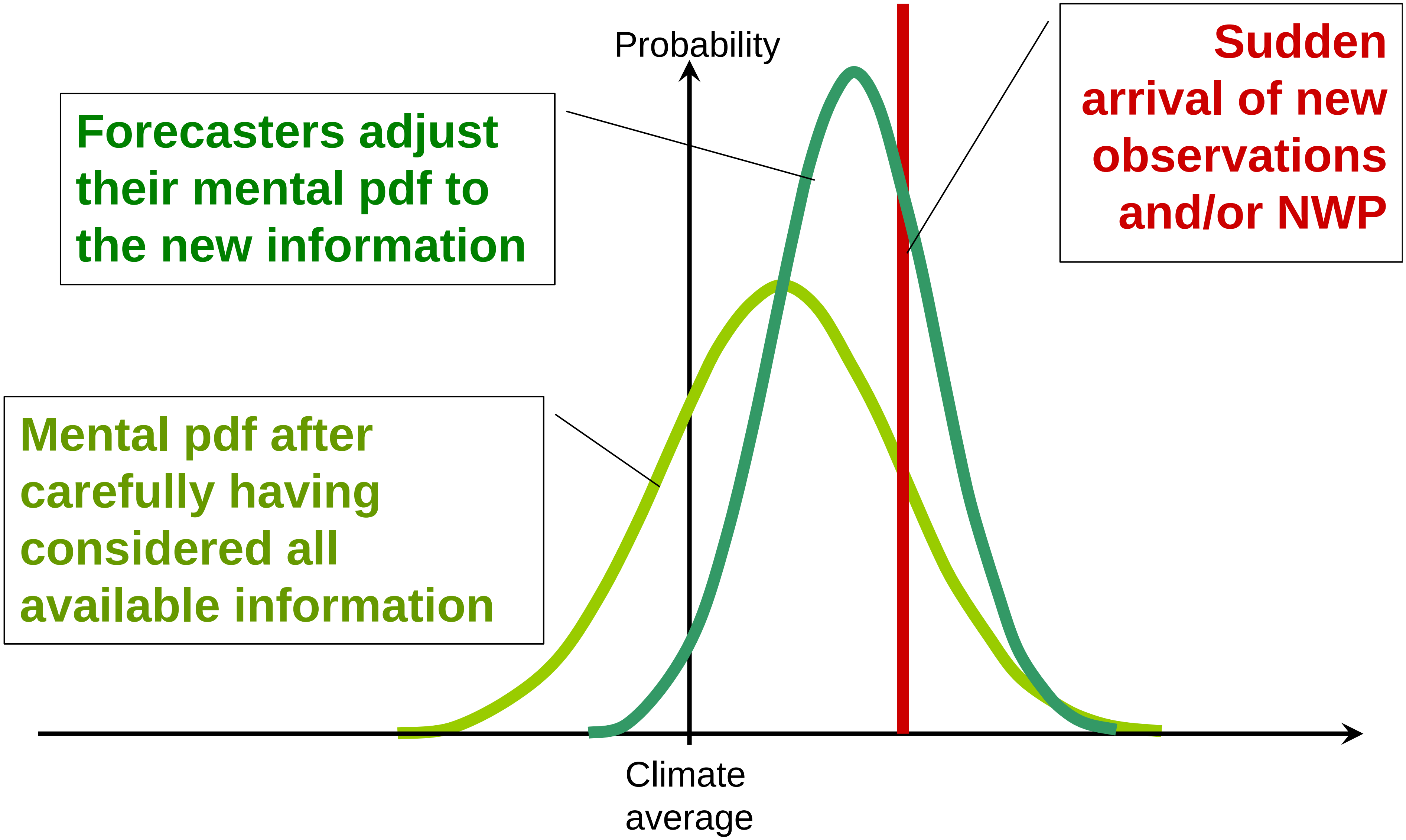
# a2) Updating of subjective probabilities



# a3) Updating of subjective probabilities



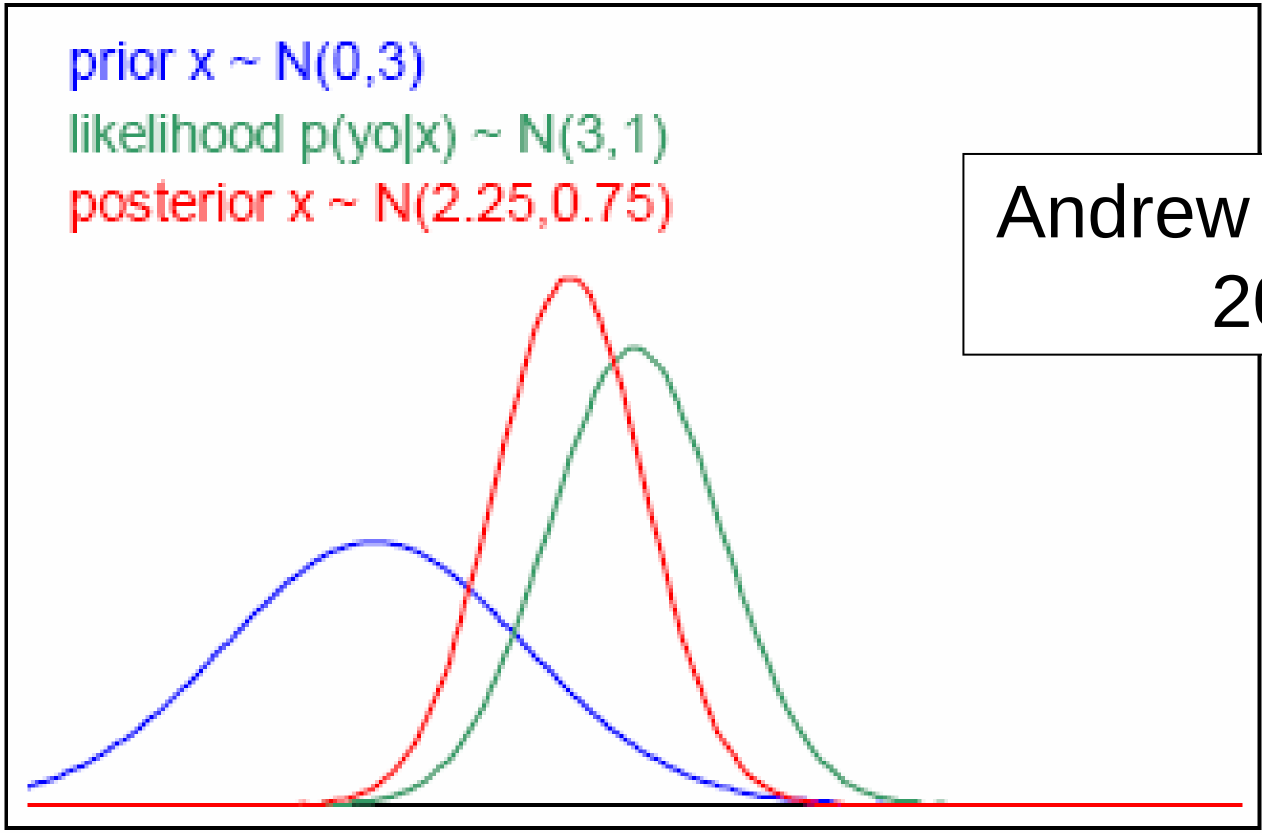
# Intuitive Bayesianism among weather forecasters



# *Bayesianism is making inroad with NWP modellers*

Met Office

Combination of Gaussian prior & observation
- Gaussian posterior,
- weights independent of values.




Andrew Lorenc, 2004

# b) Updating of preliminary probabilities

# Arthur L. Bailey

accountant

had to start some insurance activity in the US in spite of lacking statistics on accidents

Had to start by guessing and then modify them in a Bayesian way

# c) Avoids over-confident probabilities

such as Concorde before 2000 being the safest air plane

accidents

$$\frac{0}{100\,000^h} < \frac{1}{1\,000\,000^h}$$

Concorde | Some other airline

. . .after the 2000 crash the most unsafe

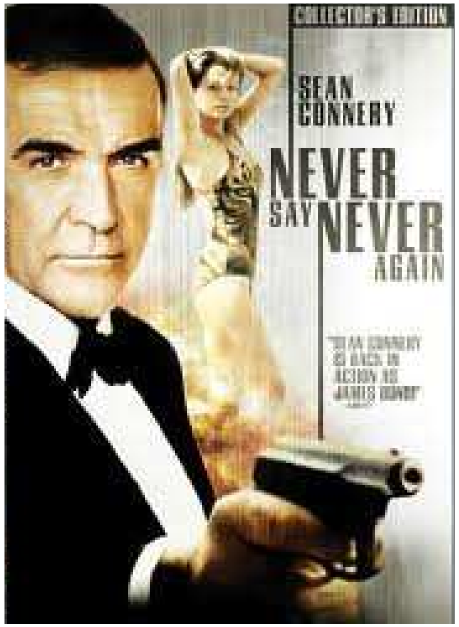
$$\frac{1}{100\,000^h} > \frac{1}{1\,000\,000^h}$$

A Bayesian would not have regarded Concord as the world's safest airplane before 2000

# “Laplace Rule of Succession” prevents us from being overconfident

since **p** never takes values 0% or 100%

$$p = \frac{1 + N_{\text{acci-dents}}}{2 + N}$$




*“-Never say never ”*

Other suggestions to "Laplace's Rule"

Laplace

$$p = \frac{1 + N_{\mathbf{rain}}}{2 + N}$$

Roulstone and Smith

$$p = \frac{\tfrac{1}{2} + N_{\mathbf{rain}}}{1 + N}$$

Neil Bowler UKMO

$$p = \frac{1 + N_{\mathbf{rain}}}{3 + N}$$

# III:2:8 Controversies over the Bayesian approach

1. Subjective probabilities (unavoidable)
2. "Creditability" intervals (accepted in weather forecasting)
3. Complicated calculations (helped by new computers and the Markov Chain Monte Carlo simulations)

# END