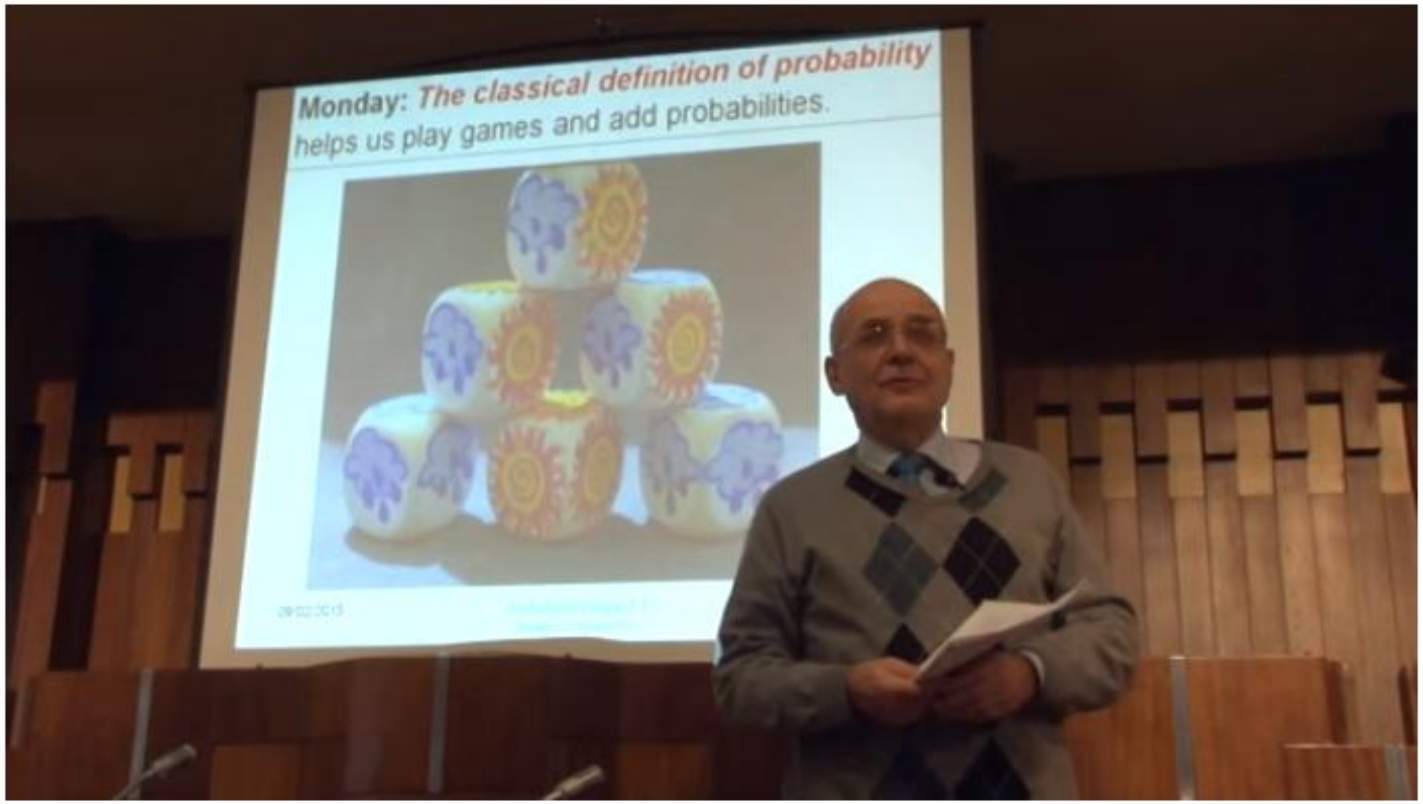
The (short) history of probability theory

1. Extracts from a week long seminar series in Bologna 9-13 February 2015

Monday 9 February

Anders Persson: Classical statistics – not so "classical" **Part 1** - **Part 2** - **Part 3** - **Part 4**



Monday: ***The classical definition of probability*** helps us play games and add probabilities.

Tuesday: ***The frequentist definition of probabilities*** involves statistical calibration and verification of probability forecasts.

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

Wednesday: ***The subjective probabilities*** measures our degree of beliefs

Thursday: ***The use of probabilities*** involve making decisions under uncertainty

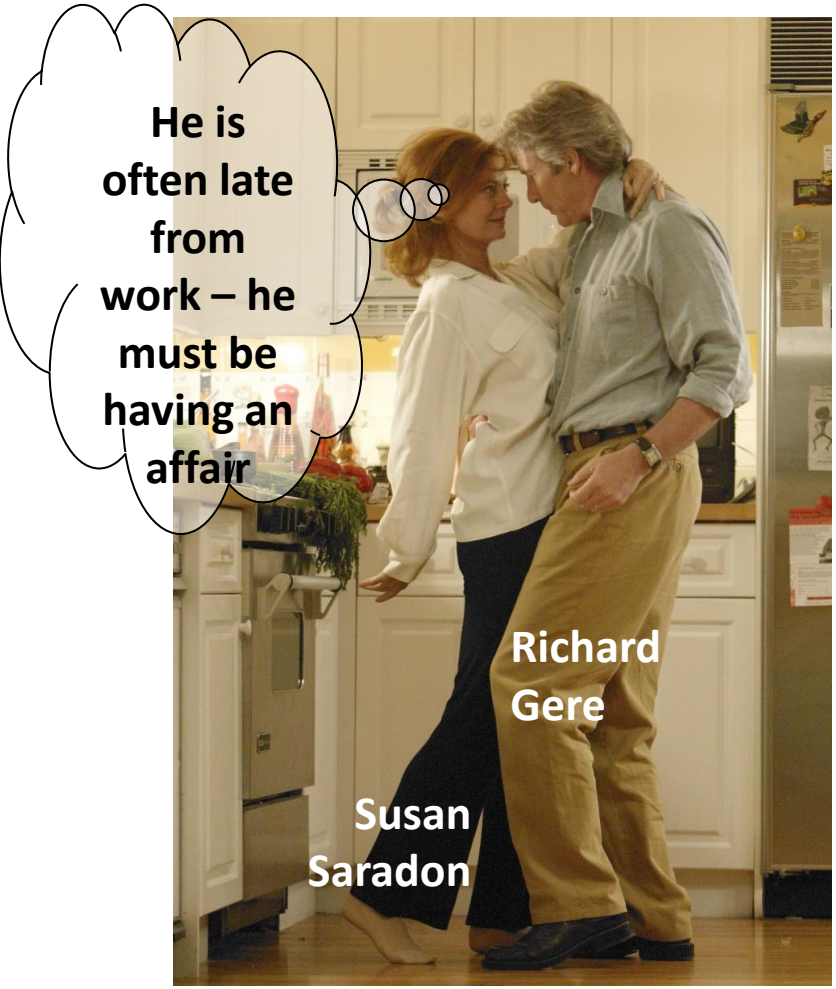
-What do you prefer?

-An 80% <u>chance</u> of winning € 1000

or

-Getting € 700 <u>directly</u> in your hand?

Friday: ***The psychology of probabilities*** deals with the understanding of uncertainty.



From the 2004 movie "Shall we dance?")

All married men in New York

Prob (being late from work while having an affair, **70%**) > Prob (having an affair while being late from work,**10%)**

P (being late | affair) > P (affair | being late)

2. What is “probability”?

Andrei Kolmogorov's definition of probabilities, 1933

1. **Probability for any event = 100%**

2. **Probability for one type of events = F/N**

3. **Probability for several mutually exclusive events = (F+G+H)/N**

There are three types of probabilities: **the classical, the frequentist and the Bayesian**

1. **<u>The classical</u> applies to the probabilities when tossing of a die (1/6) or a coin (1/2).**

2. **<u>The frequentist</u> applies to analyses of historical observation sets (to derive e.g. climatologically based probabilities).**

3. **<u>The Bayesian</u>, <u>subjective</u> or <u>degree of belief</u> is used by e.g. to summarize or update one's preliminary assessment considering new available information.**

3. Why did it take so long for probability theory to develop?

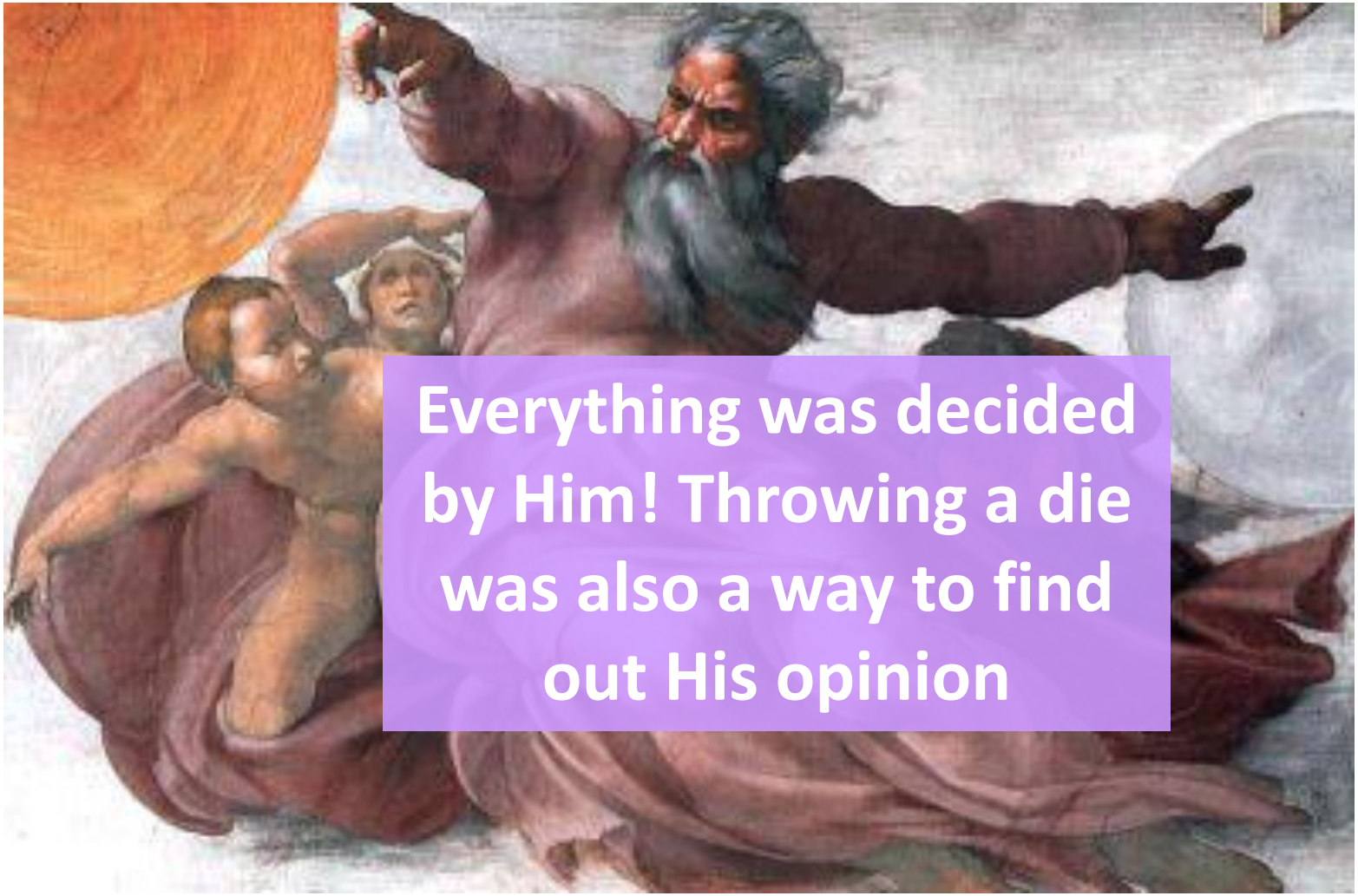
Probability theory grew out of the interest in gambling

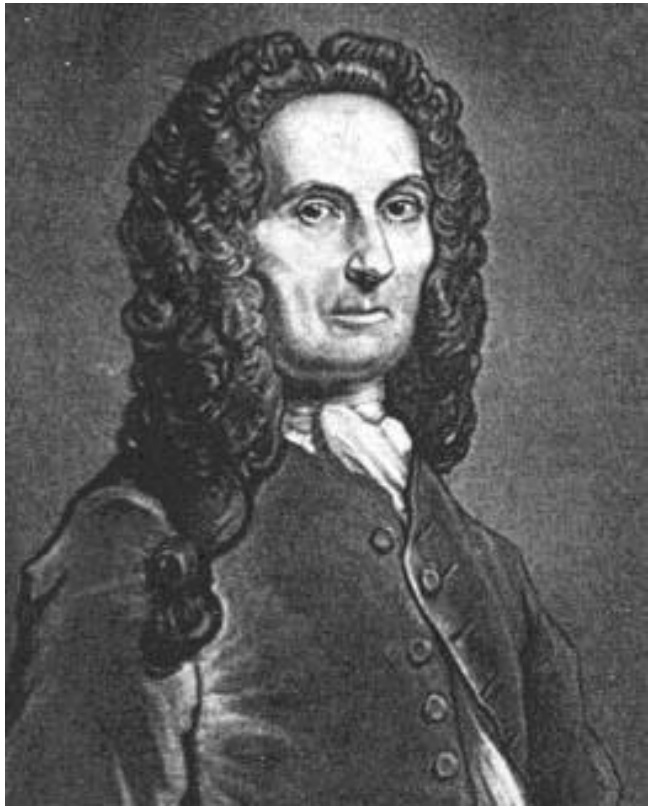
But people have gambled since the last ice age or even before that – so why the delay??

Why did this knowledge not "spill over" into science?

Because people did not have any perception of *randomness*

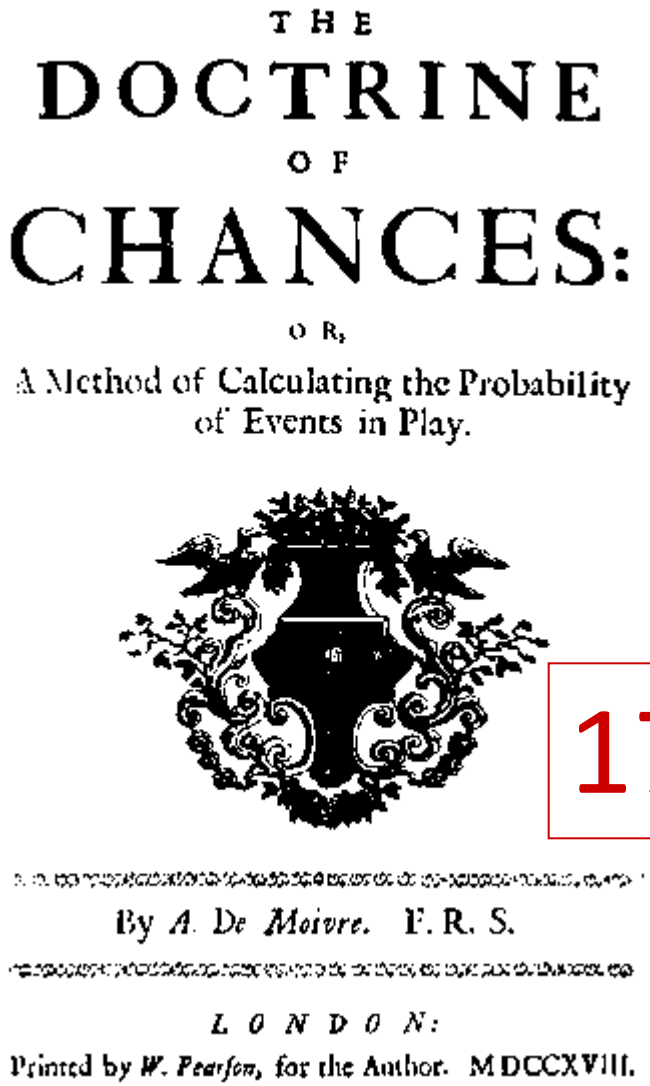
(except perhaps Cicero and some other Romans)





Abraham De Moivre
1667-1754

From causes to effects
Deduction
Direct probabilities
Combinatorics



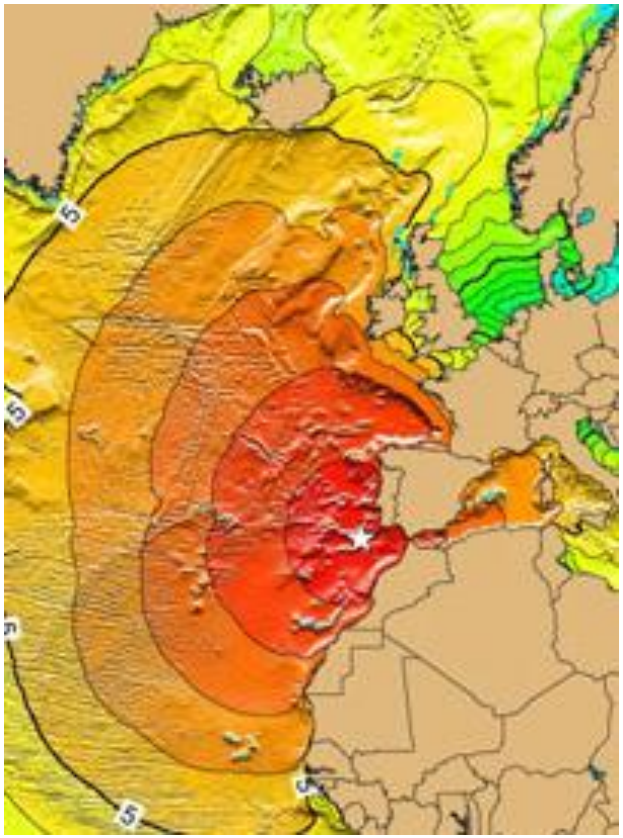
THE
DOCTRINE
OF
CHANCES:
OR,
A Method of Calculating the Probability
of Events in Play.

By *A. De Moivre.* F. R. S.

LONDON:
Printed by *W. Pearson,* for the Author. MDCCXVIII.

1718

The Lisbon earthquake and tsunami 1755

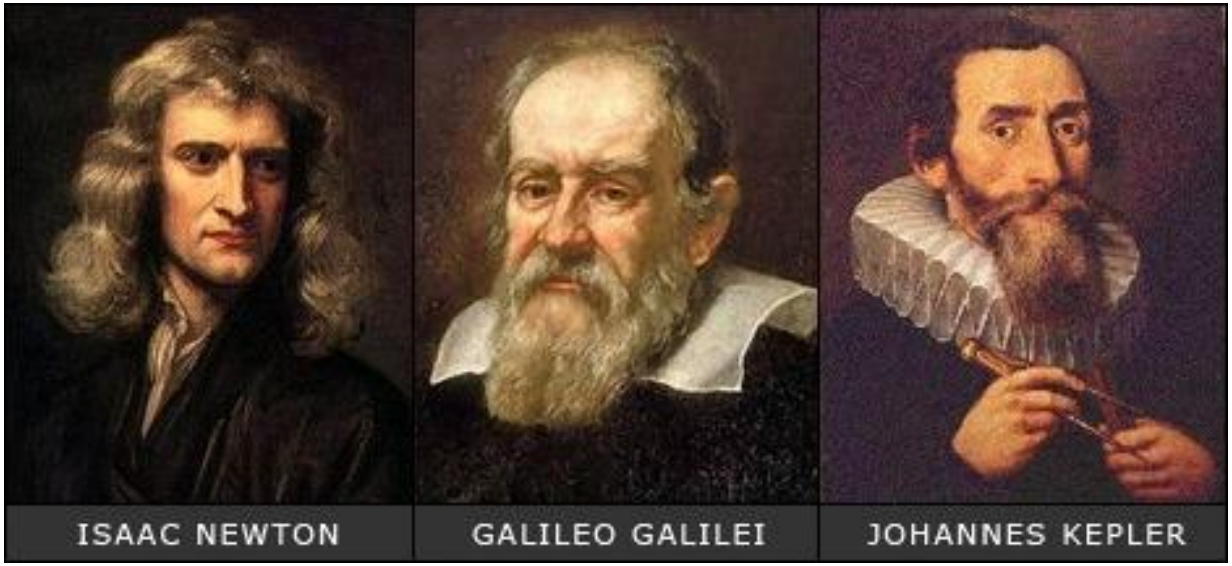


made people start doubt the existence of an all mighty God that decided everything.

From 1750's ideas about randomness in science

4. Statistical estimation theory is also quite young

Before the 1800's there was a poor understanding of randomness in measurement errors



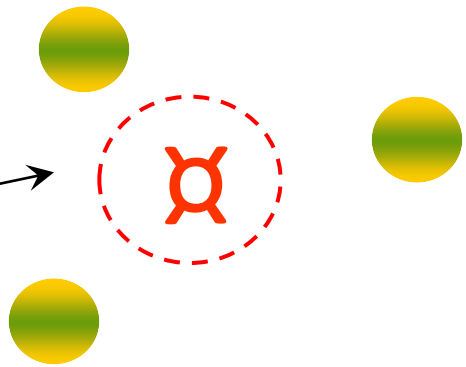
1. Scientists had the routine to select their "best" measurement
2. They didn't understand that measurement errors add up and randomly cancel out
3. **They disliked averages of observations since these did not normally agree with measured values**

18th century view on observation errors

1. Astronomers in the 1600:s and 1700:s tried to find out which of their diverging observations was the "right" one

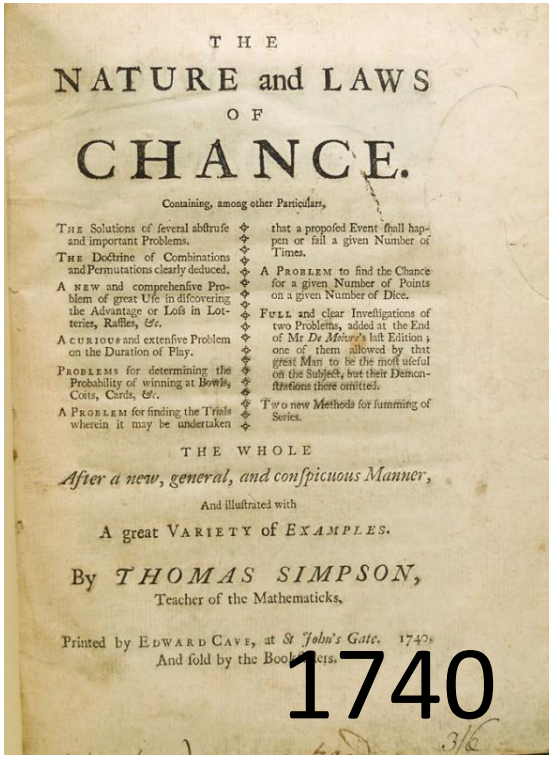
2. In the late 1700' it was realized that that the observations should be combined **even if the result did not agree with any of the observations**

Where is Jupiter?



3. The first mathematical discussion on statistical inference

Thomas Simpson
1710-61
Mathematician



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.

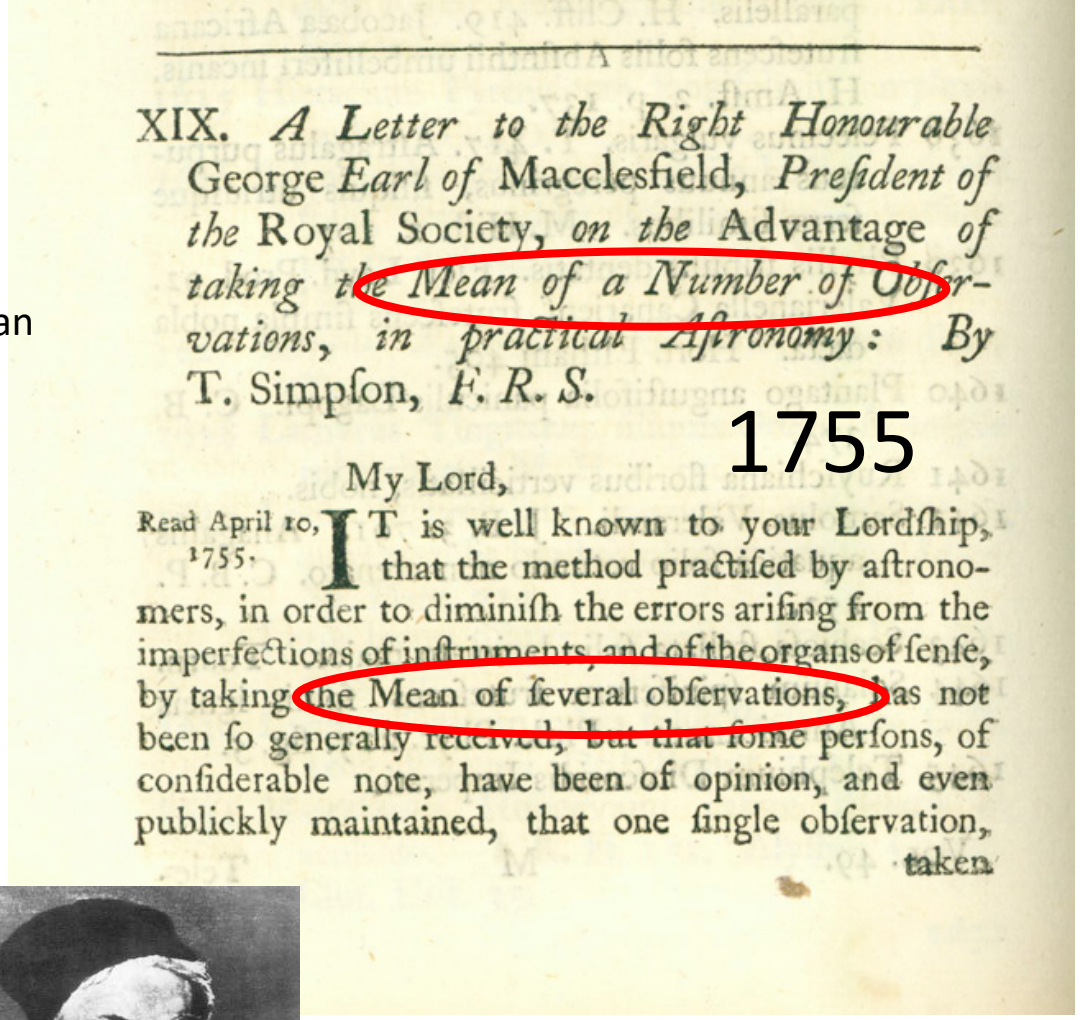
Two new Methods for ſumming of Series.

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.

1740



XIX. *A Letter to the Right Honourable* George *Earl of* Macclesfield, *Preſident of the* Royal Society, *on the* Advantage *of taking the Mean of a Number of Obſervations, in practical Aſtronomy: By* T. Simpſon, *F. R. S.*

1755

My Lord,

Read April 10, 1755.

IT is well known to your Lordſhip, that the method practiſed by aſtronomers, in order to diminiſh the errors ariſing from the imperfections of inſtruments, and of the organs of ſenſe, by taking the Mean of ſeveral obſervations, has not been ſo generally received, but that ſome perſons, of conſiderable note, have been of opinion, and even publickly maintained, that one ſingle obſervation, taken

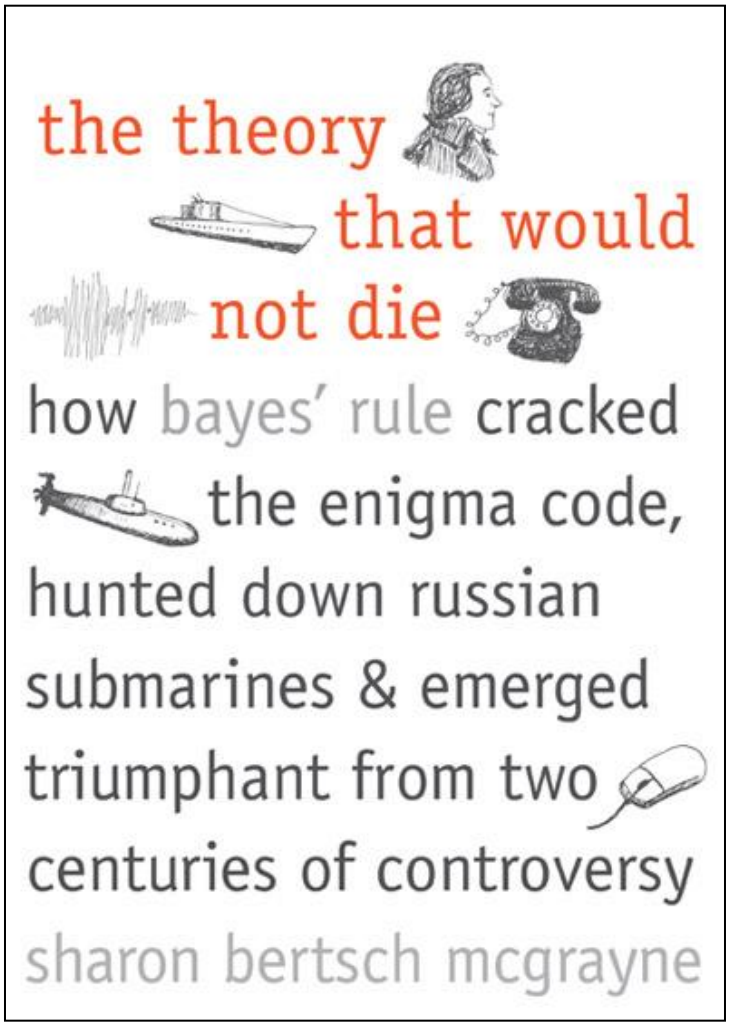
Only accepted 50-60 years later thanks to the works by Lagrange and Gauss

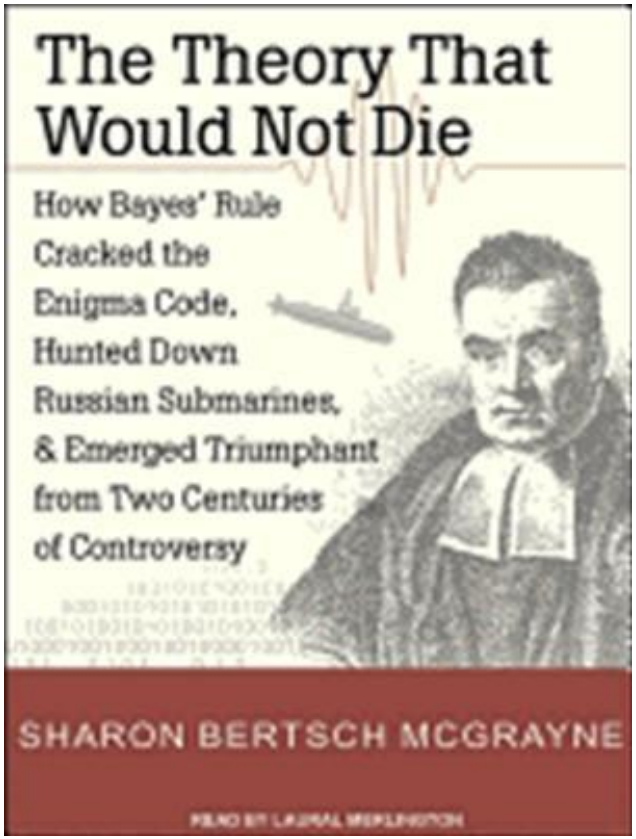
The Belgian meteorologist and statistician Adolphe Quételet (1796-1874) introduced in the mid 1800's the concept of "The Average Man" based on statistical averages from the population in Brussels.

He was criticised because there was nobody in Brussels who fitted this description

5. An old “newcomer” – Bayesian statistics

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

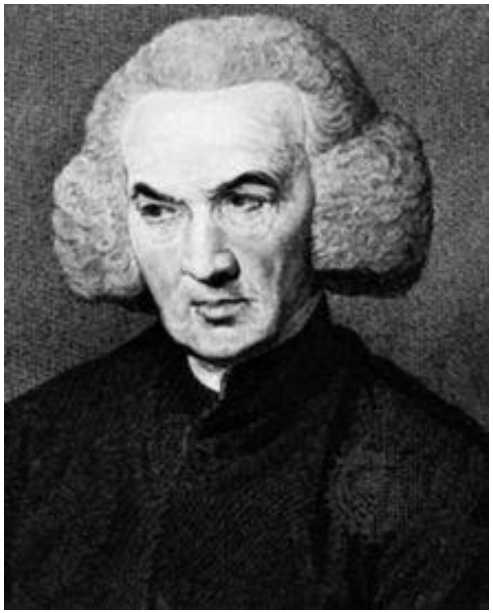




Thomas Simpson 1710-61 Mathema-tician

David Hume 1711-76

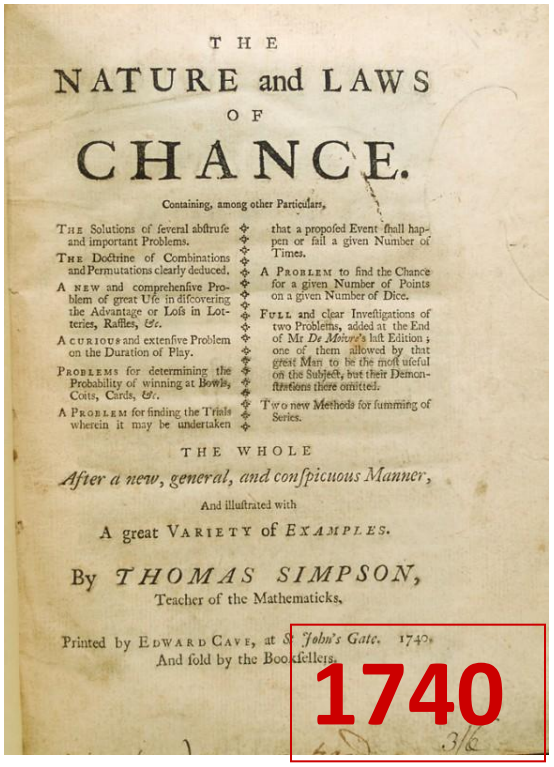
"There is no First Cause"



Richard Price 1723-91 Radical priest

Friend of the American and French revolutions

Supported women's rights



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.*

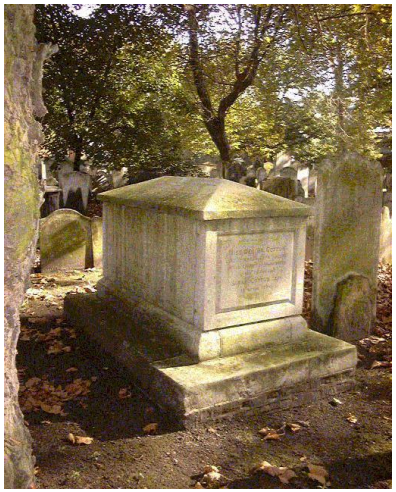
By *THOMAS SIMPSON,*

Teacher of the Mathematicks.

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

And sold by the Booksellers.

1740



Thomas Bayes

1701-61

Dissident clergyman

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

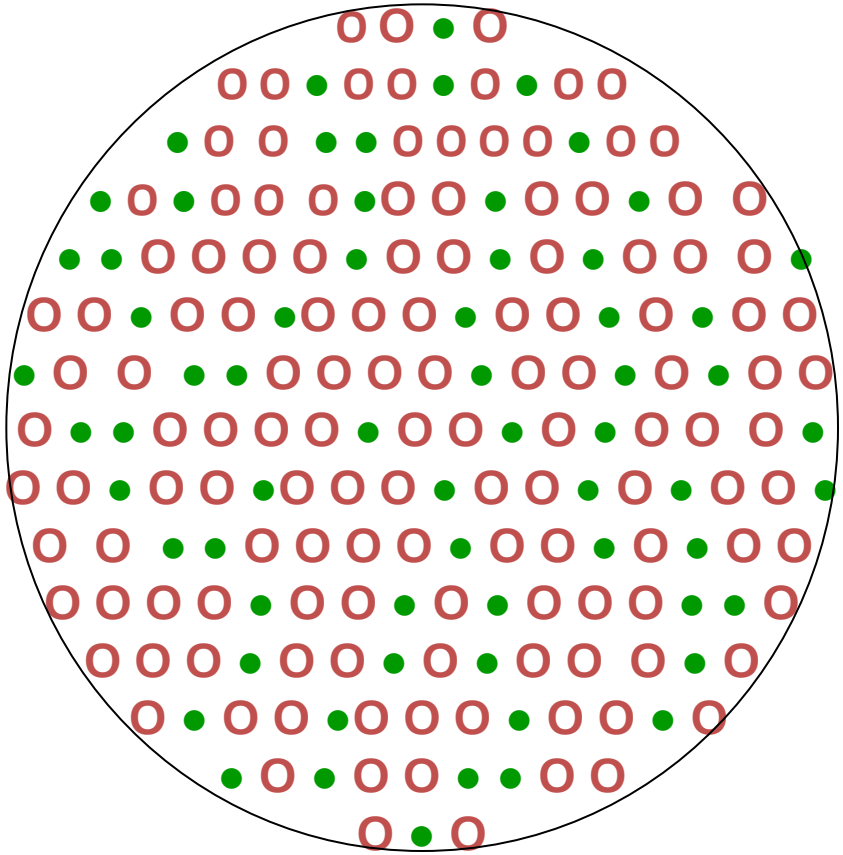
Dear Sir,

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.

Classical probabilities

rain 30% dry 70%



Selecting three balls yields

● ● ● 3% of cases

● ● o 19% - " -

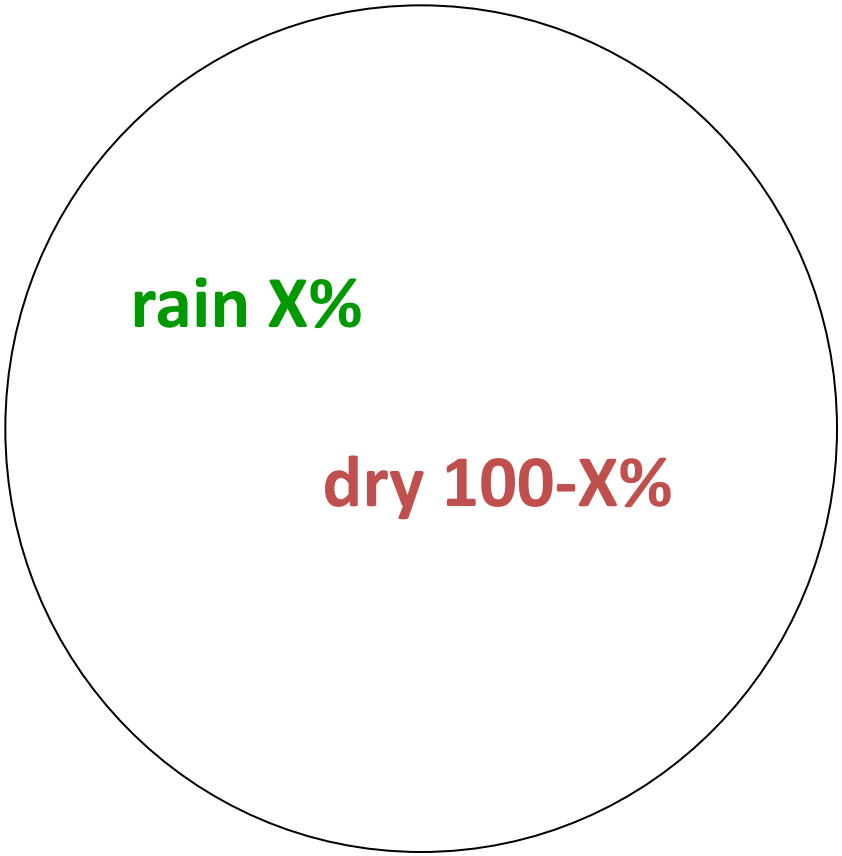
● o o 44% - " -

o o o 34% - " –

. . . with a risk of 56% of misrepresentation

Ensemble of ∞ independent NWP

Inverse or Bayesian probabilities:



Selecting three balls and getting

● ● o

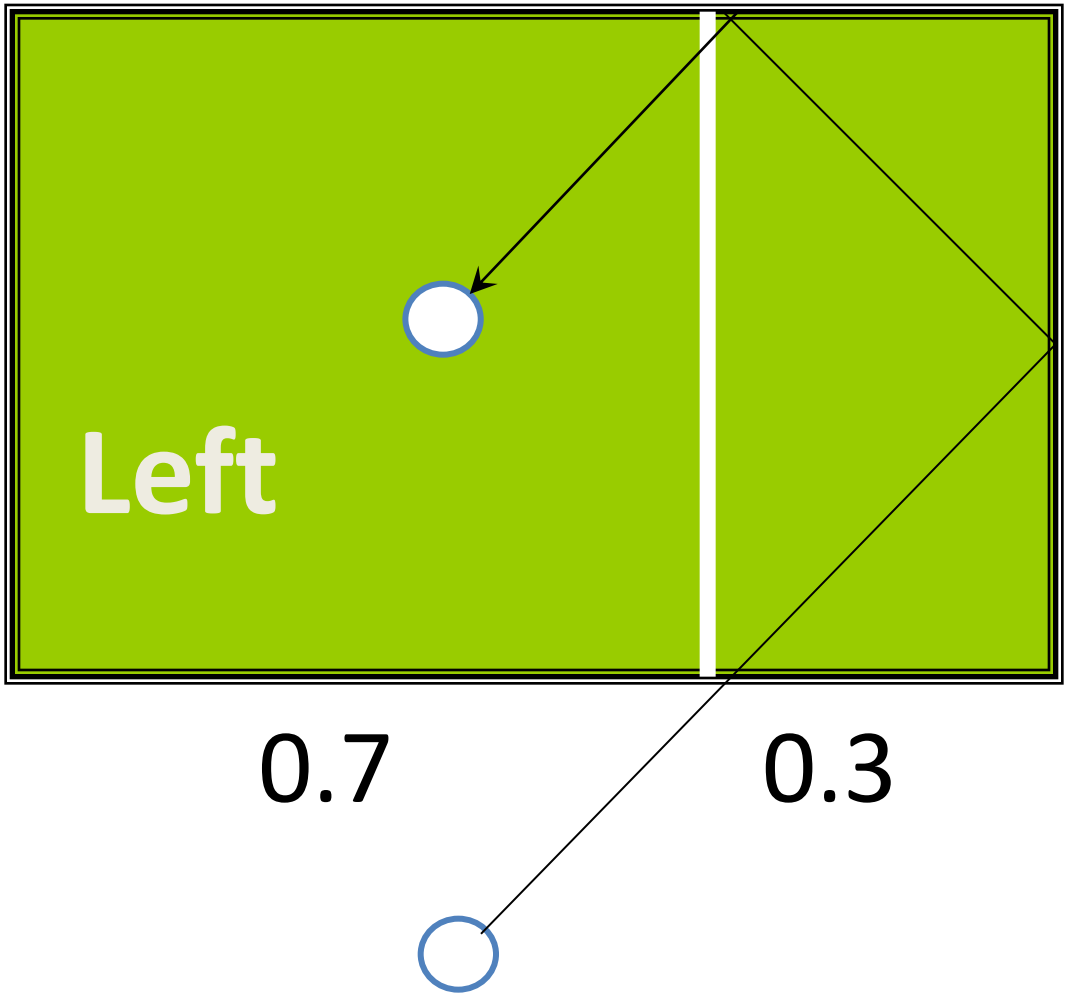
What does that tell us about the proportions

X and **100 – X**?

Ensemble of ∞ independent NWP

Bayes's billiard table experiment

Bayes's experiment as it would have been set up by (de Moivre) a classicist:



Throwing 3 balls

Prob(RRR)= 3%

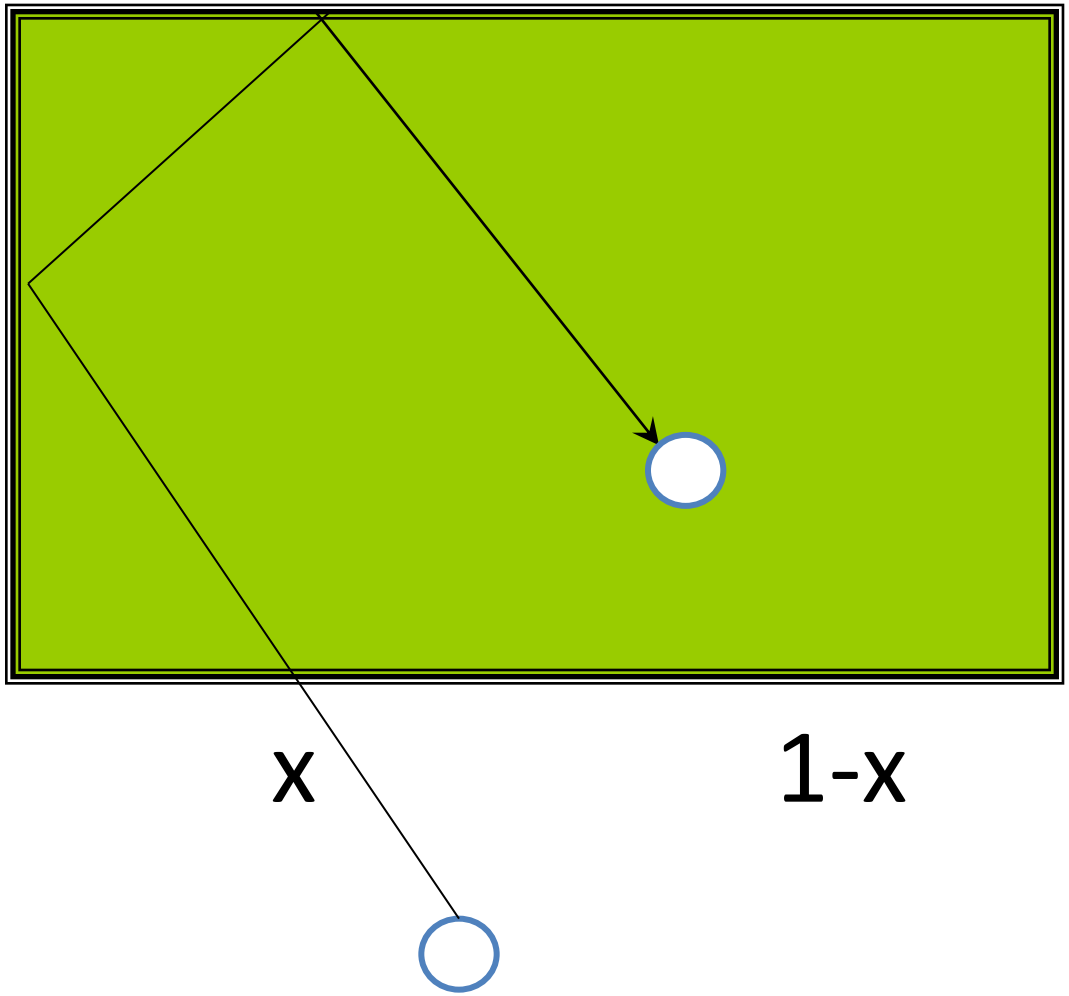
Prob(RRL,RLR,LRR)=19%

Prob(RLL,LRL,LLR)=44%

Prob(LLL)=34%

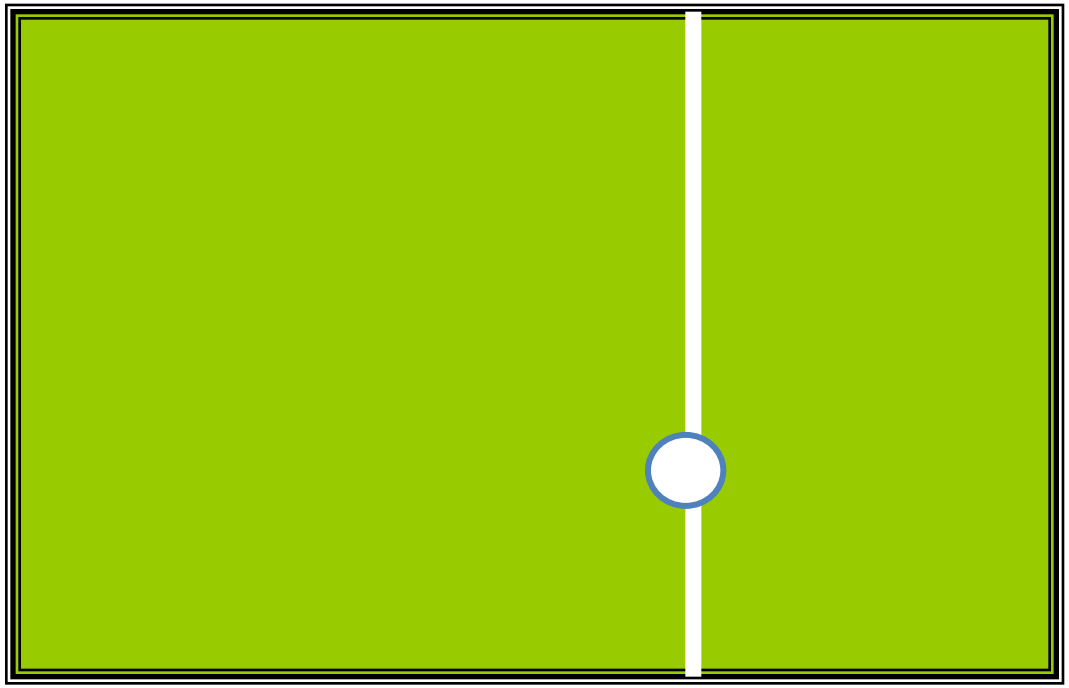
Thomas Bayes' experiment

(defining the white line)



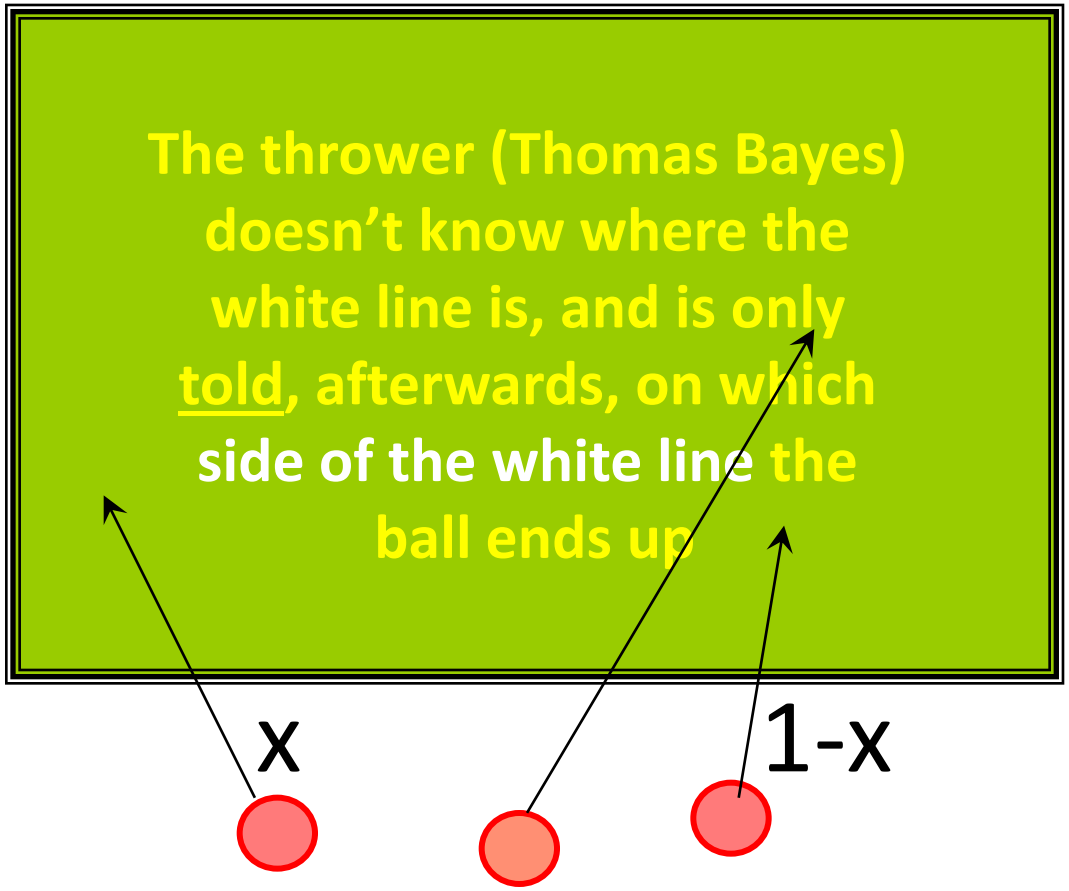
Thomas Bayes' experiment

(defining the white line)



x 1-x

Thomas Bayes' experiment



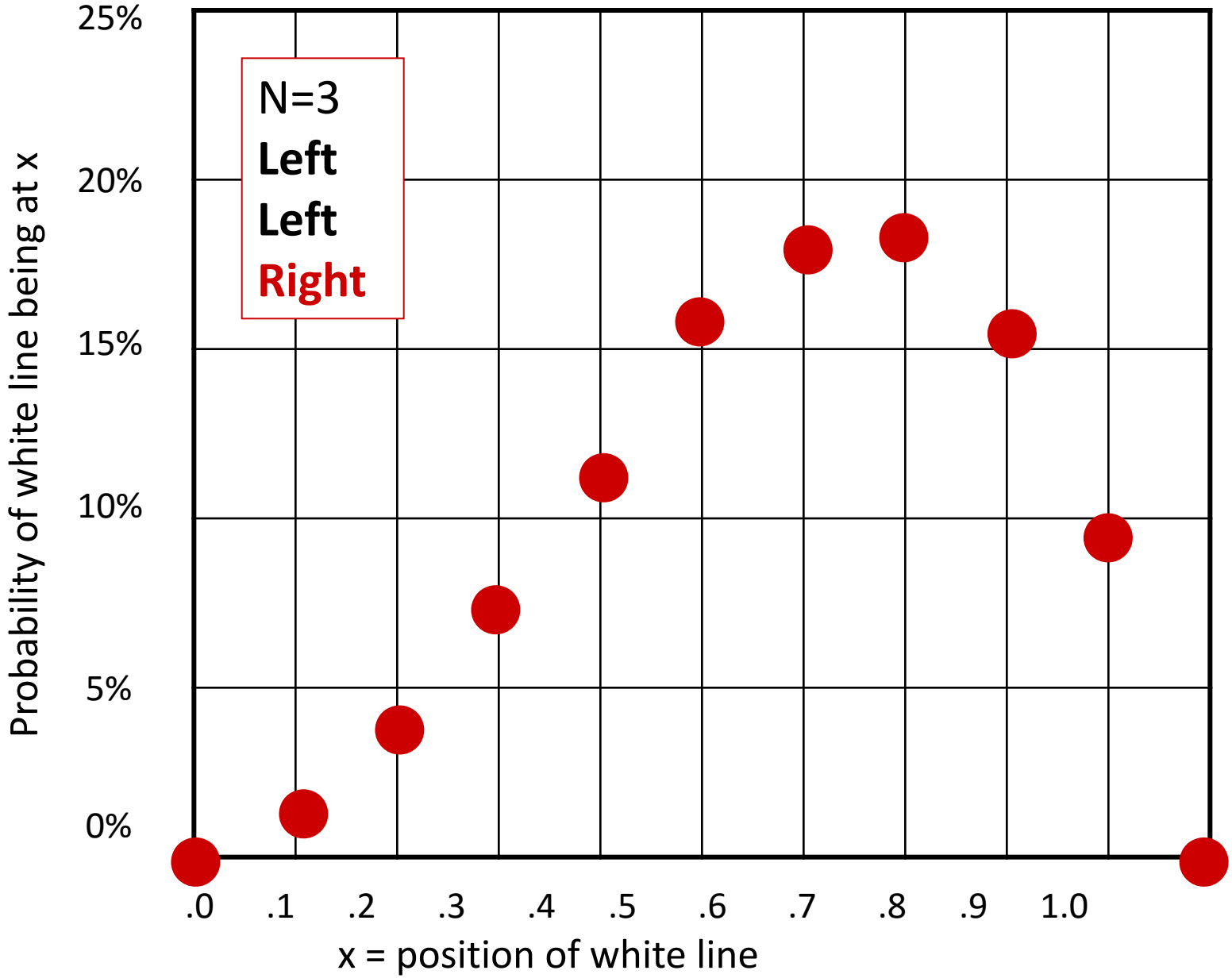
Left

Left

Right

This can be solved by using the non-controversial "Bayes Rule"

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


N=3
Left
Left
Right
25%
20%
15%
10%
5%
0%
Probability of white line being at x
.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
x = position of white line

A Bayesian approach avoids over-confident probabilities such as Concorde before 2000 being the world's safest air plane

accidents

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

Concorde — Some other airline

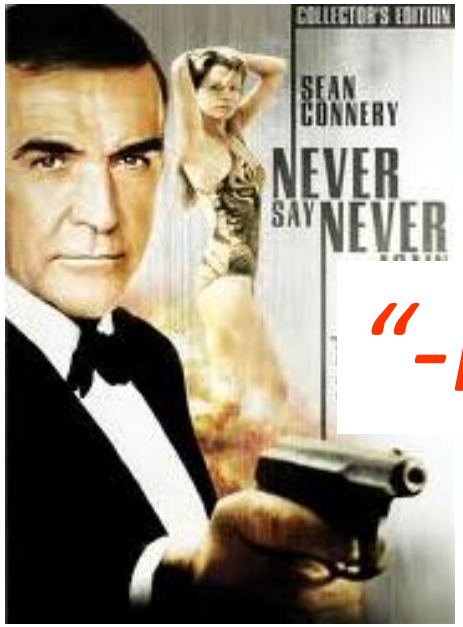
. . .after the 2000 crash the most unsafe

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

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

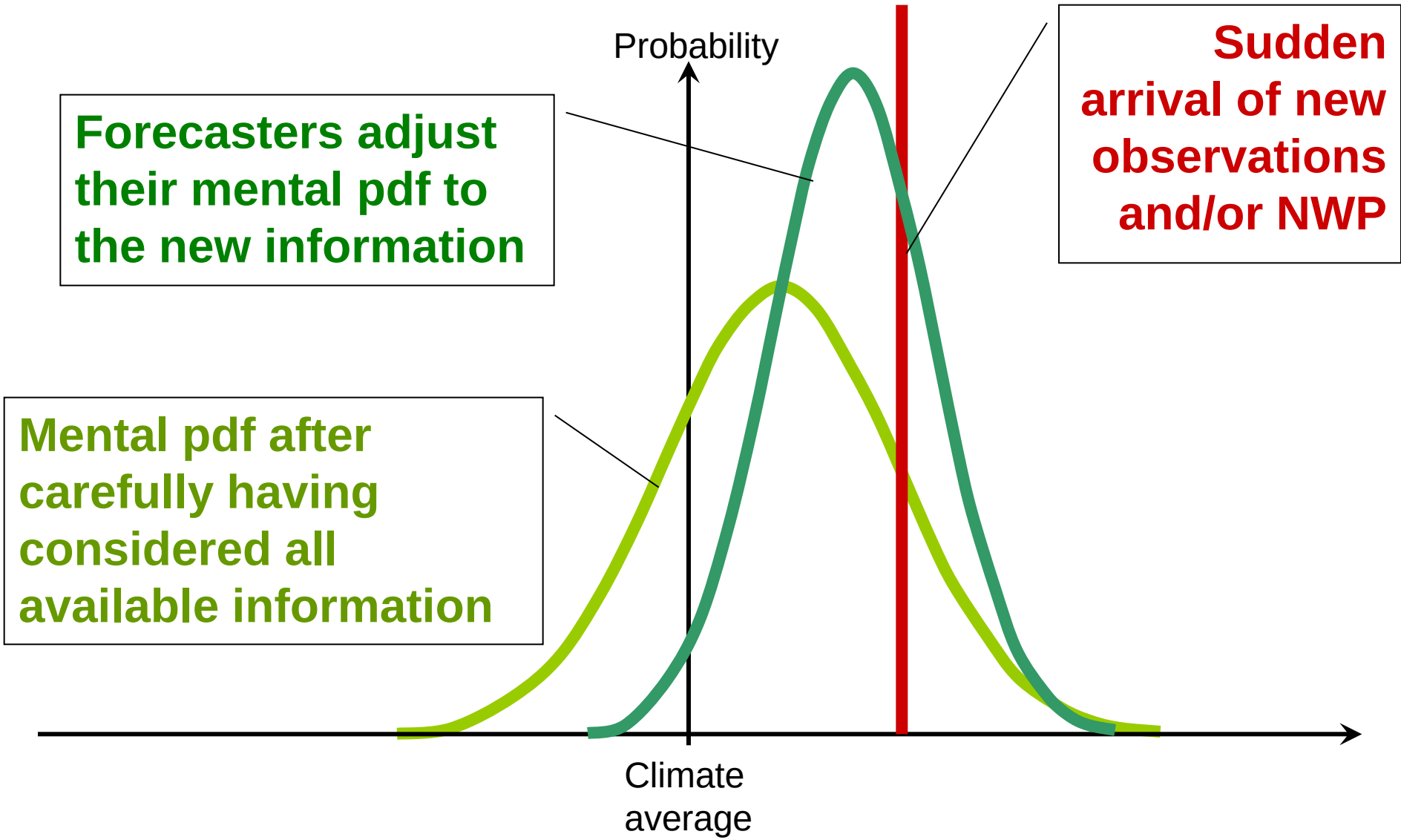
"Laplace Rule of Succession" prevents us from being overconfident since **p** never takes values 0% or 100%

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



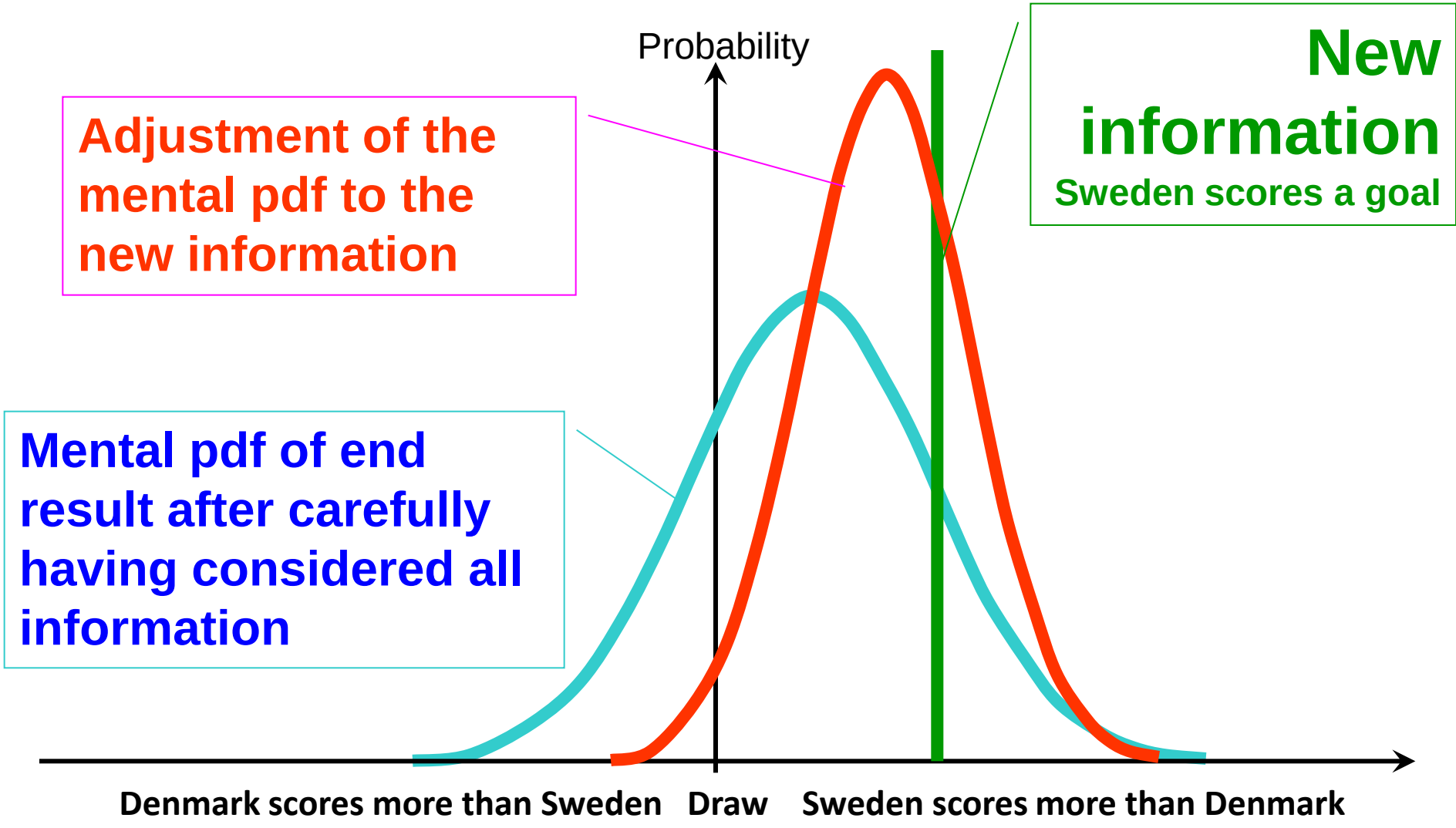
"-Never say never"

Intuitive Bayesianism among weather forecasters



Updating of subjective probabilities

Updating of subjective probabilities



Break