Telling the Truth with Statistics

Lecture 3

Giulio D’Agostini

Università di Roma La Sapienza e INFN

Roma, Italy

Overview of the contents

1st part Review of the process of learning from data

Mainly based on

- *"From observations to hypotheses: Probabilistic reasoning versus falsificationism and its statistical variations"* (Vulcano 2004, physics/0412148)
- Chapter 1 of *"Bayesian reasoning in high energy physics. Principles and applications"* (CERN Yellow Report 99-03)

Overview of the contents

1st part Review of the process of learning from data
Mainly based on

- *"From observations to hypotheses: Probabilistic reasoning versus falsificationism and its statistical variations"* (Vulcano 2004, physics/0412148)
- Chapter 1 of *"Bayesian reasoning in high energy physics. Principles and applications"* (CERN Yellow Report 99-03)

2nd part Review of the probability and 'direct probability' problems, including 'propagation of uncertainties.
Partially covered in

- First 3 sections of Chapter 3 of YR 99-03
- Chapter 4 of YR 99-03
- *"Asymmetric uncertainties: sources, treatment and possible dangers"* (physics/0403086)

Overview of the contents

3th part Probabilistic inference and applications to HEP

Much material and references in my web page. In particular, I recommend a quite concise review

- *"Bayesian inference in processing experimental data: principles and basic applications"*, Rep.Progr.Phys. 66 (2003)1383 [physics/0304102]

For a more extensive treatment:,

- *"Bayesian reasoning in data analysis – A critical introduction"*, World Scientific Publishing, 2003

  (CERN Yellow Report 99-03 updated and $\approx$ doubled in contents)

Summary of 2nd lecture

- ‘Conventional’ statistics rejects the natural concept of probability of causes, of hypotheses, etc.

Summary of 2nd lecture

- ‘Conventional’ statistics rejects the natural concept of probability of causes, of hypotheses, etc.
- but this approach causes a mismatch between natural thinking and the cultural superstructure.

Summary of 2nd lecture

- ‘Conventional’ statistics rejects the natural concept of probability of causes, of hypotheses, etc.
- but this approach causes a mismatch between natural thinking and the cultural superstructure.
- The results of ‘conventional’ statistics methods are systematically misinterpreted, with terrible consequences in scientific judgment

Summary of 2nd lecture

- 'Conventional' statistics rejects the natural concept of probability of causes, of hypotheses, etc.
- but this approach causes a mismatch between natural thinking and the cultural superstructure.
- The results of 'conventional' statistics methods are systematically misinterpreted, with terrible consequences in scientific judgment
- and, anyhow, hypotheses tests are not based on a kind of first principles; therefore their 'prescriptions' and their implementation are largely arbitrary (that is not the same as 'subjective'!)

Summary of 2nd lecture

- but this approach causes a mismatch between natural thinking and the cultural superstructure.
- The results of ‘conventional’ statistics methods are systematically misinterpreted, with terrible consequences in scientific judgment
- and, anyhow, hypotheses tests are not based on a kind of first principles; therefore their ‘prescriptions’ and their implementation are largely arbitrary (that is not the same as ‘subjective’!)
- We have then reviewed the basic question of uncertainty and, indeed, how the human mind forms naturally its degree of belief on future events both from ‘combinatoric’ evaluations and from relative frequencies of observations,

Summary of 2nd lecture

- The results of ‘conventional’ statistics methods are systematically misinterpreted, with terrible consequences in scientific judgment
- and, anyhow, hypotheses tests are not based on a kind of first principles; therefore their ‘prescriptions’ and their implementation are largely arbitrary (that is not the same as ‘subjective’!)
- We have then reviewed the basic question of uncertainty and, indeed, how the human mind forms naturally its degree of belief on future events both from ‘combinatoric’ evaluations and from relative frequencies of observations,
- though in the latter case we tend to ‘filter’ the process of transferring the past to the future.

Summary of 2nd lecture

- and, anyhow, hypotheses tests are not based on a kind of first principles; therefore their ‘prescriptions’ and their implementation are largely arbitrary (that is not the same as ‘subjective’!)
- We have then reviewed the basic question of uncertainty and, indeed, how the human mind forms naturally its degree of belief on future events both from ‘combinatoric’ evaluations and from relative frequencies of observations,
- though in the latter case we tend to ‘filter’ the process of transferring the past to the future.
- But the assessment of beliefs in the inferential problem belongs to neither of the above two kinds of reasoning, but it is a problem of ‘probability inversion’ that needs formal logic

Summary of 2nd lecture

- We have then reviewed the basic question of uncertainty and, indeed, how the human mind forms naturally its degree of belief on future events both from 'combinatoric' evaluations and from relative frequencies of observations,
- though in the latter case we tend to 'filter' the process of transferring the past to the future.
- But the assessment of beliefs in the inferential problem belongs to neither of the above two kinds of reasoning, but it is a problem of 'probability inversion' that needs formal logic
- Sources of uncertainty in measurements and 'standard' methods to handle 'statistical' and 'systematic' errors.
- Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$, naive probability inversions and the dog-hunter analogy.

Summary of 2nd lecture

- though in the latter case we tend to ‘filter’ the process of transferring the past to the future.
- But the assessment of beliefs in the inferential problem belongs to neither of the above two kinds of reasoning, but it is a problem of ‘probability inversion’ that needs formal logic
- Sources of uncertainty in measurements and ‘standard’ methods to handle ‘statistical’ and ‘systematic’ errors.
- Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$, naive probability inversions and the dog-hunter analogy.
- FORWARD TO THE PAST: restart from reviewing the very concept of probability: not bound to the text book ‘definitions’ of probability

Summary of 2nd lecture

- But the assessment of beliefs in the inferential problem belongs to neither of the above two kinds of reasoning, but it is a problem of ‘probability inversion’ that needs formal logic
- Sources of uncertainty in measurements and ‘standard’ methods to handle ‘statistical’ and ‘systematic’ errors.
- Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$, naive probability inversions and the dog-hunter analogy.
- FORWARD TO THE PAST: restart from reviewing the very concept of probability: not bound to the text book ‘definitions’ of probability
- $\rightarrow$ Intrinsic subjective nature of probability

Summary of 2nd lecture

- Sources of uncertainty in measurements and 'standard' methods to handle 'statistical' and 'systematic' errors.
- Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$, naive probability inversions and the dog-hunter analogy.
- FORWARD TO THE PAST: restart from reviewing the very concept of probability: not bound to the text book 'definitions' of probability
- $\rightarrow$ Intrinsic subjective nature of probability
- and importance of the state of information in the evaluations of probability: '$P(E)$' $\longrightarrow$ $P(E\,|\,I)$ $\longrightarrow$ $P(E\,|\,I(t))$

Unifying role of subjective probability

- Wide range of applicability

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \le m_Z/\text{GeV} \le 91.1897) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \le m_Z/\text{GeV} \le 91.1897) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the same confidence on something

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \le m_Z/\text{GeV} \le 91.1897) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the same confidence on something
- I can agree or disagree, but at least I know what this person has in mind (and this does not happens with the "C.L.'s")

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \leq m_Z/\text{GeV} \leq 91.1897) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$

  They all convey unambiguously the same confidence on something
- I can agree or disagree, but at least I know what this person has in mind (and this does not happens with the "C.L.'s")
- If a person has these beliefs and he/she has the chance to win a rich prize bound to one of these events, he/she has no reason to chose an event instead than the others.

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \le m_Z/\text{GeV} \le 91.1897) = 68\%$
  - $P(\text{free neutron decays before } 17\,\text{s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$
- probability not bound to a single evaluation rule

Unifying role of subjective probability

- Wide range of applicability
- Probability statements all have the same meaning no matter to what they refer and how the number has been evaluated.
  - $P(\text{rain tomorrow}) = 68\%$
  - $P(\text{Juventus will win Italian champion league}) = 68\%$
  - $P(91.1855 \le m_Z/\text{GeV} \le 91.1897) = 68\%$
  - $P(\text{free neutron decays before 17 s}) = 68\%$
  - $P(\text{White ball from a box with 68W+32B}) = 68\%$
- probability not bound to a single evaluation rule
- In particular, combinatorial and frequency based ‘definitions’ are easily recovered as evaluation rules under well defined hypotheses.
- Keep separate concept from evaluation rule

From the concept of probability to the probability theory

Ok, it looks nice, . . . but "how do we deal with 'numbers'?"

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to ‘propagate’ probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Coherent bet (de Finetti, Ramsey - 'Dutch book argument')

It is well understood that bet odds can express confidence†

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Coherent bet $\longrightarrow$ A bet acceptable in both directions:

- You state your confidence fixing the bet odds
- . . . but somebody else chooses the direction of the bet
- best way to honestly assess beliefs.

$\longrightarrow$ see later for details, examples, objections, etc

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Consistency arguments (Cox, + Good, Lucas)

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

  Consistency arguments (Cox, + Good, Lucas)
- Similar approach by Schrödinger (much less known)
- Supported by Jaynes and Maximum Entropy school

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Consistency arguments (Cox, + Good, Lucas)

- Similar approach by Schrödinger (much less known)
- Supported by Jaynes and Maximum Entropy school

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Lindley's 'calibration' against 'standards'

→ analogy to measures (we need to measure 'befiefs')

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

Lindley's 'calibration' against 'standards'

$\longrightarrow$ analogy to measures (we need to measure 'befiefs')

$\Rightarrow$ reference probabilities provided by simple cases in which equiprobability applies (coins, dice, turning wheels,. . .).

- Example: You are offered to options to receive a price: a) if E happens, b) if a coin will show head. Etc. . . .

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

  Lindley's 'calibration' against 'standards'

$\rightarrow$ Rational under everedays expressions like "there are 90 possibilities in 100" to state beliefs in situations in which the real possibilities are indeed only 2 (e.g. dead or alive)

- Example: a question to a student that has to pass an exam: a) normal test; b) pass it is a uniform random x will be ≤ 0.8.

From the concept of probability to the probability theory

- Formal structure: we need a mathematical structure in order to 'propagate' probability values to other, logically connected events:
  - basic rules
  - logic (mathematics)
- Assess probability: The formal structure is an empty box. In which we have to insert some numbers.
  - Is there a very general rule ?

  Lindley's 'calibration' against 'standards'
- Also based on coherence, but it avoids the 'repulsion' of several person when they are asked to think directly in terms of bet (it is proved that many person have reluctance to bet money).

Basic rules of probability

They all lead to

$$
\begin{array}{ll}
1. & 0 \leq P(A) \leq 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \quad [\,\text{if } P(A \cap B) = \emptyset\,] \\
 & \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

where

- Ω stands for ‘tautology’ (a proposition that is certainly true $\rightarrow$ referring to an event that is certainly true) and $\emptyset = \overline{\Omega}$.
- $A \cap B$ is true only when both A and B are true (logical AND) (shorthands ‘A, B’ or $A\,B$ often used $\rightarrow$ logical product)
- $A \cup B$ is true when at least one of the two propositions is true (logical OR)

Basic rules of probability

Remember that probability is always conditional probability!

1. $0 \le P(A \,|\, I) \le 1$
2. $P(\Omega \,|\, I) = 1$
3. $P(A \cup B \,|\, I) = P(A \,|\, I) + P(B \,|\, I) \quad [\text{if } P(A \cap B \,|\, I) = \emptyset\,]$
4. $P(A \cap B \,|\, I) = P(A \,|\, B, I) \cdot P(B \,|\, I) = P(B \,|\, A, I) \cdot P(A \,|\, I)$

I is the background condition (related to information I)

$\longrightarrow$ usually implicit (we only care on 're-conditioning')

Meaning of the basic rules

Have we recovered the famous axioms?

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,]$

4. $P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)$

Meaning of the basic rules

More or less yes, at least formally

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,]$

4. $P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)$

Meaning of the basic rules

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,]$

4. $P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)$

- In the axiomatic approach
  - 'probability' is just a real number that satisfies 1-3
  - rule 4 comes straight from the <u>definition</u> of conditional probability as

$$P(A \,|\, B) = \frac{P(A \cap B)}{P(B)} \quad [\text{if } P(B) > 0\,]$$

Meaning of the basic rules

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,]$

4. $P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)$

- In the subjective approach
  - the intuitive meaning of 'probability' is recovered
  - rules 1-4 derive from more basic assumptions (e.g. the coherent bet)
  - $P(A \,|\, B) = P(A \cap B)/P(B)$ does not define $P(A \,|\, B)$
    $\rightarrow$ conditional probability is an intuitive concept!
    (Remember Schrödinger quote)

Meaning of the basic rules

1. $0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ [if $P(A \cap B) = \emptyset$]

4. $P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)$

- In the subjective approach
  - the intuitive meaning of ‘probability’ is recovered
  - rules 1-4 derive from more basic assumptions (e.g. the coherent bet)
  - $P(A \,|\, B) = P(A \cap B)/P(B)$ <u>does not define</u> $P(A \,|\, B)$
    $\longrightarrow$ conditional probability is an intuitive concept!

$\Rightarrow$ As we actually use it! $\rightarrow$

About the ‘conditional probability formula’

$$4. \qquad P(E\cap H) = P(E\,|\,H)\cdot P(H) = P(H\,|\,E)\cdot P(E)$$

$$4a. \qquad P(E\,|\,H) = \frac{P(E\cap H)}{P(H)} \quad [P(H)>0]$$

About the 'conditional probability formula'

$$4. \qquad P(E \cap H) = P(E \,|\, H) \cdot P(H) = P(H \,|\, E) \cdot P(E)$$

$$4a. \qquad P(E \,|\, H) = \frac{P(E \cap H)}{P(H)} \quad [P(H) > 0]$$

In the subjective approach the meaning is clear:

- Depending on the information we have, we can assess any of the three probabilities that enter the formula: $P(H)$, $P(E \,|\, H)$ or $P(E \cap H)$.
- But, once two of the three have been assessed, the third one is constraint!

  (otherwise, one can prove it is possible to imagine a set of bets, such that one certainly gains or loses – incoherent)
- 4 is more general than 4.a, valid also if $P(H) = 0$

About the 'conditional probability formula'

$$4. \qquad P(E\cap H) = P(E\,|\,H)\cdot P(H) = P(H\,|\,E)\cdot P(E)$$

$$4a. \qquad P(E\,|\,H) = \frac{P(E\cap H)}{P(H)} \quad [P(H)>0]$$

What is the chance that a 550 GeV Higgs is detected by ATLAS?

- H = "Higgs mass 550 GeV"
- E = "Decay products observed in ATLAS"

$\Rightarrow$ $P(E\,|\,H)$ is a routine task: $\rightarrow$ set $M_H = 550$ GeV in the physics generator $\rightarrow$ run the events through the full simulation chain $\rightarrow$ run analysis program $\rightarrow$ estimate $P(E\,|\,H)$ from percentage of reconstructed events.

- None would use definition 4a [what is $P(E\cap H)$?]
- Note: $P(E\,|\,H)$ is meaningful even if $P(H)=0$ (why not?).

Some comments about subjective probability and bets

I imagine many objections, e.g.

- In physics there is no room for beliefs
- 'Subjective' is 'arbitrary'
- With whom should I bet
- Subjective probability is not suited for scientific research
  $\longrightarrow$ "I want to be objective"
- Physical probabilities do not depend on our beliefs

Beliefs in physics?

A colleague, once: "I do not believe something. I assess it. This is not matter fir religion!"

I hope at least he believes what he assesses. Otherwise I don't know what to do of his assessments.

Anyhow, and apart from the jokes, Science is nothing but a collection of rational beliefs based in experimental evidences and theoretical speculations.

The statistician Don Berry has amused himself by counting how many times Stephen Hawking uses 'belief', 'to believe', or synonyms, in his *'A brief history of time'*. The book could have been entitled *'A brief history of beliefs'*, concludes Berry.

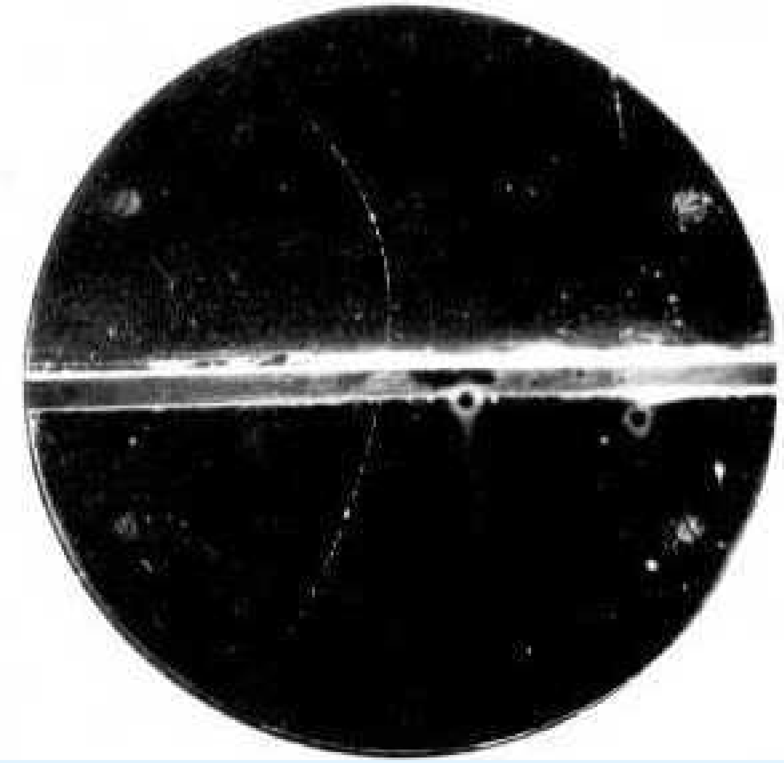
Physics: a network of beliefs

Peter Galison (*How Experiments End*): *"Experiments begin and end in a matrix of beliefs. . . . beliefs in instrument type, in programs of experiment enquiry, in the trained, individual judgments about every local behavior of pieces of apparatus."*

Physics: a network of beliefs

Peter Galison (*How Experiments End*): *"Experiments begin and end in a matrix of beliefs. . . . beliefs in instrument type, in programs of experiment enquiry, in the trained, individual judgments about every local behavior of pieces of apparatus."*

"Taken out of time there is no sense to the judgment that Anderson's track 75 is a positive electron; its textbook reproduction has been denuded of the prior experience that made Anderson confident in the cloud chamber, the magnet, the optics, and the photography."



Beliefs in physics

Pure observation does not create, or increase, knowledge without personal inputs which are needed to elaborate the information.

Beliefs in physics

Pure observation does not create, or increase, knowledge without personal inputs which are needed to elaborate the information.

There is nothing really objective in physics, if by objective we mean that something follows necessarily from observation, like the proof of a theorem.

Beliefs in physics

Pure observation does not create, or increase, knowledge without personal inputs which are needed to elaborate the information.

There is nothing really objective in physics, if by objective we mean that something follows necessarily from observation, like the proof of a theorem.

Nevertheless, physics is objective, or at least that part of it that is at present well established, if we mean by 'objective', that a rational individual cannot avoid believing it.

Beliefs in physics

Pure observation does not create, or increase, knowledge without personal inputs which are needed to elaborate the information.

There is nothing really objective in physics, if by objective we mean that something follows necessarily from observation, like the proof of a theorem.

Nevertheless, physics is objective, or at least that part of it that is at present well established, if we mean by ‘objective’, that a rational individual cannot avoid believing it.

This is the reason why we can talk in a relaxed way about beliefs in physics without even remotely thinking that it is at the same level as the stock exchange, or betting on football scores

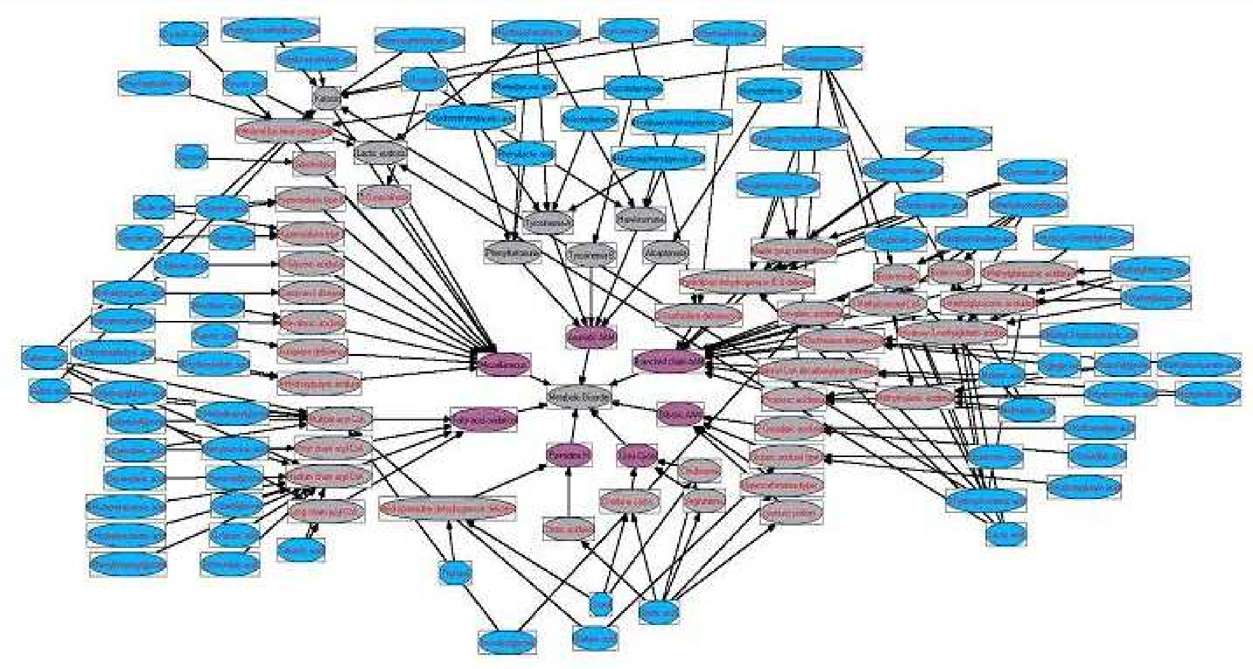
A solid core surrounded by fuzzy borders

The reason the ‘perceived objectivity’ in physics is that, after centuries of experimentation, theoretical work and successful predictions, there is such a consistent network of beliefs, it has acquired the status of an objective construction:

A solid core surrounded by fuzzy borders

The reason the ‘perceived objectivity’ in physics is that, after centuries of experimentation, theoretical work and successful predictions, there is such a consistent network of beliefs, it has acquired the status of an objective construction:

→ one cannot mistrust one of the elements of the network without contradicting many others.



A solid core surrounded by fuzzy borders

The reason the ‘perceived objectivity’ in physics is that, after centuries of experimentation, theoretical work and successful predictions, there is such a consistent network of beliefs, it has acquired the status of an objective construction:

$\longrightarrow$ one cannot mistrust one of the elements of the network without contradicting many others.

Around this solid core of objective knowledge there are fuzzy borders which correspond to areas of present investigations, where the level of intersubjectivity is still very low.

A solid core surrounded by fuzzy borders

The reason the ‘perceived objectivity’ in physics is that, after centuries of experimentation, theoretical work and successful predictions, there is such a consistent network of beliefs, it has acquired the status of an objective construction:

$\rightarrow$ one cannot mistrust one of the elements of the network without contradicting many others.

Around this solid core of objective knowledge there are fuzzy borders which correspond to areas of present investigations, where the level of intersubjectivity is still very low.

Nevertheless, when one proposes a new theory or model, one has to check immediately whether it contradicts some well-established beliefs.

A classical check of new models: “Does it influence $g-2$?”

Before speaking about objectivity

My preferred motto on this matter:

"no one should be allowed to speak about objectivity unless he/she has had 10–20 years working experience in frontier science, economics, or any other applied field"

Objectivity?

What is objective – I believe – is that the external world does exist.

But Science – that means "to know" – is entirely inside our brain, and their there is plenty of room for divergences.

Fortunately, when rational people, sharing the same scientific education have a lot of solid experimental information, they tend to reach an agreement, at least in the general aspects:

$$\text{subjectivity} \longrightarrow \text{inter-subjectivity } [= \text{objectivity}]$$

Ask practical questions and evaluate the probability in specific cases, instead of seeking refuge in abstract questions

$\rightarrow$ *Probability is objective as long as I am not asked to evaluate it*

An ‘objective’ evaluation of probability

Q. What is the probability that a molecule of nitrogen at room temperature has a velocity between 400 and 500 m/s?

A. Easy! Take the Maxwell distribution formula from a textbook, calculate an integral and get a number.

Q. I give you a vessel containing nitrogen and a detector capable of measuring the speed of a single molecule and you set up the apparatus (or you let a person you trust do it). Now, what is the probability that the first molecule that hits the detector has a velocity between 400 and 500 m/s?

A. Uhm...
$\longrightarrow$ study the problem carefully and perform preliminary measurements and checks (Where did I buy the gas? How am I sure about temperature? etc.).

What about probabilistic laws of physics?

Quantum mechanics? Hidden quantities? Statistical mechanic?

Personally very pragmatical, not engaged approach, but there are around people claiming that subjective probability clarifies QM interpretation (see e.g. Christopher A. Fuchs in quant-ph, and cited work).

- Probability deals with probability that an event may happen, given a certain state of information
- It does not matter if the fundamental laws are 'intrinsically probabilistic' or probability is just due to our ignorance.
- Extending Hume's statement:
  "Though there be no such thing as Chance in the world; our ignorance of the real cause of any event has the same influence on the understanding, and begets a like species of belief or opinion" $\rightarrow$ "Even if there were ..."
- If $P(E_1) > P(E_2)$, I believe E_1 more than E_2. That's all.

Physical probability?

If we pay attention, we see that it is just
‘a number you get from a model’.

It does not necessarily convey the confidence on the occurrence of a physical event E.

- In fact, it is correct to say $P(E \,|\, \text{Model}_{\boldsymbol{\theta}} \rightarrow p) = p$,

  but our confidence on E relies on our confidence on the model and on its parameters $\boldsymbol{\theta}$!

  If we are really interested in evaluating our confidence about the occurrence of E, we have to take into account of all models and the possible values of their parameters

  Anyhow, this ‘physical probability’ p can be easily incorporated in the probabilistic framework, including our uncertainty about it

$\Rightarrow$ Don’t worry: we lose nothing of what we really need!

Subjective $\neq$ arbitrary

Crucial role of the coherent bet

- You say that this coin has 70% to show head?
  No problem with me: you place 70€ on head, I 30€ on tail and who wins take 100€
  $\Rightarrow$ If OK with you, let's start.
- You say that this coin has 30% to show head?
  $\Rightarrow$ Just reverse the bet

  (Like sharing goods, e.g. a cake with a child)

$\Rightarrow$ Take into account all available information in the most "objective way"
(Even that someone has a different opinion!)

$\Rightarrow$ It might seem paradoxically, but the 'subjectivist' is much more 'objective' than those who blindly use so-called objective methods.

A good example of arbitrariness

What is really arbitrary is to **DEFINE** ‘confidence level’ the result of an *ad-hoc* prescription, especially when one is perfectly aware (experts are) that this number does not represent “how much one is confident of a given statement, in the sense of ‘how much one believes it’ ”

A good example of arbitrariness

What is really arbitrary is to **DEFINE** ‘confidence level’ the result of an *ad-hoc* prescription, especially when one is perfectly aware (experts are) that this number does not represent “how much one is confident of a given statement, in the sense of ‘how much one believes it’ ”

While “not experts”, i.e. the large majority of those who use those prescriptions, are influenced by the name and naively use the prescriptions to get an idea of how much they should be confident, i.e. to believe, something $\Rightarrow$ very unpleasant effects!

A good example of arbitrariness

What is really arbitrary is to **DEFINE** 'confidence level' the result of an *ad-hoc* prescription, especially when one is perfectly aware (experts are) that this number does not represent "how much one is confident of a given statement, in the sense of 'how much one believes it' "

While "not experts", i.e. the large majority of those who use those prescriptions, are influenced by the name and naively use the prescriptions to get an idea of how much they should be confident, i.e. to believe, something $\Rightarrow$ very unpleasant effects!

($\longrightarrow$ The little story of the baptized savage†)

With whom should I bet?

Coherent bet

- Operational, although hypothetical — not the only one
  - 'poisonous': "lethal <u>if</u> ingested"
  - Electric field: Force on a probing charge

With whom should I bet?

Coherent bet

- Operational, although hypothetical — not the only one
  - ‘poisonous’: “lethal if ingested”
  - Electric field: Force on a probing charge
- $\rightarrow$ Oblige people to make honest assessments
  - Given the result $x \pm \Delta x$, with $\Delta x = 1\,\sigma$ and Gaussian model

  $\rightarrow$ experimenter should be ready to place or accept a 2:1 bet on the true value inside the interval
  - If he/she feels hem/her-self ready black to place, but not to accept the bet: $\rightarrow$ incoherent

    $\rightarrow$ uncertainty overestimated

    $\rightarrow$ cheating the scientific community

Our re-starting point

- Probability means how much we believe something
- Probability values obey the following basic rules

$$
\begin{array}{ll}
1. & 0 \le P(A) \le 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \quad [\text{if } P(A \cap B) = \emptyset\,] \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

- All the rest by logic (and good sense)

Our re-starting point

- Probability means how much we believe something
- Probability values obey the following basic rules

$$
\begin{array}{ll}
1. & 0 \le P(A) \le 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \qquad [\text{if } P(A \cap B) = \emptyset\,] \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

- All the rest by logic (and good sense)

+ extension to continuity
+ some convenient quantities to summarize the uncertainty
+ some computational ‘tricks’ to overcome mathematical difficulties

Our re-starting point

- Probability means how much we believe something
- Probability values obey the following basic rules

$$
\begin{array}{ll}
1. & 0 \leq P(A) \leq 1 \\
2. & P(\Omega) = 1 \\
3. & P(A \cup B) = P(A) + P(B) \qquad [\text{if } P(A \cap B) = \emptyset\,] \\
4. & P(A \cap B) = P(A \,|\, B) \cdot P(B) = P(B \,|\, A) \cdot P(A)\,,
\end{array}
$$

- All the rest by logic (and good sense)

+ extension to continuity
+ some convenient quantities to summarize the uncertainty
+ some computational 'tricks' to overcome mathematical difficulties

$\longrightarrow$ And possibly remember the coherent bet!

Events and sets

Convenient event $\leftrightarrow$ set analogy:

		Symbol
event	set	E
certain	sample space	Ω
impossible	empty	$\emptyset$
implication	inclusion	$E_1 \subseteq E_2$
opposite (complementary)	complementary	$\overline{E} \quad (E \cup \overline{E} = \Omega)$
logical product	intersection	$E_1 \cap E_2$
logical sum	union	$E_1 \cup E_2$
incompatible	disjoint	$E_1 \cap E_2 = \emptyset$
complete class	finite partition	$\begin{cases} E_i \cap E_j = \emptyset \ \ \forall\, i \neq j \\ \cup_i E_i = \Omega \end{cases}$

Rules of probability

- $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ if $E_i \cap E_j = \emptyset \quad \forall i \neq j$
  (just an extension of the basic rule 3).
- $P(E) = 1 - P(\overline{E})$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (generalization of '3')
- $P(E) = P(E \cap H) + P(E \cap \overline{H})$

Rules of probability

- $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ if $E_i \cap E_j = \emptyset \quad \forall i \neq j$
(just an extension of the basic rule 3).
- $P(E) = 1 - P(\overline{E})$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (generalization of ‘3’)
- $P(E) = P(E \cap H) + P(E \cap \overline{H})$

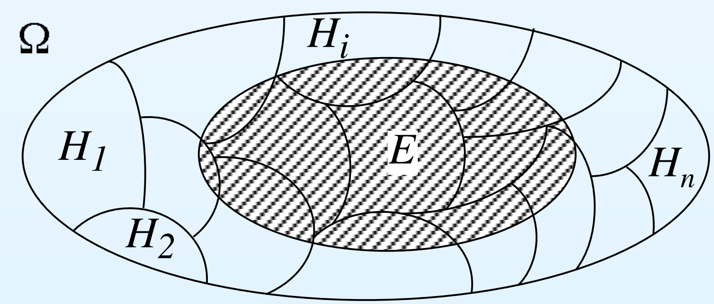
$\rightarrow$ Extension to complete class of events:

$$P(E) = P\left(\bigcup_{i=1}^{n} (E \cap H_i)\right) = \sum_{i=1}^{n} P(E \cap H_i)$$

and, applying ’4’

$$P(E) \;=\; \sum_i P(H_i) \cdot P(E \,|\, H_i)$$

(‘decomposition law’)



$E = \bigcup_{i=1}^{n} (E \cap H_i)$

$\rightarrow$ weighted average of $P(E \,|\, H_i)$

Rules of probability

- $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ if $E_i \cap E_j = \emptyset \quad \forall i \neq j$ (just an extension of the basic rule 3).
- $P(E) = 1 - P(\overline{E})$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (generalization of '3')
- $P(E) = P(E \cap H) + P(E \cap \overline{H})$

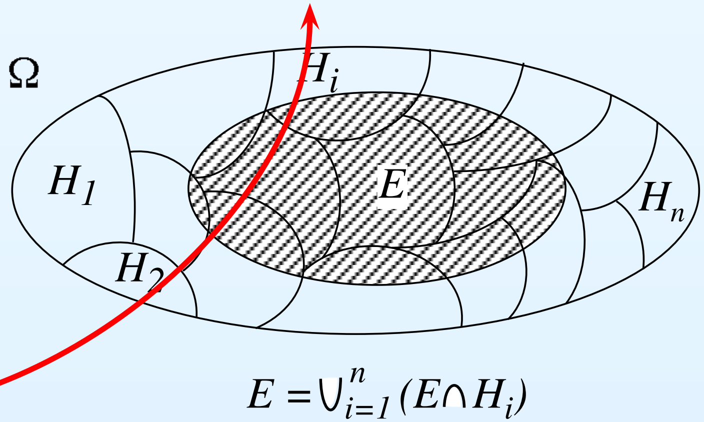
$\longrightarrow$ Extension to complete class of events:

$$P(E) = P\left(\bigcup_{i=1}^{n} (E \cap H_i)\right) = \sum_{i=1}^{n} P(E \cap H_i)$$

and, applying '4'

$$P(E) \; = \; \sum_i P(H_i) \cdot P(E \,|\, H_i)$$

('decomposition law')



$\rightarrow$ basis of 'marginalization'

Recovering the combinatorial evaluation formula

$$p = \frac{\#\,\text{favorable cases}}{\#\,\text{possible equiprobable cases}}$$

Given the ‘elementary’, equiprobable n events e_i forming a complete class, i.e. $\cup_i e_i = \Omega$,
we are interested in $P(E)$, where E = “$\cup$ m elementary events”

Recovering the combinatorial evaluation formula

$$p = \frac{\#\,\text{favorable cases}}{\#\,\text{possible equiprobable cases}}$$

Given the 'elementary', equiprobable n events e_i forming a complete class, i.e. $\cup_i e_i = \Omega$,
we are interested in $P(E)$, where E = "$\cup\, m$ elementary events"

$$\begin{aligned}
P(e_i) &= p_0 \\
P(\cup_i e_i) &= \sum_i P(e_i) = n\,p_0 = 1 \\
\rightarrow \quad p_0 &= \frac{1}{n} \\
\rightarrow \quad P(E) &= \sum_{e_i \subset E} P(e_i) = m\,p_0 = \frac{m}{n}
\end{aligned}$$

Independence

We remind that two events are called *independent* if

$$P(E \cap H) = P(E)\,P(H)\,.$$

This is equivalent to saying that

- $P(E \,|\, H) = P(E)$ and
- $P(H \,|\, E) = P(H)$,

i.e. the knowledge that one event has occurred does not change the probability of the other.

Independence

We remind that two events are called *independent* if

$$P(E \cap H) = P(E)\, P(H)\,.$$

This is equivalent to saying that

- $P(E \,|\, H) = P(E)$ and
- $P(H \,|\, E) = P(H)$,

i.e. the knowledge that one event has occurred does not change the probability of the other. If $P(E \,|\, H) \neq P(E)$, then the events E and H are *correlated*. In particular:

- if $P(E \,|\, H) > P(E)$ then E and H are *positively* correlated;
- if $P(E \,|\, H) < P(E)$ then E and H are *negatively* correlated.

Independence

We remind that two events are called *independent* if

$$P(E \cap H) = P(E)\, P(H)\,.$$

This is equivalent to saying that

- $P(E\,|\,H) = P(E)$ and
- $P(H\,|\,E) = P(H)$,

i.e. the knowledge that one event has occurred does not change the probability of the other. If $P(E\,|\,H) \neq P(E)$, then the events E and H are *correlated*. In particular:

- if $P(E\,|\,H) > P(E)$ then E and H are *positively* correlated;
- if $P(E\,|\,H) < P(E)$ then E and H are *negatively* correlated.

[By the way, $P(E \cap H\,|\,I) = P(E\,|\,I)\, P(H\,|\,I)$, and so on.]

Independence

We remind that two events are called *independent* if

$$P(E \cap H) = P(E)\, P(H)\,.$$

This is equivalent to saying that

- $P(E \,|\, H) = P(E)$ and
- $P(H \,|\, E) = P(H)$,

i.e. the knowledge that one event has occurred does not change the probability of the other. If $P(E \,|\, H) \neq P(E)$, then the events E and H are *correlated*. In particular:

- if $P(E \,|\, H) > P(E)$ then E and H are *positively* correlated;
- if $P(E \,|\, H) < P(E)$ then E and H are *negatively* correlated.

[By the way, $P(E \cap H \,|\, I) = P(E \,|\, I)\, P(H \,|\, I)$, and so on.]
$\rightarrow$ comment on definition Vs use of 'independence'

Uncertain numbers

We are often uncertain in numbers and, consistently, we quantify of belief with probability.

Uncertain number is the more general term for random variable, though the adjective random is more committing, since it rely on the concept of randomness (see von Mises).

Nevertheless, I often use the name ‘random variable’, just to mean ’uncertain number’,

Uncertain numbers

We are often uncertain in numbers and, consistently, we quantify of belief with probability.

Uncertain number is the more general term for random variable, though the adjective random is more committing, since it rely on the concept of randomness (see von Mises).

Nevertheless, I often use the name ‘random variable’, just to mean ’uncertain number’, i.e.

> *A number respect to which we are in condition of uncertainty*

Uncertain numbers

We are often uncertain in numbers and, consistently, we quantify of belief with probability.

Uncertain number is the more general term for random variable, though the adjective random is more committing, since it rely on the concept of randomness (see von Mises).

Nevertheless, I often use the name ‘random variable’, just to mean ’uncertain number’, i.e.

A number respect to which we are in condition of uncertainty

- The first number rolling a die
- The temperature at the Geneva airport tomorrow at 7:00 am
- The integrated luminosity provided by LHC in 2008
- The number of signatures of the first LHC physics paper

Uncertain numbers

We are often uncertain in numbers and, consistently, we quantify of belief with probability.

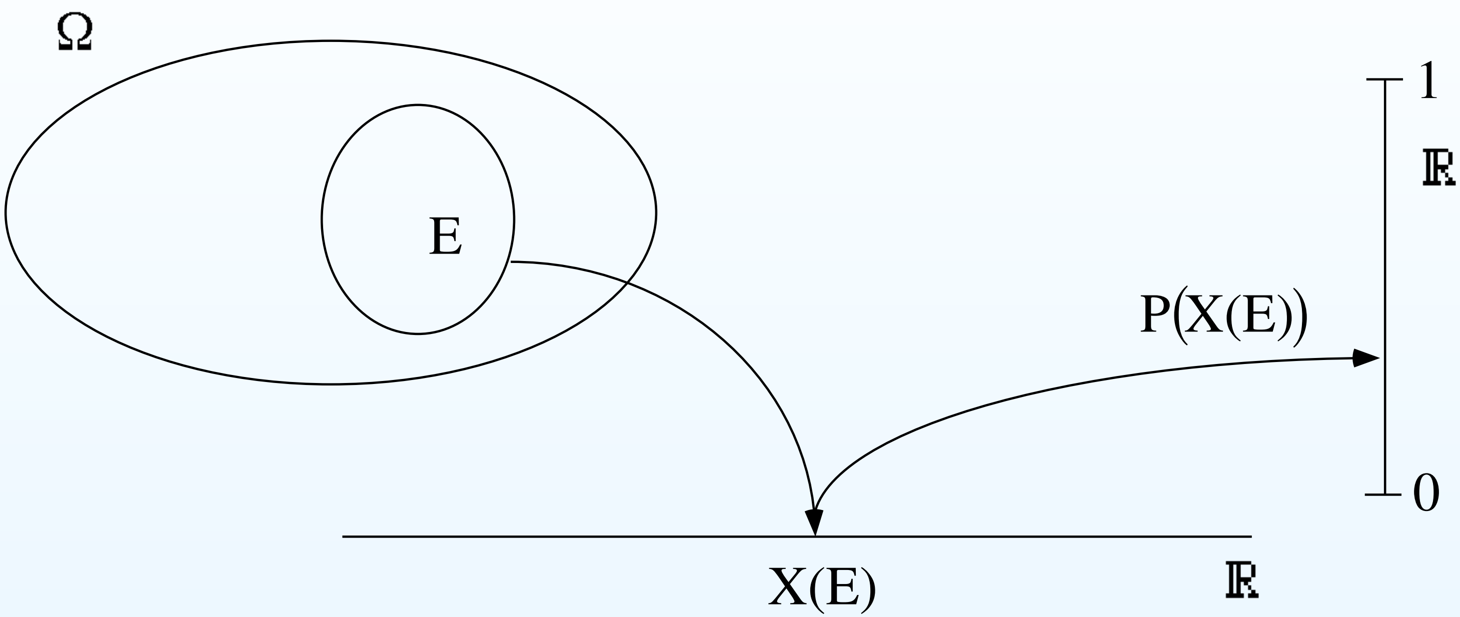
Uncertain number is the more general term for random variable, though the adjective random is more committing, since it rely on the concept of randomness (see von Mises).

Nevertheless, I often use the name ‘random variable’, just to mean ’uncertain number’, i.e.

A number respect to which we are in condition of uncertainty

- No need that the numbers can be framed in a von Mises' *collective*
- But it must be a well defined number (any uncertainty on its definition will increase our uncertainty about it)

From events to uncertain numbers

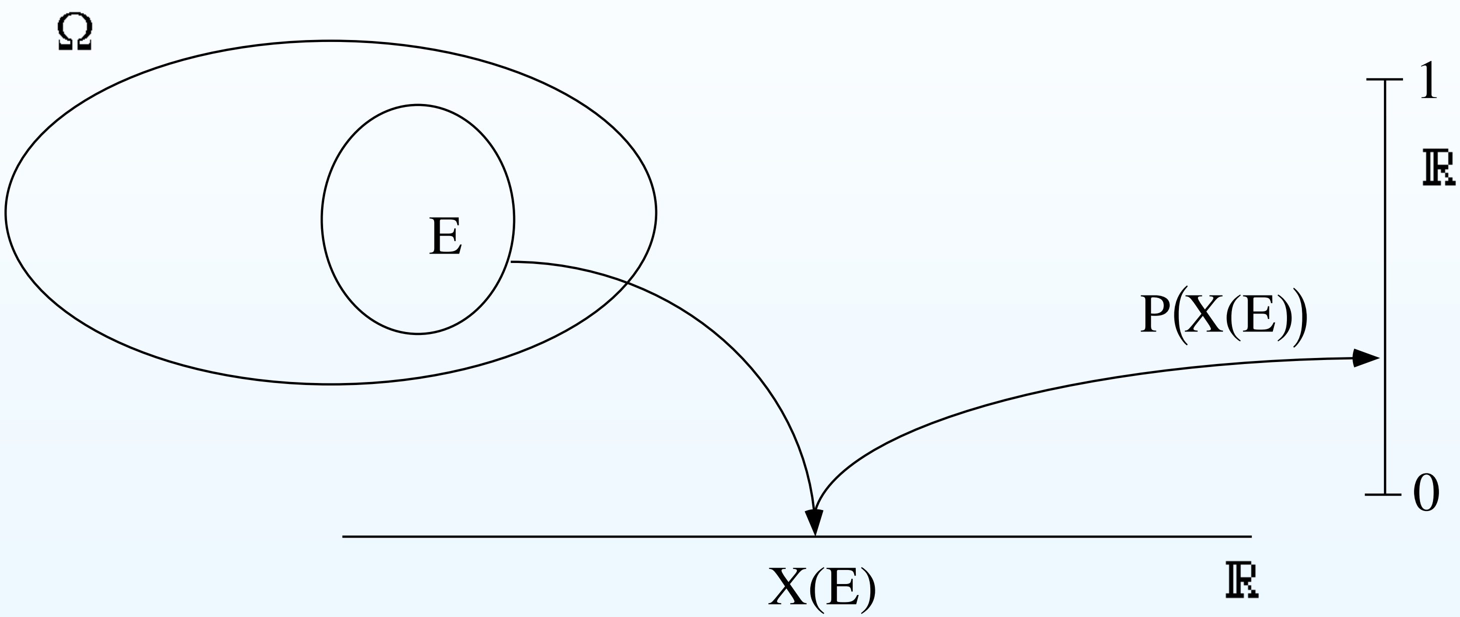


Uncertain numbers are associated to events

- Rolling one die: $X = 4 \leftrightarrow$ 'face marked with 4'
  (note: no intrinsic order in the numbers associated a die)

$\longrightarrow$ $P(X = 4) = P($'face marked with 4'$)$

From events to uncertain numbers



Uncertain numbers are associated to events

Event $\longrightarrow$ number: univocal, but not bi-univocal

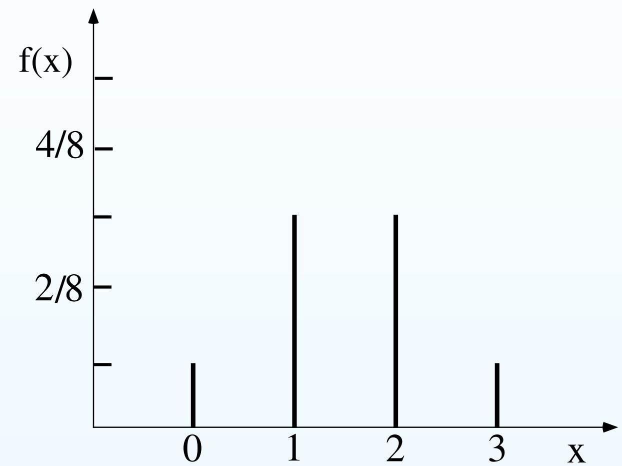
- Rolling two dice, with X 'sum of results'

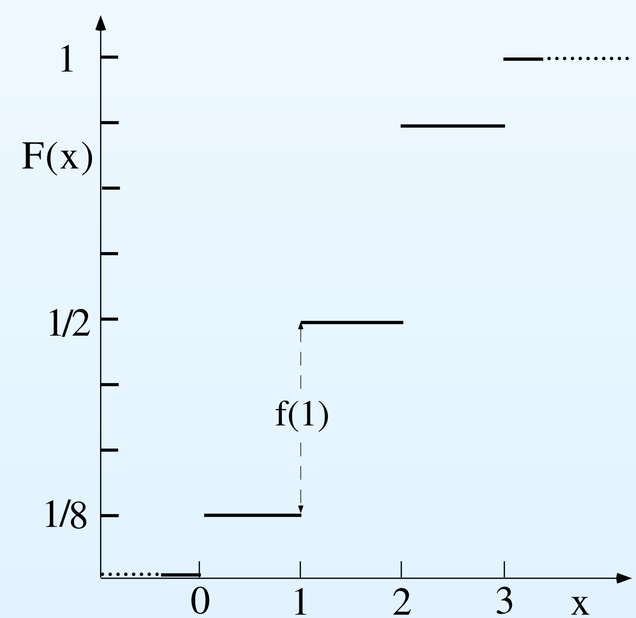
$\longrightarrow$ $P(X = 4) = \sum P($'events giving 4'$)$

Probability function (discrete numbers)

To each possible value of X we associate a degree of belief:

$$f(x) = P(X = x)\,.$$





Probability function (discrete numbers)

To each possible value of X we associate a degree of belief:

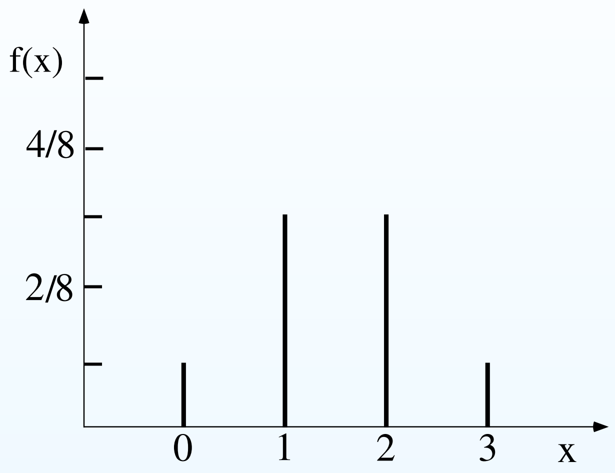
$$f(x) = P(X = x)\,.$$

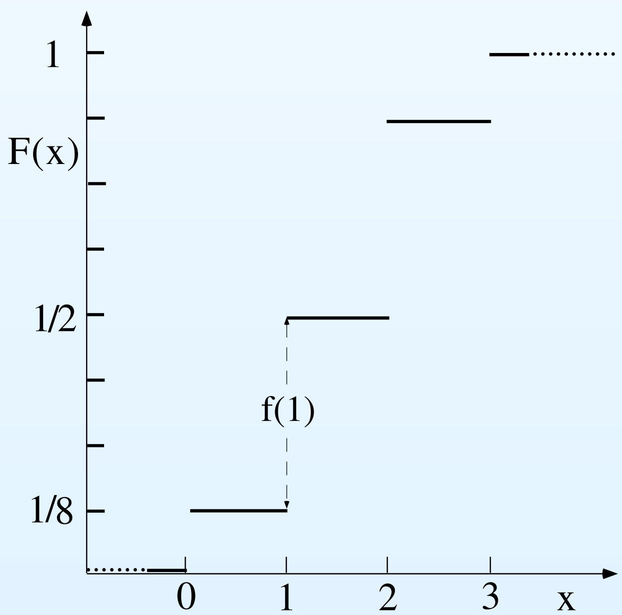
$f(x)$, being a probability, must satisfy the following properties:

$$0 \le f(x_i) \le 1\,,$$

$$P(X = x_i \cup X = x_j) = f(x_i) + f(x_j)\,,$$

$$\sum_i f(x_i) = 1\,.$$





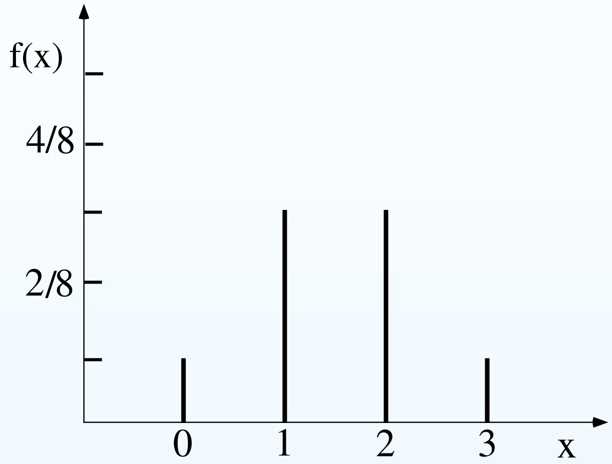
Probability function (discrete numbers)

To each possible value of X we associate a degree of belief:

$$f(x) = P(X = x)\,.$$

$f(x)$, being a probability, must satisfy the following properties:

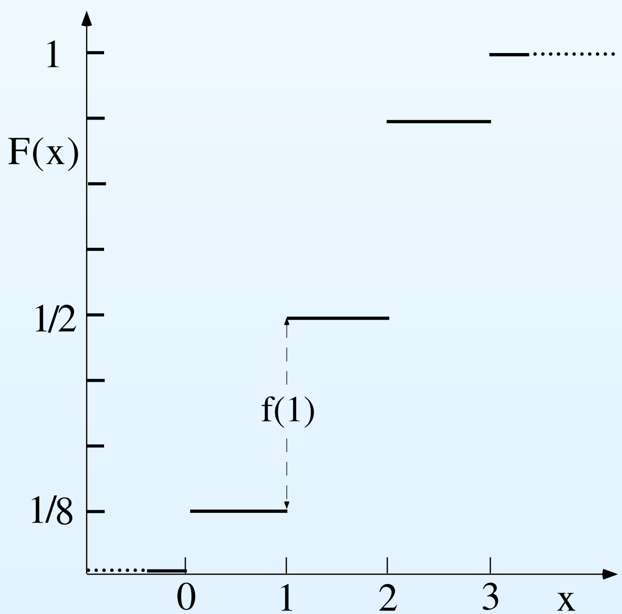
$$0 \leq f(x_i) \leq 1\,,$$
$$P(X = x_i \cup X = x_j) = f(x_i) + f(x_j)\,,$$
$$\sum_i f(x_i) = 1\,.$$



Cumulative function (defined for all x)

$$F(x_k) \equiv P(X \leq x_k) = \sum_{x_i \leq x_k} f(x_i)\,.$$

$[\,F(-\infty) = 0;\ F(+\infty) = 1;$
$F(x_i) - F(x_{i-1}) = f(x_i);$
$\lim_{\epsilon \to 0} F(x+\epsilon) = F(x)\,]$



Some simple examples

- Discrete uniform, well known $\rightarrow f(x) = 1/n \quad (1 \leq X \leq n)$

Some simple examples

- Discrete uniform, well known $\rightarrow f(x) = 1/n \quad (1 \leq X \leq n)$
- Bernoulli process
  - $X : 0, 1$ (failure/success)

    $f(0) = 1 - p$

    $f(1) = p$
  - it seems of practical irrelevance,
    $\rightarrow$ but of primary importance

Some simple examples

- Discrete uniform, well known $\rightarrow f(x) = 1/n \quad (1 \leq X \leq n)$
- Bernoulli process
  - $X : 0, 1$ (failure/success)
    $f(0) = 1 - p$
    $f(1) = p$
  - it seems of practical irrelevance,
    $\rightarrow$ but of primary importance
- The drunk man problem
  - Eight keys
  - After each trial he 'loses memory'
  - We watch him and – cynically – bet on the attempt on which he will succeed:
  - $X = 1$, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ... ?

Some simple examples

- Discrete uniform, well known $\rightarrow f(x) = 1/n \quad (1 \leq X \leq n)$
- Bernoulli process
  - $X : 0, 1$ (failure/success)

    $f(0) = 1 - p$

    $f(1) = p$
  - it seems of practical irrelevance,
    $\rightarrow$ but of primary importance
- The drunk man problem
  - Eight keys
  - After each trial he 'loses memory'
  - We watch him and – cynically – bet on the attempt on which he will succeed:
  - $X = 1$, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ... ?

$\rightarrow$ On which number would you bet?

Propagating probability values

We cannot say any number at random,

“because All attempts are equally likely”

Propagating probability values

We cannot say any number at random,

"because All attempts are equally likely"

$\longrightarrow$ 'half true', i.e. wrong...

- what is constant is $P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, where $E_i \to X = i$.

Propagating probability values

We cannot say any number at random,

"because All attempts are equally likely"

$\rightarrow$ 'half true', i.e. wrong. . .

- what is constant is $P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, where $E_i \rightarrow X = i$.

$\Rightarrow$ Beliefs are framed in a network!

- Once we assess something, we are implicitly making an infinity of assessments concerning logically connected events!
- We only need to make them explicit, using logic (trivial in principle, though it can be sometimes hard)

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$f(1) \;=\; P(E_1) = p$$

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned} f(1) &= P(E_1) = p \\ f(2) &= P(E_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)\,p \end{aligned}$$

Building up $f(x)$ of the drunk man problem

$P(E_i\,|\,\overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned} f(1) &= P(E_1) = p \\ f(2) &= P(E_2\,|\,\overline{E}_1)\cdot P(\overline{E}_1) = (1-p)\,p \\ f(3) &= P(E_3\,|\,\overline{E}_1 \cap \overline{E}_1)\cdot P(\overline{E}_2\,|\,\overline{E}_1)\cdot P(\overline{E}_1) = (1-p)^2\,p \end{aligned}$$

Building up $f(x)$ of the drunk man problem

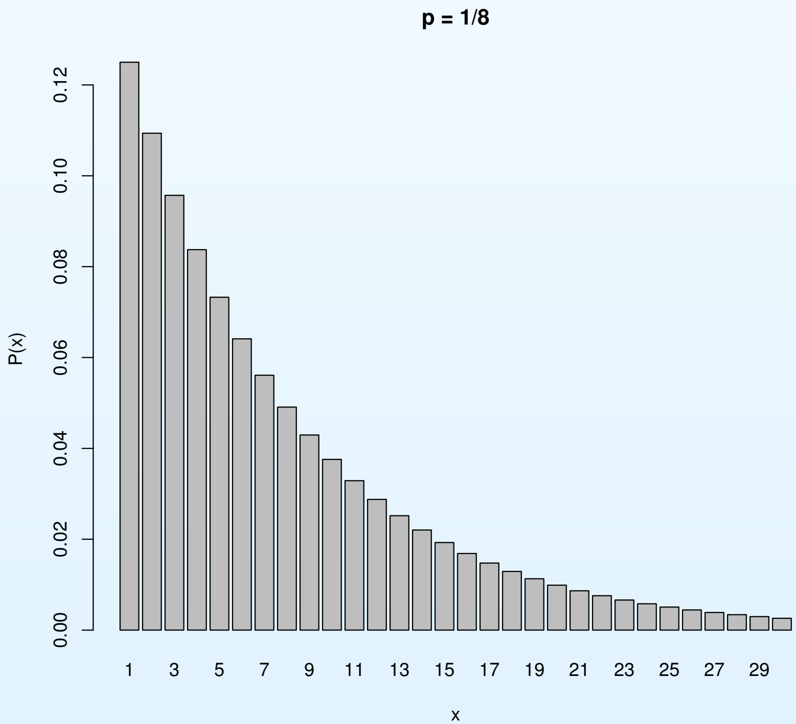
$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned}
f(1) &= P(E_1) = p \\
f(2) &= P(E_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)\,p \\
f(3) &= P(E_3 \,|\, \overline{E}_1 \cap \overline{E}_1) \cdot P(\overline{E}_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)^2\,p \\
f(x) &= p\,(1-p)^{x-1}
\end{aligned}$$

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned}
f(1) &= P(E_1) = p \\
f(2) &= P(E_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)\,p \\
f(3) &= P(E_3 \,|\, \overline{E}_1 \cap \overline{E}_1) \cdot P(\overline{E}_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)^2\,p \\
f(x) &= p\,(1-p)^{x-1}
\end{aligned}$$



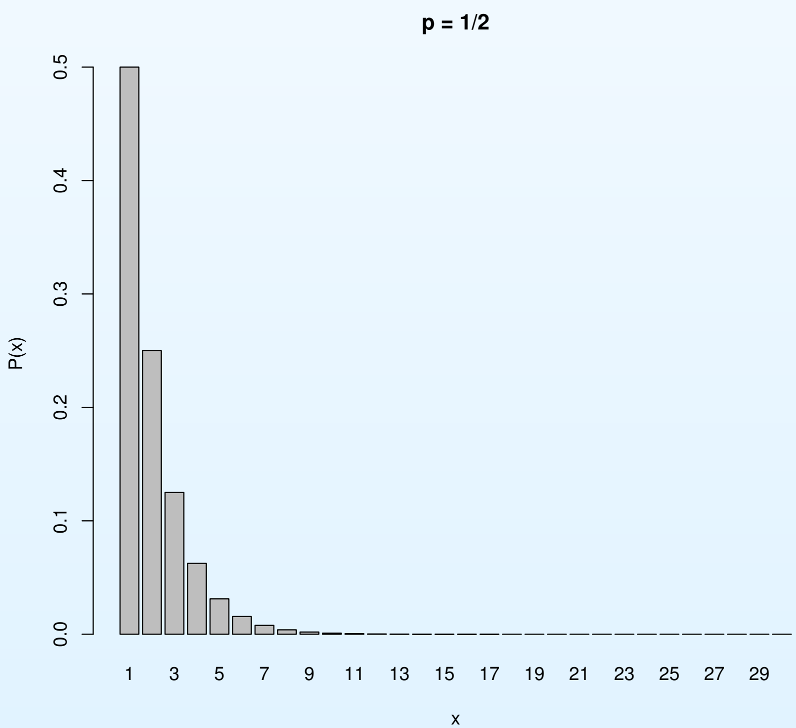
Beliefs decrease geometrically

$\Rightarrow$ Geometric distribution

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned}
f(1) &= P(E_1) = p \\
f(2) &= P(E_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)\,p \\
f(3) &= P(E_3 \,|\, \overline{E}_1 \cap \overline{E}_1) \cdot P(\overline{E}_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)^2\,p \\
f(x) &= p\,(1-p)^{x-1}
\end{aligned}$$



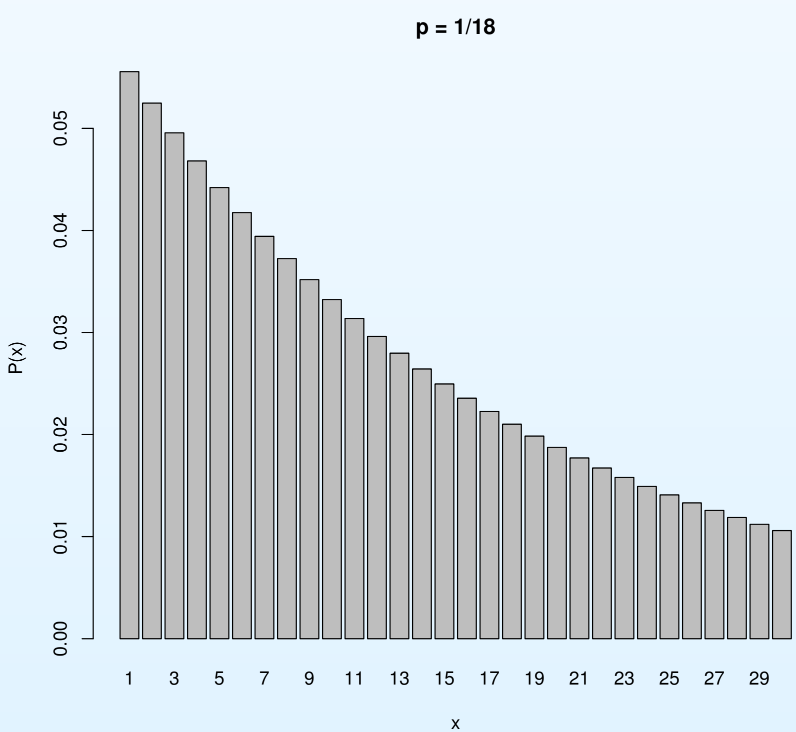
$p = 1/2 \rightarrow$ tossing a coin

Building up $f(x)$ of the drunk man problem

$P(E_i \,|\, \overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned}
f(1) &= P(E_1) = p \\
f(2) &= P(E_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)\,p \\
f(3) &= P(E_3 \,|\, \overline{E}_1 \cap \overline{E}_1) \cdot P(\overline{E}_2 \,|\, \overline{E}_1) \cdot P(\overline{E}_1) = (1-p)^2\,p \\
f(x) &= p\,(1-p)^{x-1}
\end{aligned}$$

$p = 1/18 \rightarrow$ a particular number at the Italian lotto ($p = 5/90$)

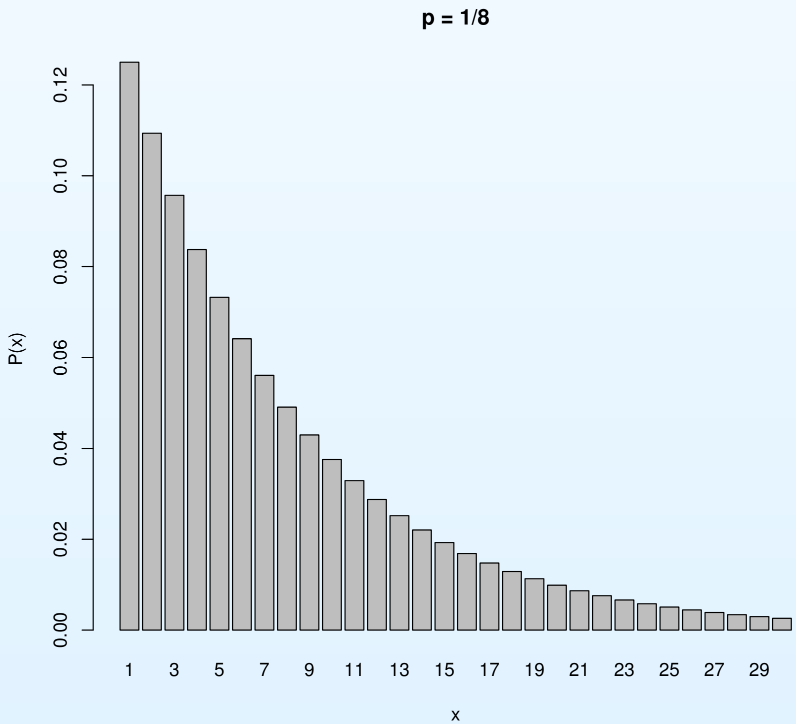


Building up $f(x)$ of the drunk man problem

$P(E_i\,|\,\overline{\bigcup_{j<i} E_j}) = p$, with $p = 1/8$:

$$\begin{aligned}
f(1) &= P(E_1) = p \\
f(2) &= P(E_2\,|\,\overline{E}_1)\cdot P(\overline{E}_1) = (1-p)\,p \\
f(3) &= P(E_3\,|\,\overline{E}_1 \cap \overline{E}_1)\cdot P(\overline{E}_2\,|\,\overline{E}_1)\cdot P(\overline{E}_1) = (1-p)^2\,p \\
f(x) &= p\,(1-p)^{x-1}
\end{aligned}$$

Most probable value does not depend on p.
Not a suitable indicator to state our expectation
The same is true for the range of possibilities: $X : 1, 2, \ldots, \infty$



Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic ‘expectation’ and its uncertainty:

Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic ‘expectation’ and its uncertainty:

$$\mathsf{E}[X] \; = \; \sum_x x\, f(x)$$

Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic ‘expectation’ and its uncertainty:

$$\begin{aligned} \mathsf{E}[X] &= \sum_x x\, f(x) \\ \mathsf{Variance}(X) &= \sum_x (x - \mathsf{E}[X])^2\, f(x) \quad \longrightarrow \sigma^2(X) \rightarrow \sigma = \sqrt{\sigma} \end{aligned}$$

Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic ‘expectation’ and its uncertainty:

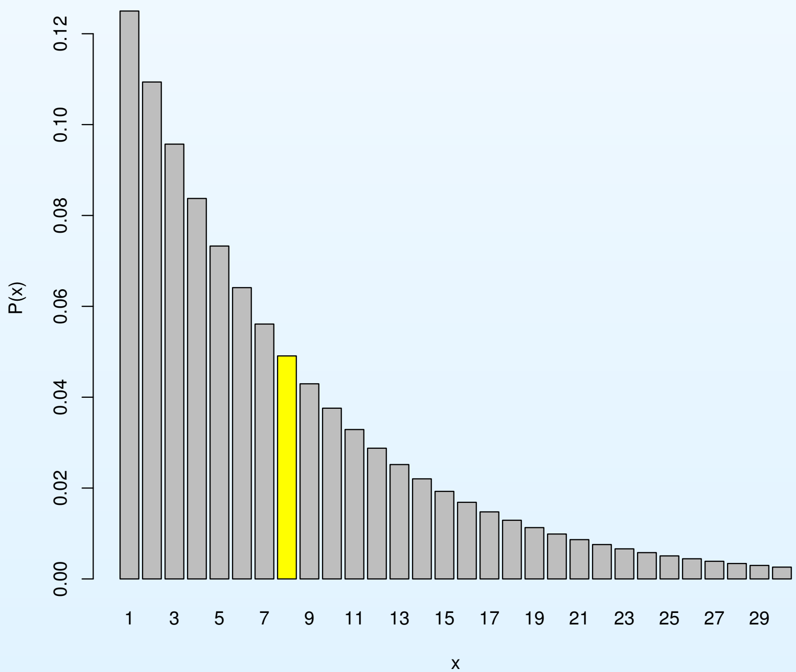
$$\mathsf{E}[X] = \sum_x x\, f(x)$$

$$\text{Variance}(X) = \sum_x (x - \mathsf{E}[X])^2 f(x) \longrightarrow \sigma^2(X) \rightarrow \sigma = \sqrt{\sigma}$$

$\mathsf{E}[X] = 1/p$
$\sigma(X) = \sqrt{1-p}/p$

$p = 1/8$:

$$\mathsf{E}[X] = 8$$
$$\sigma(X) = 7.5$$



Prevision and prevision uncertainty

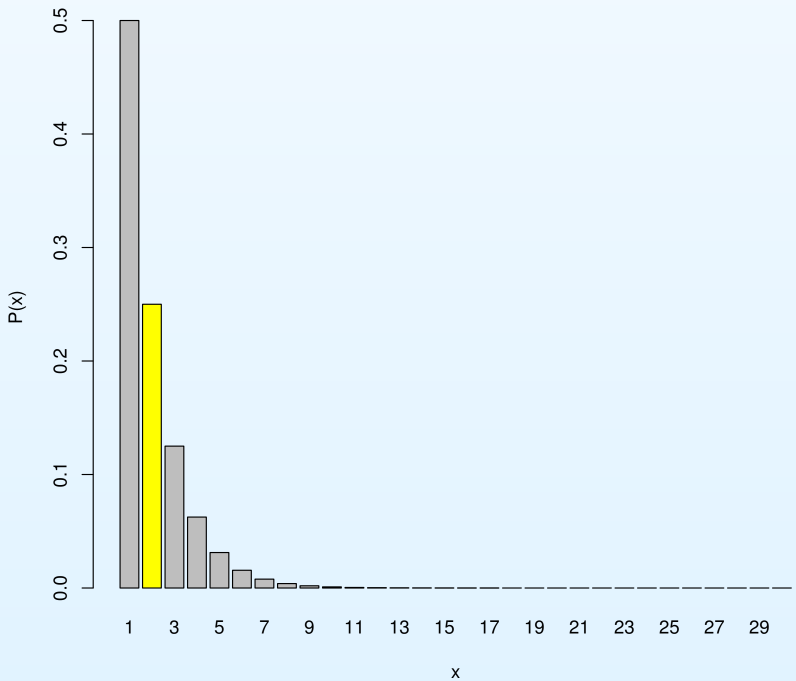
More suitable quantity two summarize in two numbers the our probabilistic 'expectation' and its uncertainty:

$$\begin{aligned} \mathsf{E}[X] &= \sum_x x\, f(x) \\ \mathsf{Variance}(X) &= \sum_x (x - \mathsf{E}[X])^2\, f(x) \quad \longrightarrow \sigma^2(X) \rightarrow \sigma = \sqrt{\sigma} \end{aligned}$$

$\mathsf{E}[X] = 1/p$
$\sigma(X) = \sqrt{1-p}/p$

$p = 1/2$:

$$\begin{aligned} \mathsf{E}[X] &= 2 \\ \sigma(X) &= 1.4 \end{aligned}$$



Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic ‘expectation’ and its uncertainty:

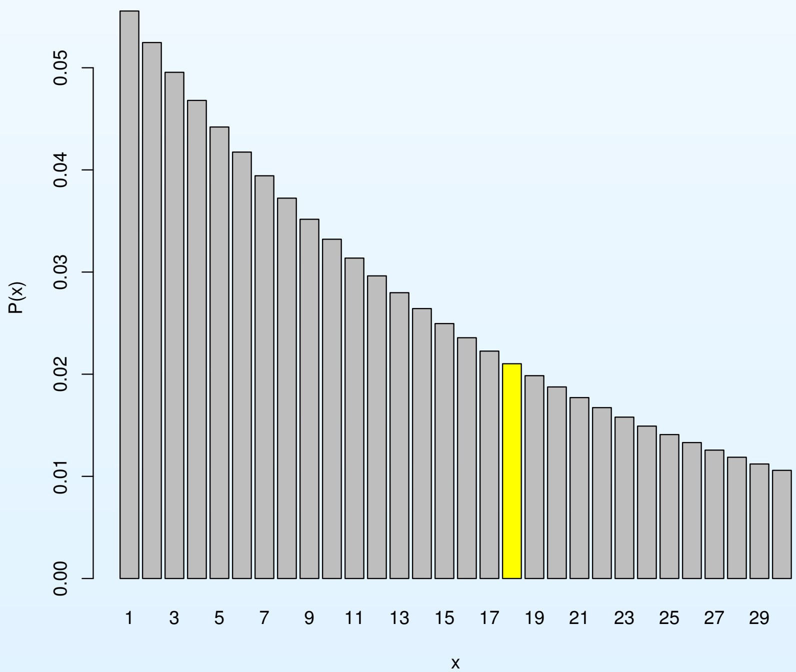
$$\mathrm{E}[X] = \sum_x x\, f(x)$$

$$\mathrm{Variance}(X) = \sum_x (x - \mathrm{E}[X])^2\, f(x) \longrightarrow \sigma^2(X) \rightarrow \sigma = \sqrt{\sigma}$$

$$\mathrm{E}[X] = 1/p$$
$$\sigma(X) = \sqrt{1-p}/p$$

$p = 1/18$:

$$\mathrm{E}[X] = 18$$
$$\sigma(X) = 17.5$$



Prevision and prevision uncertainty

More suitable quantity two summarize in two numbers the our probabilistic 'expectation' and its uncertainty:

$$\mathsf{E}[X] = \sum_x x\, f(x)$$

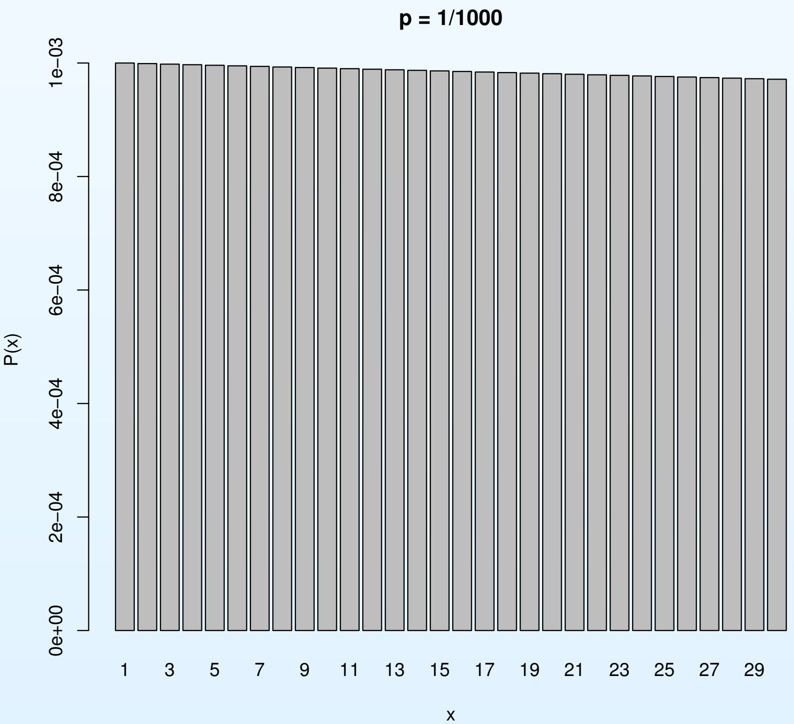
$$\text{Variance}(X) = \sum_x (x - \mathsf{E}[X])^2\, f(x) \;\longrightarrow \sigma^2(X) \rightarrow \sigma = \sqrt{\sigma}$$

$\mathsf{E}[X] = 1/p$

$\sigma(X) = \sqrt{1-p}/p \xrightarrow[p\rightarrow 0]{} 1/p$

→ rare events might happen at any moment!

(Though they have 'zero' probability to happen at any given moment!)



Expected value and ‘standard uncertainty’

The detail on the uncertainty is provided by $f(x)$.

- E$[X]$ and $\sigma(X)$ are just convenient summaries.
- In the general case they do not convey a precise confidence that X will occur in the range E$[X] \pm \sigma(X)$, though this probability is rather ‘high’ for typical $f(x)$ of interest.

Expected value and 'standard uncertainty'

The detail on the uncertainty is provided by $f(x)$.

- E$[X]$ and $\sigma(X)$ are just convenient summaries.
- In the general case they do not convey a precise confidence that X will occur in the range E$[X] \pm \sigma(X)$, though this probability is rather 'high' for typical $f(x)$ of interest.
- Another location summary (that statisticians like much) is given by the median, while the 'quantiles' provide (left open) intervals in which the variable is expected to fall with some probability (typically 10%, 20%, etc.).

Expected value and ‘standard uncertainty’

The detail on the uncertainty is provided by $f(x)$.

- E$[X]$ and $\sigma(X)$ are just convenient summaries.
- In the general case they do not convey a precise confidence that X will occur in the range E$[X] \pm \sigma(X)$, though this probability is rather ‘high’ for typical $f(x)$ of interest.
- Another location summary (that statisticians like much) is given by the median, while the ’quantiles’ provide (left open) intervals in which the variable is expected to fall with some probability (typically 10%, 20%, etc.).
- Anyway, it is important to prepared to $f(x)$ of any kind, because – fortunately! – nature is not boring. . .

Expected value and ‘standard uncertainty’

The detail on the uncertainty is provided by $f(x)$.

- E$[X]$ and $\sigma(X)$ are just convenient summaries.
- In the general case they do not convey a precise confidence that X will occur in the range E$[X] \pm \sigma(X)$, though this probability is rather ‘high’ for typical $f(x)$ of interest.
- Another location summary (that statisticians like much) is given by the median, while the ’quantiles’ provide (left open) intervals in which the variable is expected to fall with some probability (typically 10%, 20%, etc.).
- Anyway, it is important to prepared to $f(x)$ of any kind, because – fortunately! – nature is not boring. . .
- In particular, $f(x)$ might be asymmetric or, ‘multinomial’, i.e. with more than one local maximum.

Probability distributions and ‘statistical’ distributions

It is important to stress the difference between

- Probability distribution
  - To each possible outcome we associate how much we are confident on it:

$$x \longleftrightarrow f(x)$$

- Statistical distribution
  - To each observed outcome we associated its (relative) frequency

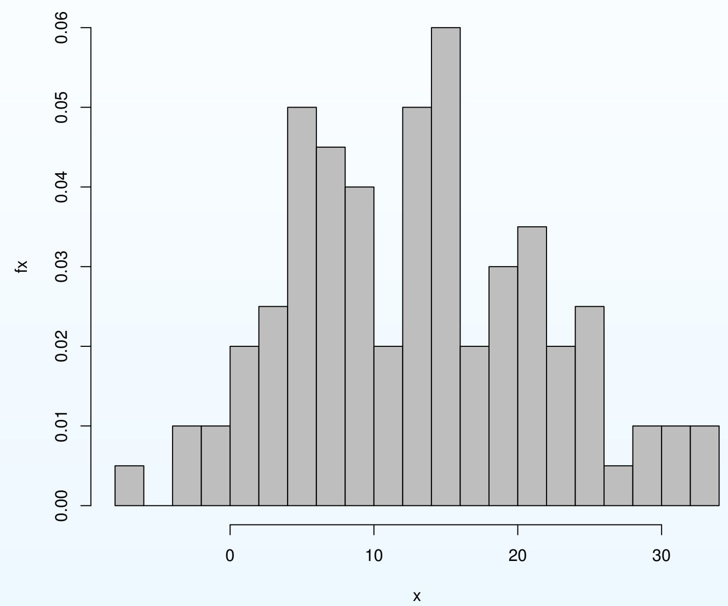
$$x \longleftrightarrow f_x$$

(e.g. an histogram of experimental observations)

Summaries (‘mean’, variance, ’σ’, ’skewness’, etc) have similar names and analogous definitions, but conceptual different meaning.

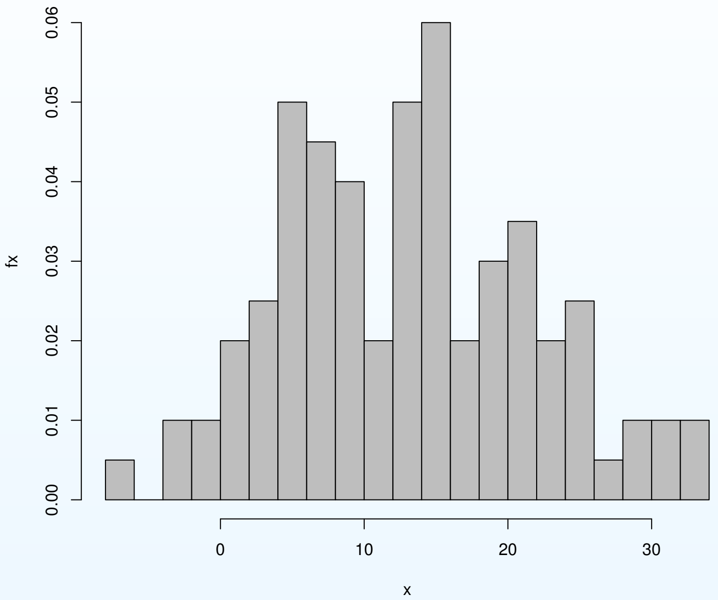
A histogram is not, usually, a probability distribution

In particular a histogram of experimental data <u>is not</u> a probability distribution (unless one reshuffles those events, and extracts one of them at random).



A histogram is not, usually, a probability distribution

In particular a histogram of experimental data <u>is not</u> a probability distribution (unless one reshuffles those events, and extracts one of them at random).



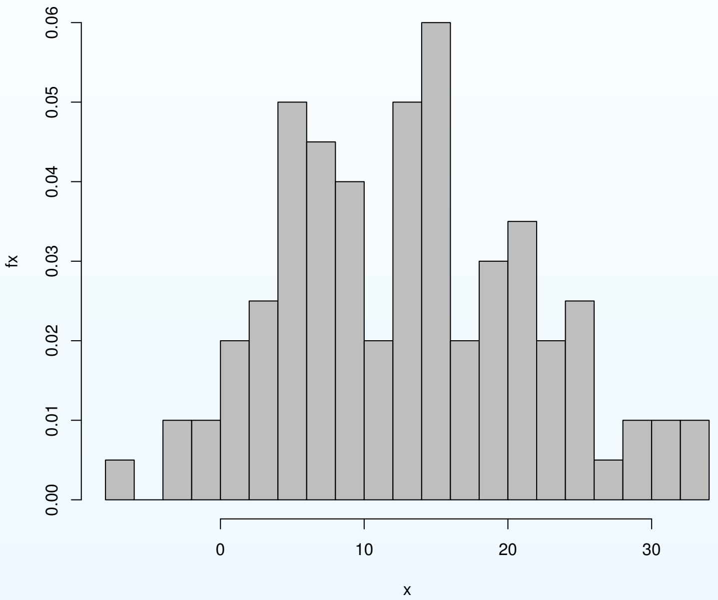
Average and variance

$$\overline{x} = \sum_x x\, f_x$$

$$\sigma^2 = \sum_x (x-\overline{x})^2\, f_x$$

A histogram is not, usually, a probability distribution

In particular a histogram of experimental data <u>is not</u> a probability distribution (unless one reshuffles those events, and extracts one of them at random).



Average and variance

$$\overline{x} = \sum_x x\, f_x$$

$$\sigma^2 = \sum_x (x - \overline{x})^2\, f_x$$

$\longrightarrow$ Just a rough empirical description of the shape
$\Rightarrow$ center of mass and momentum of inertia!
(Famous '$n/(n-1)$' correction: interference descriptive $\leftrightarrow$ inferential statistics.)

Distributions derived from the Bernoulli process

Bernoulli

Geometric

(trial of
1st success)

Pascal

(trial of
k-th success)

Binomial

(# of successes
in ind. n trials)

(Binomial well known. We shall not use the Pascal)

End of lecture

End of lecture 3

Notes

The following slides should be reached by hyper-links, clicking on words with the symbol †

Determinism/indeterminism

Pragmatically, as far as uncertainty and inference matter, it doesn't really matter.

"Though there be no such thing as Chance in the world; our ignorance of the real cause of any event has the same influence on the understanding, and begets a like species of belief or opinion" (Hume)

Go Back

Processo di Biscardi

A single quote gives an idea of the talk show:

"Please, don't speak more than two or three at the same time!"

Go Back

Hume's view about 'combinatoric evaluation'

"There is certainly a probability, which arises from a superiority of chances on any side; and according as this superiority increases, and surpasses the opposite chances, the probability receives a proportionable increase, and begets still a higher degree of belief or assent to that side, in which we discover the superiority."

Hume’s view about ‘combinatoric evaluation’

“There is certainly a probability, which arises from a superiority of chances on any side; and according as this superiority increases, and surpasses the opposite chances, the probability receives a proportionable increase, and begets still a higher degree of belief or assent to that side, in which we discover the superiority. If a dye were marked with one figure or number of spots on four sides, and with another figure or number of spots on the two remaining sides, it would be more probable, that the former would turn up than the latter; though, if it had a thousand sides marked in the same manner, and only one side different, the probability would be much higher, and our belief or expectation of the event more steady and secure.” (David Hume)

Go Back

Hume's view about 'frequency based evaluation'

> *"Being determined by custom to transfer the past to the future, in all our inferences; where the past has been entirely regular and uniform, we expect the event with the greatest assurance, and leave no room for any contrary supposition."*

Hume's view about 'frequency based evaluation'

"Being determined by custom to transfer the past to the future*, in all our inferences; where the past has been entirely regular and uniform, we expect the event with the greatest assurance, and leave no room for any contrary supposition. But where* different effects *have been found to* follow from causes*, which are to appearance exactly similar,* all these various effects must occur to the mind *in transferring the past to the future, and enter into our consideration, when we determine the probability of the event."*

Hume's view about 'frequency based evaluation'

"Being determined by custom to transfer the past to the future*, in all our inferences; where the past has been entirely regular and uniform, we expect the event with the greatest assurance, and leave no room for any contrary supposition. But where* different effects *have been found to* follow from causes*, which are to appearance exactly similar,* all these various effects must occur to the mind *in transferring the past to the future, and enter into our consideration, when we determine the probability of the event." Though we give the preference to that which has been found most usual, and believe that this effect will exist, we must not overlook the other effects, but must* assign to each of them a particular weight and authority, in proportion as we have found it to be more or less frequent*."* (David Hume)

Go Back

Bet odds to express confidence

“The best way to explain it is, I’ll bet you fifty to one that you don’t find anything”
(Feynman)

Bet odds to express confidence

"The best way to explain it is, I'll bet you fifty to one that you don't find anything" (Feynman)

"It is a bet of 11,000 to 1 that the error on this result (the mass of Saturn) *is not 1/100th of its value"* (Laplace)

Bet odds to express confidence

"The best way to explain it is, I'll bet you fifty to one that you don't find anything" (Feynman)

"It is a bet of 11,000 to 1 that the error on this result (the mass of Saturn) *is not 1/100th of its value"* (Laplace)

$\longrightarrow$ 99.99% confidence on the result

Bet odds to express confidence

“The best way to explain it is, I’ll bet you fifty to one that you don’t find anything” (Feynman)

“It is a bet of 11,000 to 1 that the error on this result (the mass of Saturn) *is not 1/100th of its value”* (Laplace)

$\longrightarrow$ 99.99% confidence on the result

$\Rightarrow$ Is a 95% C.L. upper/lower limit a ‘19 to 1 bet’?

Go Back

Batpizing prescriptions

A missionary converts the savage Agu-bu Bu-gu and while he baptizes him, pouring some water on his head, gives him a cristiam name: *"From now you no longer Agu-bu Bu-gu. Now you Antonio"*.

Then he explains him how to be a good christian to gain the Heaven, respecting the Ten Commandaments and the various precepts, including *"no meet on Friday"* [the story is a bit old].

After several weeks, a Friday the missionary finds the converted savage eating a lamb. *"Antonio, my friend, why are you not observing the Friday precept?"*. *"No problem father. I poured water over lamb's head and told: you no longer lamb, now you fish "*.

Go Back