Telling the Truth with Statistics
Lecture 2

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 1st lecture

- The main interest in ‘statistics’ of physicists is inference, i.e. how to learn from data

Summary of 1st lecture

- The main interest in ‘statistics’ of physicists is inference, i.e. how to learn from data
- but the ‘prescriptions’ of ‘conventional statistics’ are not satisfactory

Summary of 1st lecture

- The main interest in ‘statistics’ of physicists is inference, i.e. how to learn from data
- but the ‘prescriptions’ of ‘conventional statistics’ are not satisfactory
- They produce confusions in those who want to understand the sense of what they do

Summary of 1st lecture

- The main interest in ‘statistics’ of physicists is inference, i.e. how to learn from data
- but the ‘prescriptions’ of ‘conventional statistics’ are not satisfactory
- They produce confusions in those who want to understand the sense of what they do
- and, anyhow, they are responsible of severe errors in scientific judgments.

Summary of 1st lecture

- The main interest in ‘statistics’ of physicists is inference, i.e. how to learn from data
- but the ‘prescriptions’ of ‘conventional statistics’ are not satisfactory
- They produce confusions in those who want to understand the sense of what they do
- and, anyhow, they are responsible of severe errors in scientific judgments.
- Trying to start from the very beginning, and focusing on ‘hypotheses tests’, we have seen that

Summary of 1st lecture

- but the ‘prescriptions’ of ‘conventional statistics’ are not satisfactory
- They produce confusions in those who want to understand the sense of what they do
- and, anyhow, they are responsible of severe errors in scientific judgments.
- Trying to start from the very beginning, and focusing on ‘hypotheses tests’, we have seen that
- the very source of uncertainty is the uncertainty in the causal connections:

Summary of 1st lecture

- They produce confusions in those who want to understand the sense of what they do
- and, anyhow, they are responsible of severe errors in scientific judgments.
- Trying to start from the very beginning, and focusing on ‘hypotheses tests’, we have seen that
- the very source of uncertainty is the uncertainty in the causal connections:
- the standard statistical methods to treat the problem, can be seen as the practical attempt to implement the ideal of falsification;

Summary of 1st lecture

- and, anyhow, they are responsible of severe errors in scientific judgments.
- Trying to start from the very beginning, and focusing on ‘hypotheses tests’, we have seen that
- the very source of uncertainty is the uncertainty in the causal connections:
- the standard statistical methods to treat the problem, can be seen as the practical attempt to implement the ideal of falsification;
- falsificationism is a kind of extension of the ‘proof by contradiction’ to the natural science.

Summary of 1st lecture

- Trying to start from the very beginning, and focusing on ‘hypotheses tests’, we have seen that
- the very source of uncertainty is the uncertainty in the causal connections:
- the standard statistical methods to treat the problem, can be seen as the practical attempt to implement the ideal of falsification;
- falsificationism is a kind of extension of the ‘proof by contradiction’ to the natural science.
- But strict falsificationism is just naive,

Summary of 1st lecture

- the standard statistical methods to treat the problem, can be seen as the practical attempt to implement the ideal of falsification;
- falsificationism is a kind of extension of the ‘proof by contradiction’ to the natural science.
- But strict falsificationism is just naive,
- while its statistical implementations are logically flawed.

Summary of 1st lecture

- falsificationism is a kind of extension of the 'proof by contradiction' to the natural science.
- But strict falsificationism is just naive,
- while its statistical implementations are logically flawed.
- We ended with some examples from HEP that had quite some resonance in the past years, where fake claims of discoveries can be easily attributed to the universal inability of physicists to handle the probability inversion problem, *"the essential problem of the the experimental method"* (Poincaré)

Conflict natural thinking $\Leftrightarrow$ cultural superstructure

Why? 'Who' is responsible?

- Since beginning of '900 it is dominant an unnatural approach to probability, in contrast to that of the founding fathers (Poisson, Bernoulli, Bayes, Laplace, Gauss, . . .).

Conflict natural thinking $\Leftrightarrow$ cultural superstructure

Why? ‘Who’ is responsible?

- Since beginning of ’900 it is dominant an unnatural approach to probability, in contrast to that of the founding fathers (Poisson, Bernoulli, Bayes, Laplace, Gauss, . . .).
- In this, still dominant approach (frequentism) it is forbidden to speak about probability of hypotheses, probability of causes, probability of values of physical quantities, etc.

Conflict natural thinking $\Leftrightarrow$ cultural superstructure

Why? 'Who' is responsible?

- Since beginning of '900 it is dominant an unnatural approach to probability, in contrast to that of the founding fathers (Poisson, Bernoulli, Bayes, Laplace, Gauss, ...).
- In this, still dominant approach (frequentism) it is forbidden to speak about probability of hypotheses, probability of causes, probability of values of physical quantities, etc.
- The concept of probability of causes [*"The essential problem of the experimental method"* (Poincaré)] has been surrogated by the mechanism of hypothesis test and 'p-values'. (And of 'confidence intervals' in parametric inference)

Conflict natural thinking $\Leftrightarrow$ cultural superstructure

Why? ‘Who’ is responsible?

- Since beginning of ’900 it is dominant an unnatural approach to probability, in contrast to that of the founding fathers (Poisson, Bernoulli, Bayes, Laplace, Gauss, . . .).
- In this, still dominant approach (frequentism) it is forbidden to speak about probability of hypotheses, probability of causes, probability of values of physical quantities, etc.
- The concept of probability of causes [*“The essential problem of the experimental method”* (Poincaré)] has been surrogated by the mechanism of hypothesis test and ‘p-values’. (And of ‘confidence intervals’ in parametric inference)

$\Rightarrow$ BUT people think naturally in terms of probability of causes, and use p-values as if they were probabilities of null hypotheses.

Conflict natural thinking $\Leftrightarrow$ cultural superstructure

Why? ‘Who’ is responsible?

- Since beginning of ’900 it is dominant an unnatural approach to probability, in contrast to that of the founding fathers (Poisson, Bernoulli, Bayes, Laplace, Gauss, . . .).
- In this, still dominant approach (frequentism) it is forbidden to speak about probability of hypotheses, probability of causes, probability of values of physical quantities, etc.
- The concept of probability of causes [*“The essential problem of the experimental method”* (Poincaré)] has been surrogated by the mechanism of hypothesis test and ‘p-values’. (And of ‘confidence intervals’ in parametric inference)

$\Rightarrow$ BUT people think naturally in terms of probability of causes, and use p-values as if they were probabilities of null hypotheses. $\Rightarrow$ Terrible mistakes in judgment!

... indeed not a very solid superstructure

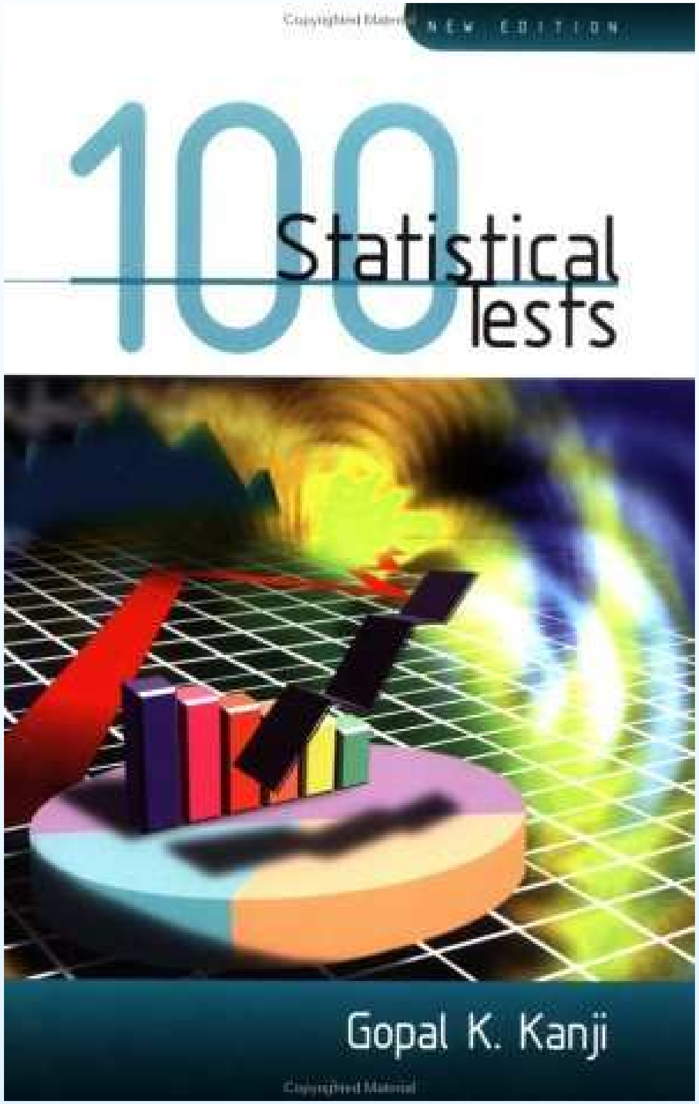
Moreover, a part the ‘philosophical’ problem of interpretation, there are plenty of ‘practical’ problems, since ‘statistical test’ are based on authority principle and not grounded on solid bases (probabilistic ‘first principles’).

. . . indeed not a very solid superstructure

Moreover, a part the ‘philosophical’ problem of interpretation, there are plenty of ‘practical’ problems, since ‘statistical test’ are based on authority principle and not grounded on solid bases (probabilistic ‘first principles’).

- Rich choice

. . . indeed not a very solid superstructure



Not exhaustive compilation. . .

. . . indeed not a very solid superstructure

Moreover, a part the 'philosophical' problem of interpretation, there are plenty of 'practical' problems, since 'statistical test' are based on authority principle and not grounded on solid bases (probabilistic 'first principles').

- Rich choice $\rightarrow$ > '100 tests'

. . . indeed not a very solid superstructure

Moreover, a part the 'philosophical' problem of interpretation, there are plenty of 'practical' problems, since 'statistical test' are based on authority principle and not grounded on solid bases (probabilistic 'first principles').

- Rich choice $\rightarrow$ > '100 tests'
- Discussions about which test to use it and how to use it are not deeper that discussions in pubs among soccer fans (Italians might think to the 'Processo di Biscardi' †)

. . . indeed not a very solid superstructure

Moreover, a part the 'philosophical' problem of interpretation, there are plenty of 'practical' problems, since 'statistical test' are based on authority principle and not grounded on solid bases (probabilistic 'first principles').

- Rich choice $\longrightarrow$ > '100 tests'
- Discussions about <u>which test</u> to use it and <u>how to use it</u> are not deeper that discussions in pubs among soccer fans (Italians might think to the 'Processo di Biscardi' †)

$\Rightarrow$ Tendency to look for the test that gives the result one wants

. . . indeed not a very solid superstructure

Moreover, a part the 'philosophical' problem of interpretation, there are plenty of 'practical' problems, since 'statistical test' are based on authority principle and not grounded on solid bases (probabilistic 'first principles').

- Rich choice $\rightarrow$ > '100 tests'
- Discussions about which test to use it and how to use it are not deeper that discussions in pubs among soccer fans (Italians might think to the 'Processo di Biscardi' †)

$\Rightarrow$ Tendency to look for the test that gives the result one wants

- My personal prejudice: *The fancier the name of the test is, the less believable the claim is*, because I am pretty sure that other, more common tests were discarded, because 'they did not work' .

. . . indeed not a very solid superstructure

Moreover, a part the ‘philosophical’ problem of interpretation, there are plenty of ‘practical’ problems, since ‘statistical test’ are based on authority principle and not grounded on solid bases (probabilistic ‘first principles’).

- Rich choice $\rightarrow$ > ‘100 tests’
- Discussions about which test to use it and how to use it are not deeper that discussions in pubs among soccer fans (Italians might think to the ‘Processo di Biscardi’ †)

$\Rightarrow$ Tendency to look for the test that gives the result one wants

- My personal prejudice: *The fancier the name of the test is, the less believable the claim is*, because I am pretty sure that other, more common tests were discarded, because ~~‘they did not work’~~ $\rightarrow$ ‘they did not support what the guy wanted the data to prove’

  $\chi^2 \rightarrow$ run-test $\rightarrow$ Kolmogorov $\rightarrow$ . . . ? . . . $\Rightarrow$ Lourdes .

Is statistics something serious?

Last, but not least, standard statistical methods, essentially a contradictory collection of *ad-hoc-eries*, induce scientists, and physicists in particular, to think that

> *'statistics' is something 'not scientific'*.

$\Rightarrow$ 'creative' behavior is encouraged

Is statistics something serious?

Last, but not least, standard statistical methods, essentially a contradictory collection of *ad-hoc-eries*, induce scientists, and physicists in particular, to think that

'statistics' is something 'not scientific'.

$\Rightarrow$ 'creative' behavior is encouraged

Last invention I have learned:

- imagine a χ^2 test where three models are compared to data
- $\nu = 40$; resulting χ^2 of 37.9, 49.1 and 52.4
- What would you say ('classically')?

Is statistics something serious?

Last, but not least, standard statistical methods, essentially a contradictory collection of *ad-hoc-eries*, induce scientists, and physicists in particular, to think that

'statistics' is something 'not scientific'.

$\Rightarrow$ 'creative' behavior is encouraged

Last invention I have learned:

- imagine a χ^2 test where three models are compared to data
- $\nu = 40$; resulting χ^2 of 37.9, 49.1 and 52.4
- What would you say ('classically')?

  $\Rightarrow$ Being $\mathrm{E}[\chi^2_{40}] = 40$ and $\sigma[\chi^2_{40}] = \sqrt{2 \times 40} \approx 9$:
  $\rightarrow$ none of the model gives us reasons to worry!

Is statistics something serious?

Last, but not least, standard statistical methods, essentially a contradictory collection of *ad-hoc-eries*, induce scientists, and physicists in particular, to think that

'statistics' is something 'not scientific'.

$\Rightarrow$ 'creative' behavior is encouraged

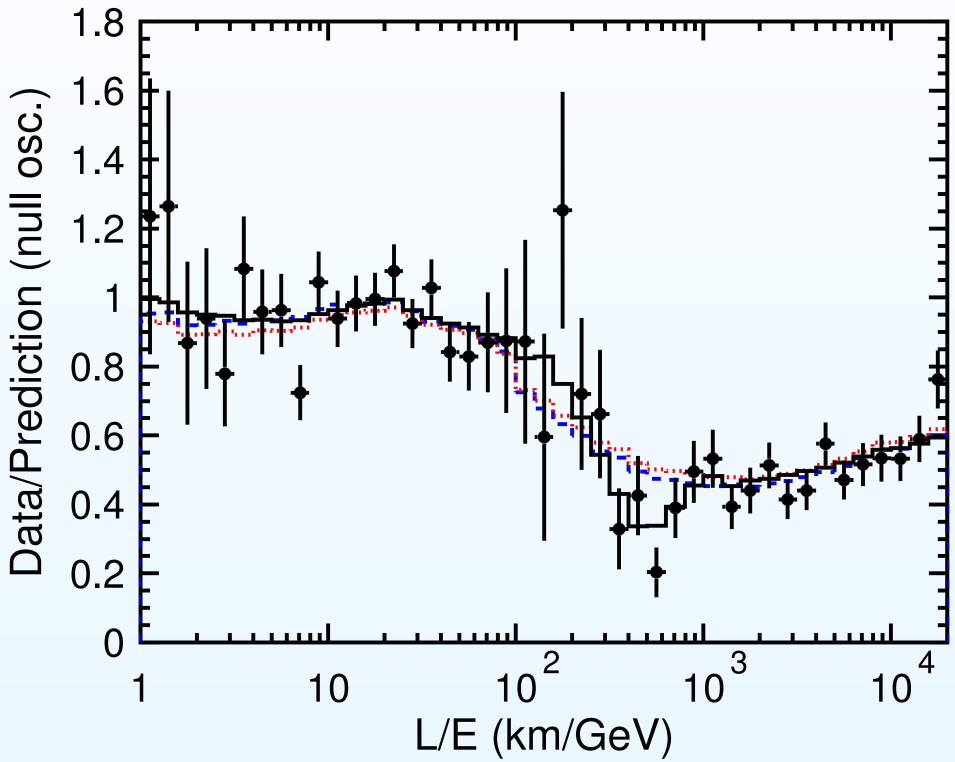
Last invention I have learned:

- imagine a χ^2 test where three models are compared to data
- $\nu = 40$; resulting χ^2 of 37.9, 49.1 and 52.4
- What would you say ('classically')?

  $\Rightarrow$ Being $\mathrm{E}[\chi^2_{40}] = 40$ and $\sigma[\chi^2_{40}] = \sqrt{2 \times 40} \approx 9$:
  $\rightarrow$ none of the model gives us reasons to worry!

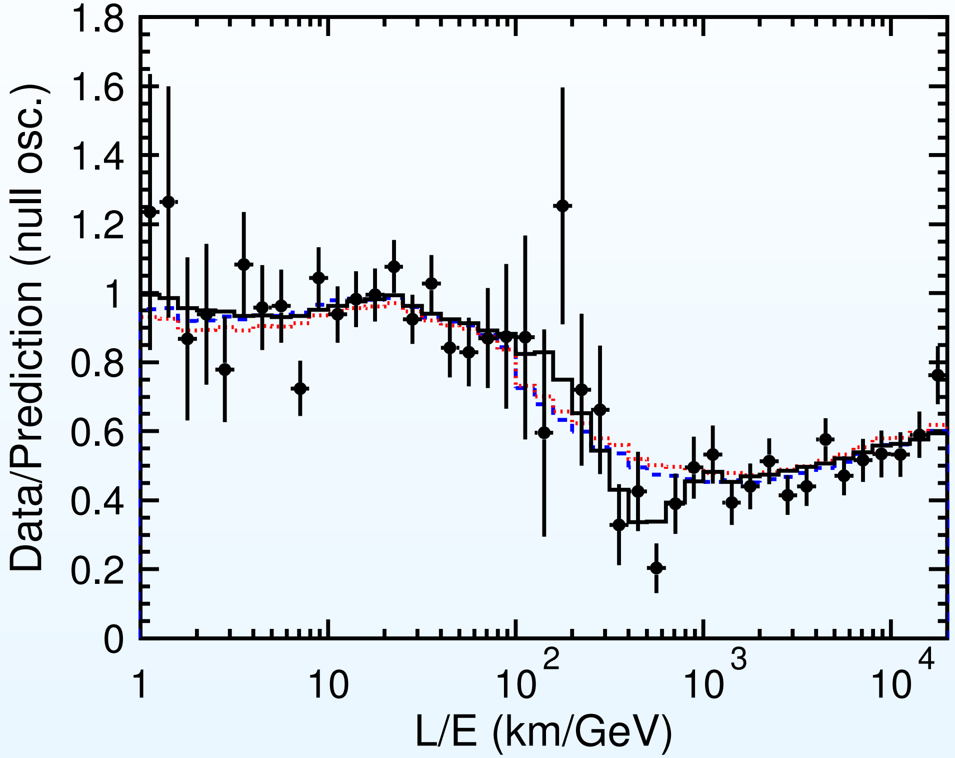
  $\Rightarrow$ p-values of 56%, 15% and 9.1%:
  $\rightarrow$ no model is below the customary p-value thresholds! (5%, 1%, or less.)

Super-Kamiokande, PRL 93 (1004) 101801-1



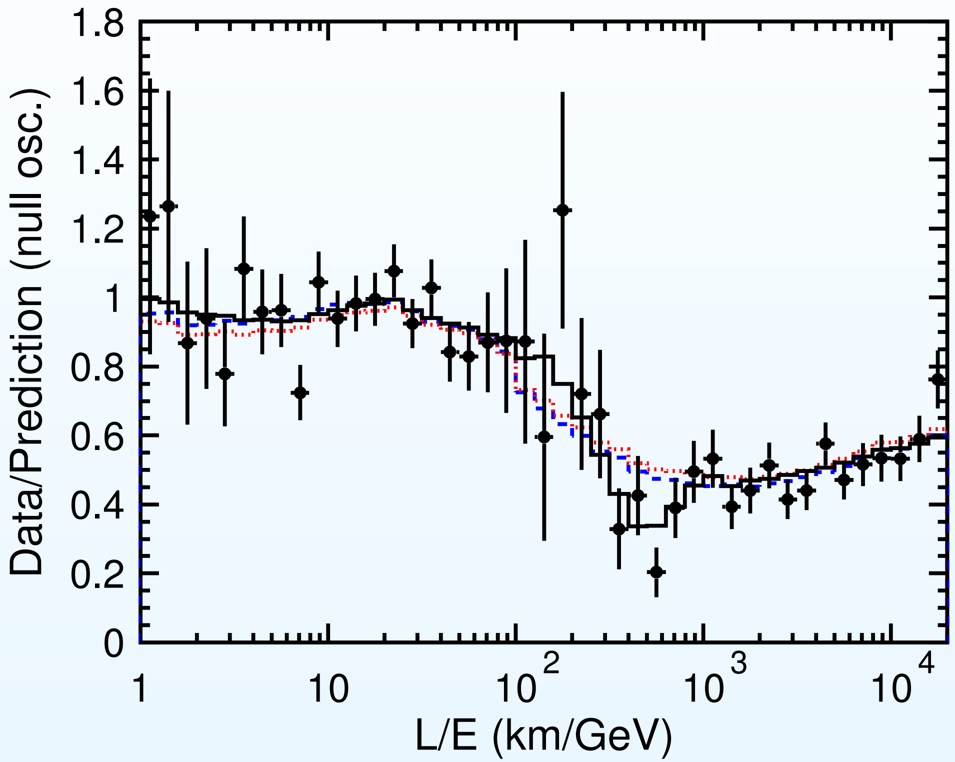
If you put <u>some attention</u>, you will realize that there are indeed **three** model expectations under the experimental points.

Super-Kamiokande, PRL 93 (1004) 101801-1



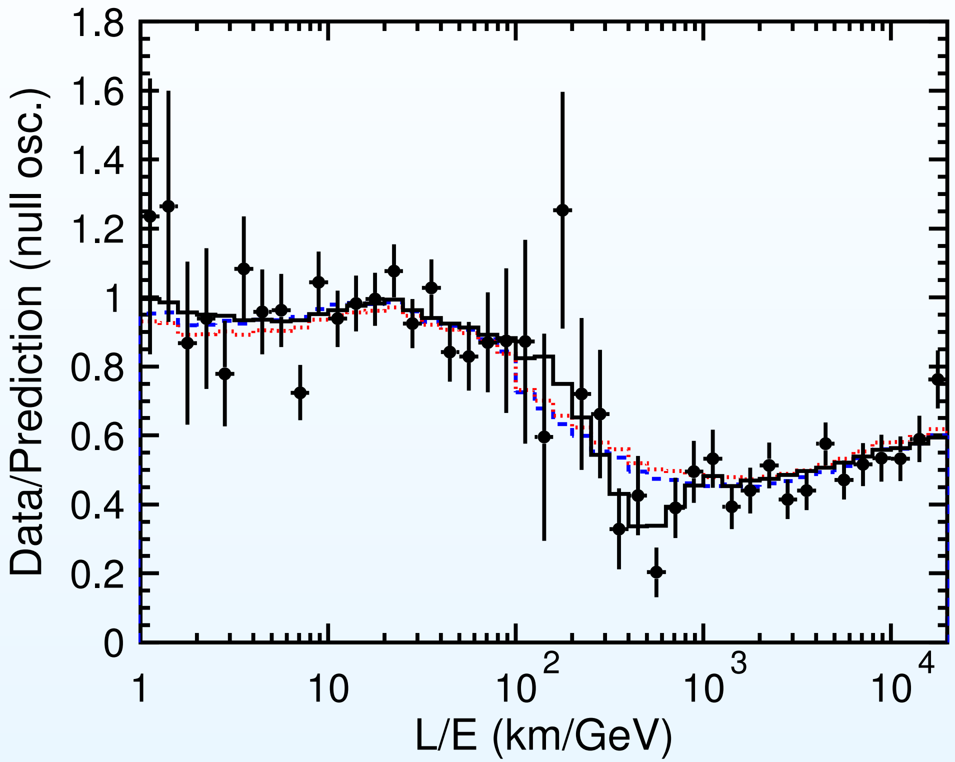
Believe it or not, the Collaboration claims the red and the blue models are excluded at 3.4 and 3.6 σ's.

Super-Kamiokande, PRL 93 (1004) 101801-1



Believe it or not, the Collaboration claims the red and the blue models are excluded at 3.4 and 3.6 σ's.
$\Rightarrow$ Personally, it seems to me that the 'excluded' two models are not believed much

Super-Kamiokande, PRL 93 (1004) 101801-1



Believe it or not, the Collaboration claims the red and the blue models are excluded at 3.4 and 3.6 σ's.

$\Rightarrow$ Personally, it seems to me that the 'excluded' two models are not believed much *a priori*: $\rightarrow$ nothing to do with the statistical numerology to get the 3.4 and 3.6 σ's!

Uncertainty: restart from scratch

Roll a die:

1, 2, 3, 4, 5, 6: ?

Toss a coin:

Head/Tail: ?

Having to perform a measurement:

Which numbers shall come out from our device ?

Having performed a measurement:

What have we learned about the value of the quantity of interest ?

Many other examples from real life:

Football, weather, tests/examinations, . . .

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

We are not uncertain in the same way about E_1, E_2 and E_3:

Rolling a die

Let us consider three outcomes:

$E_1 =$ ‘6’

$E_2 =$ ‘even number’

$E_3 =$ ‘≥ 2’

We are not uncertain in the same way about E_1, E_2 and E_3:

- Which event do you consider more likely, possible, credible, believable, plausible?

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

We are not uncertain in the same way about E_1, E_2 and E_3:

- Which event do you consider more likely, possible, credible, believable, plausible?
- You will get a price if the event you chose will occur. On which event would you bet?

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

We are not uncertain in the same way about E_1, E_2 and E_3:

- Which event do you consider more likely, possible, credible, believable, plausible?
- You will get a price if the event you chose will occur. On which event would you bet?
- On which event are you more confident? Which event you trust more, you believe more? etc

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

We are not uncertain in the same way about E_1, E_2 and E_3:

- Which event do you consider more likely, possible, credible, believable, plausible?
- You will get a price if the event you chose will occur. On which event would you bet?
- On which event are you more confident? Which event you trust more, you believe more? etc
- Imagine to repeat the experiment: which event do you expect to occur mostly? (More frequently)

Rolling a die

Let us consider three outcomes:

$$E_1 = \text{‘6’}$$

$$E_2 = \text{‘even number’}$$

$$E_3 = \text{‘}\geq 2\text{’}$$

$\Rightarrow$ Many expressions to state our preference

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

$\Rightarrow$ Many expressions to state our preference

Which reasoning have we applied to prefer E_3?

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

$\Rightarrow$ Many expressions to state our preference

Which reasoning have we applied to prefer E_3?
Can we use it for all other events of our interest?
($\rightarrow$ two envelop 'paradox')

Rolling a die

Let us consider three outcomes:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

$\Rightarrow$ Many expressions to state our preference

Which reasoning have we applied to prefer E_3?
Can we use it for all other events of our interest?
($\rightarrow$ two envelop 'paradox')

Indeed, using David Hume's words,† *"this process of the thought or reasoning may seem trivial and obvious"*

A counting experiment

Imagine a small scintillation counter, with suitable threshold, placed

here

now

Fix the measuring time (e.g. 5 second each) and perform 20 measurements: 0, 0, 1, 0, 0, 0, 1, 2, 0, 0, 1, 1, 0.

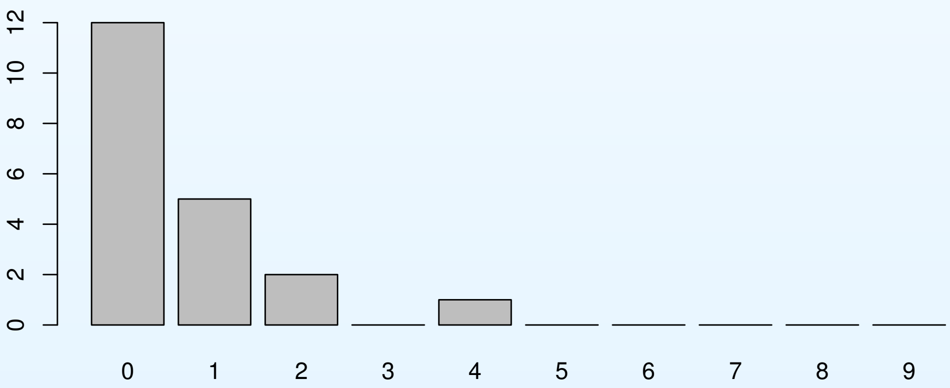
A counting experiment

Imagine a small scintillation counter, with suitable threshold, placed

here

now

Fix the measuring time (e.g. 5 second each) and perform 20 measurements: 0, 0, 1, 0, 0, 0, 1, 2, 0, 0, 1, 1, 0.



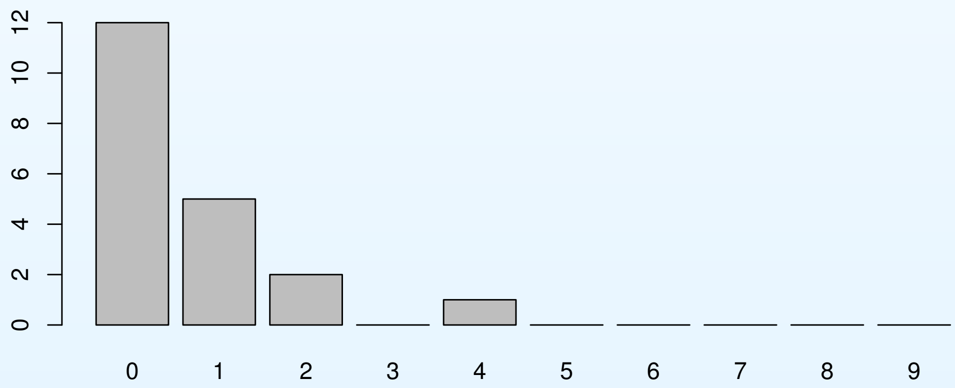
A counting experiment

Imagine a small scintillation counter, with suitable threshold, placed

here

now

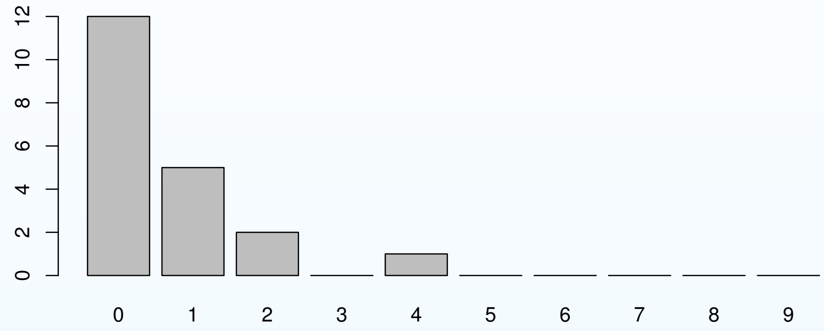
Fix the measuring time (e.g. 5 second each) and perform 20 measurements: 0, 0, 1, 0, 0, 0, 1, 2, 0, 0, 1, 1, 0.



Think at the 21^{st} measurement:

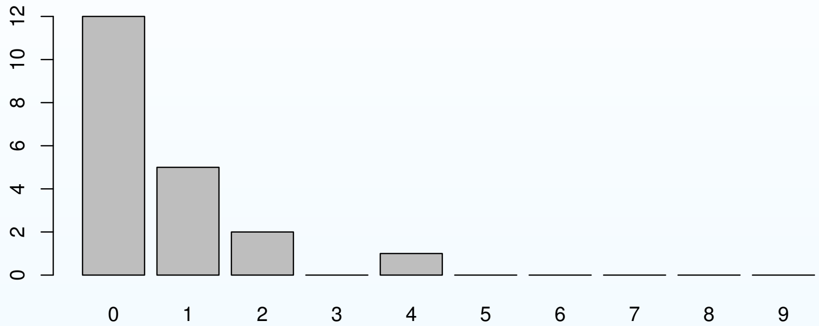
- Which outcome do you consider more likely? (0, 1, 2, 3, . . .)
- Why?

A counting experiment



$\Rightarrow$ Next ?

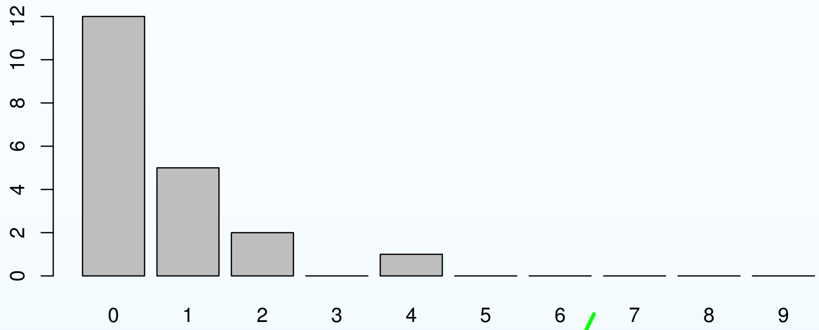
A counting experiment



$\Rightarrow$ Next ?

$P(0) > P(1) > P(2)$ $\checkmark$

A counting experiment

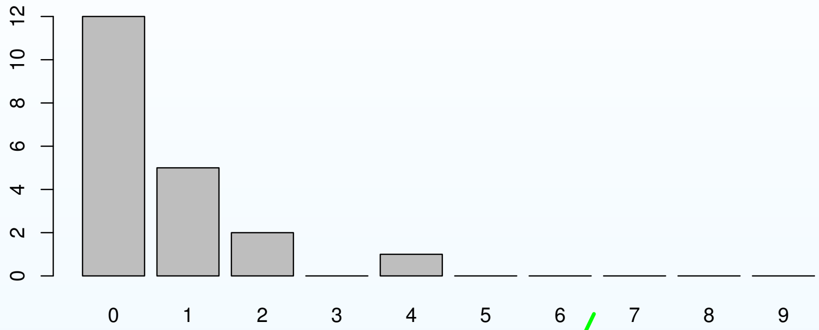


$P(0) > P(1) > P(2)$ $\surd$

$P(3) < P(4)$, or $P(3) \geq P(4)$?

$\Rightarrow$ Next ?

A counting experiment



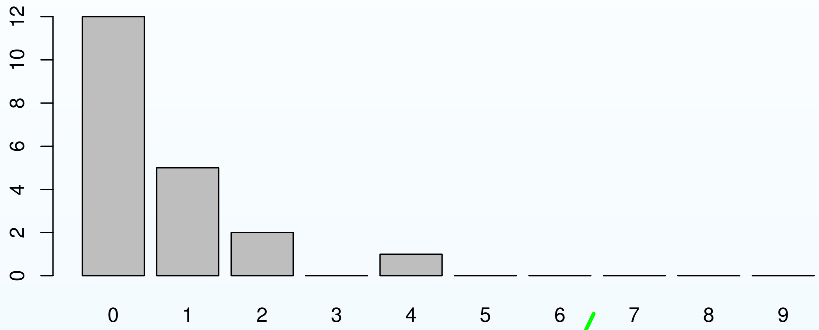
$\Rightarrow$ Next ?

$P(0) > P(1) > P(2)$ $\surd$

$P(3) < P(4)$, or $P(3) \geq P(4)$?

$P(3) = 0$, or $P(5) = 0$?

A counting experiment



$\Rightarrow$ Next ?

$P(0) > P(1) > P(2)$ $\surd$

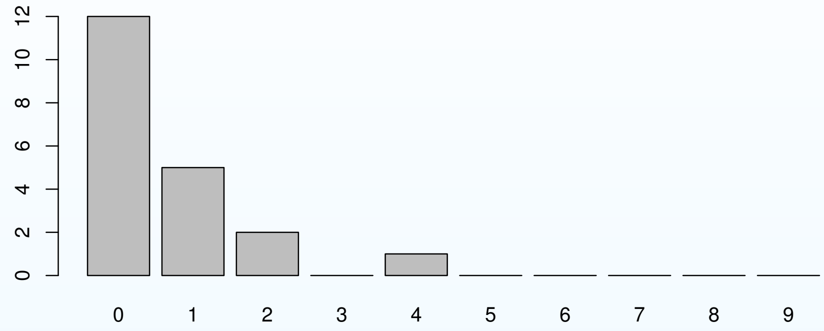
$P(3) < P(4)$, or $P(3) \geq P(4)$?

$P(3) = 0$, or $P(5) = 0$?

Not correct to say *"we cannot do it"*, or *"let us do other measurements and see"*:

> In real life we are asked to make assessments (and take decisions) with the information we have NOW. If, later, the information changes, we can (must!) use the update one (and perhaps update our opinion).

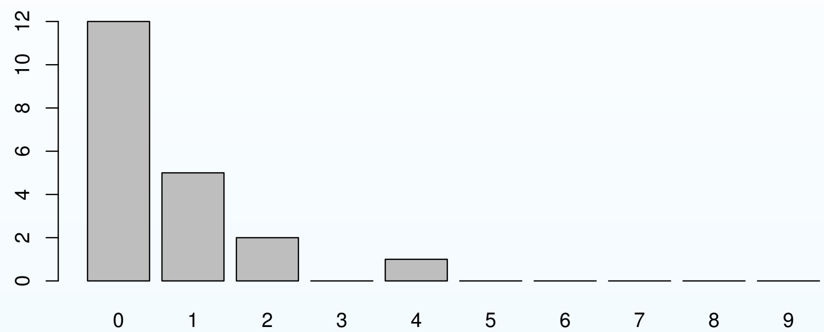
A counting experiment



$\Rightarrow$ Next ?

Why we, as physicists, tend to state $P(3) > P(4)$ and $P(5) > 0$?

A counting experiment



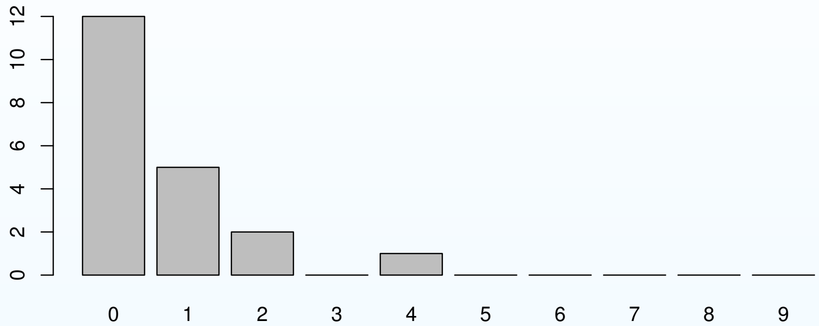
$\Rightarrow$ Next ?

Why we, as physicists, tend to state $P(3) > P(4)$ and $P(5) > 0$?

Given our ‘experience’, ‘education’, ‘mentality’ (. . .)

We ‘know’ ‘assume’ ‘hope’ ‘guess’ ‘postulate’ regularity of nature

A counting experiment



$\Rightarrow$ Next ?

Why we, as physicists, tend to state $P(3) > P(4)$ and $P(5) > 0$?

Given our 'experience', 'education', 'mentality' (...)

We

'know'

'assume'

'hope' regularity of nature

'guess'

'postulate'

'believe'

A philosopher, physicist and mathematician joke

A philosopher, a physicist and a mathematician travel by train through Scotland.

The train is going slowly and they see a cow walking along a country road parallel to the railway.

A philosopher, physicist and mathematician joke

A philosopher, a physicist and a mathematician travel by train through Scotland.

The train is going slowly and they see a cow walking along a country road parallel to the railway.

- Philosopher: *"In Scotland cows are black"*

A philosopher, physicist and mathematician joke

A philosopher, a physicist and a mathematician travel by train through Scotland.

The train is going slowly and they see a cow walking along a country road parallel to the railway.

- Philosopher: *"In Scotland cows are black"*
- Physicist: *"In Scotland there is at least a black cow"*

A philosopher, physicist and mathematician joke

A philosopher, a physicist and a mathematician travel by train through Scotland.

The train is going slowly and they see a cow walking along a country road parallel to the railway.

- Philosopher: *“In Scotland cows are black”*
- Physicist: *“In Scotland there is at least a black cow”*
- Mathematician: *“In Scotland at least a cow has a black side”*

A philosopher, physicist and mathematician joke

A philosopher, a physicist and a mathematician travel by train through Scotland.

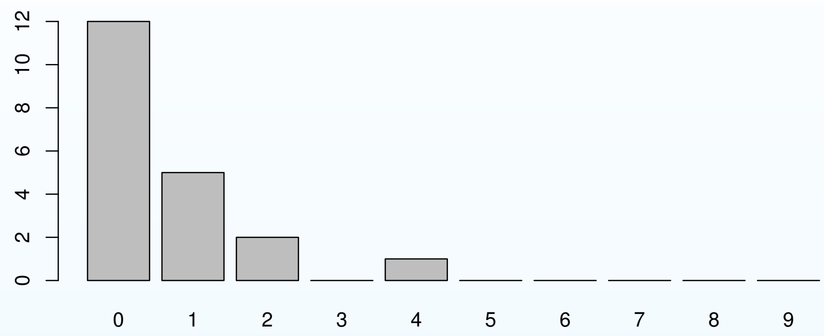
The train is going slowly and they see a cow walking along a country road parallel to the railway.

- Philosopher: *“In Scotland cows are black”*
- Physicist: *“In Scotland there is at least a black cow”*
- Mathematician: *“In Scotland at least a cow has a black side”*

Physicists’ statements about reality have plenty of tacit – mostly very reasonable! — assumptions that derive from experience and rationality.

$\Rightarrow$ We constantly use theory/models to link past and future!.

Transferring past to future

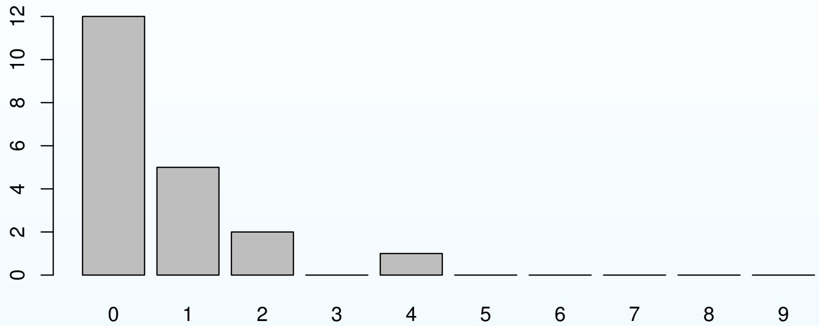


$\Rightarrow$ Next ?

Basic reasoning: assuming regularity of nature and a regular flow from the past to the future, we tend to believe that the effects that happened more frequently in the past will also occur more likely in the future.

Again, well expressed by Hume.†

Transferring past to future



$\Rightarrow$ Next ?

Basic reasoning: assuming regularity of nature and a regular flow from the past to the future, we tend to believe that the effects that happened more frequently in the past will also occur more likely in the future.

Again, well expressed by Hume.†

We physicists tend to filter the process of transferring the past to the future by 'laws'.

$\Rightarrow$ an experimental histogram shows a relative-frequency distribution, and not a probability distribution!

Relative frequencies *might* become probabilities, but only after they have been processed by our mind.

Uncertainties in measurements

Having to perform a measurement:

Which numbers shall come out from our device?

Having performed a measurement:

What have we learned about the value of the quantity of interest?

Uncertainties in measurements

Having to perform a measurement:

Which numbers shall come out from our device?

Having performed a measurement:

What have we learned about the value of the quantity of interest?

How to quantify these kinds of uncertainty?

Uncertainties in measurements

Having to perform a measurement:

Which numbers shall come out from our device?

Having performed a measurement:

What have we learned about the value of the quantity of interest?

How to quantify these kinds of uncertainty?

Under well controlled conditions (calibration) we can make use of past frequencies to evaluate 'somehow' the detector response $f(x \,|\, \mu)$.

Uncertainties in measurements

Having to perform a measurement:

Which numbers shall come out from our device?

Having performed a measurement:

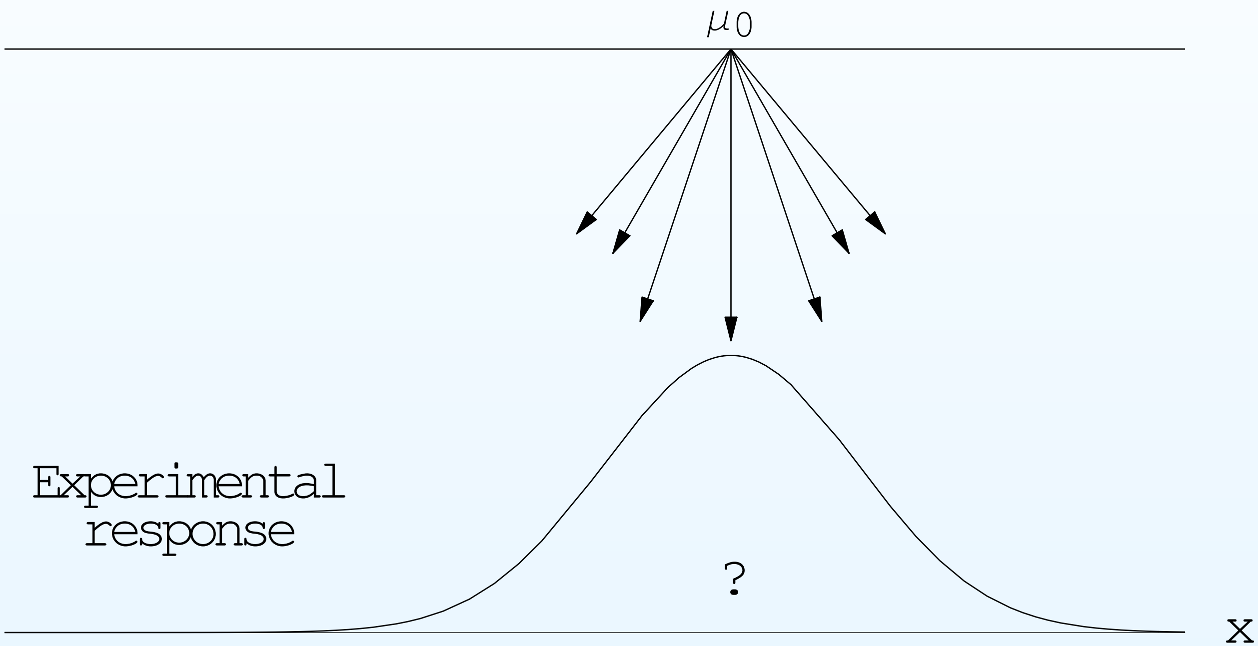
What have we learned about the value of the quantity of interest?

How to quantify these kinds of uncertainty?

Under well controlled conditions (calibration) we can make use of past frequencies to evaluate 'somehow' the detector response $f(x \mid \mu)$.

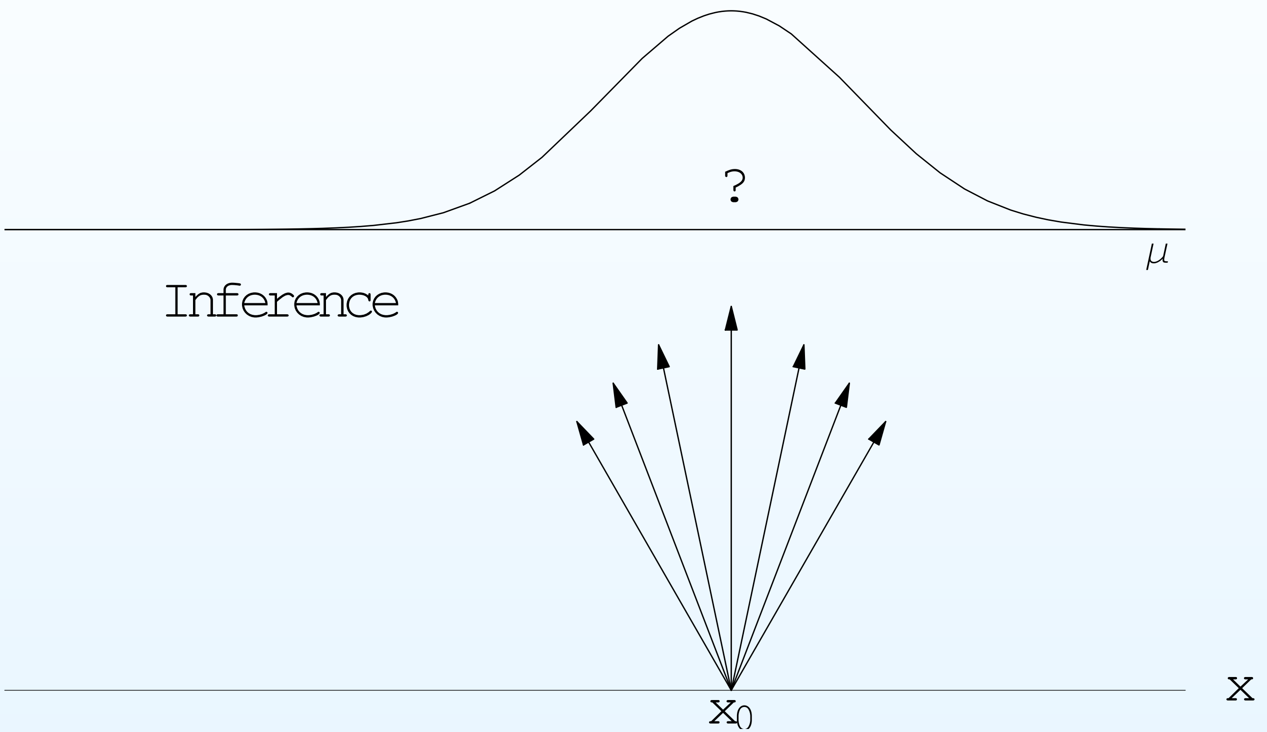
There is (in most cases) no way to get *directly* hints about $f(\mu \mid x)$.

Uncertainties in measurements



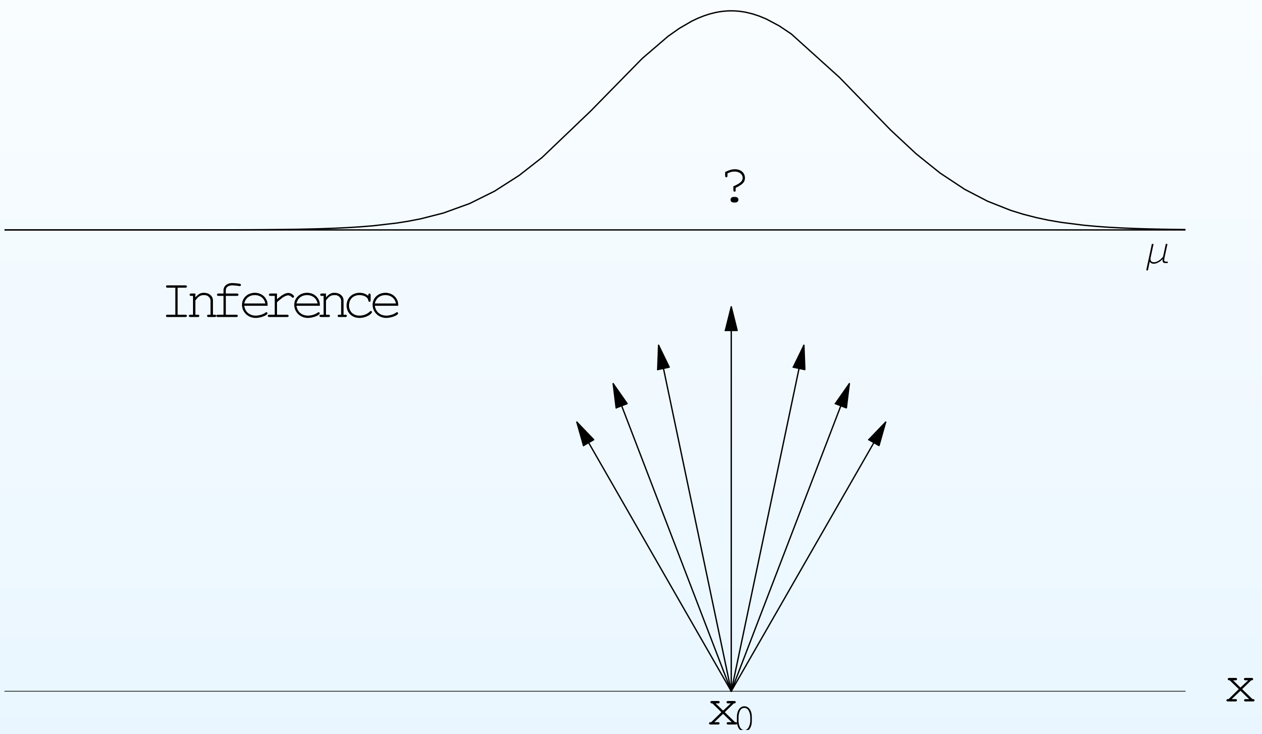
$f(x \mid \mu)$ experimentally accessible (though 'model filtered')

Uncertainties in measurements



$f(\mu \,|\, x)$ experimentally inaccessible

Uncertainties in measurements

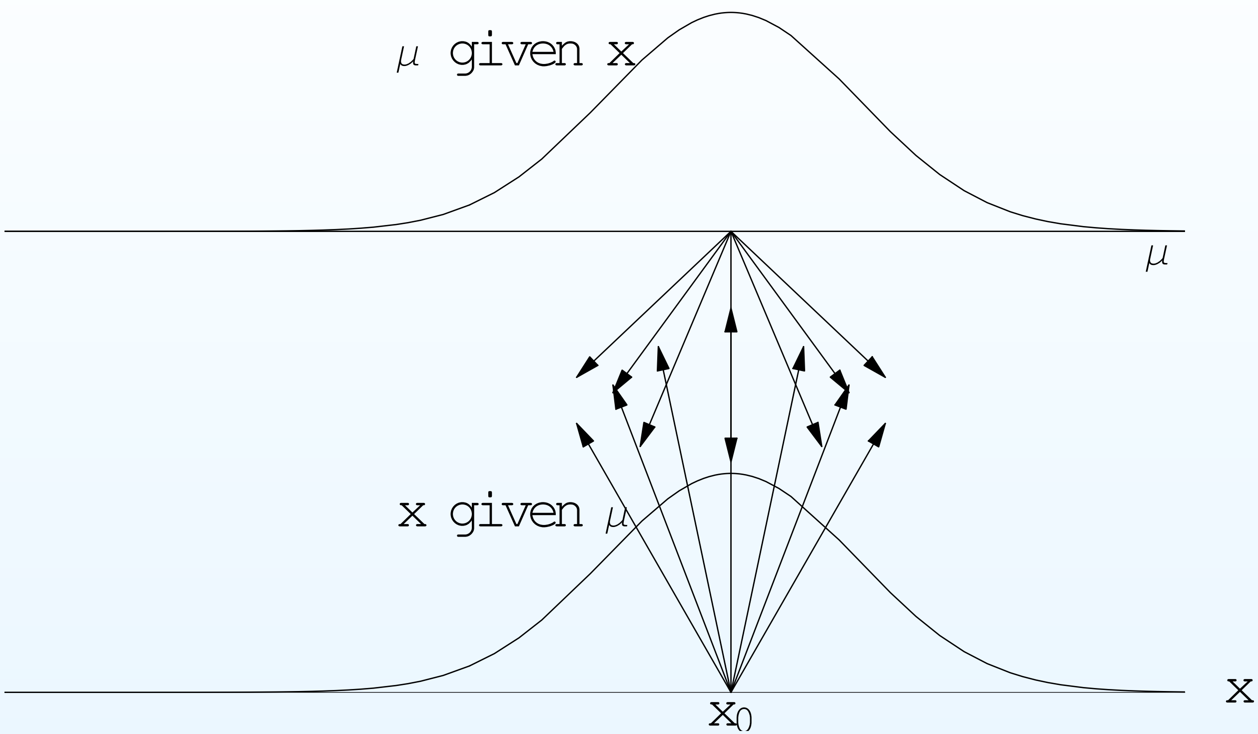


$f(\mu \,|\, x)$ experimentally inaccessible

but logically accessible!

$\rightarrow$ we need to learn how to do it

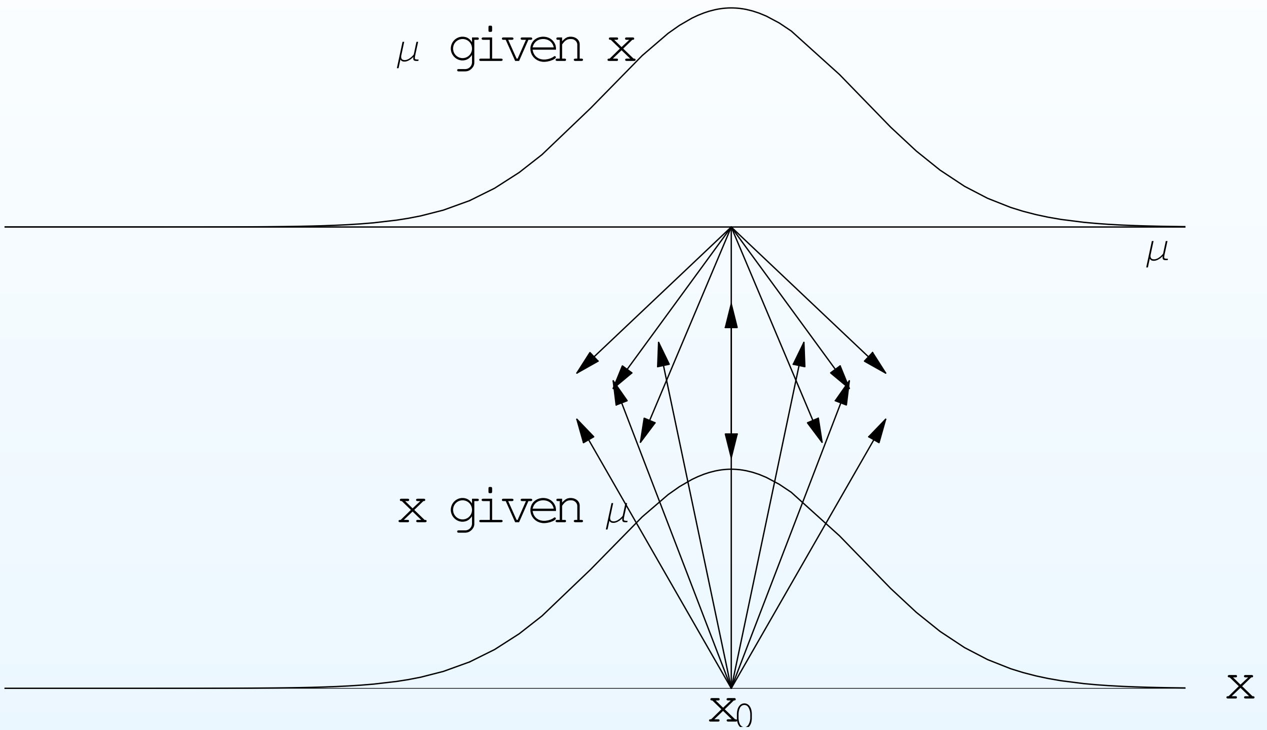
Uncertainties in measurements



- Review sources of uncertainties
- How measurement uncertainties are currently treated
- How to treat them logically using probability theory

  (But we also need to review what we mean by ‘probability’!)

Uncertainties in measurements



- Review sources of uncertainties ⟶ See next
- How measurement uncertainties are currently treated
- How to treat them logically using probability theory

  (But we also need to review what we mean by 'probability'!)

Sources of uncertainties (ISO Guide)

1 *incomplete definition of the measurand;*

$\longrightarrow$ g

$\longrightarrow$where?

$\longrightarrow$inertial effects subtracted?

2 *imperfect realization of the definition of the measurand;*

$\longrightarrow$ scattering on neutron

$\longrightarrow$how to realize a neutron target?

3 *non-representative sampling — the sample measured may not represent the measurand;*

4 *inadequate knowledge of the effects of environmental conditions on the measurement, or imperfect measurement of environmental conditions;*

5 *personal bias in reading analogue instruments;*

Sources of uncertainties (ISO Guide)

6 *finite instrument resolution or discrimination threshold;*

7 *inexact values of measurement standards and reference materials;*

8 *inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithm;*

9 *approximations and assumptions incorporated in the measurement method and procedure;*

10 *variations in repeated observations of the measurand under apparently identical conditions.*

$\longrightarrow$ *"statistical errors"*

Note

- Sources not necessarily independent
- In particular, sources 1-9 may contribute to 10

Usual handling of measurement uncertainties

Uncertainties due to statistical errors are currently treated using the frequentistic concept of 'confidence interval',

Usual handling of measurement uncertainties

Uncertainties due to **statistical errors** are currently treated using the frequentistic concept of ‘confidence interval’, although

- there are well-know cases — of great relevance in frontier physics — in which the approach is not applicable (e.g. small number of observed events, or measurement close to the edge of the physical region);

Usual handling of measurement uncertainties

Uncertainties due to statistical errors are currently treated using the frequentistic concept of 'confidence interval', although

- there are well-know cases — of great relevance in frontier physics — in which the approach is not applicable (e.g. small number of observed events, or measurement close to the edge of the physical region);
- the procedure is rather unnatural, and in fact the interpretation of the results is unconsciously (intuitively) probabilistic (see later).
  $\rightarrow$ Intuitive reasoning $\Longleftrightarrow$ statistics education

Usual handling of measurement uncertainties

Uncertainties due to statistical errors are currently treated using the frequentistic concept of 'confidence interval', although

- there are well-know cases — of great relevance in frontier physics — in which the approach is not applicable (e.g. small number of observed events, or measurement close to the edge of the physical region);
- the procedure is rather unnatural, and in fact the interpretation of the results is unconsciously (intuitively) probabilistic (see later).
  $\rightarrow$ Intuitive reasoning $\Longleftrightarrow$ statistics education

These cases have not to be seen as "the exception that confirms the rule" [in physics exceptions falsify laws!], but as symptoms of something flawed in the reasoning, that could seriously effects also results that are not as self-evidently paradoxical as in these cases!

Usual handling of measurement uncertainties

There is no satisfactory theory or model to treat uncertainties due to systematic errors:

- *“my supervisor says . . . ”*
- *“add them linearly”;*
- *“add them linearly if . . . , else add them quadratically”*;
- *“don’t add them at all”.*

Usual handling of measurement uncertainties

There is no satisfactory theory or model to treat uncertainties due to systematic errors:

- *"my supervisor says . . . "*
- *"add them linearly";*
- *"add them linearly if . . . , else add them quadratically";*
- *"don't add them at all".*

The modern *fashion*: add them quadratically if they are considered to be independent, or build a covariance matrix of statistical and systematic contributions in the general case.

Usual handling of measurement uncertainties

There is no satisfactory theory or model to treat uncertainties due to systematic errors:

- *"my supervisor says . . . "*
- *"add them linearly";*
- *"add them linearly if . . . , else add them quadratically";*
- *"don't add them at all".*

The modern *fashion*: add them quadratically if they are considered to be independent, or build a covariance matrix of statistical and systematic contributions in the general case.
In my opinion, simply the reluctance to combine linearly 10, 20 or more contributions to a global uncertainty, as the (out of fashion) 'theory' of maximum bounds would require.

Usual handling of measurement uncertainties

There is no satisfactory theory or model to treat uncertainties due to systematic errors:

- *"my supervisor says ..."*
- *"add them linearly";*
- *"add them linearly if ..., else add them quadratically"*;
- *"don't add them at all".*

The modern *fashion*: add them quadratically if they are considered to be independent, or build a covariance matrix of statistical and systematic contributions in the general case.
In my opinion, simply the reluctance to combine linearly 10, 20 or more contributions to a global uncertainty, as the (out of fashion) 'theory' of maximum bounds would require.
$\rightarrow$ Right in most cases!

$\rightarrow$ Good sense of physicists $\Longleftrightarrow$ cultural background

A simple case

n independent measurements of the same quantity μ (with n large enough and no systematic effects, to avoid, for the moment, extra complications).

Evaluate $\overline{x}$ and σ from the data

report result: $\rightarrow \mu = \overline{x} \pm \sigma/\sqrt{n}$

- what does it mean?

A simple case

n independent measurements of the same quantity μ (with n large enough and no systematic effects, to avoid, for the moment, extra complications).

Evaluate $\overline{x}$ and σ from the data

report result: $\rightarrow \mu = \overline{x} \pm \sigma/\sqrt{n}$

- what does it mean?

1. For the large majority of physicists $P(\overline{x} - \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$

A simple case

n independent measurements of the same quantity μ (with n large enough and no systematic effects, to avoid, for the moment, extra complications).

Evaluate $\overline{x}$ and σ from the data

report result: $\rightarrow \mu = \overline{x} \pm \sigma/\sqrt{n}$

- what does it mean?

1. For the large majority of physicists $P(\overline{x} - \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$
2. And many explain (also to students!) that *"this means that, if I repeat the experiment a great number of times, then I will find that in roughly 68% of the cases the observed average will be in the interval* $[\overline{x} - \sigma/\sqrt{n},\ \overline{x} + \sigma/\sqrt{n}]$*."*

A simple case

n independent measurements of the same quantity μ (with n large enough and no systematic effects, to avoid, for the moment, extra complications).

Evaluate $\overline{x}$ and σ from the data

report result: $\rightarrow \mu = \overline{x} \pm \sigma/\sqrt{n}$

- what does it mean?

1. For the large majority of physicists
$P(\overline{x} - \frac{\sigma}{\sqrt{n}} \le \mu \le \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$
2. And many explain (also to students!) that *“this means that, if I repeat the experiment a great number of times, then I will find that in roughly 68% of the cases the observed average will be in the interval* $[\overline{x} - \sigma/\sqrt{n},\ \overline{x} + \sigma/\sqrt{n}]$*.”*
3. Statistics experts tell that the interval $[\overline{x} - \sigma/\sqrt{n},\ \overline{x} + \sigma/\sqrt{n}]$ covers the true μ in 68% of cases

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

1 $P(\overline{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%$

OK to me, and perhaps no objections by many of you

- But it depends on what we mean by probability
- If probability is the "limit of the frequency", this statement is meaningless, because the 'frequency based' probability theory only speak about

$$P(\mu - \frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + \frac{\sigma}{\sqrt{n}}) = 68\%\,,$$

(that is a probabilistic statement about $\overline{X}$: probabilistic statements about μ are not allowed by the theory).

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

2 *"if I repeat the experiment a great number of times, then I will find that in roughly 68% of the cases the observed average will be in the interval* $[\overline{x} - \sigma/\sqrt{n},\ \overline{x} + \sigma/\sqrt{n}]$*."*

- Nothing wrong in principle (in my opinion)
- but a $\sqrt{2}$ mistake in the width of the interval

$\rightarrow$ $P(\overline{x} - \sigma/\sqrt{n}\ \leq\ \overline{x}_f\ \leq\ \overline{x} + \sigma/\sqrt{n}) = 52\%$,
where $\overline{x}_f$ stands for future averages;

or $P(\overline{x} - \sqrt{2}\,\sigma/\sqrt{n}\ \leq\ \overline{x}_f\ \leq\ \overline{x} + \sqrt{2}\,\sigma/\sqrt{n}) = 68\%$,

as we shall see later ($\rightarrow$ 'predictive distributions').

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ “several problems”

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)
- It fails in frontier cases
  - 'technically' [see e.g. G. Zech, *Frequentistic and Bayesian confidence limits*, EPJdirect C12 (2002) 1]
  - 'in terms of performance' $\longrightarrow$ 'very strange' that no quantities show in 'other side' of a 95% C.L. bound !

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)
- It fails in frontier cases
  - 'technically' [see e.g. G. Zech, *Frequentistic and Bayesian confidence limits*, EPJdirect C12 (2002) 1]
  - 'in terms of performance' $\longrightarrow$ 'very strange' that no quantities show in 'other side' of a 95% C.L. bound !
- Not suited to express our confidence! Simply because it was not invented for that purpose!

  The peculiar characteristic of frequentistic coverage is not to express confidence, but, when it works, to 'ensure' that, when applied a great number of times, in a defined percentage of the report the coverage statement is true. (See e.g. P. Clifford, 2000 CERN Workshop on C.L.'s.)

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)
- It fails in frontier cases
  - 'technically' [see e.g. G. Zech, *Frequentistic and Bayesian confidence limits*, EPJdirect C12 (2002) 1]
  - 'in terms of performance' $\longrightarrow$ 'very strange' that no quantities show in 'other side' of a 95% C.L. bound !
- Not suited to express our confidence! Simply because it was not invented for that purpose!

  The ultimate 68.3% C.L. confidence interval calculator: a random number generator that gives
  - $[-10^{+9999}, +10^{+9999}]$ with 68.3% probability
  - $[1.00000001 \times 10^{-300}, 1.00000002 \times 10^{-300}]$ with 31.7% probability.

  Great! (No experiment required!)

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)
- It fails in frontier cases
  - 'technically' [see e.g. G. Zech, *Frequentistic and Bayesian confidence limits*, EPJdirect C12 (2002) 1]
  - 'in terms of performance' $\longrightarrow$ 'very strange' that no quantities show in 'other side' of a 95% C.L. bound !
- Not suited to express our confidence! Simply because it was not invented for that purpose!

  If you do not like it, it might be you do not really care about 'coverage'. You, as a physicist who care about your physical quantity, think in terms of 'confidence':

$\Rightarrow$ How much you are confident that the value of your quantity of interest is in a given interval. We do not play a lottery!

Meaning of $\mu = \overline{x} \pm \sigma/\sqrt{n}$

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\longrightarrow$ taken by most users as if it were a probability interval (not just semantic!)
- It fails in frontier cases
  - 'technically' [see e.g. G. Zech, *Frequentistic and Bayesian confidence limits*, EPJdirect C12 (2002) 1]
  - 'in terms of performance' $\longrightarrow$ 'very strange' that no quantities show in 'other side' of a 95% C.L. bound !
- Not suited to express our confidence! Simply because it was not invented for that purpose!

  "that technological and commercial apparatus" (Fisher)

Arbitrary probability inversions

As with hypotheses tests, problem arises from arbitrary probability inversions.

How do we turn, just 'intuitively'

$$P(\mu - \frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + \frac{\sigma}{\sqrt{n}}) = 68\%$$

into

$$P(\overline{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%\,?$$

Arbitrary probability inversions

As with hypotheses tests, problem arises from arbitrary probability inversions.

How do we turn, just 'intuitively'

$$P(\mu - \frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + \frac{\sigma}{\sqrt{n}}) = 68\%$$

into

$$P(\overline{x} - \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + \frac{\sigma}{\sqrt{n}}) = 68\%\,?$$

We can paraphrase as

"the dog and the hunter"

The dog and the hunter

We know that a dog has a 50% probability of being 100 m from the hunter

$\Rightarrow$ if we observe the dog, what can we say about the hunter?

The dog and the hunter

We know that a dog has a 50% probability of being 100 m from the hunter

$\Rightarrow$ if we observe the dog, what can we say about the hunter?

The terms of the analogy are clear:

$$\begin{aligned} \text{hunter} &\quad\leftrightarrow\quad \text{true value} \\ \text{dog} &\quad\leftrightarrow\quad \text{observable}\,. \end{aligned}$$

The dog and the hunter

We know that a dog has a 50% probability of being 100 m from the hunter

$\Rightarrow$ if we observe the dog, what can we say about the hunter?

The terms of the analogy are clear:

$$\begin{aligned} \text{hunter} &\quad\leftrightarrow\quad \text{true value} \\ \text{dog} &\quad\leftrightarrow\quad \text{observable}\,. \end{aligned}$$

Intuitive and reasonable answer:

"The hunter is, with 50% probability, within 100 m of the position of the dog."

The dog and the hunter

We know that a dog has a 50% probability of being 100 m from the hunter

$\Rightarrow$ if we observe the dog, what can we say about the hunter?

The terms of the analogy are clear:

$$\begin{aligned} \text{hunter} \quad &\leftrightarrow \quad \text{true value} \\ \text{dog} \quad &\leftrightarrow \quad \text{observable}\,. \end{aligned}$$

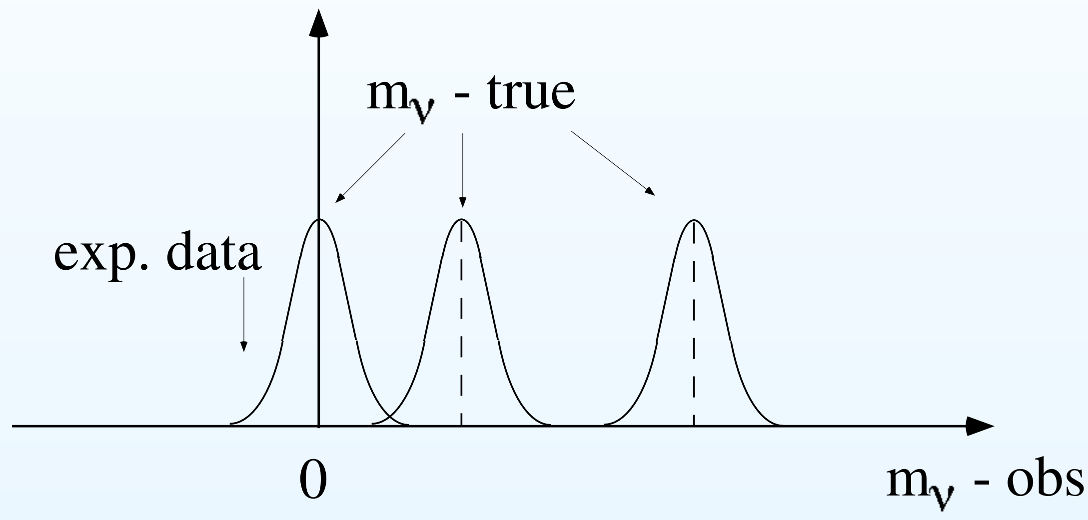
Easy to understand that this conclusion is based on some tacit assumptions:

- the hunter can be anywhere around the dog
- the dog has no preferred direction of arrival at the point where we observe him.

$\longrightarrow$ not always valid!

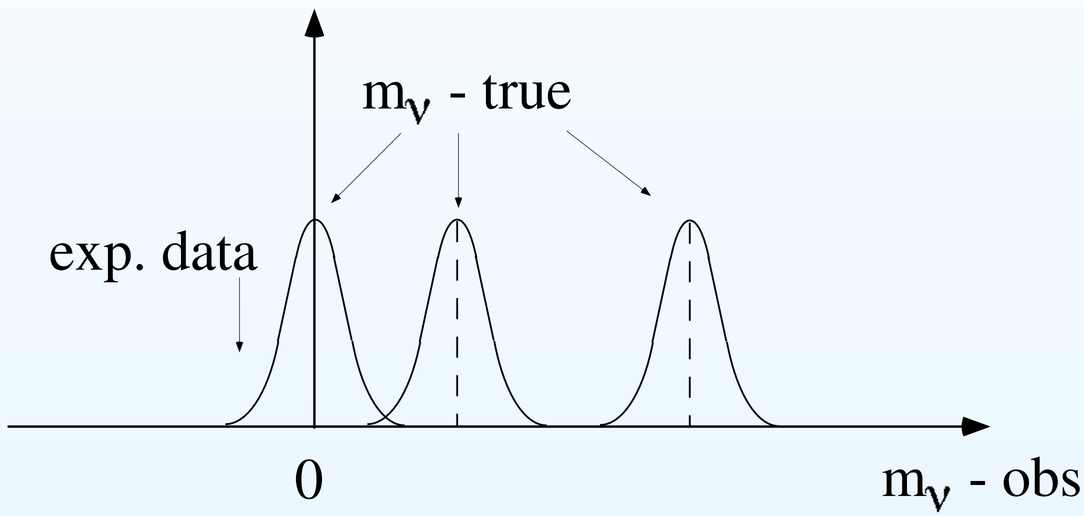
Measurement at the edge of a physical region

Electron-neutrino experiment, mass resolution $\sigma = 2\,\mathrm{eV}$, independent of m_ν.



Measurement at the edge of a physical region

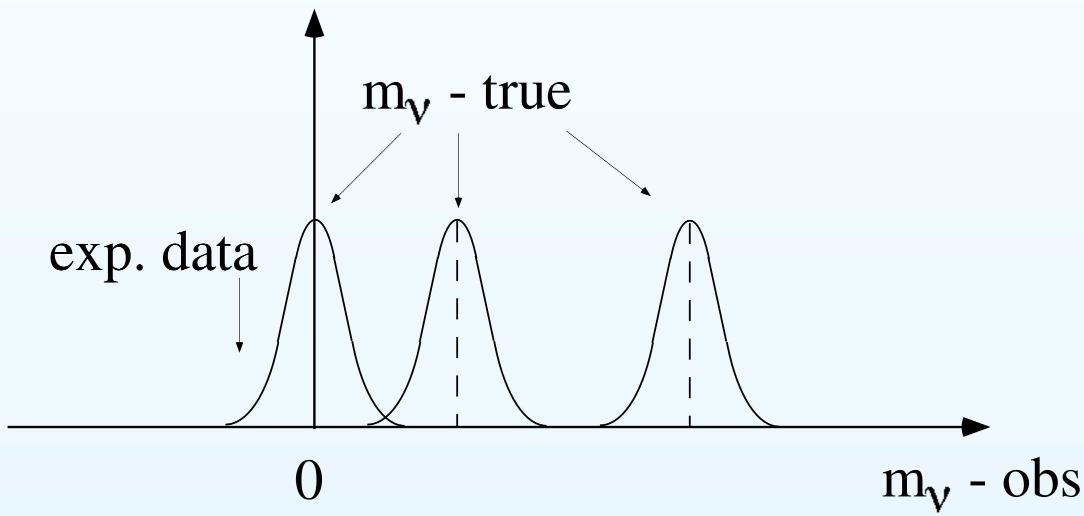
Electron-neutrino experiment, mass resolution $\sigma = 2\,$eV, independent of m_ν.



Observation: $-4\,$eV.
What can we tell about m_ν?

Measurement at the edge of a physical region

Electron-neutrino experiment, mass resolution $\sigma = 2$ eV, independent of m_ν.

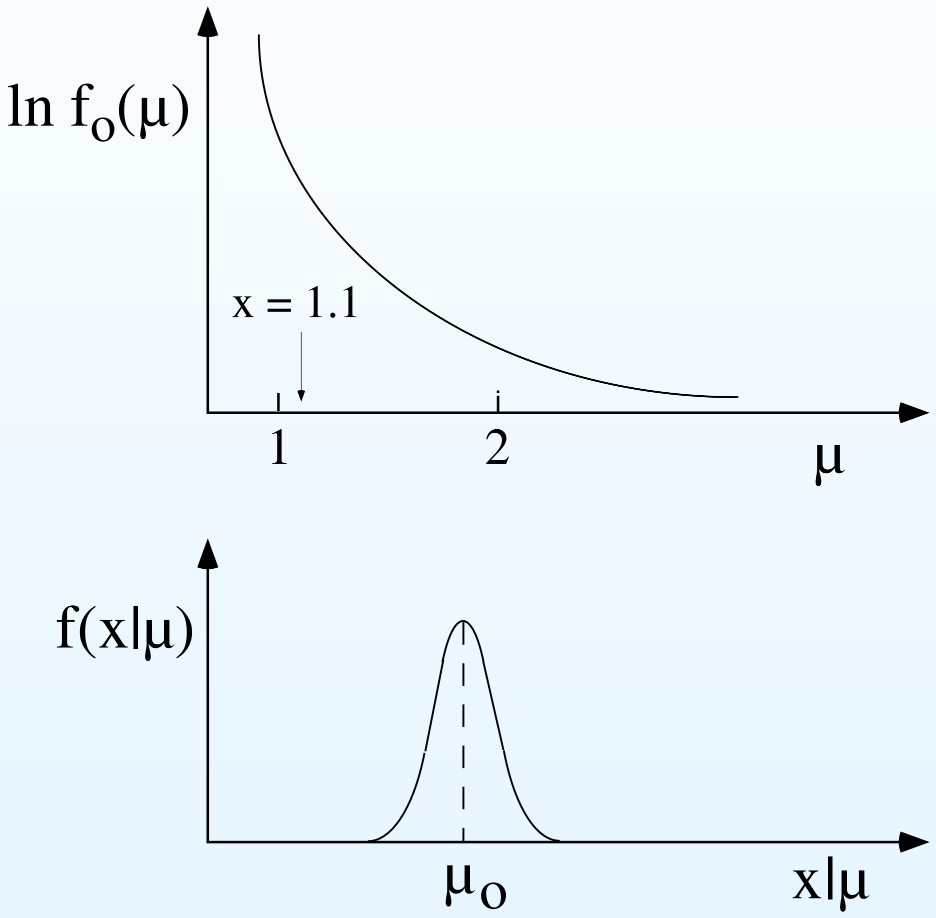


Observation: -4 eV.
What can we tell about m_ν?

$m_\nu = -4 \pm 2$ eV ?
$P(-6 \le m_\nu/\text{eV} \le -2) = 68\%$?
$P(m_\nu \le 0 \text{ eV}) = 98\%$?

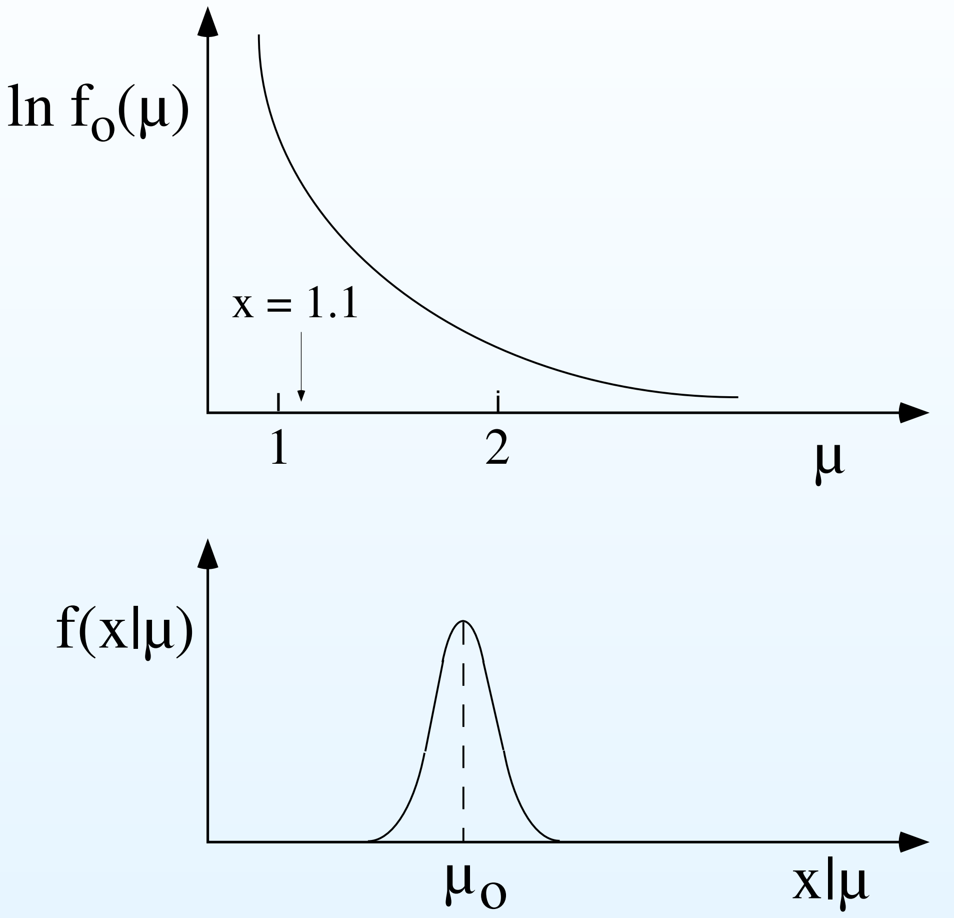
Non-flat distribution of a physical quantity

Imagine a cosmic ray particle or a bremsstrahlung γ.



Non-flat distribution of a physical quantity

Imagine a cosmic ray particle or a bremsstrahlung γ.
Observed $x = 1.1$.
What can we say about the true value μ that has caused this observation?



Non-flat distribution of a physical quantity

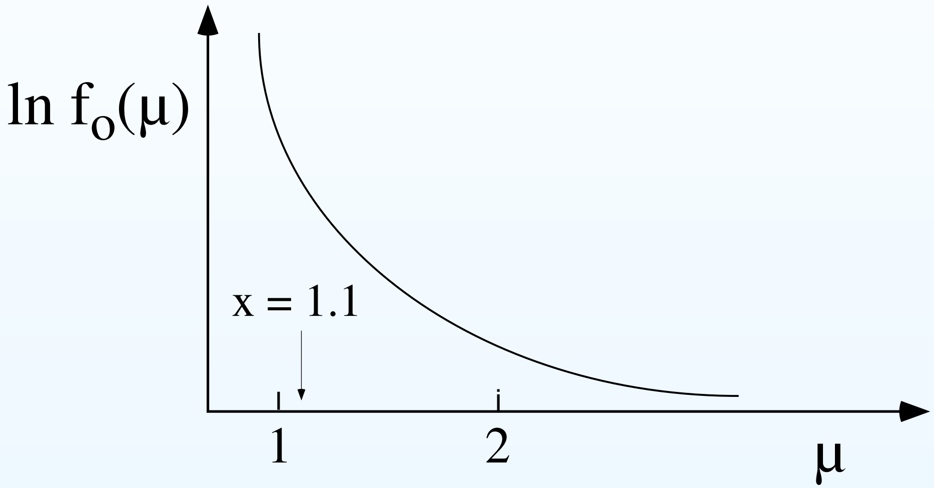
Imagine a cosmic ray particle or a bremsstrahlung γ.

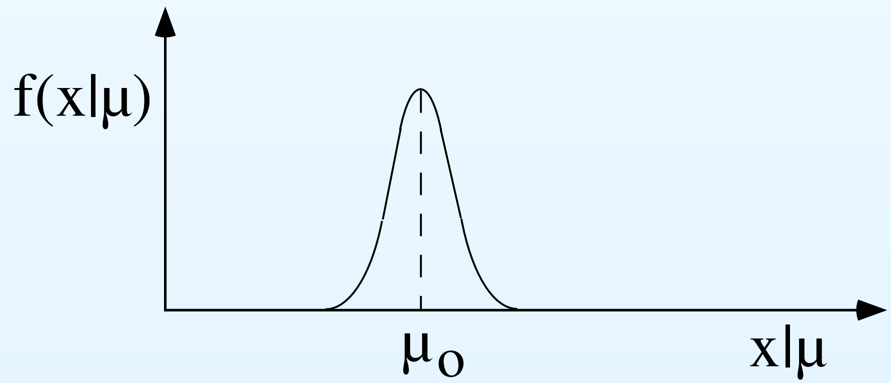
Observed $x = 1.1$.

What can we say about the true value μ that has caused this observation?

Also in this case the formal definition of the confidence interval does not work.

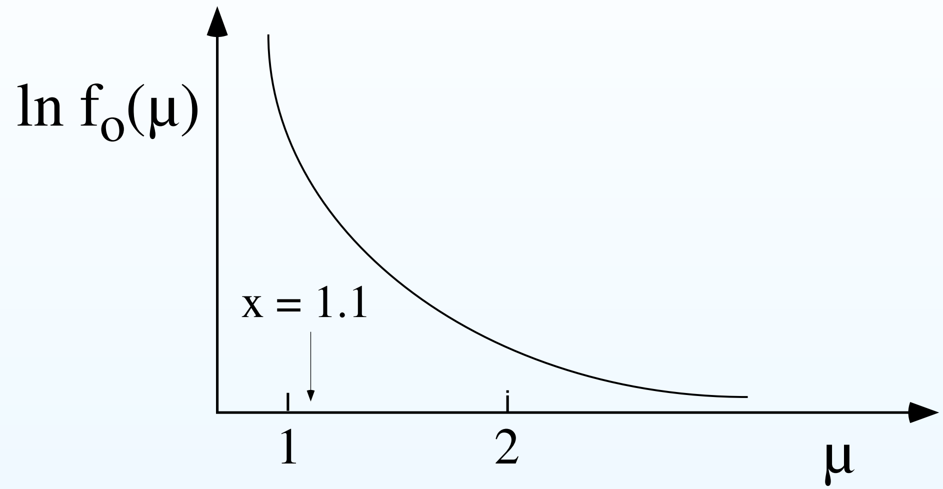
Intuitively, we feel that there is more chance that μ is on the left of 1.1 than on the right. In the jargon of the experimentalists, *"there are more migrations from left to right than from right to left"*.

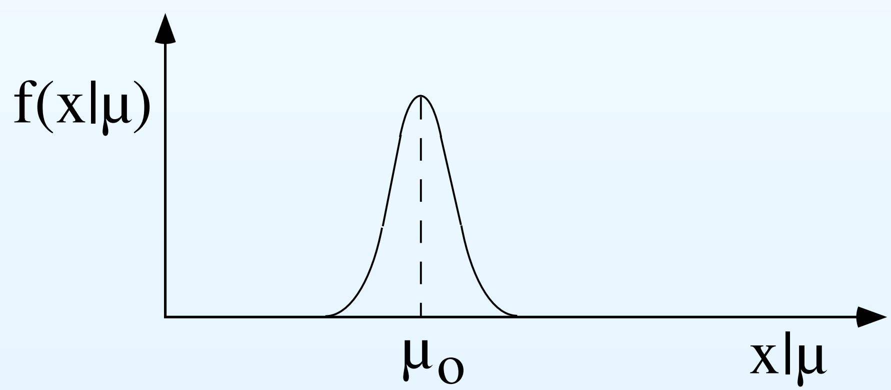




Non-flat distribution of a physical quantity

These two examples deviates from the dog-hunter picture only because of an asymmetric possible position of the 'hunter', i.e our expectation about μ is not uniform. But there are also interesting cases in which the response of the apparatus $f(x \mid \mu)$ is not symmetric around μ, e.g. the reconstructed momentum in a magnetic spectrometer.



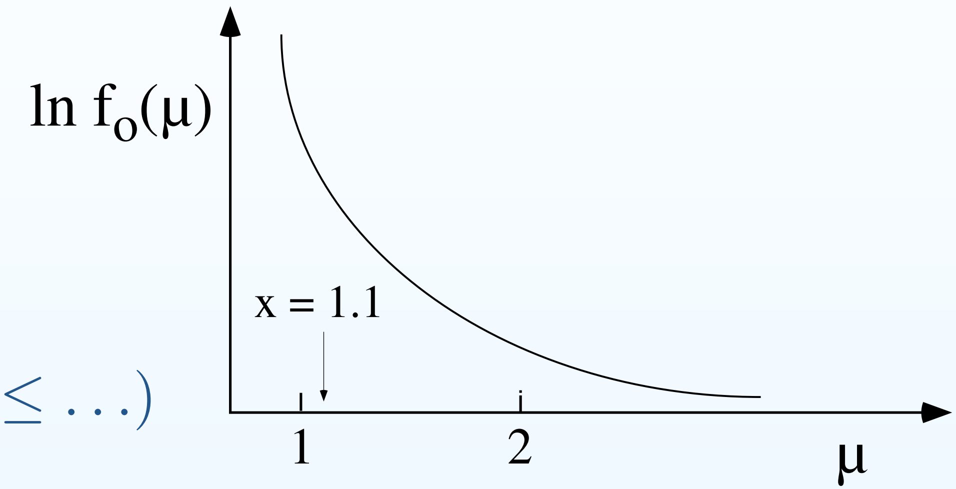


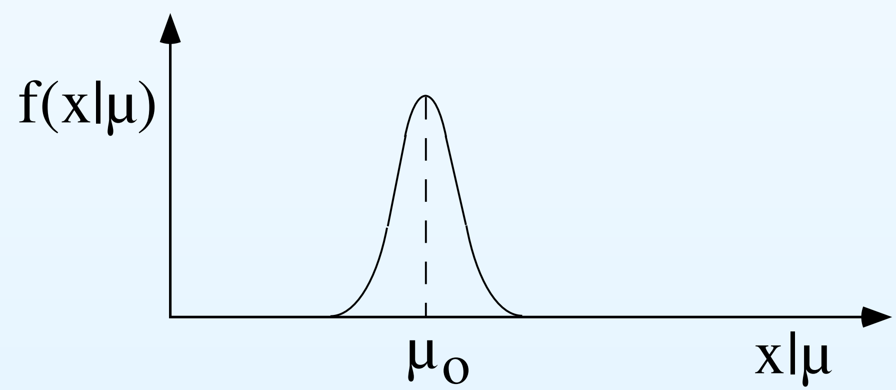
Non-flat distribution of a physical quantity

Summing up:
"the intuitive inversion of probability

$$P(\ldots \leq \overline{X} \leq \ldots) \Longrightarrow P(\ldots \leq \mu \leq \ldots)$$

besides being theoretically unjustifiable, yields results which are numerically correct only in the case of symmetric problems."





Summary about standard methods

Situation is not satisfactory in the critical situations that often occur in HEP, both in

- hypotheses tests
- confidence intervals

Summary about standard methods

Situation is not satisfactory in the critical situations that often occur in HEP, both in

- hypotheses tests
- confidence intervals

Plus there are issues not easy to treat in that frame
[and I smile at the heroic effort to get some result :-)]

- systematic errors
- background

Implicit assumptions

We have seen clearly what are the hidden assumptions in the ‘naive probability inversion’ (that corresponds more or less to the prescriptions to build confidence intervals).

> We shall see that, similarly, there are hidden assumptions behind the naive probabilistic inversions in problems like the AIDS one.

Implicit assumptions

We have seen clearly what are the hidden assumptions in the 'naive probability inversion' (that corresponds more or less to the prescriptions to build confidence intervals).

We shall see that, similarly, there are hidden assumptions behind the naive probabilistic inversions in problems like the AIDS one.

Curiously enough, these methods are advertised as objective because they do not need as input our scientific expectations of where the value of the quantity might lie, or of which physical hypothesis seems more reasonable!

Implicit assumptions

We have seen clearly what are the hidden assumptions in the 'naive probability inversion' (that corresponds more or less to the prescriptions to build confidence intervals).

We shall see that, similarly, there are hidden assumptions behind the naive probabilistic inversions in problems like the AIDS one.

Curiously enough, these methods are advertised as objective because they do not need as input our scientific expectations of where the value of the quantity might lie, or of which physical hypothesis seems more reasonable!

But if we are convinced (by logic, or by the fact that neglecting that knowledge paradoxical results can be achieved) that prior expectation is relevant in inferences, we cannot accept methods which systematically neglect it and that, for that reason, solve problems different from those we are interested in!

What to do?

$\Rightarrow$ Back to the past

What to do?

$\Rightarrow$ Back to the past

But benefiting of

- Theoretical progresses in probability theory
- Advance in computation (both symbolic and numeric)

  $\longrightarrow$ many frequentistic ideas had their *raison d'être* in the computational barrier (and many simplified – often simplistic – methods were ingeniously worked out)

  $\longrightarrow$ no longer an excuse

What to do?

$\Rightarrow$ Back to the past

But benefiting of

- Theoretical progresses in probability theory
- Advance in computation (both symbolic and numeric)
  - $\rightarrow$ many frequentistic ideas had their *raison d'être* in the computational barrier (and many simplified – often simplistic – methods were ingeniously worked out)
    - $\rightarrow$ no longer an excuse

$\Rightarrow$ Use consistently probability theory

What to do?

$\Rightarrow$ Back to the past

But benefiting of

- Theoretical progresses in probability theory
- Advance in computation (both symbolic and numeric)

  $\rightarrow$ many frequentistic ideas had their *raison d'être* in the computational barrier (and many simplified – often simplistic – methods were ingeniously worked out)

  $\rightarrow$ no longer an excuse

$\Rightarrow$ Use consistently probability theory

- ◦ It's easy if you try
- ◦ But first you have to recover the intuitive idea of probability.

Probability

What is probability?

Standard textbook definitions

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

$$p = \frac{\#\,\text{times the event has occurred}}{\#\,\text{independent trials under same conditions}}$$

Standard textbook definitions

It is easy to check that 'scientific' definitions suffer of circularity

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

$$p = \frac{\#\,\text{times the event has occurred}}{\#\,\text{independent}\ \ \text{trials under same conditions}}$$

Standard textbook definitions

It is easy to check that 'scientific' definitions suffer of circularity

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

$$p = \frac{\#\,\text{times the event has occurred}}{\#\,\text{independent}\quad \text{trials under same conditions}}$$

Laplace: *"lorsque rien ne porte à croire que l'un de ces cas doit arriver plutot que les autres"*

Pretending that replacing 'equi-probable' by 'equi-possible' is just cheating students (as I did in my first lecture on the subject…).

Standard textbook definitions

It is easy to check that 'scientific' definitions suffer of circularity , plus other problems

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

$$p \quad = \lim_{n\to\infty} \frac{\#\,\text{times the event has occurred}}{\#\ \text{independent trials under same condition}}$$

Future $\Leftrightarrow$ Past (believed so)

$n \longrightarrow \infty$: $\longrightarrow$ *"usque tandem?"*

$\longrightarrow$ *" in the long run we are all dead"*

$\longrightarrow$ It limits the range of applications

Definitions $\longrightarrow$ evaluation rules

Very useful evaluation rules

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

$$B) \quad p \quad = \quad \frac{\#\,\text{times the event has occurred}}{\#\text{independent trials under same condition}}$$

If the implicit beliefs are well suited for each case of application.

Definitions $\longrightarrow$ evaluation rules

Very useful evaluation rules

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

$$B) \quad p \quad = \quad \frac{\#\,\text{times the event has occurred}}{\#\text{independent trials under same condition}}$$

If the implicit beliefs are well suited for each case of application.

In the probabilistic approach we are going to see

- Rule A will be recovered immediately (under the same assumption of equiprobability).
- Rule B will result from a theorem (under well defined assumptions).

Probability

What <u>is</u> probability?

Probability

What is probability?

It is what everybody knows what it is before going at school

Probability

What is probability?

It is what everybody knows what it is before going at school

- how much we are confident that something is true

What is probability?

It is what everybody knows what it is before going at school

- how much we are confident that something is true
- how much we believe something

Probability

What is probability?

It is what everybody knows what it is before going at school

- how much we are confident that something is true
- how much we believe something
- “A measure of the degree of belief that an event *will* occur”

→ ‘will’ does not imply future, but only uncertainty

Or perhaps you prefer this way. . .

"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1] *. . . ,*

Or perhaps you prefer this way. . .

"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1] *. . . ,*

[1] *While in ordinary speech "to come true" usually refers to an event that is envisaged before it has happened, we use it here in the general sense, that the verbal description turns out to agree with actual facts.*

Or perhaps you prefer this way...

"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1] *..., the numerical probability* p *of this event is to be a real number by the indication of which we try* in some cases *to setup a quantitative measure of the strength of our conjecture or anticipation, founded on the said knowledge, that the event comes true"*

[1] *While in ordinary speech "to come true" usually refers to an event that is envisaged before it has happened, we use it here in the general sense, that the verbal description turns out to agree with actual facts.*

Or perhaps you prefer this way...

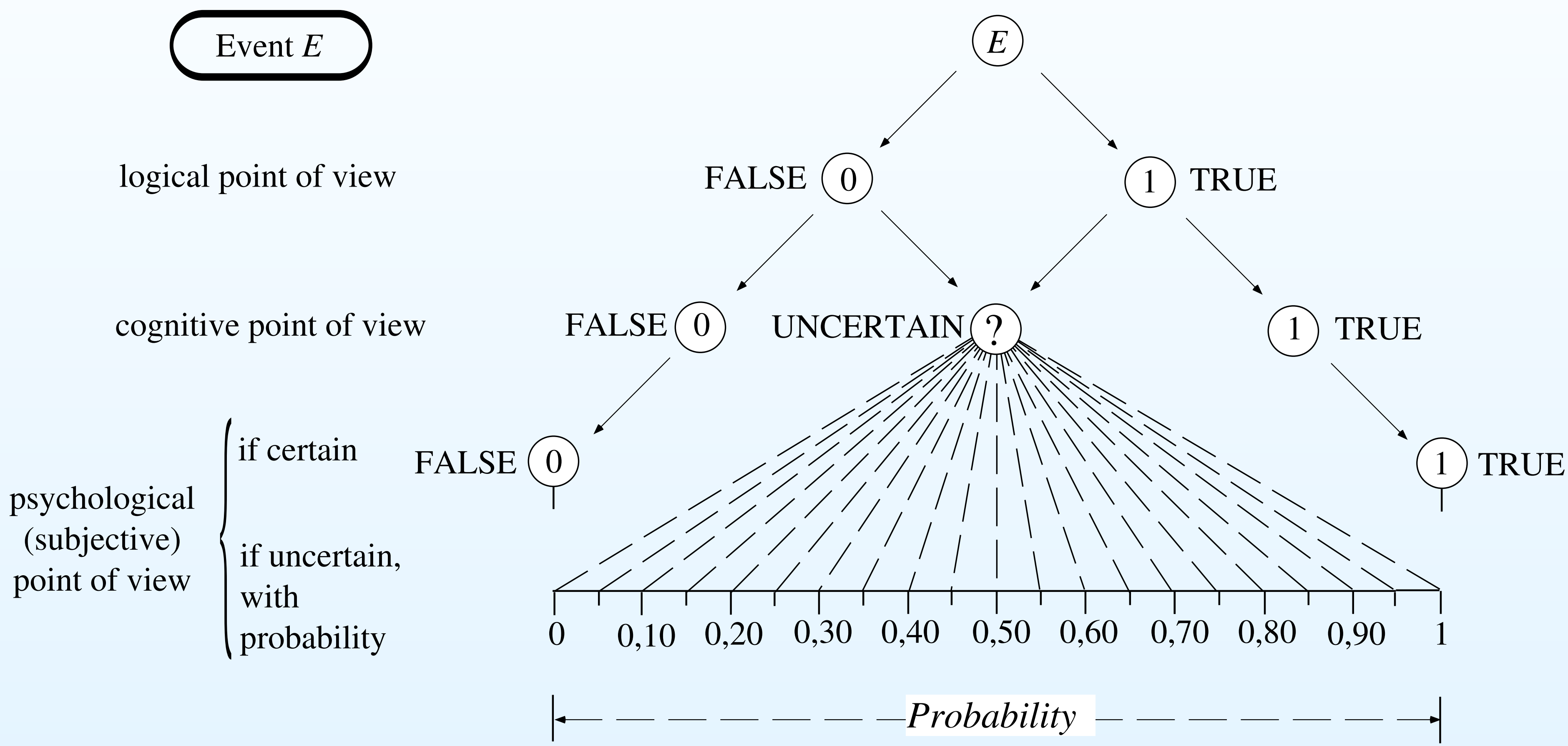
"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1] *..., the numerical probability p of this event is to be a real number by the indication of which we try* in some cases *to setup a quantitative measure of the strength of our conjecture or anticipation, founded on the said knowledge, that the event comes true"*
(E. Schrödinger, *The foundation of the theory of probability - I*, Proc. R. Irish Acad. 51A (1947) 51)

[1] *While in ordinary speech "to come true" usually refers to an event that is envisaged before it has happened, we use it here in the general sense, that the verbal description turns out to agree with actual facts.*

. . . or this other one

"In order to cope with this situation Weizsäcker has introduced the concept of 'degree of Truth.' For any simple statement in an alternative like 'The atom is in the left (or in the right) half of the box' a complex number is defined as a measure for its 'degree of Truth.' If the number is 1, it means that the statement is true; if the number is 0, it means that it is false. But other values are possible. The absolute square of the complex number gives the probability for the statement being true." (Heisenberg)

False, True and probable



Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

"If we were not ignorant there would be no probability, there could only be certainty. But our ignorance cannot be absolute, for then there would be no longer any probability at all. Thus the problems of probability may be classed. according to the greater or less depth of our ignorance." (Poincaré)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

The state of information can be different from subject to subject

$\Rightarrow$ intrinsic subjective nature.

- No negative meaning: only an acknowledgment that several persons might have different information and, therefore, necessarily different opinions.

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

The state of information can be different from subject to subject

$\Rightarrow$ intrinsic subjective nature.

- No negative meaning: only an acknowledgment that several persons might have different information and, therefore, necessarily different opinions.
- *"Since the knowledge may be different with different persons or with the same person at different times, they may anticipate the same event with more or less confidence, and thus different numerical probabilities may be attached to the same event"* (Schrödinger)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$\text{`}P(E)\text{'} \;\longrightarrow\; P(E\,|\,I) \;\longrightarrow\; P(E\,|\,I(t))$$

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$\text{‘}P(E)\text{’} \quad \longrightarrow \quad P(E\,|\,I) \quad \longrightarrow \quad P(E\,|\,I(t))$$

- *“Thus whenever we speak loosely of ‘the probability of an event,’ it is always to be understood: probability with regard to a certain given state of knowledge*” (Schrödinger)

Uncertainty $\longrightarrow$ probability

Probability is related to uncertainty and not (only) to the results of repeated experiments

Probability is always conditional probability

$$\text{`}P(E)\text{'} \quad \longrightarrow \quad P(E\,|\,I) \quad \longrightarrow \quad P(E\,|\,I(t))$$

- *"Thus whenever we speak loosely of 'the probability of an event,' it is always to be understood: probability with regard to a certain given state of knowledge"* (Schrödinger)
- Some examples:
  - box with 5 balls;
  - 'three box problems'.

End of lecture

End of lecture 2

Notes

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

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