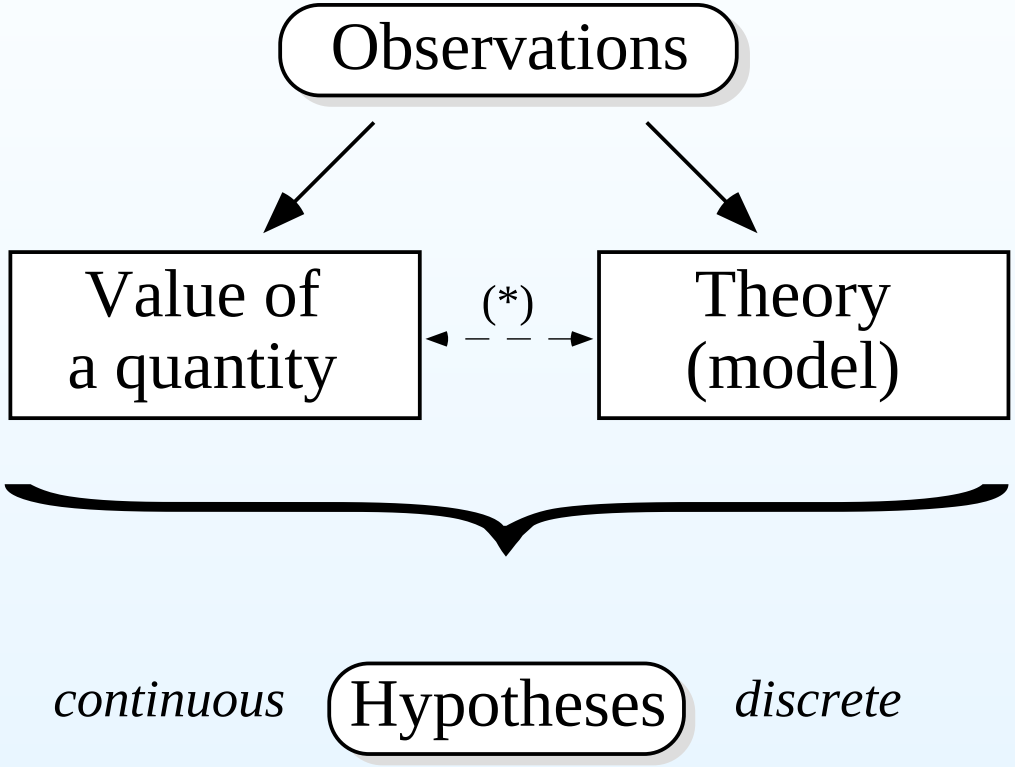
Dottorato 2015 – 1

Giulio D'Agostini

Università di Roma La Sapienza e INFN

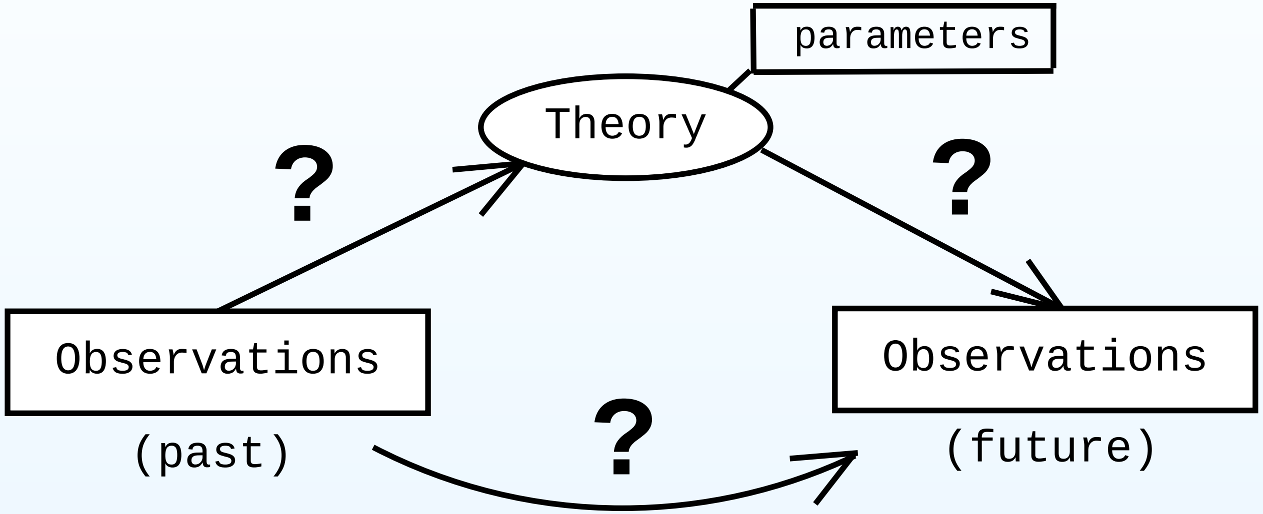
Roma, Italy

Physics



* A quantity might be meaningful only within a theory/model

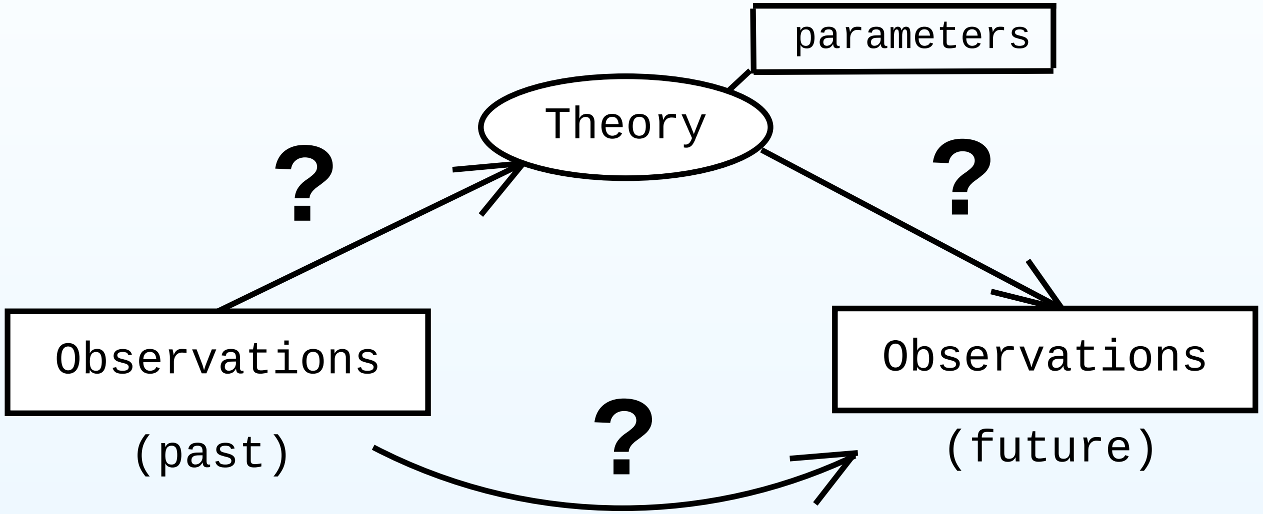
From the past to the future



Uncertainty:

Theory	— **?** ⟶	
Past observations	— **?** ⟶	
Theory	— **?** ⟶	Future observations

From the past to the future



Uncertainty:

Theory — **?** $\longrightarrow$

Past observations — **?** $\longrightarrow$

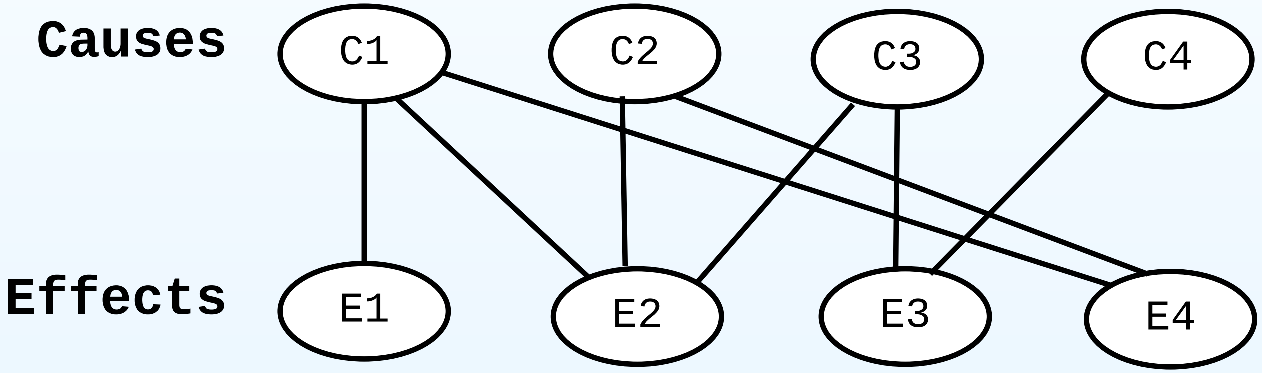
Theory — **?** $\longrightarrow$ Future observations

$\Longrightarrow$ **Uncertainty about causal connections**

CAUSE $\Longleftrightarrow$ **EFFECT**

Causes $\longrightarrow$ effects

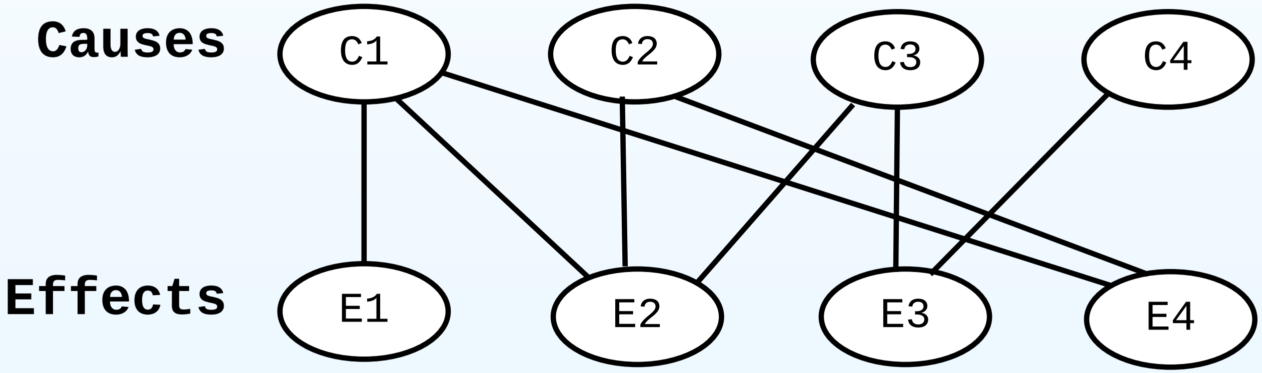
The same *apparent* cause might produce several,different effects



Given an observed effect, we are not sure about the exact cause that has produced it.

Causes → effects

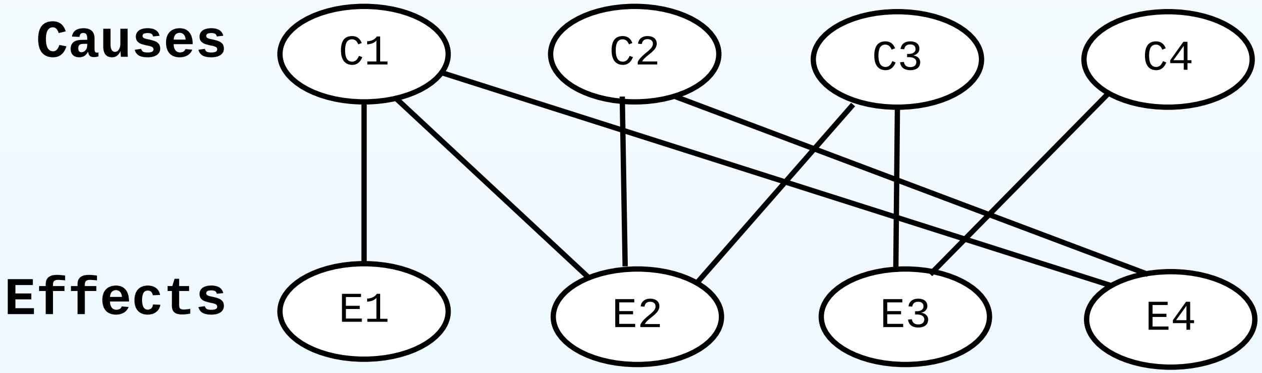
The same *apparent* cause might produce several,different effects



Given an observed effect, we are not sure about the exact cause that has produced it.

Causes $\longrightarrow$ effects

The same *apparent* cause might produce several,different effects



Given an observed effect, we are not sure about the exact cause that has produced it.

$$\mathbf{E_2} \Rightarrow \{C_1,\, C_2,\, C_3\}?$$

The “essential problem” of the experimental method

“Now, these problems are classified as *probability of causes*, and are most interesting of all their scientific applications. I play at *écarté* with a gentleman whom I know to be perfectly honest. What is the chance that he turns up the king? It is 1/8. This is a problem of the *probability of effects*.

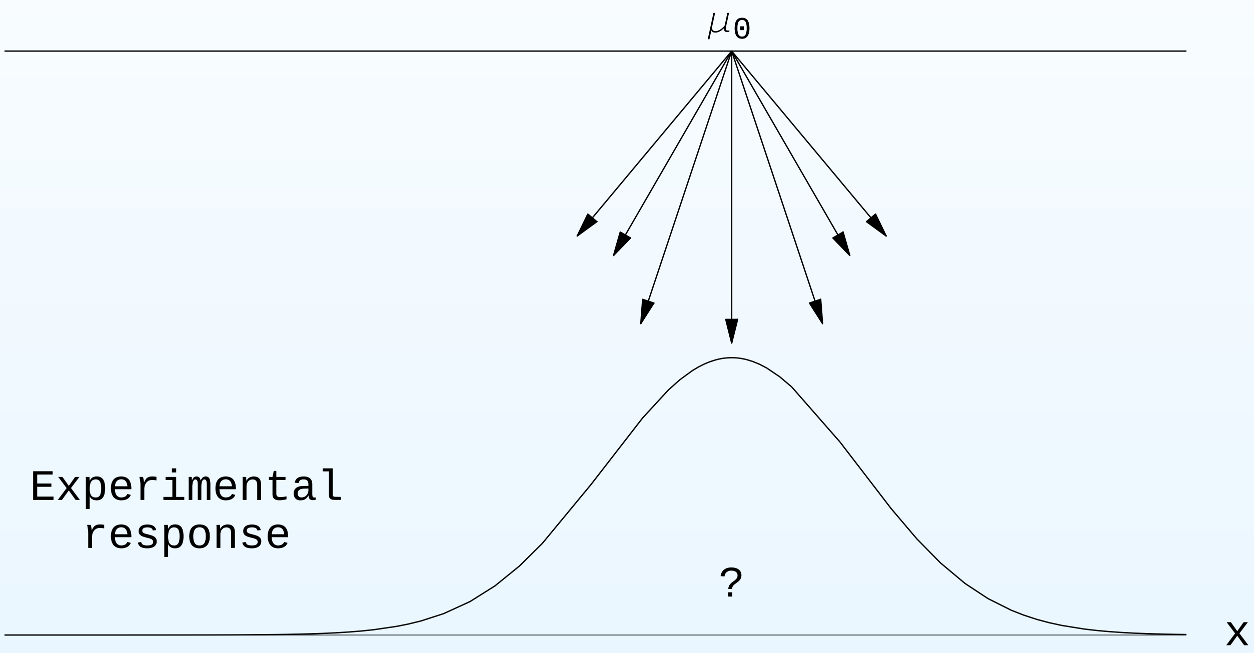
The “essential problem” of the experimental method

“Now, these problems are classified as *probability of causes*, and are most interesting of all their scientific applications. I play at *écarté* with a gentleman whom I know to be perfectly honest. What is the chance that he turns up the king? It is 1/8. This is a problem of the *probability of effects*.

I play with a gentleman whom I do not know. He has dealt ten times, and he has turned the king up six times. What is the chance that he is a sharper? This is a problem in the probability of causes. It may be said that it is the essential problem of the experimental method.”

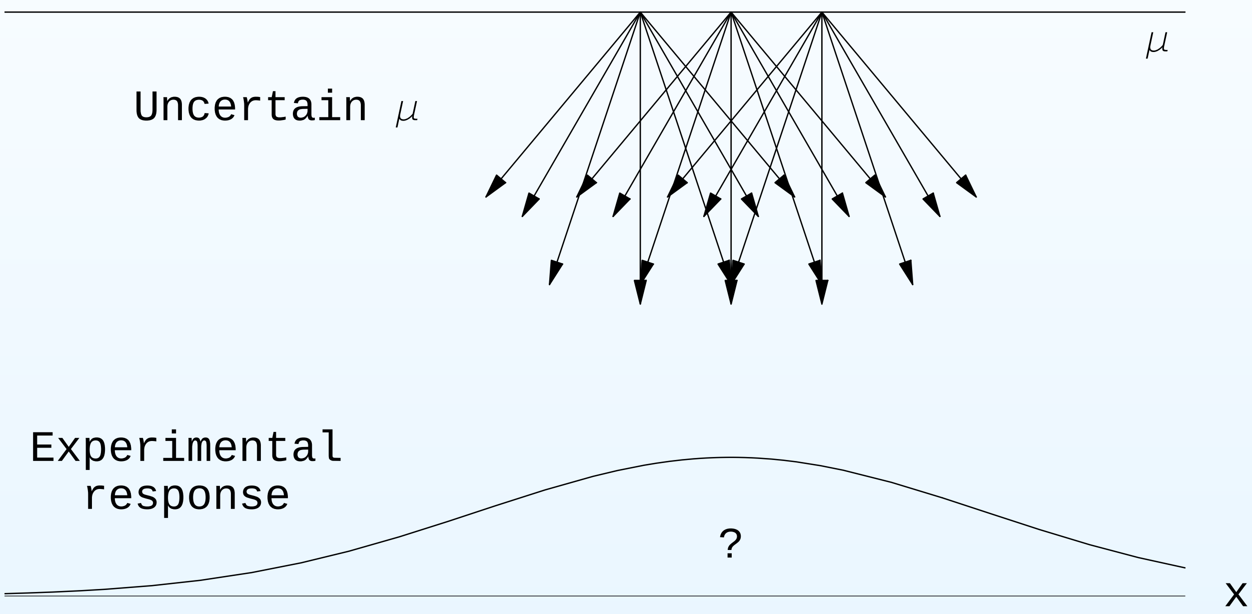
(H. Poincaré – *Science and Hypothesis*)

From ‘true value’ to observations



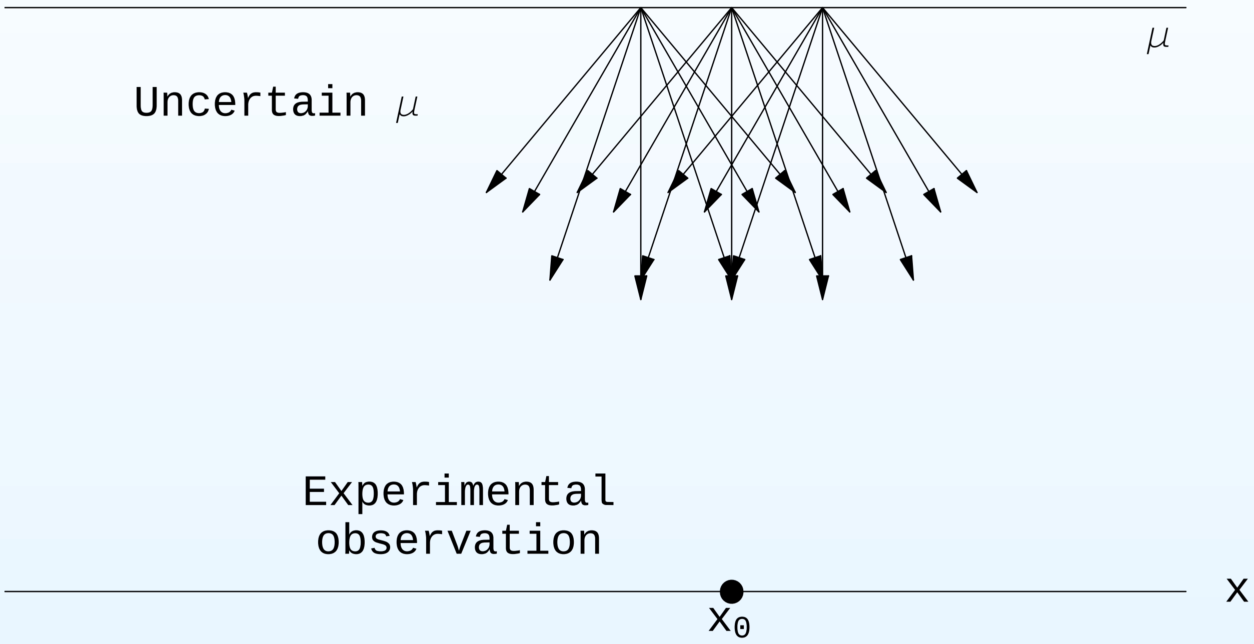
Given μ (exactly known) we are uncertain about x

From ‘true value’ to observations



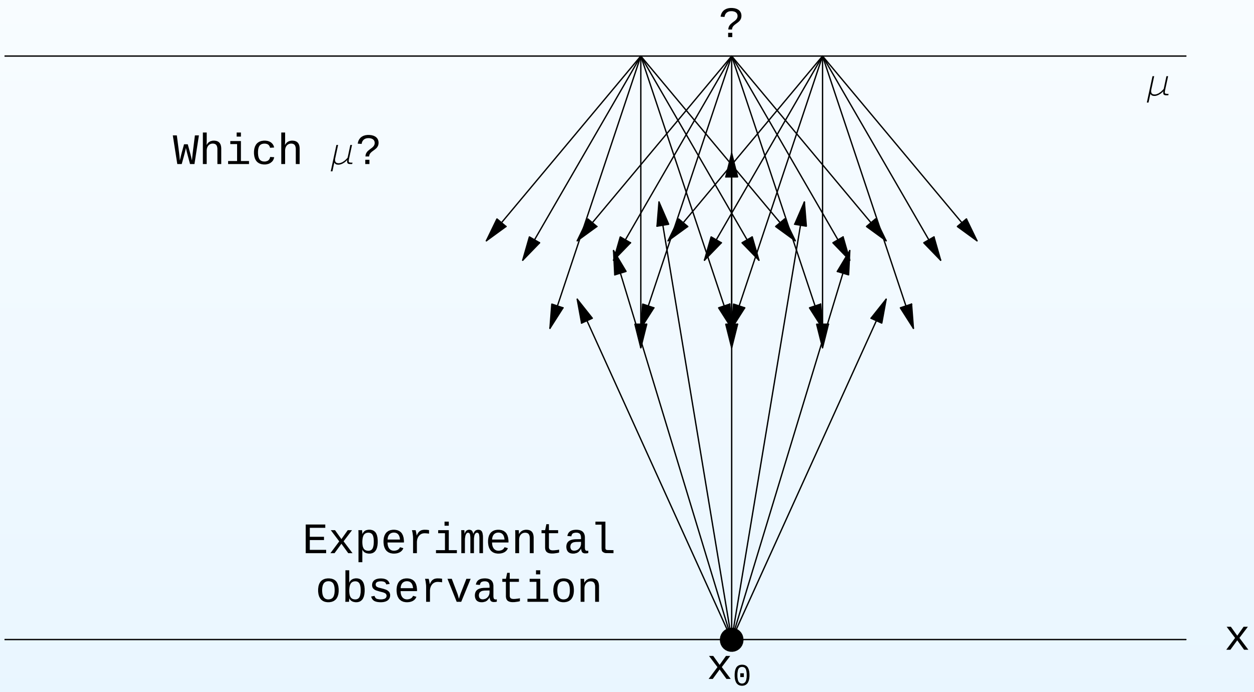
Uncertainty about μ makes us more uncertain about x

Inferring a true value



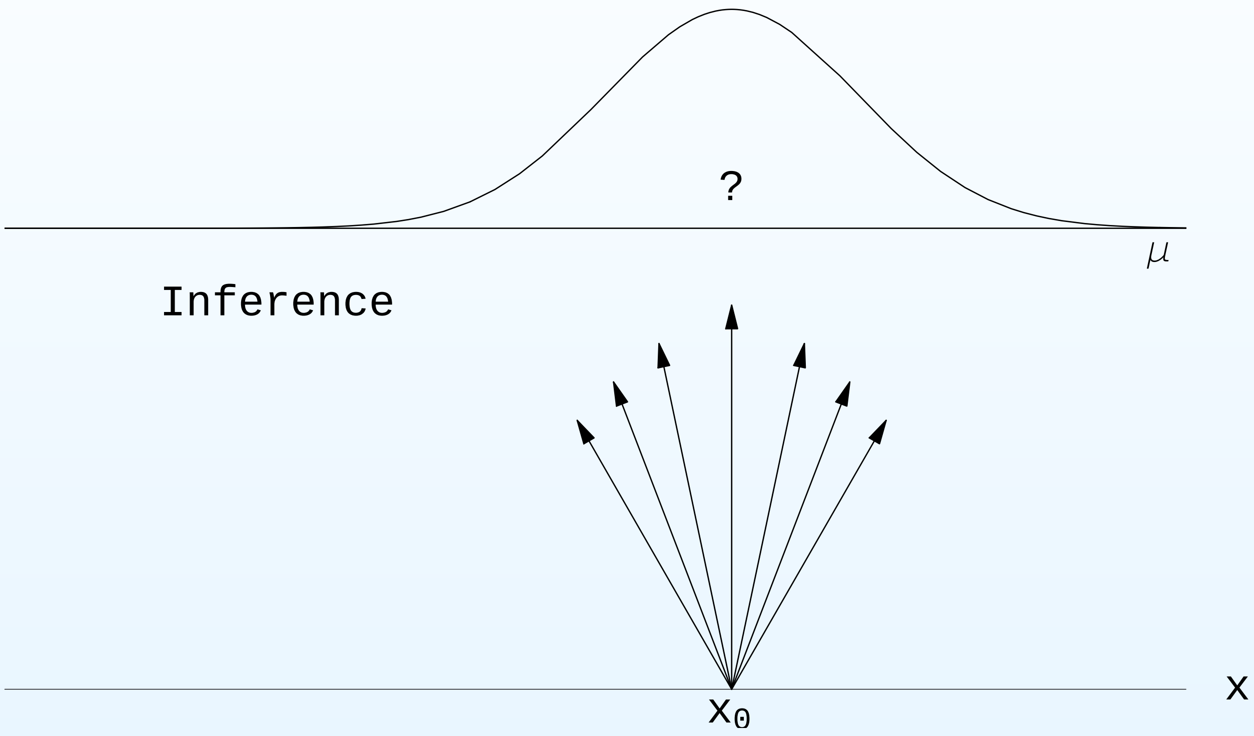
The observed data is <u>certain</u>: $\rightarrow$ 'true value' uncertain.

Inferring a true value



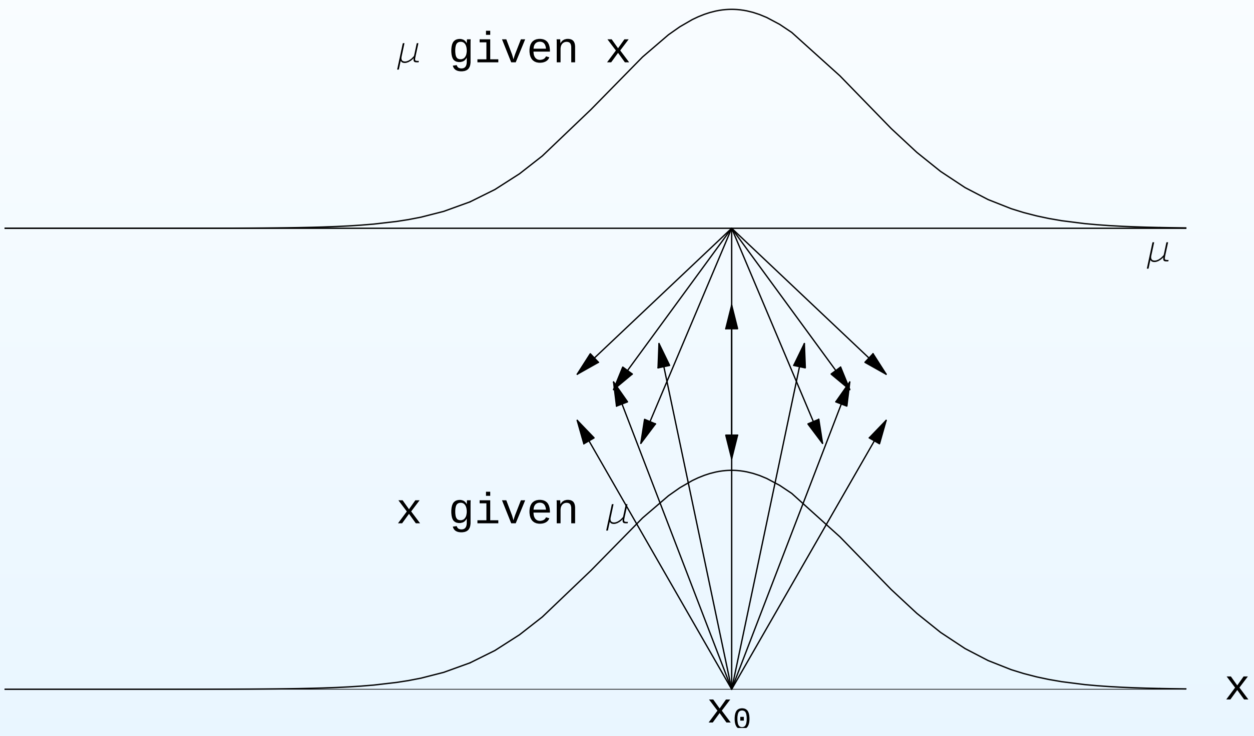
Where does the observed value of x comes from?

Inferring a true value



We are now uncertain about μ, given x.

Inferring a true value



Note the symmetry in reasoning.

Uncertainty

The human mind is used to live — and survive — in conditions of uncertainty and has developed mental categories to handle it.

Uncertainty

The human mind is used to live — and survive — in conditions of uncertainty and has developed mental categories to handle it.

As a matter of fact, although we are in a constant state of uncertainty about many events which might or might not occur,

- we can be "more or less *sure* — or *confident* — on something than on something else";
- "we consider something more or less *probable* (or *likely*)";
- or "we *believe* something more or less than something else".

Uncertainty

The human mind is used to live — and survive — in conditions of uncertainty and has developed mental categories to handle it.

As a matter of fact, although we are in a constant state of uncertainty about many events which might or might not occur,

- we can be "more or less *sure* — or *confident* — on something than on something else";
- "we consider something more or less *probable* (or *likely*)";
- or "we *believe* something more or less than something else".

We can use similar expressions, all referring to the intuitive idea of probability.

Uncertainty and probability

We, as physicists, consider absolutely natural and meaningful statements of the following kind

- $P(-10 < \epsilon'/\epsilon \times 10^4 < 50) >> P(\epsilon'/\epsilon \times 10^4 > 100)$
- $P(170 \le m_{top}/\text{GeV} \le 180) \approx 70\%$
- $P(M_H < 200\,\text{GeV}) > P(M_H > 200\,\text{GeV})$

(Statements of 2004!)

Uncertainty and probability

We, as physicists, consider absolutely natural and meaningful statements of the following kind

- $P(-10 < \epsilon'/\epsilon \times 10^4 < 50) >> P(\epsilon'/\epsilon \times 10^4 > 100)$
- $P(170 \le m_{top}/\text{GeV} \le 180) \approx 70\%$
- $P(M_H < 200\,\text{GeV}) > P(M_H > 200\,\text{GeV})$

(Statements of 2004!)

…thus, such statements are considered blaspheme to statistics gurus

Doing Natural Science in conditions of uncertainty

The constant status of uncertainty does not prevent us from doing Science (in the sense of Natural Science and not just Mathematics)

Doing Natural Science in conditions of uncertainty

The constant status of uncertainty does not prevent us from doing Science (in the sense of Natural Science and not just Mathematics)

Indeed

"It is scientific only to say what is more likely and what is less likely" (Feynman)

How to quantify all that?

(Let us start with the issue of ‘hypothesis tests’)

How to quantify all that?

(Let us start with the issue of ‘hypothesis tests’)

- **Falsificationist approach**

  [and statistical variations over the theme].

How to quantify all that?

(Let us start with the issue of ‘hypothesis tests’)

- **Falsificationist approach**

  [and statistical variations over the theme].

- **Probabilistic approach**

  [In the sense that probability theory is used throughly]

From Falsificationism to p-values

- Proof by contradiction of standard logic;

$\rightarrow$ Extension to the experimental sciences (impossible $\rightarrow$ impossible;

$\rightarrow$ P-values;

$\rightarrow$ Fake claims of discoveries
(Much money from tax payes spend in the experiment, wrong conclusions due to trivial mistakes in logic – quite sad...)

Example: Has the student made a mistake?

Homework: calculate the average of 300 random numbers, uniformly distributed between 0 and 1.

Example: Has the student made a mistake?

Homework: calculate the average of 300 random numbers, uniformly distributed between 0 and 1.

- Teacher expectation:

$$
\begin{aligned}
\mathsf{E}\left[\overline{X}_{300}\right] &= \frac{1}{2} \\
\sigma\left[\overline{X}_{300}\right] &= \frac{1}{\sqrt{12}} \cdot \frac{1}{\sqrt{300}} = 0.017\,,
\end{aligned}
$$

Example: Has the student made a mistake?

Homework: calculate the average of 300 random numbers, uniformly distributed between 0 and 1.

- Teacher expectation:

$$\begin{aligned}
\mathrm{E}\left[\overline{X}_{300}\right] &= \frac{1}{2} \\
\sigma\left[\overline{X}_{300}\right] &= \frac{1}{\sqrt{12}} \cdot \frac{1}{\sqrt{300}} = 0.017\,,
\end{aligned}$$

- 99% probability interval

$$P(0.456 \le \overline{X}_{300} \le 0.544) = 99\%\,.$$

Example: Has the student made a mistake?

Homework: calculate the average of 300 random numbers, uniformly distributed between 0 and 1.

- Teacher expectation:

$$\begin{aligned} \mathrm{E}\left[\overline{X}_{300}\right] &= \frac{1}{2} \\ \sigma\left[\overline{X}_{300}\right] &= \frac{1}{\sqrt{12}} \cdot \frac{1}{\sqrt{300}} = 0.017\,, \end{aligned}$$

- 99% probability interval

$$P(0.456 \le \overline{X}_{300} \le 0.544) = 99\%\,.$$

- Student gets a value outside the interval, e.g. $\overline{x} = 0.550$.

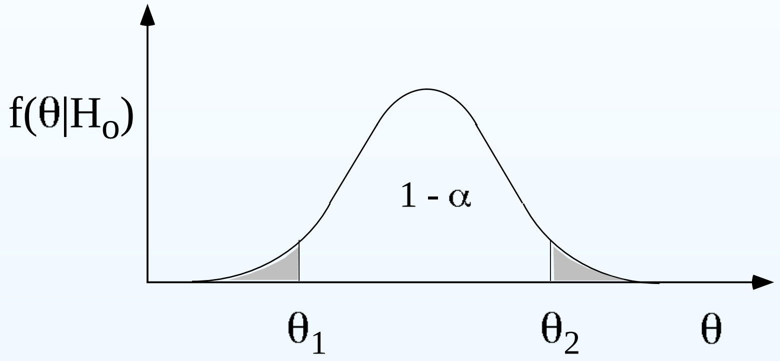
$\Rightarrow$ Has the student made a mistake?

Example: Has the student made a mistake?

Conventional statistician solution:
$\Rightarrow$ test the hypothesis $H_0 =$ ‘no mistakes’

Example: Has the student made a mistake?

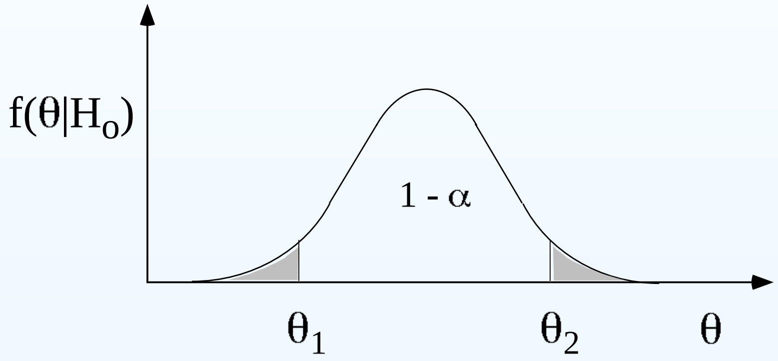
Conventional statistician solution:
$\Rightarrow$ test the hypothesis $H_0 =$ 'no mistakes'



- Test variable θ is $\overline{X}_{300}$.
- Acceptance interval $[\theta_1, \theta_2]$ is $[0.456, 0.544]$.
  We are 99% confident that $\overline{X}_{300}$ will fall inside it:
  $\rightarrow \alpha = 1\%$.

Example: Has the student made a mistake?

Conventional statistician solution:
$\Rightarrow$ test the hypothesis $H_0 =$ 'no mistakes'

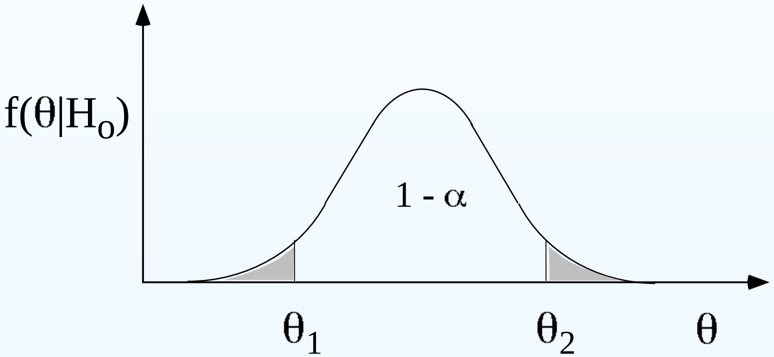


- Test variable θ is $\overline{X}_{300}$.
- Acceptance interval $[\theta_1, \theta_2]$ is $[0.456, 0.544]$.
  We are 99% confident that $\overline{X}_{300}$ will fall inside it:
  $\rightarrow \alpha = 1\%$.
- $\overline{x} = 0.550$ lies outside the acceptance interval

$\Rightarrow$ Hypothesis H_0 is rejected at 1% significance.

Example: Has the student made a mistake?

Conventional statistician solution:
$\Rightarrow$ test the hypothesis $H_0 =$ 'no mistakes'



- Test variable θ is $\overline{X}_{300}$.
- Acceptance interval $[\theta_1, \theta_2]$ is $[0.456, 0.544]$.
  We are 99% confident that $\overline{X}_{300}$ will fall inside it:
  $\rightarrow \alpha = 1\%$.
- $\overline{x} = 0.550$ lies outside the acceptance interval

$\Rightarrow$ Hypothesis H_0 is rejected at 1% significance.

$\Rightarrow$ What does it mean?

Meaning of the hypothesis test

Conclusion from test:

"the hypothesis $H_\circ$ = 'no mistakes' is rejected at the 1% level of significance".

Meaning of the hypothesis test

Conclusion from test:

"the hypothesis $H_\circ$ *= 'no mistakes' is rejected at the 1% level of significance".*

What does it mean?

"there is only a 1% probability that the average falls outside the selected interval, if the calculations were done correctly".

Meaning of the hypothesis test

Conclusion from test:

> *"the hypothesis $H_\circ$ = 'no mistakes' is rejected at the 1% level of significance".*

What does it mean?

> *"there is only a 1% probability that the average falls outside the selected interval, if the calculations were done correctly".*

So what?

Meaning of the hypothesis test

Conclusion from test:

> *"the hypothesis $H_\circ$ = 'no mistakes' is rejected at the 1% level of significance".*

What does it mean?

> *"there is only a 1% probability that the average falls outside the selected interval, if the calculations were done correctly".*

So what?

- It does not reply our natural question, i.e. that concerning the probability of mistake – quite impolite, by the way.
- The statement sounds as if one would be 99% sure that the student has made a mistake! (Mostly interpreted in this way).

$\Rightarrow$ Highly misleading!

Something is missing in the reasoning

If you ask the students (before they take a standard course in hypothesis tests) you will realize of a crucial ingredient extraneous to the logic of hypothesis tests:

Something is missing in the reasoning

If you ask the students (before they take a standard course in hypothesis tests) you will realize of a crucial ingredient extraneous to the logic of hypothesis tests:

“It all depends on whom has made the calculation!”

Something is missing in the reasoning

If you ask the students (before they take a standard course in hypothesis tests) you will realize of a crucial ingredient extraneous to the logic of hypothesis tests:

> *"It all depends on whom has made the calculation!"*

In fact, if the calculation was done by a well-tested program, the probability of mistake would be zero.
And students know rather well their tendency to do or not mistakes.

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, <u>but</u>

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, <u>but</u>

we cannot say

- “the value X has 0.27% probability of coming from that generator”

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, <u>but</u>

we cannot say

- “the value X has 0.27% probability of coming from that generator”
- “the probability that the observation is a statistical fluctuation is 0.27%”

'Something is missing': another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, but

we cannot say

- "the value X has 0.27% probability of coming from that generator"
- "the probability that the observation is a statistical fluctuation is 0.27%"

$\Rightarrow$ the value comes with 100% probability from that generator!

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, but

we cannot say

- “the value X has 0.27% probability of coming from that generator”
- “the probability that the observation is a statistical fluctuation is 0.27%”

$\Rightarrow$ the value comes with 100% probability from that generator!

$\Rightarrow$ it is at 100% a statistical fluctuation

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, but

we cannot say

- “the value X has 0.27% probability of coming from that generator”
- “the probability that the observation is a statistical fluctuation is 0.27%”

$\Rightarrow$ the value comes with 100% probability from that generator!

$\Rightarrow$ it is at 100% a statistical fluctuation

Logical bug of the reasoning:

$\Rightarrow$ One cannot tell how much one is confident in generator A only if another generator B is not taken into account.

‘Something is missing’: another example

The value $x = 3.01$ is extracted from a Gaussian random number generator having $\mu = 0$ and $\sigma = 1$.
It is well known that $P(|X| > 3) = 0.27\%$, but

we cannot say

- “the value X has 0.27% probability of coming from that generator”
- “the probability that the observation is a statistical fluctuation is 0.27%”

$\Rightarrow$ the value comes with 100% probability from that generator!

$\Rightarrow$ it is at 100% a statistical fluctuation

Logical bug of the reasoning:

$\Rightarrow$ This is the original sin of conventional hypothesis test methods

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 ⇔ 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!

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:

$$E_1 = \text{‘6’}$$

$$E_2 = \text{‘even number’}$$

$$E_3 = \text{‘}\geq 2\text{’}$$

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:

$$E_1 = \text{'6'}$$

$$E_2 = \text{'even number'}$$

$$E_3 = \text{'}\geq 2\text{'}$$

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:

$E_1 =$ '6'

$E_2 =$ 'even number'

$E_3 =$ '≥ 2'

$\Rightarrow$ 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')

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)

Hume’s point of view

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)

Hume’s point of view

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)

- Same as Laplace: World instrinsecally deterministic, but we are ignorant about exact donditions.

Hume's point of view

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)

- Same as Laplace: World instrinsecally deterministic, but we are ignorant about exact donditions.
- Big issues related to Quantum Mechanics!

Hume’s point of view

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)

- Same as Laplace: World instrinsecally deterministic, but we are ignorant about exact donditions.
- Big issues related to Quantum Mechanics!
- Personal opinion:

  “Probability does not exist”,
  Chance does!

Hume’s point of view

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)

- Same as Laplace: World instrinsecally deterministic, but we are ignorant about exact donditions.
- Big issues related to Quantum Mechanics!
- Personal opinion:

  “Probability does not exist”, Chance does!

  (It will be clear during the course)

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.

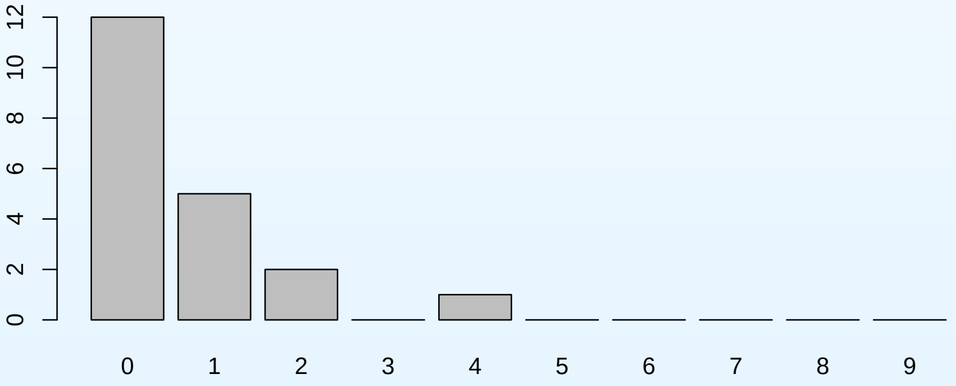
A counting experiment

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.



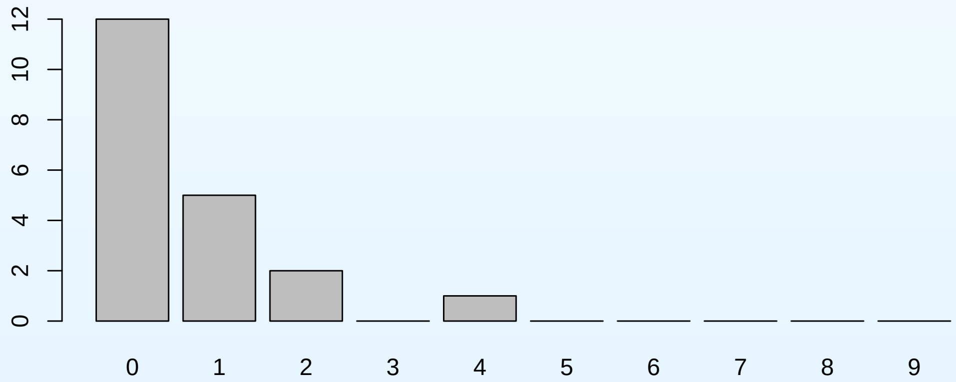
A counting experiment

Imagine a small scintillation counter, with suitable threshold, placed

here

now

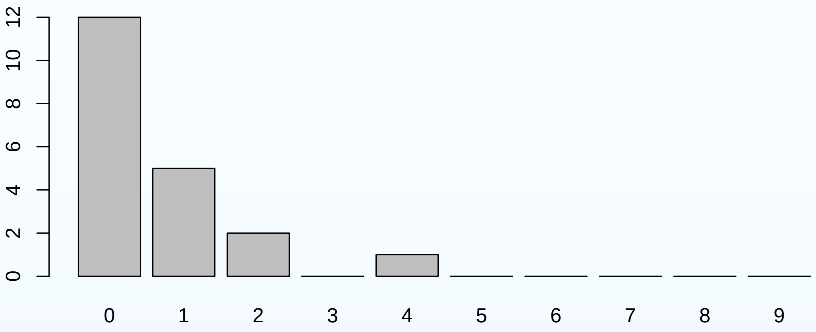
Fix the measuring time (e.g. 5 second each) and perform 20 measurements: 0, 0, 1, 0, 0, 0, 1, 2, 0, 0, 1, 1, 0.



Think at the 21^{st} measurement:

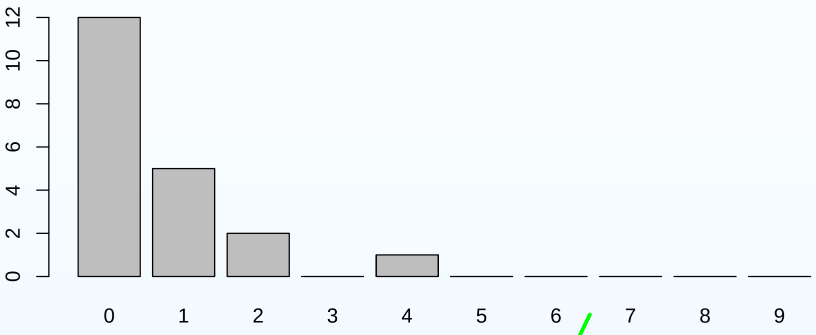
- Which outcome do you consider more likely? (0, 1, 2, 3, . . .)
- Why?

A counting experiment



$\Rightarrow$ Next ?

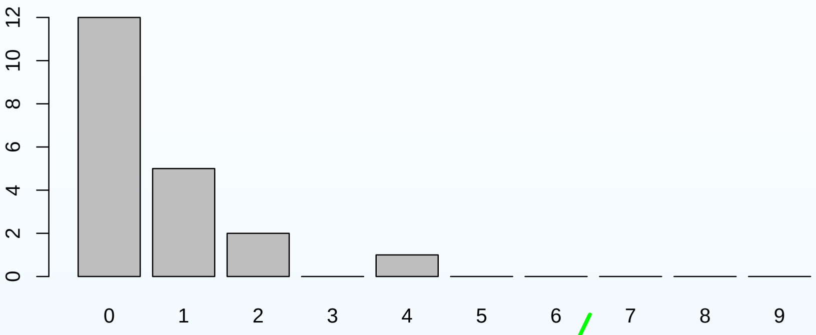
A counting experiment



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

$\Rightarrow$ Next ?

A counting experiment

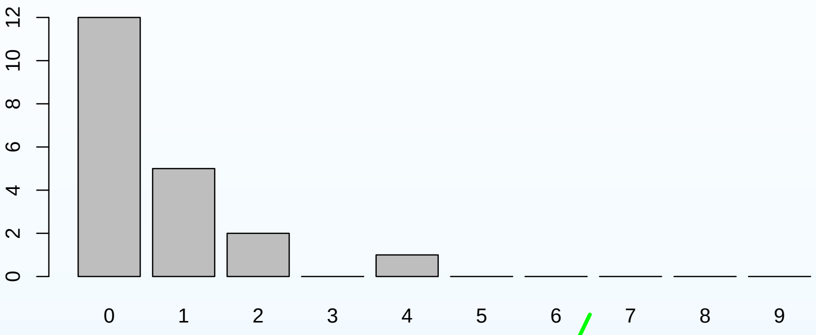


$\Rightarrow$ Next ?

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

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

A counting experiment



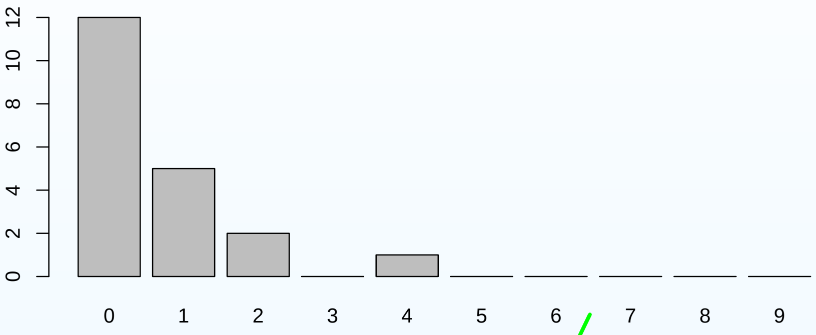
$\Rightarrow$ Next ?

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

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

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

A counting experiment



$\Rightarrow$ Next ?

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

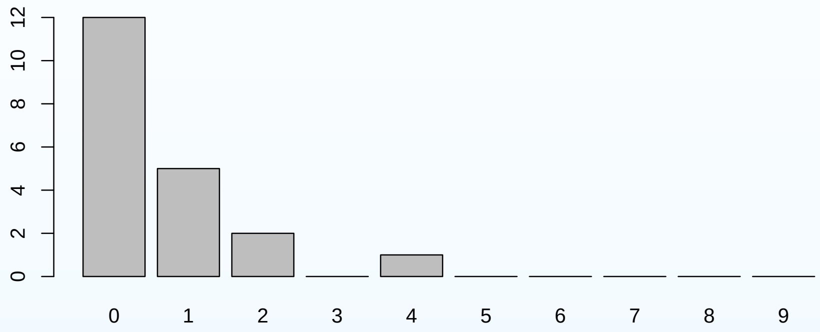
$P(3) < P(4)$, or $P(3) \geq P(4)$?

$P(3) = 0$, 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).

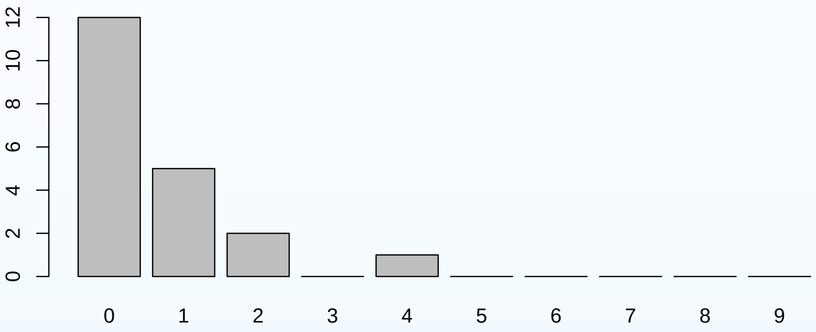
A counting experiment



$\Rightarrow$ Next ?

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

A counting experiment



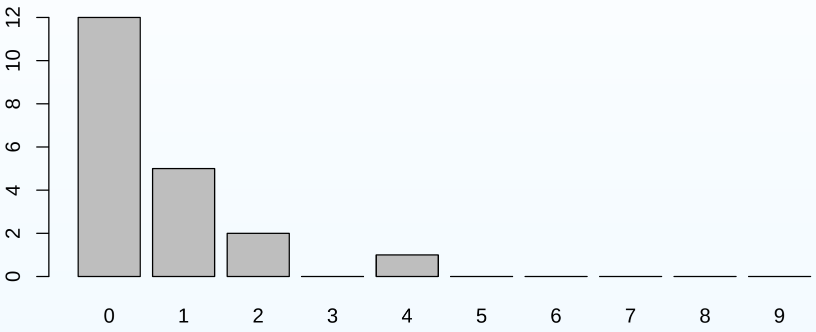
$\Rightarrow$ Next ?

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'

'guess'

'postulate'

'believe'

regularity of nature

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.

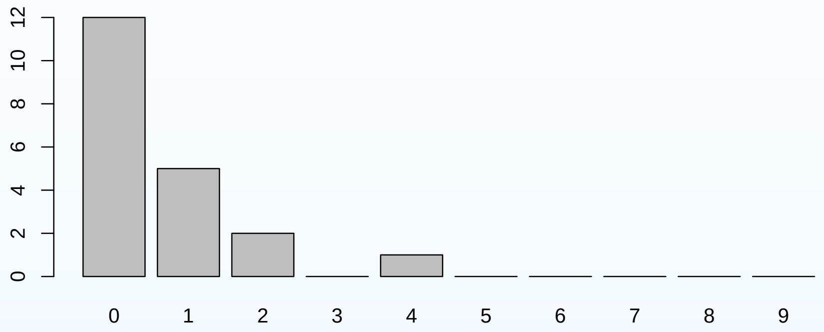
The train is going slowly and they see a cow walking along a country road parallel to the railway.

- 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!.

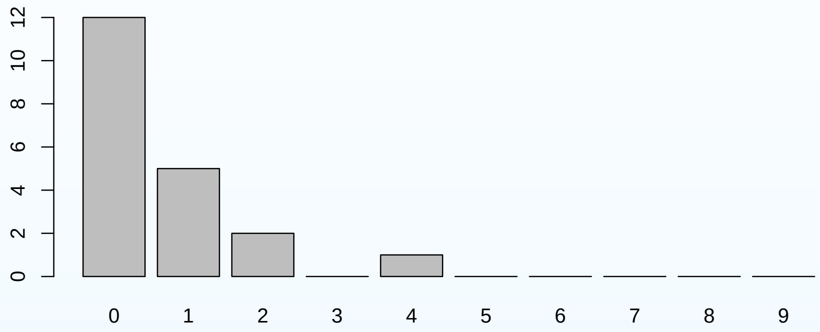
Transferring past to future



$\Rightarrow$ 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.

Transferring past to future

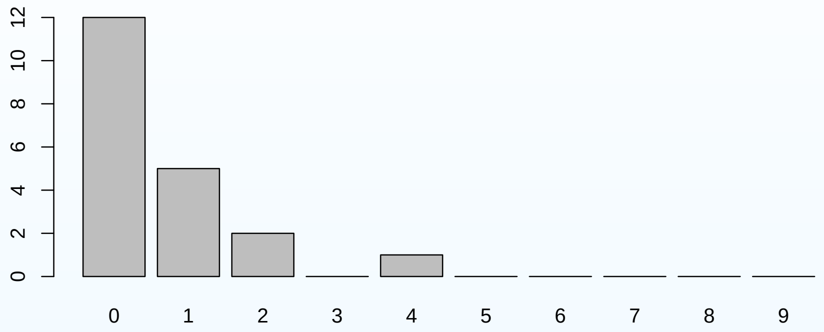


$\Rightarrow$ Next ?

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: *"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."*

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.

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:

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 $f(x \mid \mu)$.

Uncertainties in measurements

Having to perform a measurement:

Which numbers shall come out from our device?

Having performed a measurement:

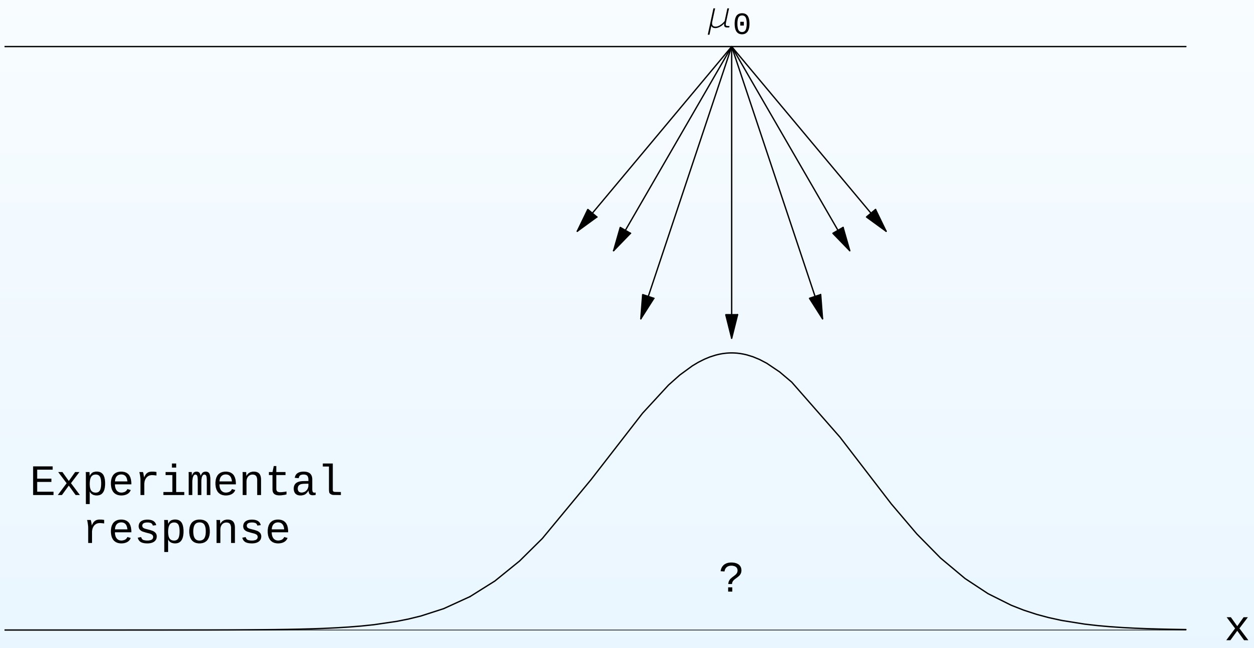
What have we learned about the value of the quantity of interest?

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 $f(x \,|\, \mu)$.

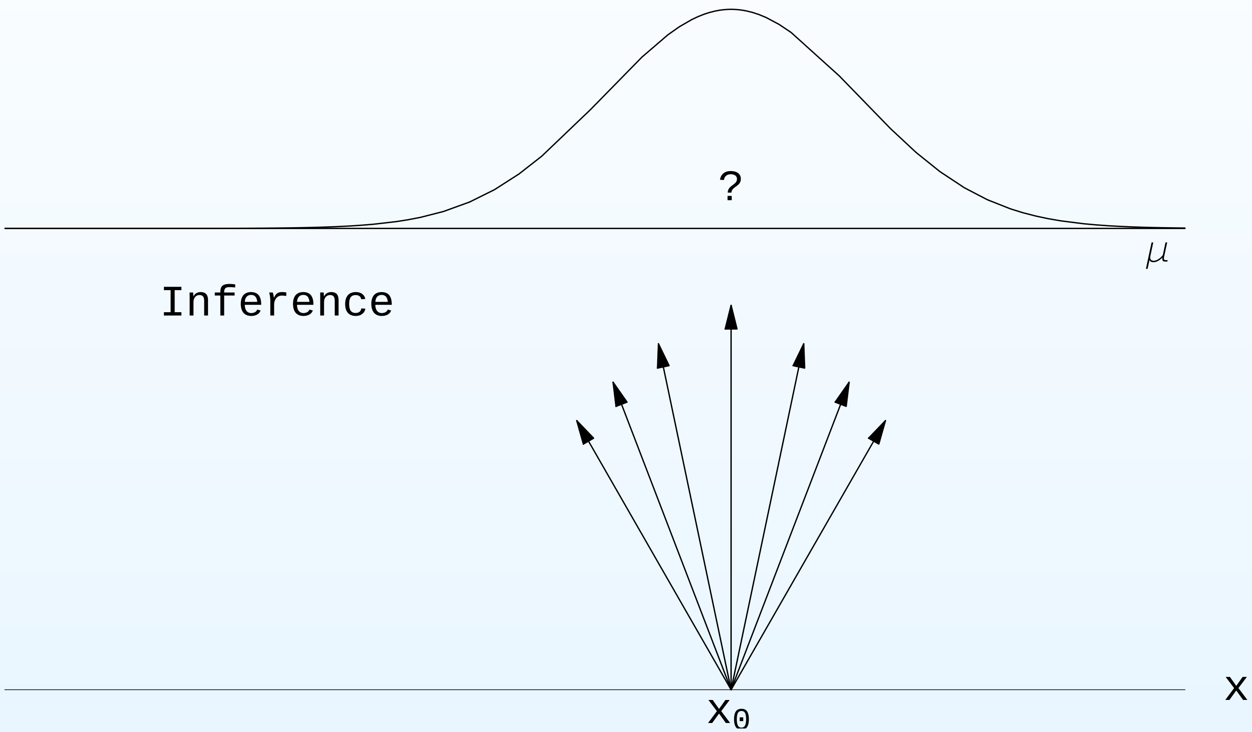
There is (in most cases) no way to get *directly* hints about $f(\mu \,|\, x)$.

Uncertainties in measurements



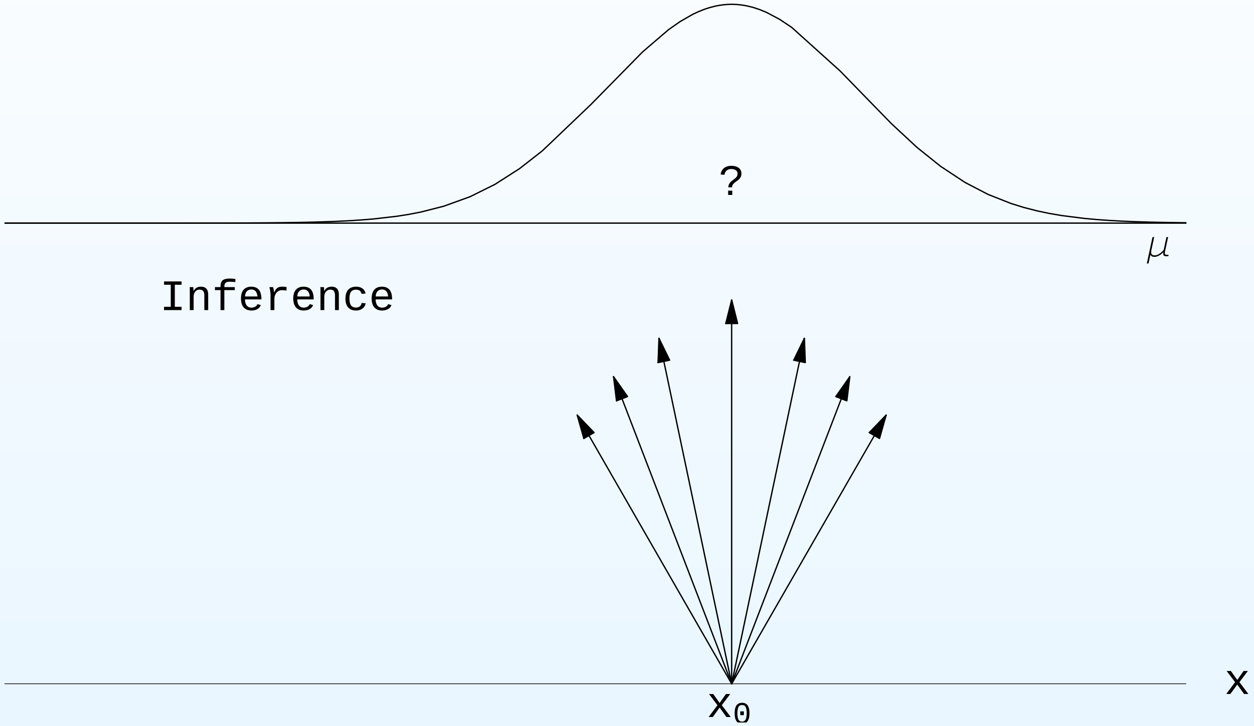
$f(x \mid \mu)$ experimentally accessible (though 'model filtered')

Uncertainties in measurements



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

Uncertainties in measurements

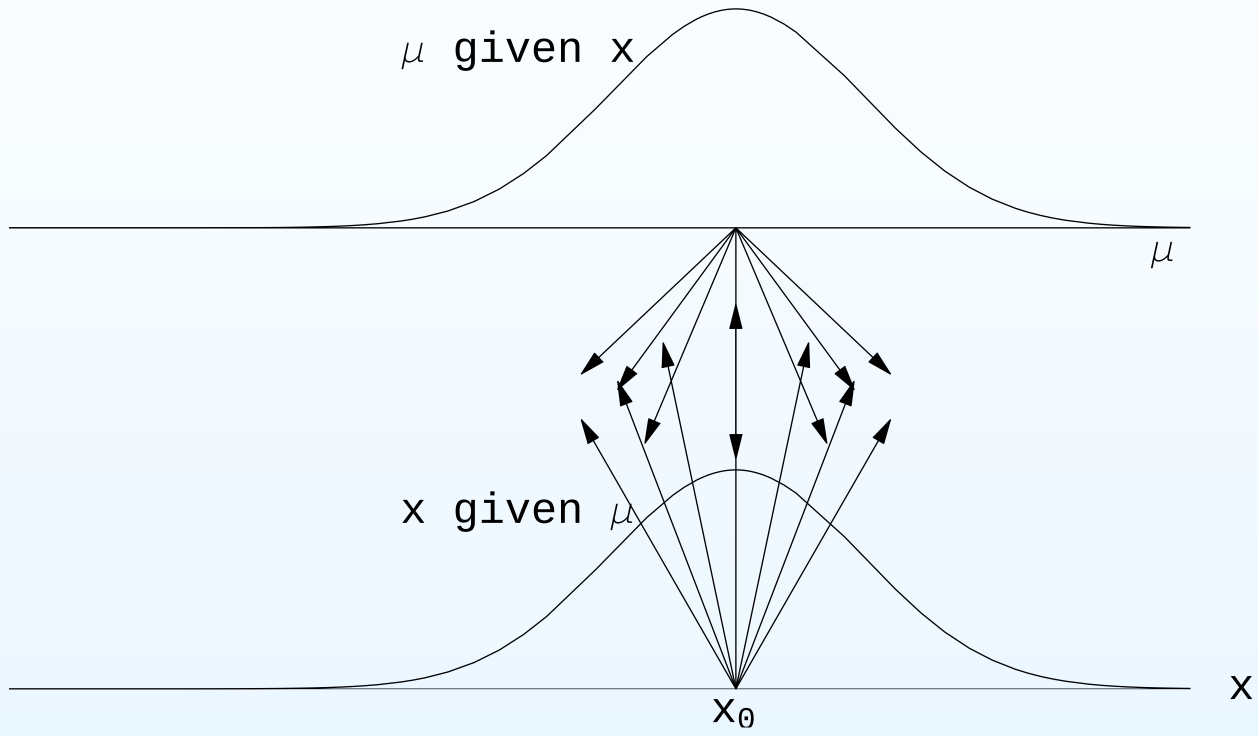


$f(\mu \,|\, x)$ experimentally inaccessible

but logically accessible!

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

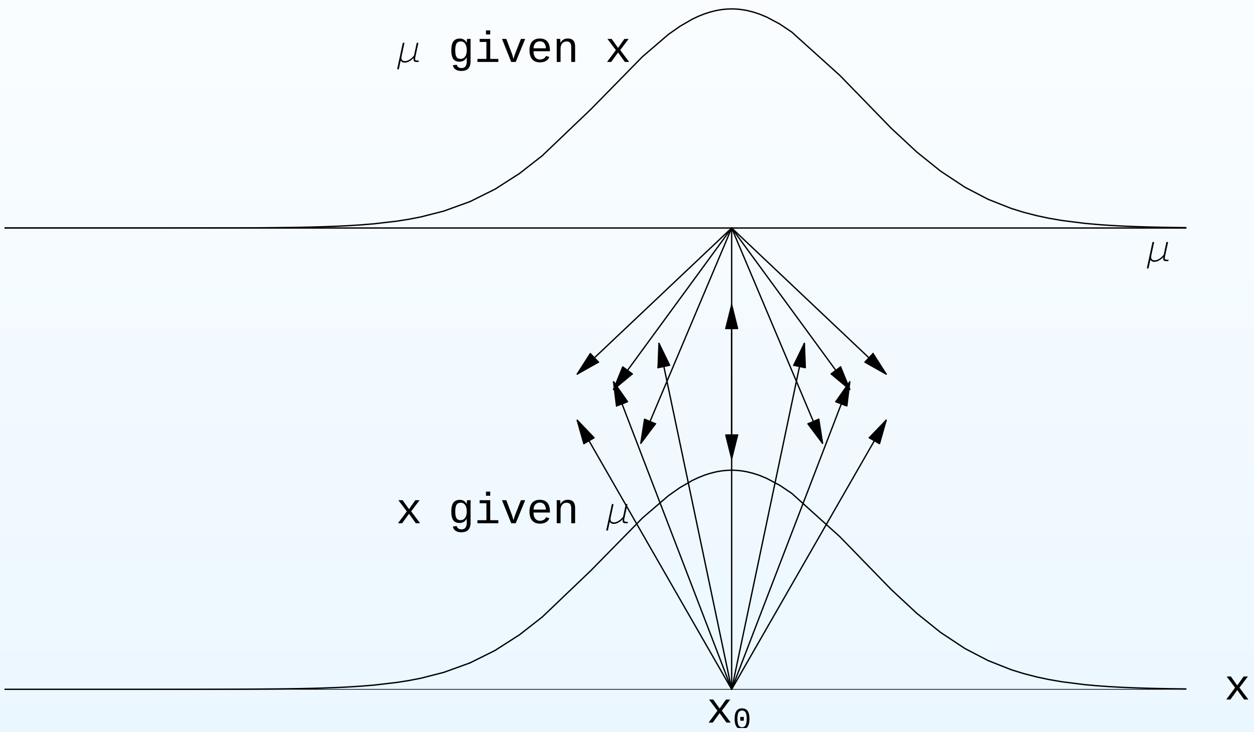
Uncertainties in measurements



- 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 $\longrightarrow$ 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.*

$\rightarrow$ *"statistical errors"*

Note

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

“Error” or “uncertainty”?

. . . almost synomimous among physicists,

“Error” or “uncertainty”?

…almost synomimous among physicists, although they aren’t.

“Error” or “uncertainty”?

. . . almost synomimous among physicists, although they aren’t.

ISO (BIPM, INST, etc.):

Uncertainty: *“a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurement.”*

“Error” or “uncertainty”?

…almost synomimous among physicists, although they aren’t.

ISO (BIPM, INST, etc.):

Uncertainty: *“a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurement.”*

Error: *“the result of a measurement minus a true value of the measurand.”*

“Error” or “uncertainty”?

. . . almost synomimous among physicists, although they aren’t.

ISO (BIPM, INST, etc.):

Uncertainty: *“a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurement.”*

Error: *“the result of a measurement minus a true value of the measurand.”*

??

“Error” or “uncertainty”?

. . . almost synomimous among physicists, although they aren’t.

ISO (BIPM, INST, etc.):

Uncertainty: *“a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurement.”*

Error: *“the result of a measurement minus a true value of the measurand.”*

True value: *“a value compatible with the definition of a given particular quantity.”*

“Error” or “uncertainty”?

... almost synomimous among physicists, although they aren’t.

ISO (BIPM, INST, etc.):

Uncertainty: *“a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurement.”*

Error: *“the result of a measurement minus a true value of the measurand.”*

True value: *“a value compatible with the definition of a given particular quantity.”*

$\Rightarrow$ Since we usually do not knoe *the* true value, we also do not know the error

$\rightarrow$ otherwise we would correct for it!!

“True value”

True value: *“a value compatible with the definition of a given particular quantity.”*

“True value”

True value: *“a value compatible with the definition of a given particular quantity.”*
This definition may seem vague, but it is more practical and pragmatic, and of more general use, than *“the value obtained after an infinite series of measurements performed under the same conditions with an instrument not affected by systematic errors.”*

“True value”

True value: *“a value compatible with the definition of a given particular quantity.”*

This definition may seem vague, but it is more practical and pragmatic, and of more general use, than *“the value obtained after an infinite series of measurements performed under the same conditions with an instrument not affected by systematic errors.”*

For instance, it holds also for quantities for which it is not easy to repeat the measurements, and even for those cases in which it makes no sense to speak about repeated measurements under the same conditions.

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

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

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? Objections?

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 $\rightarrow$ "several problems"

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

3 Frequentistic coverage $\longrightarrow$ "several problems"

- 'Trivial' interpretation problem: $\rightarrow$ 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: $\rightarrow$ 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’ $\rightarrow$ ‘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 $\rightarrow$ "several problems"

- 'Trivial' interpretation problem: $\rightarrow$ 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' $\rightarrow$ '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 $\rightarrow$ “several problems”

- ‘Trivial’ interpretation problem: $\rightarrow$ 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’ $\rightarrow$ ‘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 $\rightarrow$ "several problems"

- 'Trivial' interpretation problem: $\rightarrow$ 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' $\rightarrow$ '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 $\rightarrow$ “several problems”

- ‘Trivial’ interpretation problem: $\rightarrow$ 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’ $\rightarrow$ ‘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)

What had Neyman in mind?

"Carry out your experiment, calculate the confidence interval, and *state* that *c* belong to this interval. If you are asked whether you 'believe' that *c* belongs to the confidence interval you must refuse to answer. In the long run your assertions, if independent of each other, will be right in approximately a proportion α of cases" (J. Neyman)
(Quoted by I.J. Good in *Probability and the weighing of Evidence*)

What had Neyman in mind?

"Carry out your experiment, calculate the confidence interval, and *state* that *c* belong to this interval. If you are asked whether you 'believe' that *c* belongs to the confidence interval you must refuse to answer. In the long run your assertions, if independent of each other, will be right in approximately a proportion α of cases" (J. Neyman)
(Quoted by I.J. Good in *Probability and the weighing of Evidence*)

A 'confidence' that has nothing to do with how much we have to be confident on something!

What had Neyman in mind?

"Carry out your experiment, calculate the confidence interval, and *state* that c belong to this interval. If you are asked whether you 'believe' that c belongs to the confidence interval you must refuse to answer. In the long run your assertions, if independent of each other, will be right in approximately a proportion α of cases" (J. Neyman)
(Quoted by I.J. Good in *Probability and the weighing of Evidence*)

A 'confidence' that has nothing to do with how much we have to be confident on something!

If you name yourself SUPERMAN you do not become *ipso facto* a super-hero!

Arbitrary probability inversions

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

How do we turn, just 'intuitively'

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

into

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

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:

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

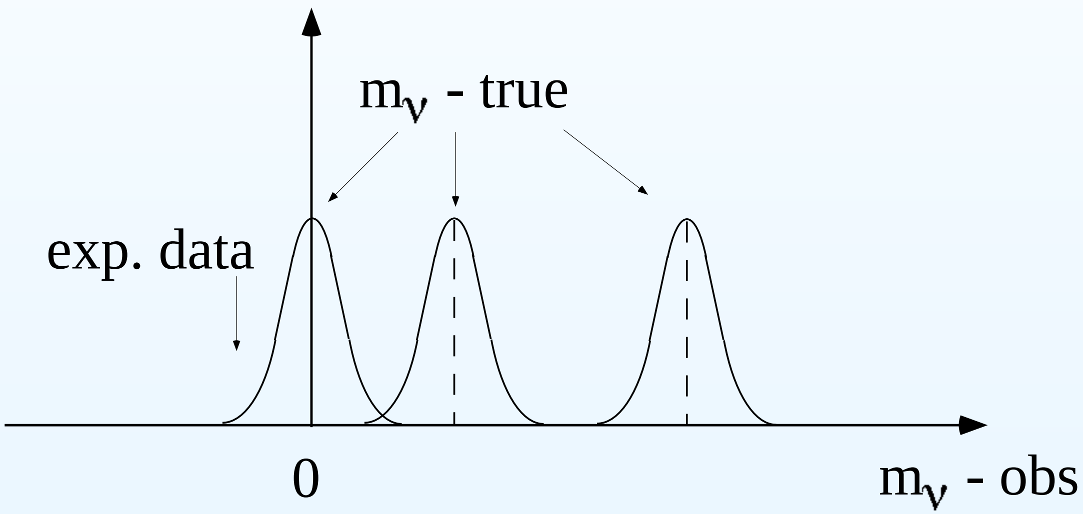
Easy to understand that this conclusion is based on some tacit assumptions:

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

$\rightarrow$ not always valid!

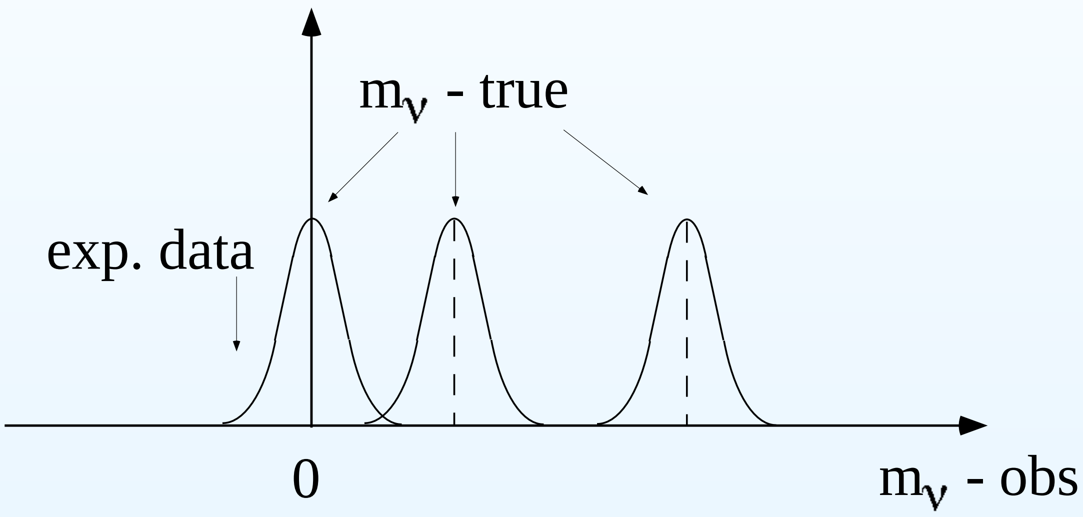
Measurement at the edge of a physical region

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



Measurement at the edge of a physical region

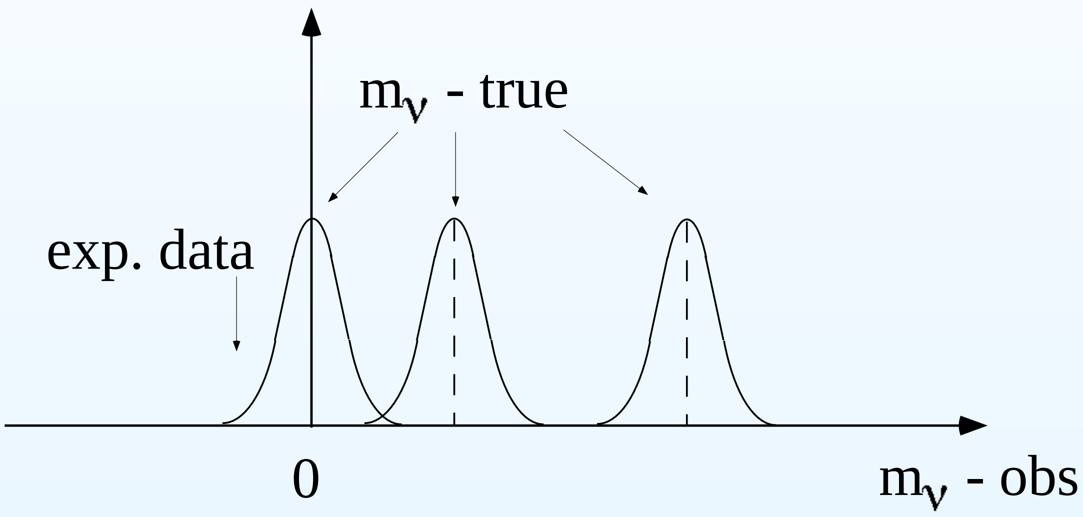
Electron-neutrino experiment, mass resolution $\sigma = 2$ eV, independent of m_ν.



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

Measurement at the edge of a physical region

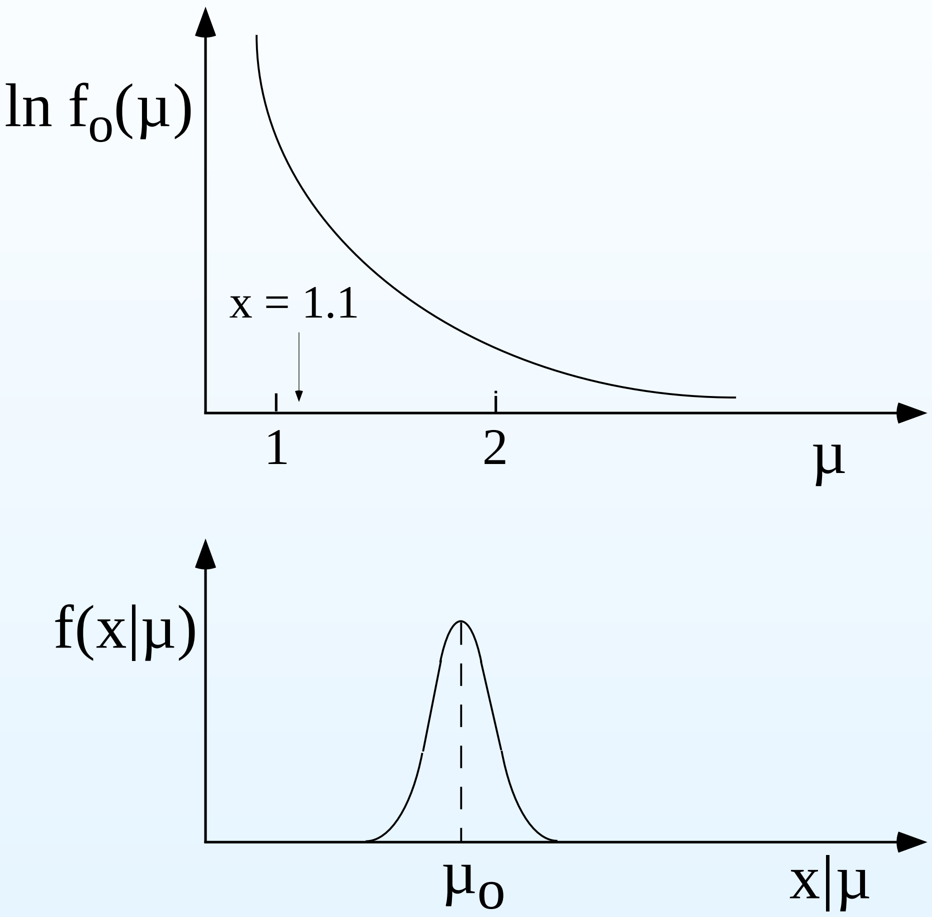
Electron-neutrino experiment, mass resolution $\sigma = 2$ eV, independent of m_ν.



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

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

Non-flat distribution of a physical quantity



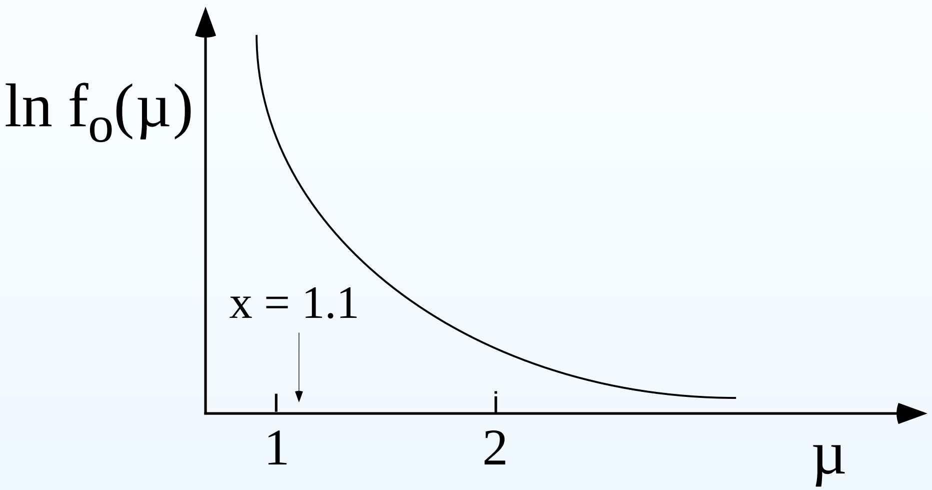
Imagine a cosmic ray particle or a bremsstrahlung γ.

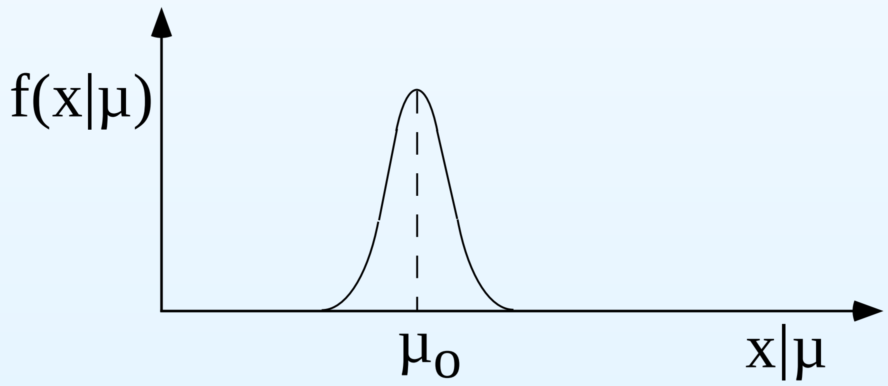
Non-flat distribution of a physical quantity

Imagine a cosmic ray particle or a bremsstrahlung γ.

Observed $x = 1.1$.

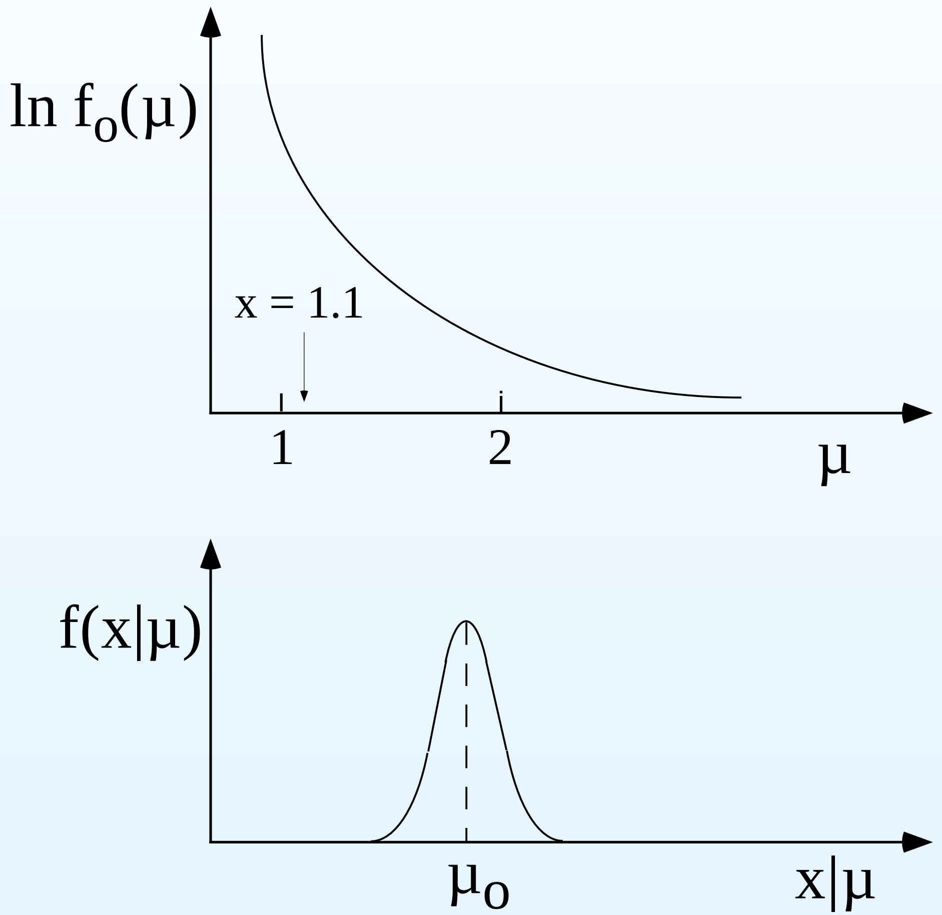
What can we say about the true value μ that has caused this observation?





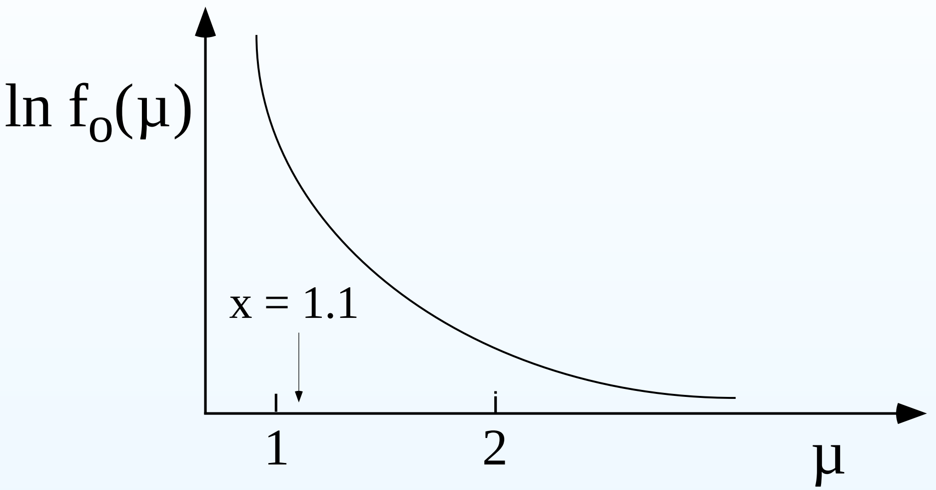
Non-flat distribution of a physical quantity

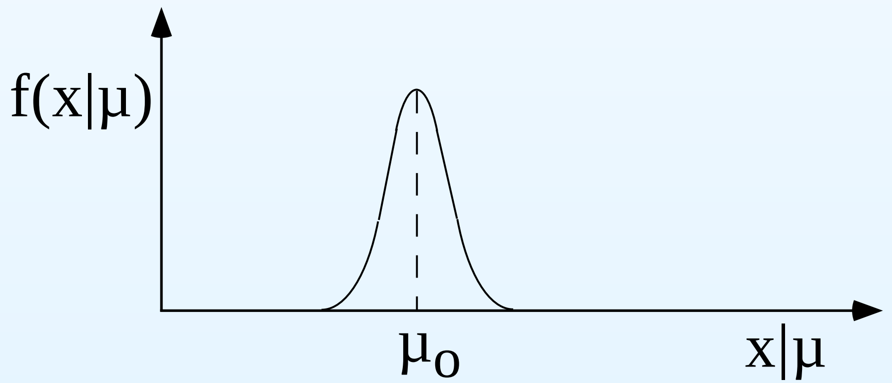
Imagine a cosmic ray particle or a bremsstrahlung γ.
Observed $x = 1.1$.
What can we say about the true value μ that has caused this observation?
Also in this case the formal definition of the confidence interval does not work.
Intuitively, we feel that there is more chance that μ is on the left of 1.1 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.



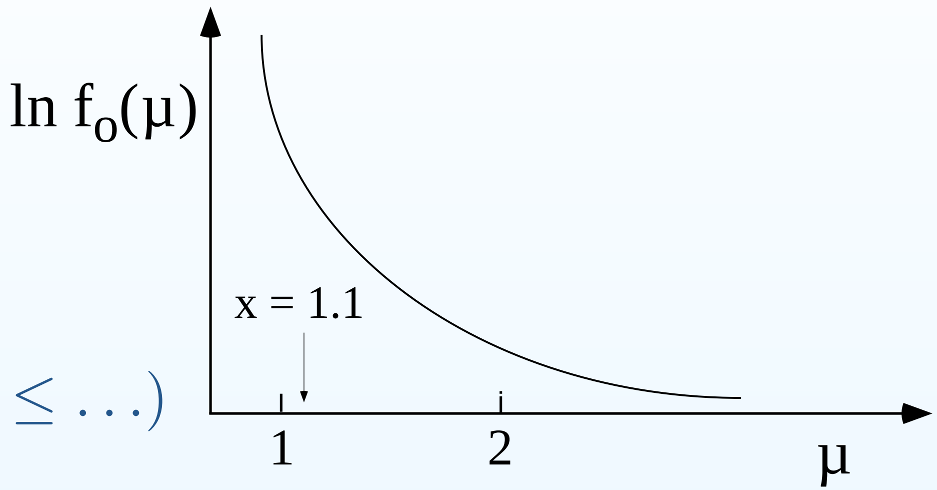


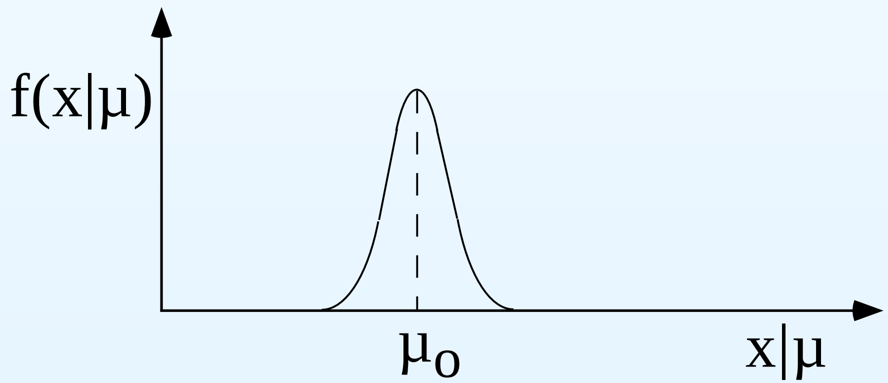
Non-flat distribution of a physical quantity

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).

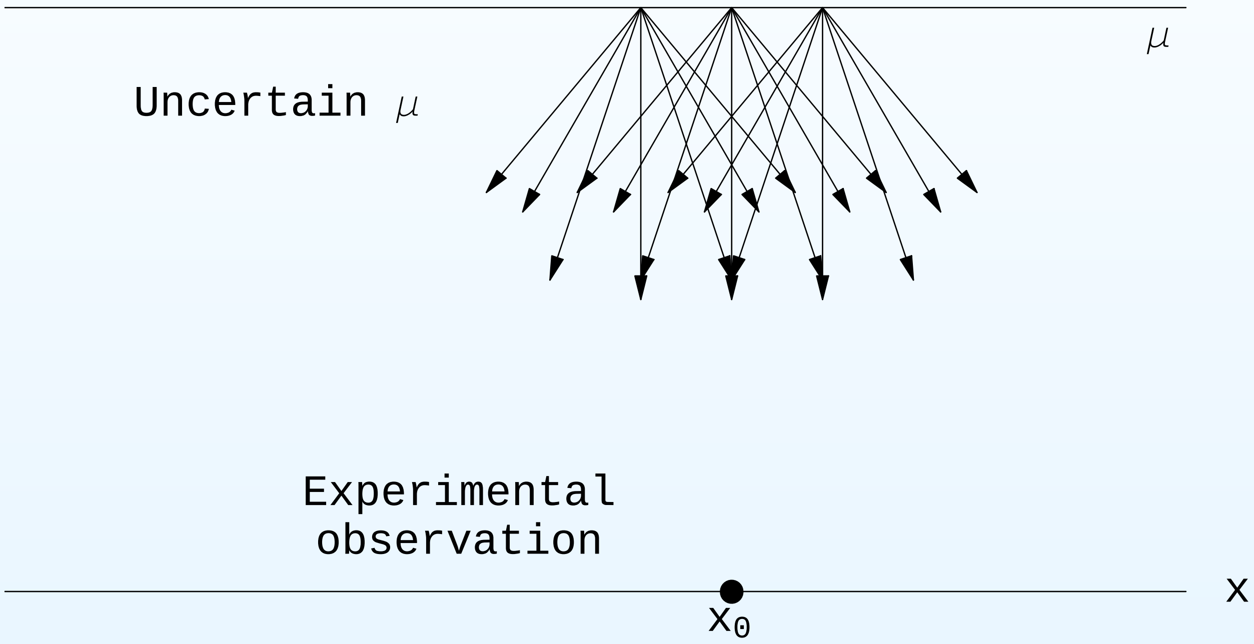
> We shall see that, similarly, there are hidden assumptions behind the naive probabilistic inversions in problems like the AIDS one.
>
> Curiously enough, these methods are advertised as objective because they do not need as input our scientific expectations of where the value of the quantity might 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).

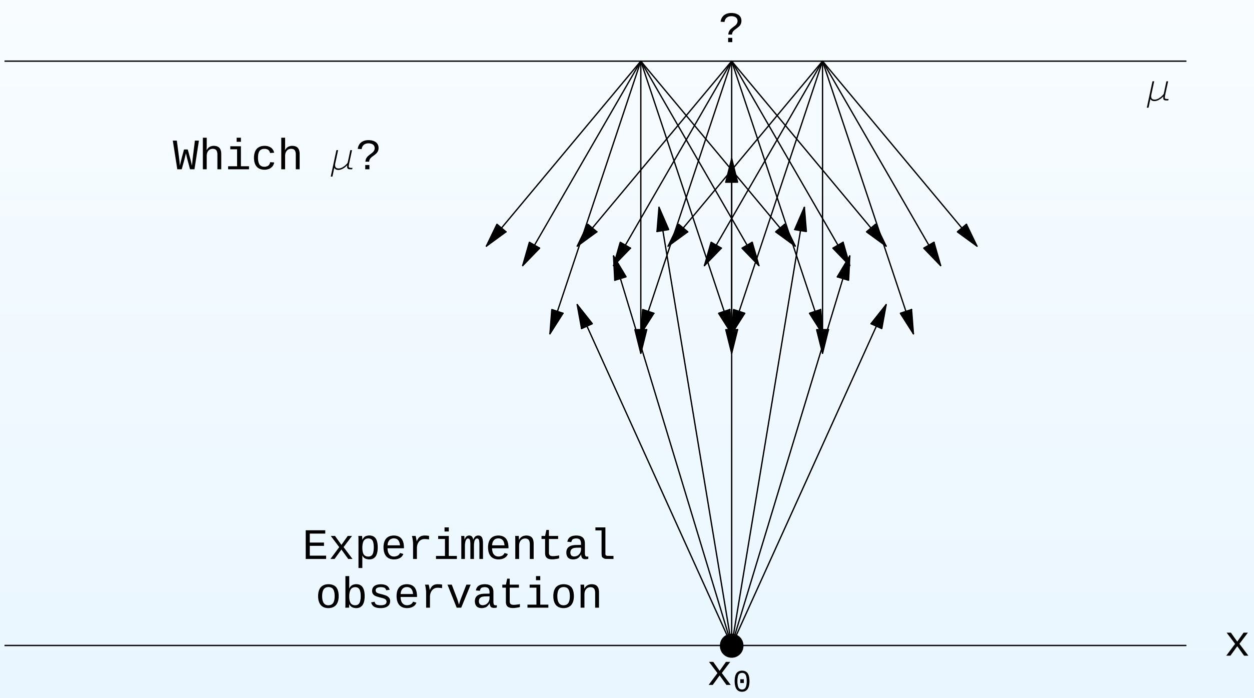
> We shall see that, similarly, there are hidden assumptions behind the naive probabilistic inversions in problems like the AIDS one.
>
> Curiously enough, these methods are advertised as objective because they do not need as input our scientific expectations of where the value of the quantity might 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!

Inferring a true value (repetita. . .)



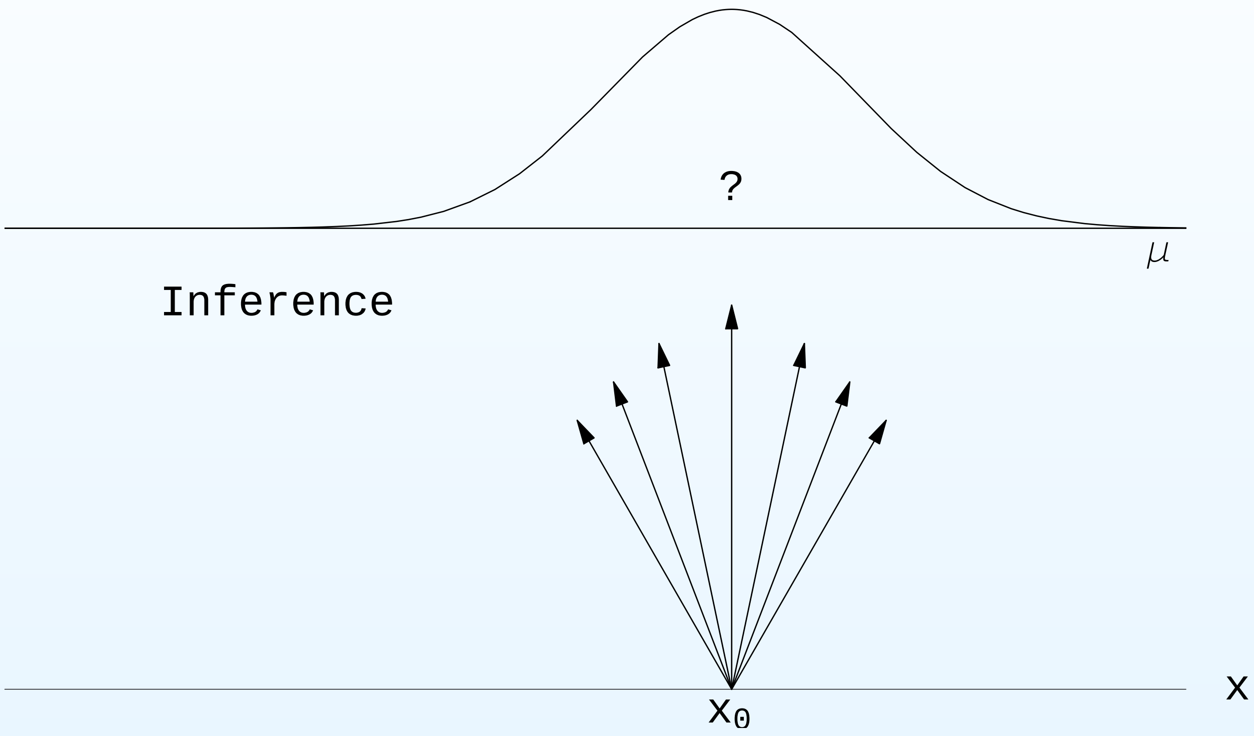
The observed data is <u>certain</u>: → ‘true value’ <u>uncertain</u>.

Inferring a true value (repetita. . .)



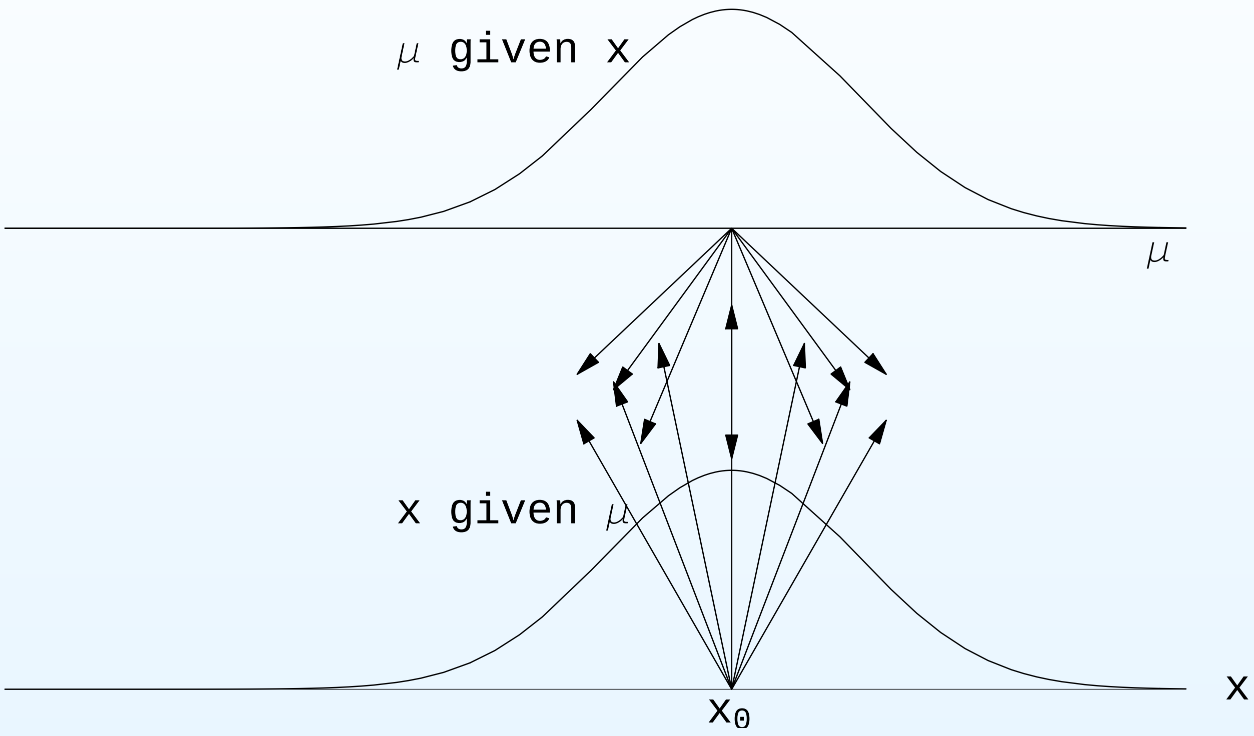
Where does the observed value of x comes from?

Inferring a true value (repetita. . .)



We are now uncertain about μ, given x.

Inferring a true value (repetita. . .)



Note the symmetry in reasoning.

A very simple experiment

Let’s make an experiment

A very simple experiment

Let’s make an experiment

- Here
- Now

A very simple experiment

Let's make an experiment

- Here
- Now

For simplicity

- μ can assume only six possibilities:

$$\mathbf{0}, \mathbf{1}, \ldots, \mathbf{5}$$

- x is binary:

$$\mathbf{0}, \mathbf{1}$$

$$[\,(1,2);\ \text{Black/White};\ \text{Yes/Not};\ \ldots\,]$$

A very simple experiment

Let's make an experiment

- Here
- Now

For simplicity

- μ can assume only six possibilities:

$$\mathbf{0}, \mathbf{1}, \ldots, \mathbf{5}$$

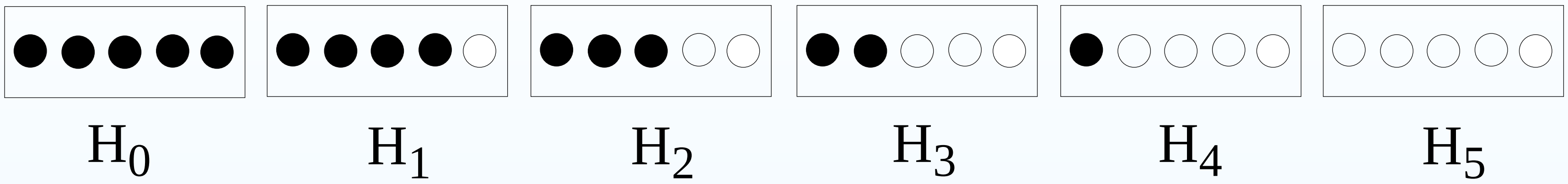
- x is binary:

$$\mathbf{0}, \mathbf{1}$$

$$[\,(1, 2);\ \text{Black/White};\ \text{Yes/Not};\ \ldots\,]$$

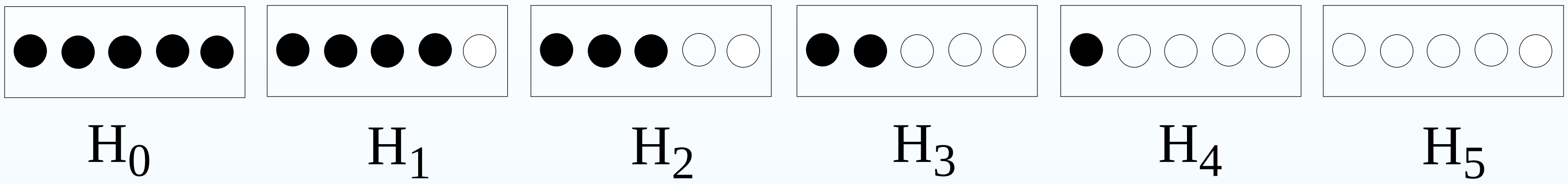
$\Rightarrow$ Later we shall make μ continous.

Which box? Which ball?



Let us take randomly one of the boxes.

Which box? Which ball?



Let us take randomly one of the boxes.

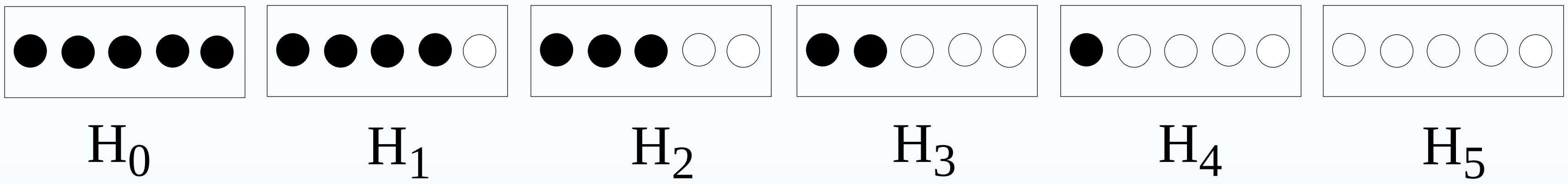
We are in a state of uncertainty concerning several *events*, the most important of which correspond to the following questions:

(a) Which box have we chosen, $H_0, H_1, \ldots, H_5$?

(b) If we extract randomly a ball from the chosen box, will we observe a white ($E_W \equiv E_1$) or black ($E_B \equiv E_2$) ball?

Our certainties:
$$\begin{aligned} \cup_{j=0}^{5} H_j &= \Omega \\ \cup_{i=1}^{2} E_i &= \Omega\,. \end{aligned}$$

Which box? Which ball?



Let us take randomly one of the boxes.

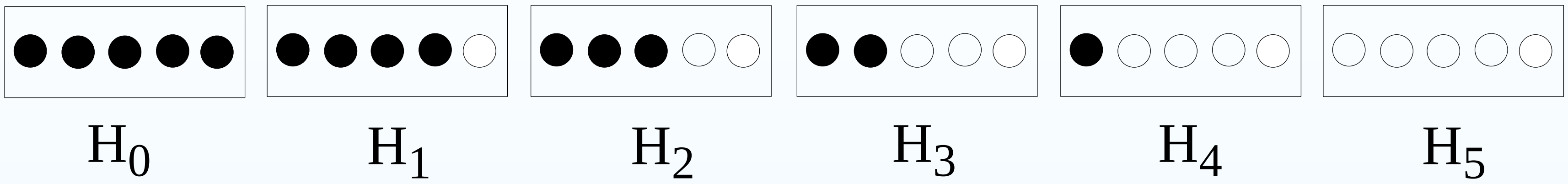
- What happens after we have extracted one ball and looked its color?
  - Intuitively feel *how to* roughly *change* our opinion about
    - the possible cause
    - a future observation

In general, we are uncertain about all the combinations of E_i and H_j: $E_1 \cap H_0$, $E_1 \cap H_1, \ldots, E_2 \cap H_5$, and these 12 *constituents* are not equiprobable.

Our certainty:

$$\begin{aligned} \cup_{j=0}^{5} H_j &= \Omega \\ \cup_{i=1}^{2} E_i &= \Omega \,. \end{aligned}$$

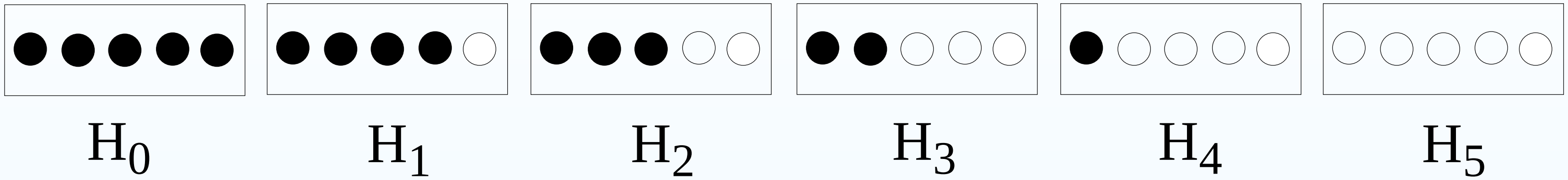
Which box? Which ball?



Let us take randomly one of the boxes.

- What happens after we have extracted one ball and looked its color?
  - Intuitively feel *how to* roughly *change* our opinion about
    - the possible cause
    - a future observation
  - Can we do it *quantitatively*, in an ‘objective way’?

Which box? Which ball?



Let us take randomly one of the boxes.

- What happens after we have extracted one ball and looked its color?
  - Intuitively feel *how to* roughly *change* our opinion about
    - the possible cause
    - a future observation
  - Can we do it *quantitatively*, in an ‘objective way’?
- And after a sequence of extractions?

The toy inferential experiment

The aim of the experiment will be to guess the content of the box without looking inside it, only extracting a ball, record its color and reintroducing in the box

The toy inferential experiment

The aim of the experiment will be to guess the content of the box without looking inside it, only extracting a ball, record its color and reintroducing in the box

This toy experiment is conceptually very close to what we do in Physics

$\Rightarrow$ try to guess what we cannot see (the electron mass, a branching ratio, etc)

. . . from what we can see (somehow) with our senses.

The rule of the game is that we are not allowed to watch inside the box! (As we cannot open an electron and read its properties, unlike we read the MAC address of a PC interface.)

Where is probability?

We all agree that the experimental results change

- the probabilities of the box compositions;
- the probabilities of a future outcomes,

Where is probability?

We all agree that the experimental results change

- the probabilities of the box compositions;
- the probabilities of a future outcomes,

although the box composition remains unchanged ('extractions followed by reintroduction').

Where is probability?

We all agree that the experimental results change

- the probabilities of the box compositions;
- the probabilities of a future outcomes,

although the box composition remains unchanged ('extractions followed by reintroduction').

Where is the probability?

Where is probability?

We all agree that the experimental results change

- the probabilities of the box compositions;
- the probabilities of a future outcomes,

although the box composition remains unchanged ('extractions followed by reintroduction').

Where is the probability?

Certainly <u>not</u> *in* the box!

Probability

What is probability?

$\Longrightarrow$ CERN Lectures (nr 2, from p. 30)

Probability

What is probability?

$\Longrightarrow$ CERN Lectures (nr 2, from p. 30)

Probability

What is probability?

$\Longrightarrow$ CERN Lectures (nr 2, from p. 30)

Probability

What is probability?

$\Rightarrow$ CERN Lectures (nr 2, from p. 30)

Probability

What is probability?

Standard textbook definitions

Standard textbook definitions

It is easy to check that ‘scientific’ definitions suffer of circularity

Standard textbook definitions

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

Standard textbook definitions

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

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

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

Future $\Leftrightarrow$ Past (believed so)

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

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

$\rightarrow$ It limits the range of applications

Definitions $\to$ evaluation rules

Very useful evaluation rules

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

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

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

Definitions $\longrightarrow$ evaluation rules

Very useful evaluation rules

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

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

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 is 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

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

Probability

What <u>is</u> 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"

$\rightarrow$ '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] *…,*

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"*

Or perhaps you prefer this way…

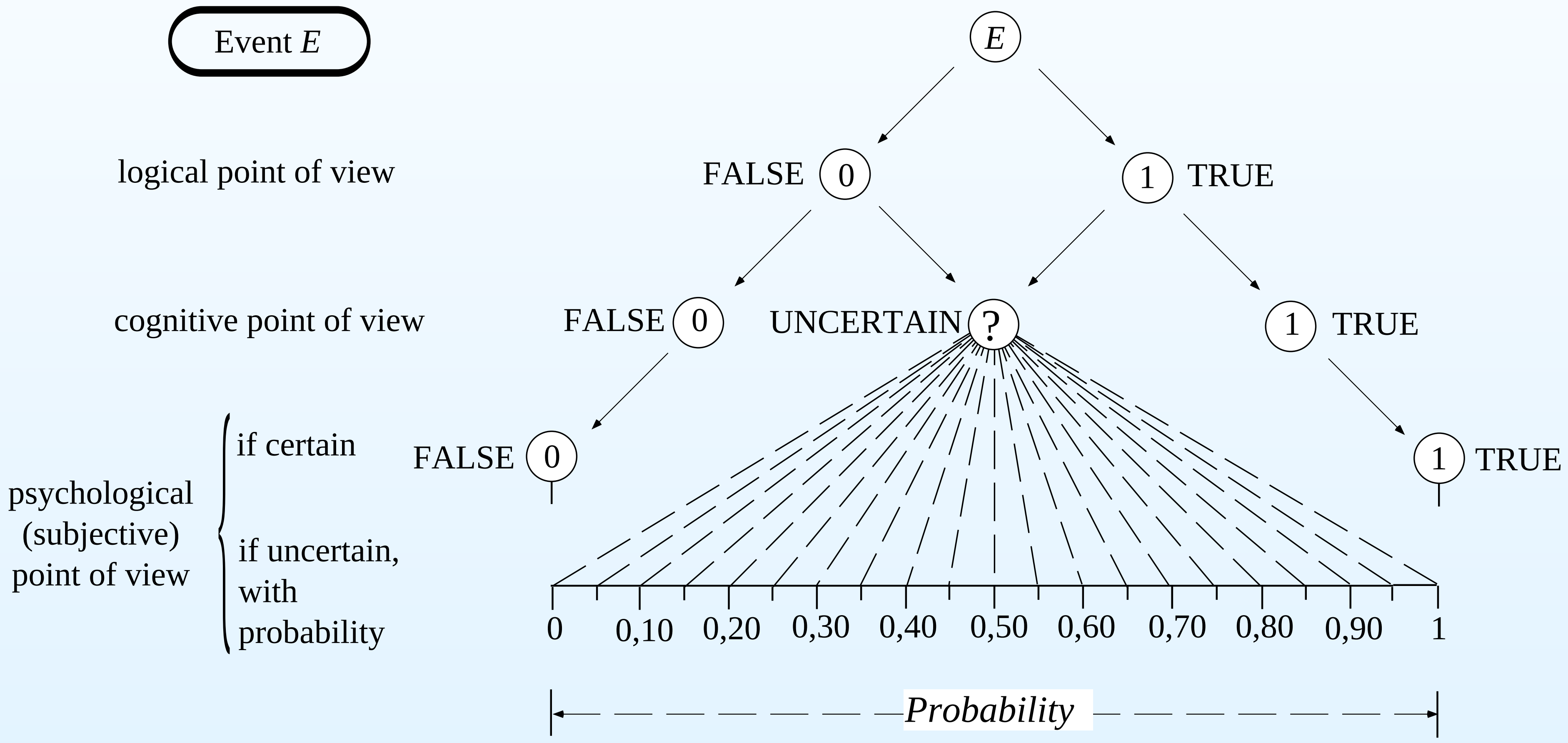
"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1]*. . . , the numerical probability p of this event is to be a real number by the indication of which we try* in some cases *to setup a quantitative measure of the strength of our conjecture or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

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

- *“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 $\to$ probability

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

Probability is always conditional probability

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

- *"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:
  - coin trown in the air;
  - 'three box problems'.