Probabilità e incertezze di misura

Giulio D'Agostini

Dipartimento di Fisica

Università di Roma La Sapienza

Piano dei due incontri

1. Rassegna critica e introduzione all'inferenza probabilistica
   - Quanto sono sensate e ben fondate le regolette per la valutazione dei cosiddetti "errori di misura"?
   - Per imparare dall'esperienza in modo quantitativo, facendo uso della logica dell'incerto, dobbiamo
     - rivedere il concetto di probabilità;
     - imparare ad . . . imparare dall'esperienza.
2. Stima delle incertezze in misure dirette e indirette
   - Sorgenti delle incertezze di misura (*decalogo ISO*).
   - Applicazione dell'inferenza probabilistica alle misure sperimentali (semplice caso di errori gaussiani):
     - singola osservazione
     - campione di osservazioni
     - stima dei parametri di un andamento lineare
   - Propagazione delle incertezze

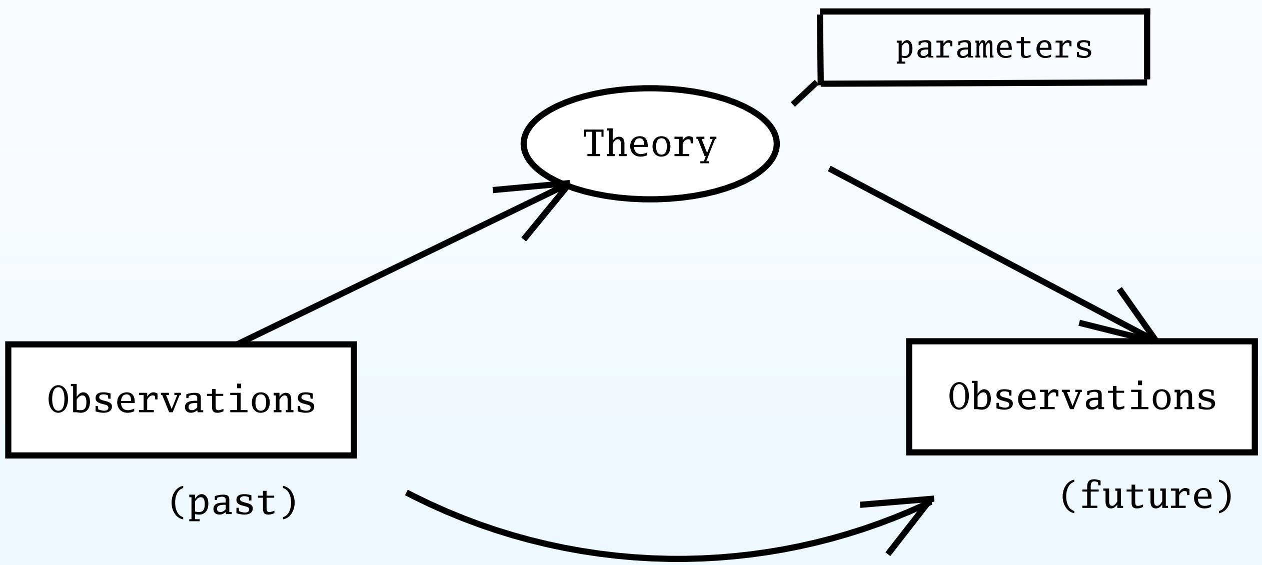
Scaletta del primo incontro

- Metodo scientifico: osservazioni e ipotesi
- Incertezza
- Cause $\longleftrightarrow$ Effetti

  "Il problema essenziale del metodo sperimentale" (Poincaré).
- L'esempio guida: il problema delle sei scatole.

  "La probabilità à riferita a casi reali o non ha alcun senso" (de Finetti).
- *Fisichettume*: una rassegna critica.
- Falsificazionismo e *variazioni statistiche* ('test').
- Approccio probabilistico.
- *Cosè la probabilità?* Regole di base della probabilità.
- Aggiornamento della probabilità alla luce delle osservazioni (formula di Bayes) $\Rightarrow$inferenza probabilistica (bayesiana)
- Conclusioni.

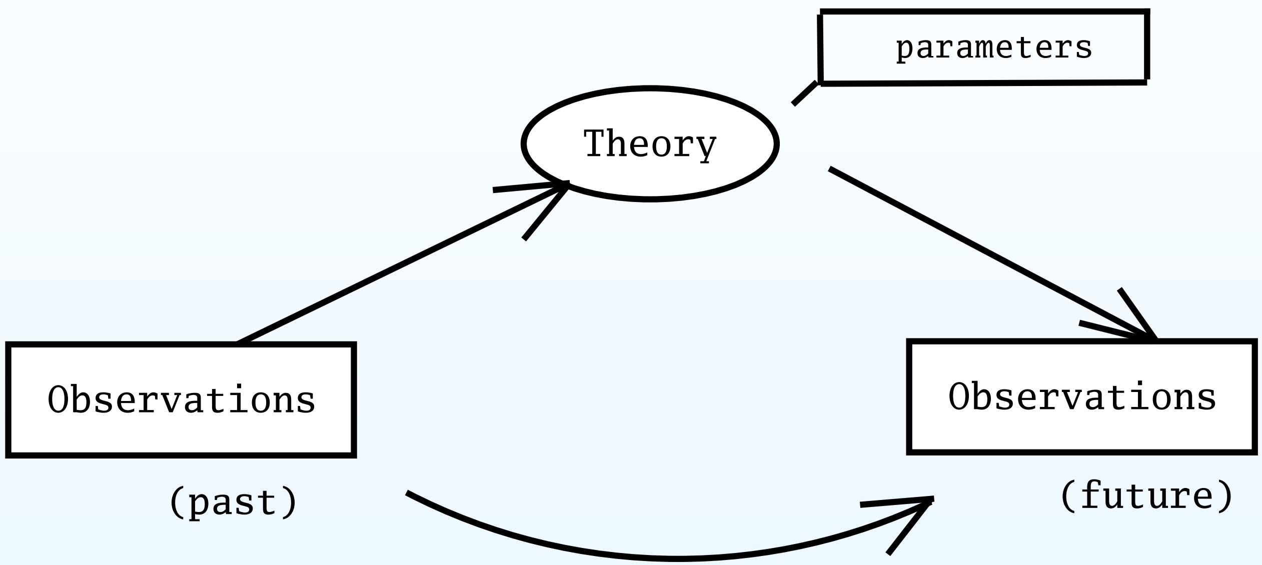
From past to future



Task of the ‘physicist’ (scientist, decision maker):

- Describe/understand the physical world
  - $\Rightarrow$ inference of laws and their parameters
- Predict observations
  - $\Rightarrow$ forecasting

From past to future



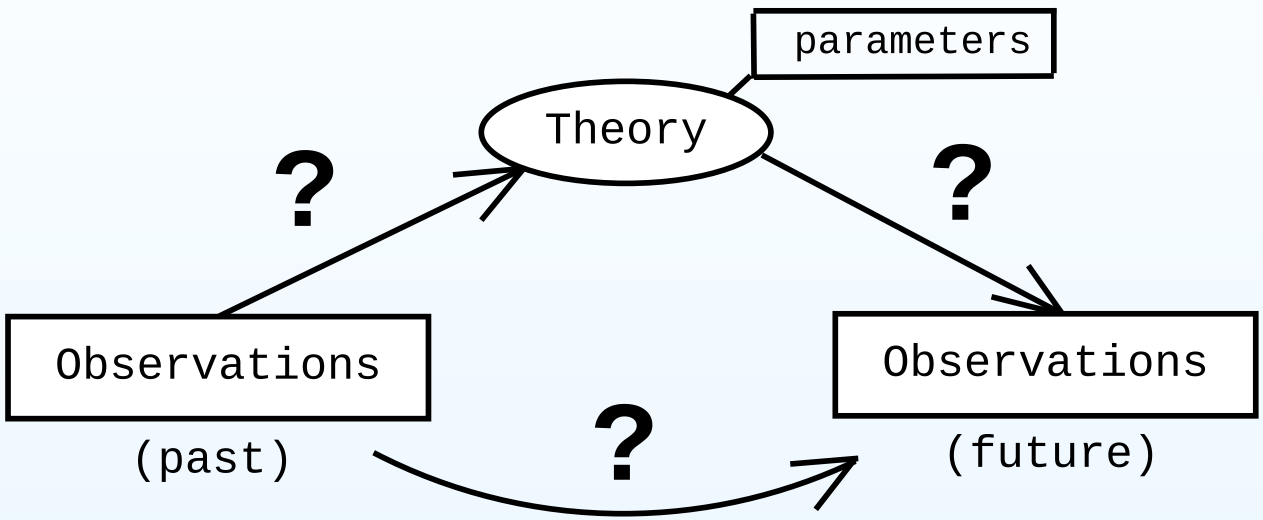
Process

- neither automatic
- nor purely contemplative

  $\longrightarrow$ ‘scientific method’

  $\longrightarrow$ planned experiments (‘actions’) $\Rightarrow$ decision.

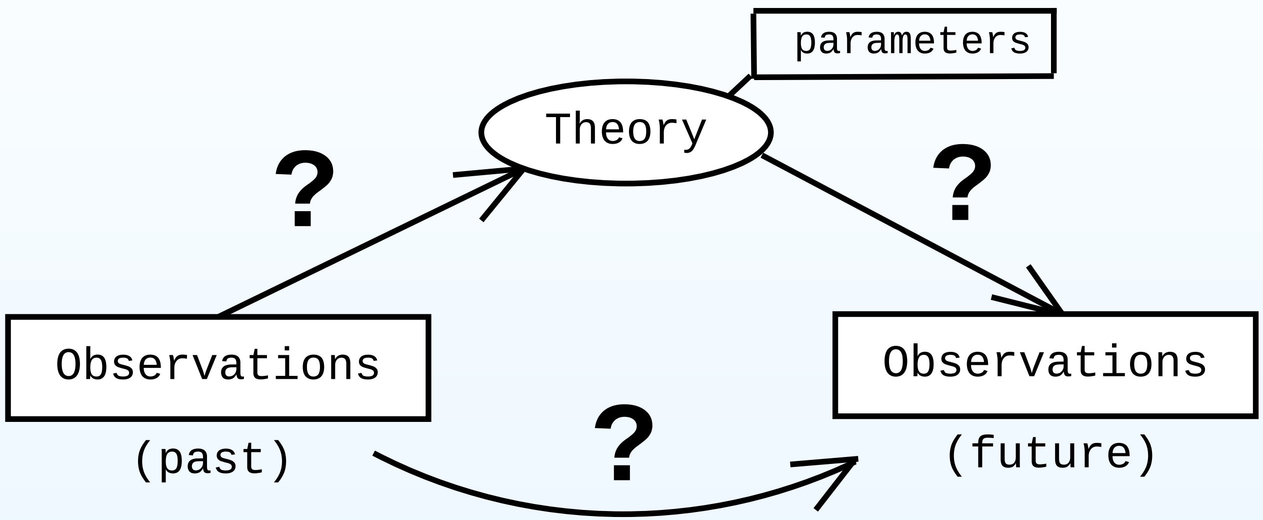
From past to future



⇒ Uncertainty:

1. Given the past observations, in general we are **not sure about** the **theory** parameter (and/or the theory itself)
2. Even if we were sure about theory and parameters, there could be internal (e.g. Q.M.) or external effects (initial/boundary conditions, ‘errors’, etc) that make the forecasting uncertain.

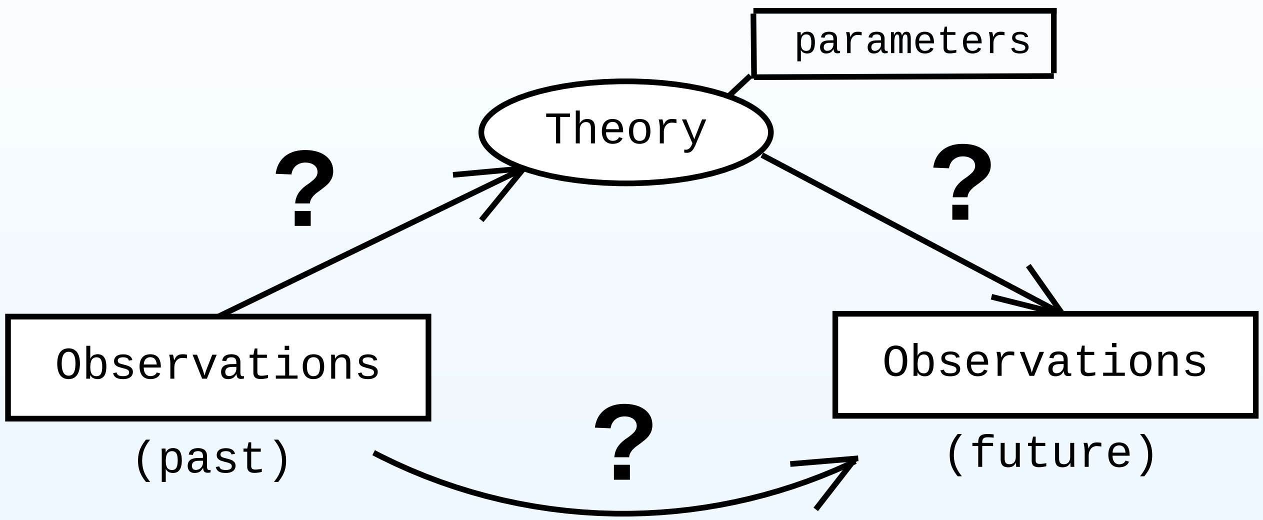
From past to future



⇒ Decision

- What is be **best action** ('experiment') to take in order **'to be confident'** that what we would like will occur?
  (Decision issues always assume uncertainty about future outcomes.)
- Before tackling problems of decision **we need to learn to reason about uncertainty**, possibly in a quantitative way.

From past to future



Deep reason of uncertainty

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

A simple example

- Three boxes each contains two balls: White-White, White-Black, Black-Black. We take randomly one of the box and extract one ball, e.g. White. We can extract the second ball from any of the three boxes.

A simple example

- Three boxes each contains two balls: White-White, White-Black, Black-Black. We take randomly one of the box and extract one ball, e.g. White. We can extract the second ball from any of the three boxes.
- Decision problem: From which box should we extract the second ball in order to have a second White?

A simple example

- Three boxes each contains two balls: White-White, White-Black, Black-Black. We take randomly one of the box and extract one ball, e.g. White. We can extract the second ball from any of the three boxes.
- Decision problem: From which box should we extract the second ball in order to have a second White?
- Uncertanty:
  - Which box have we taken?
  - What is the chance to get White from the same box, or from one of the remaining two, selected at random?

About predictions

Remember:

“Prediction is very difficult,
especially if it’s about the future” (Bohr)

About predictions

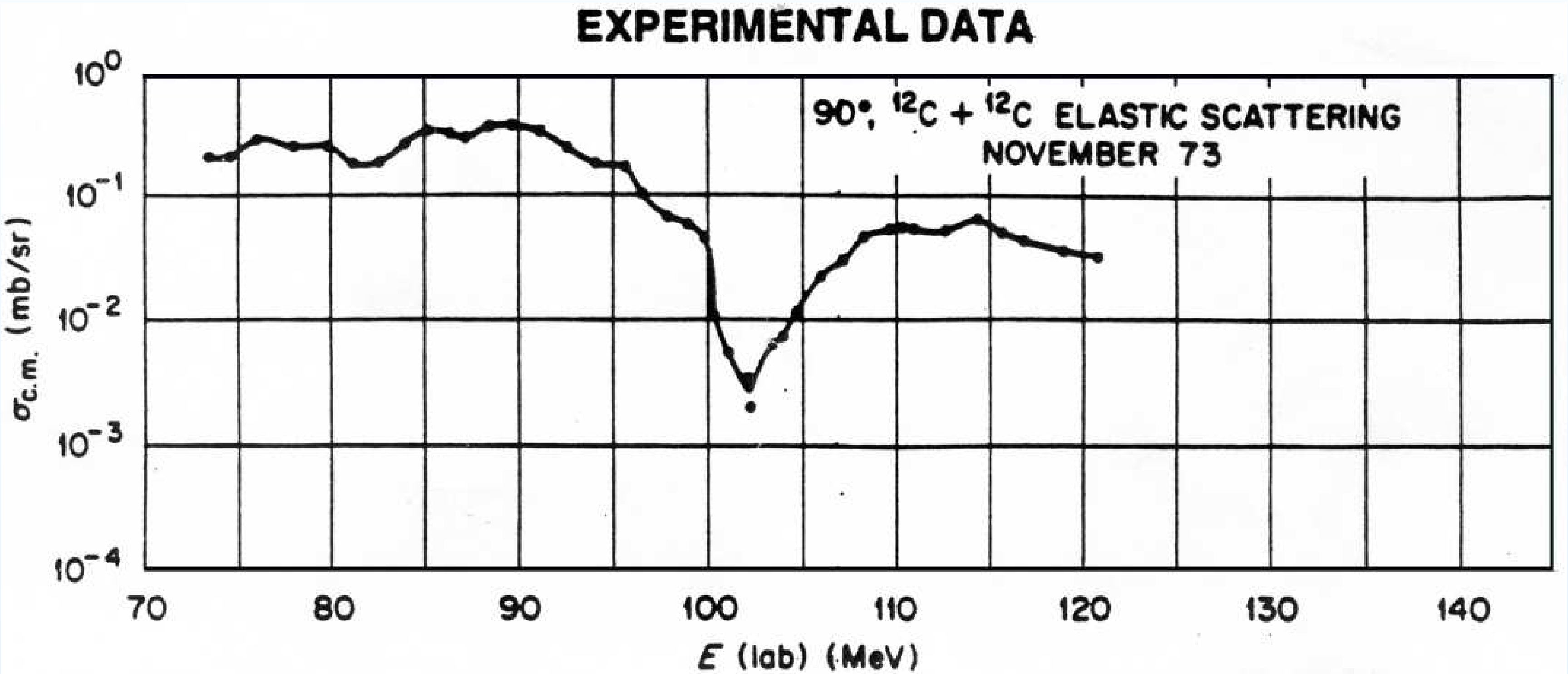
Remember:

*"Prediction is very difficult,
especially if it's about the future"* (Bohr)

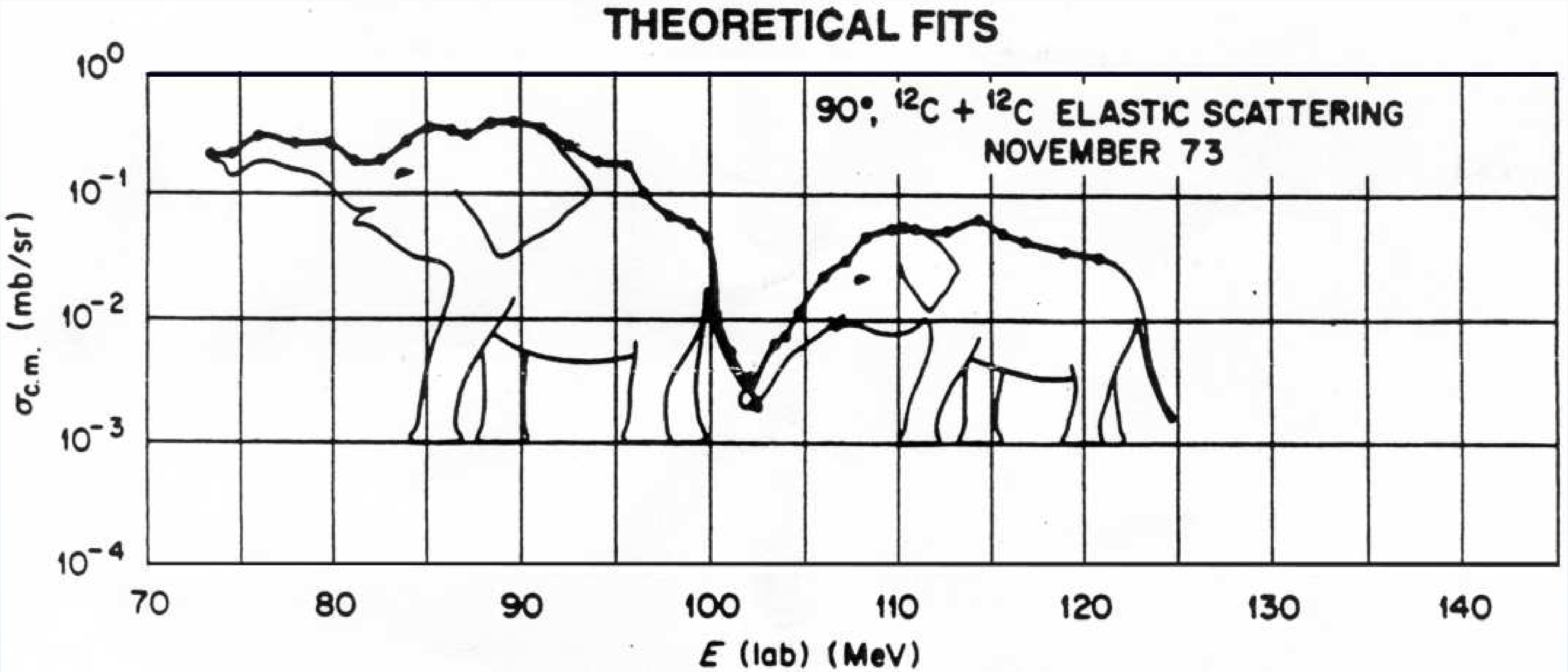
But, anyway:

"It is far better to foresee even without certainty than not to foresee at all"
(Poincaré)

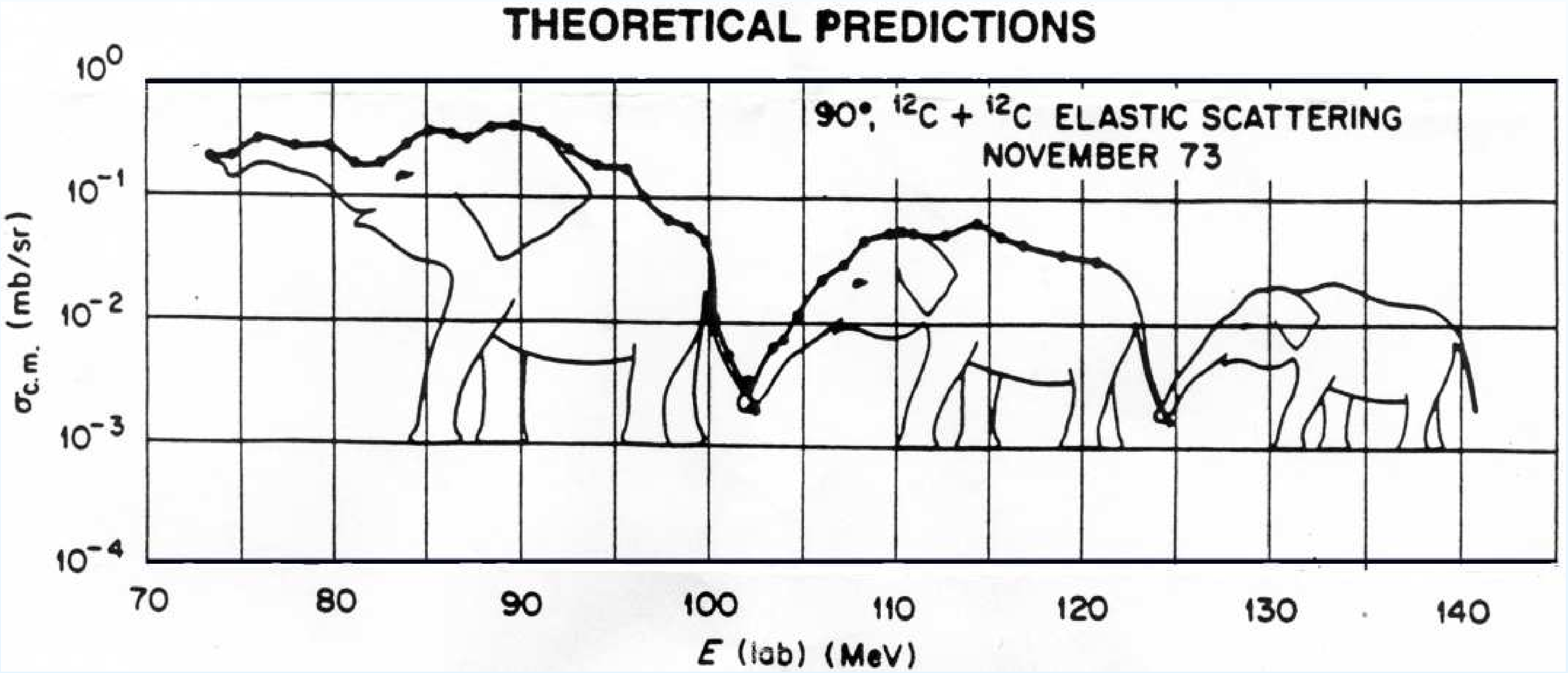
Inferential process



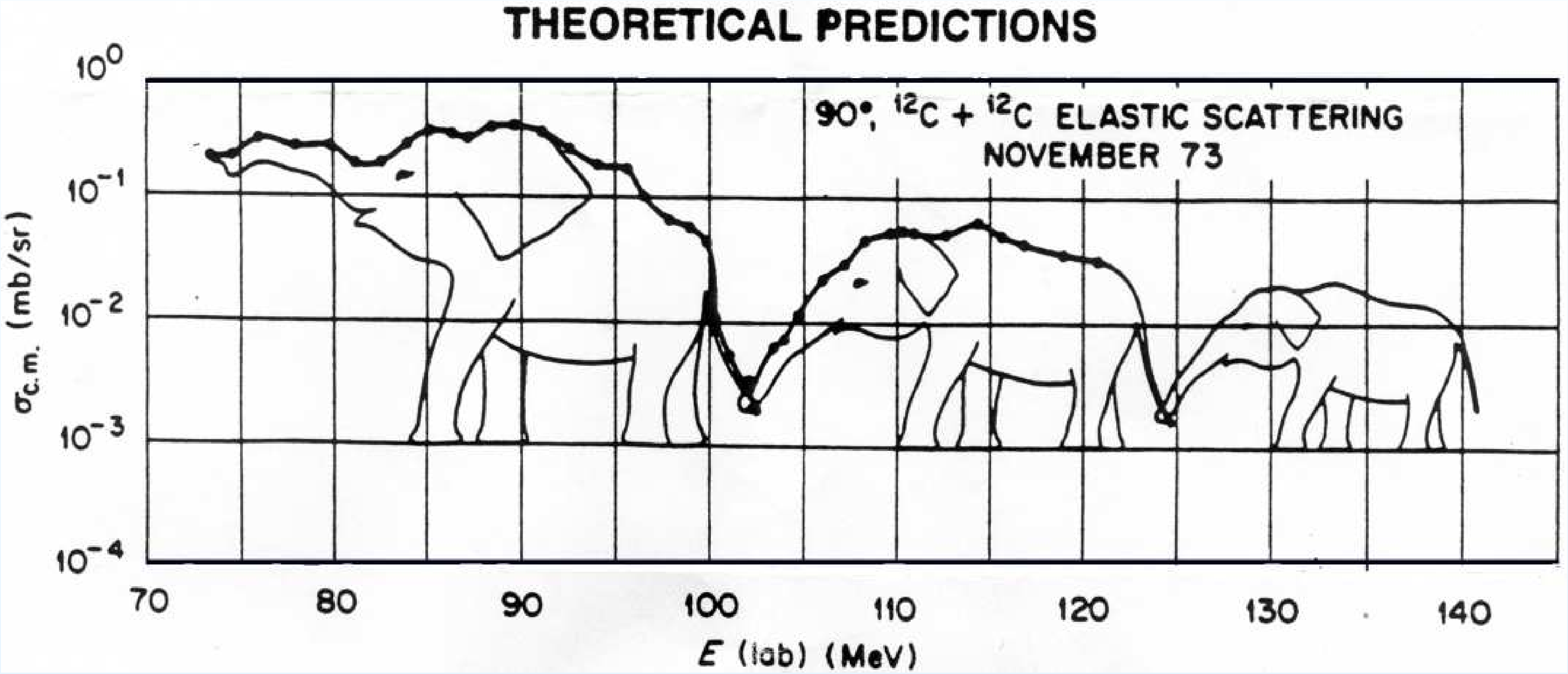
Inferential process



Inferential process

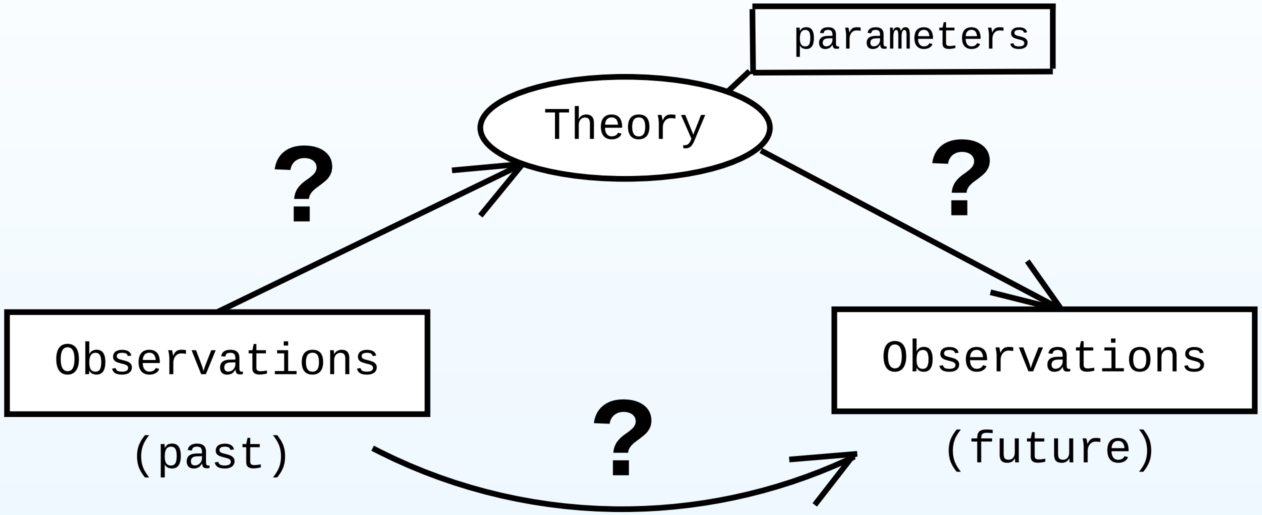


Inferential process



(S. Raman, *Science with a smile*)

Deep source of uncertainty



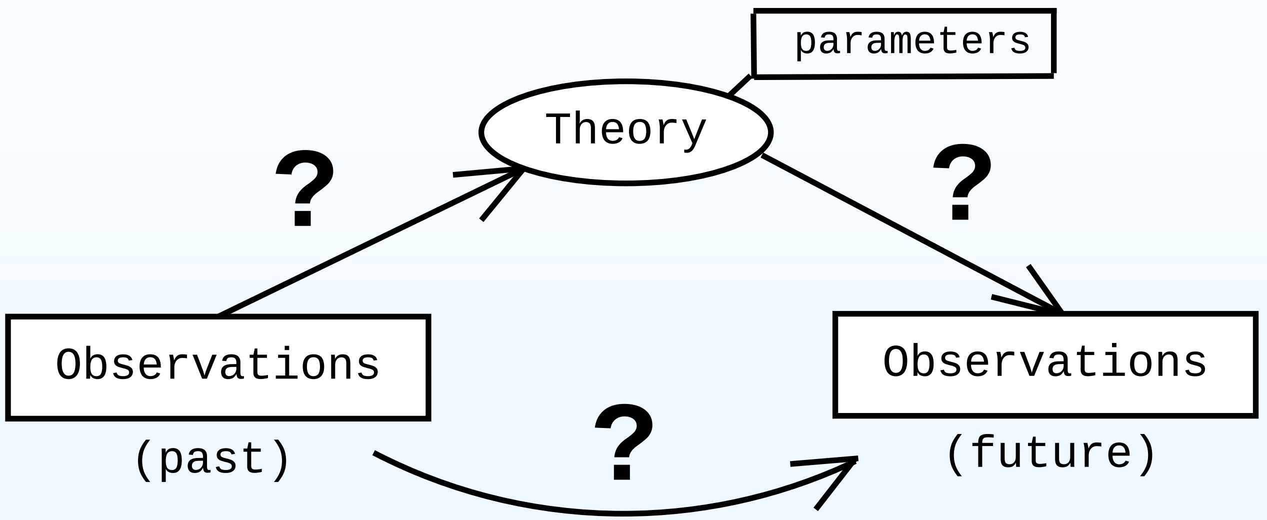
Uncertainty:

Theory $\quad - ? \longrightarrow$

Past observations $\quad - ? \longrightarrow$

Theory $\quad - ? \longrightarrow \quad$ Future observations

Deep source of uncertainty



Uncertainty:

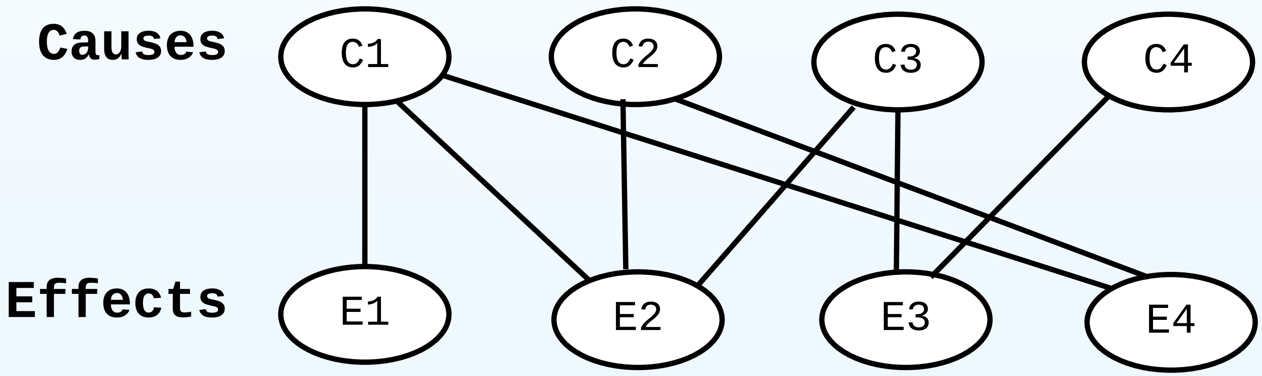
Theory	— **?** ⟶	Future observations
Past observations	— **?** ⟶	Theory
Theory	— **?** ⟶	Future observations

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

CAUSE $\Longleftrightarrow$ **EFFECT**

Causes $\longrightarrow$ effects

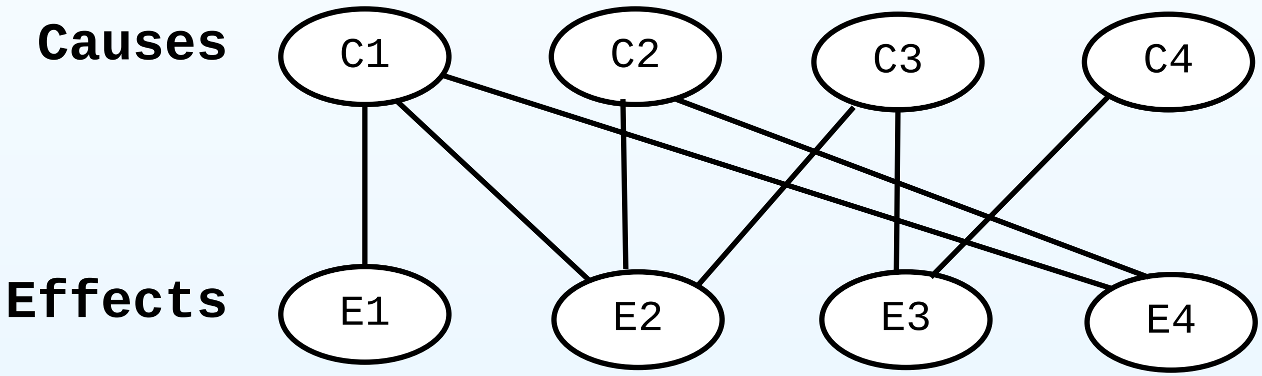
The same *apparent* cause might produce several,different effects



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

Causes $\longrightarrow$ effects

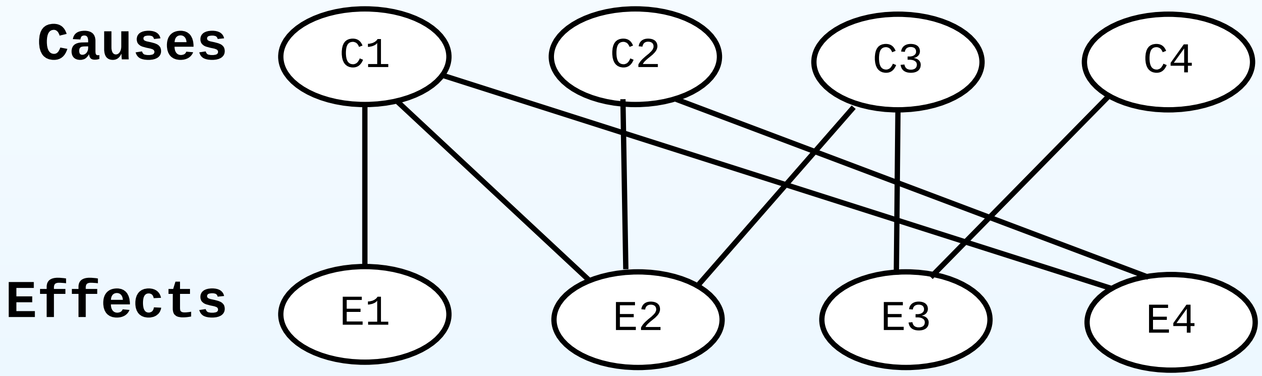
The same *apparent* cause might produce several,different effects



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

Causes $\longrightarrow$ 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."

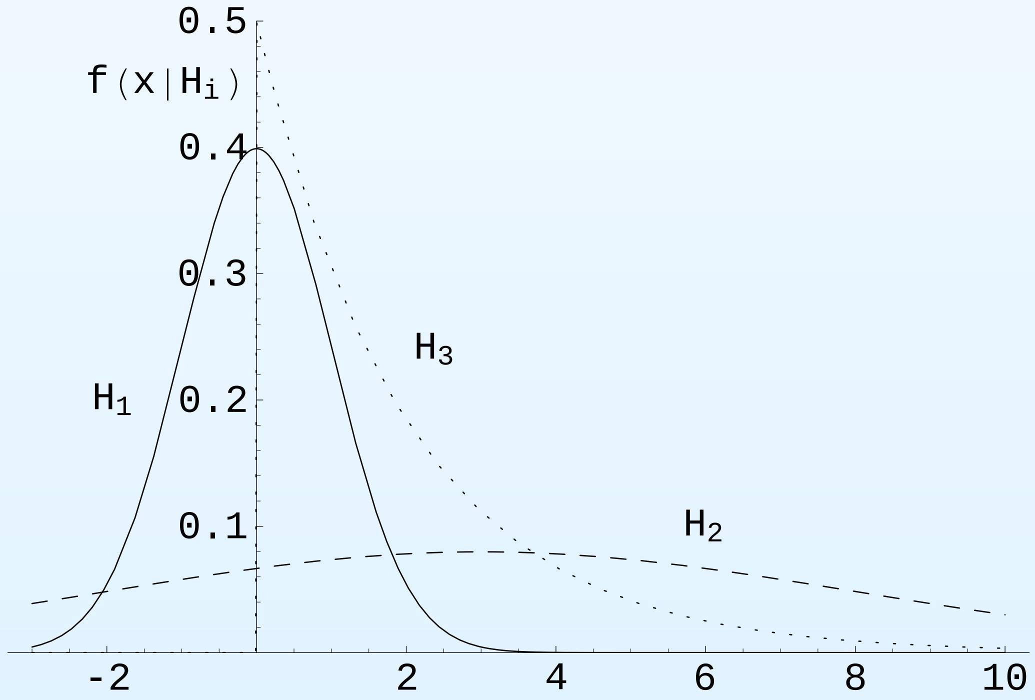
(H. Poincaré – *Science and Hypothesis*)

A numerical example

- Effect: number $x = 3$ extracted 'at random'
- Hypotheses: one of the following random generators:
  - H_1 Gaussian, with $\mu = 0$ and $\sigma = 1$
  - H_2 Gaussian, with $\mu = 3$ and $\sigma = 5$
  - H_3 Exponential, with $\tau = 2$

A numerical example

- Effect: number $x = 3$ extracted ‘at random’
- Hypotheses: one of the following random generators:
  - H_1 Gaussian, with $\mu = 0$ and $\sigma = 1$
  - H_2 Gaussian, with $\mu = 3$ and $\sigma = 5$
  - H_3 Exponential, with $\tau = 2$



A numerical example

- Effect: number $x = 3$ extracted 'at random'
- Hypotheses: one of the following random generators:
  - H_1 Gaussian, with $\mu = 0$ and $\sigma = 1$
  - H_2 Gaussian, with $\mu = 3$ and $\sigma = 5$
  - H_3 Exponential, with $\tau = 2$

$\Rightarrow$ Which one to prefer?

Note: $\Rightarrow$ none of the hypotheses of this example can be excluded and, therefore, there is no way to reach a boolean conclusion. We can only state, somehow, our *rational preference*, based on the experimental result and our best knowledge of the behavior of each *model*.

A numerical example

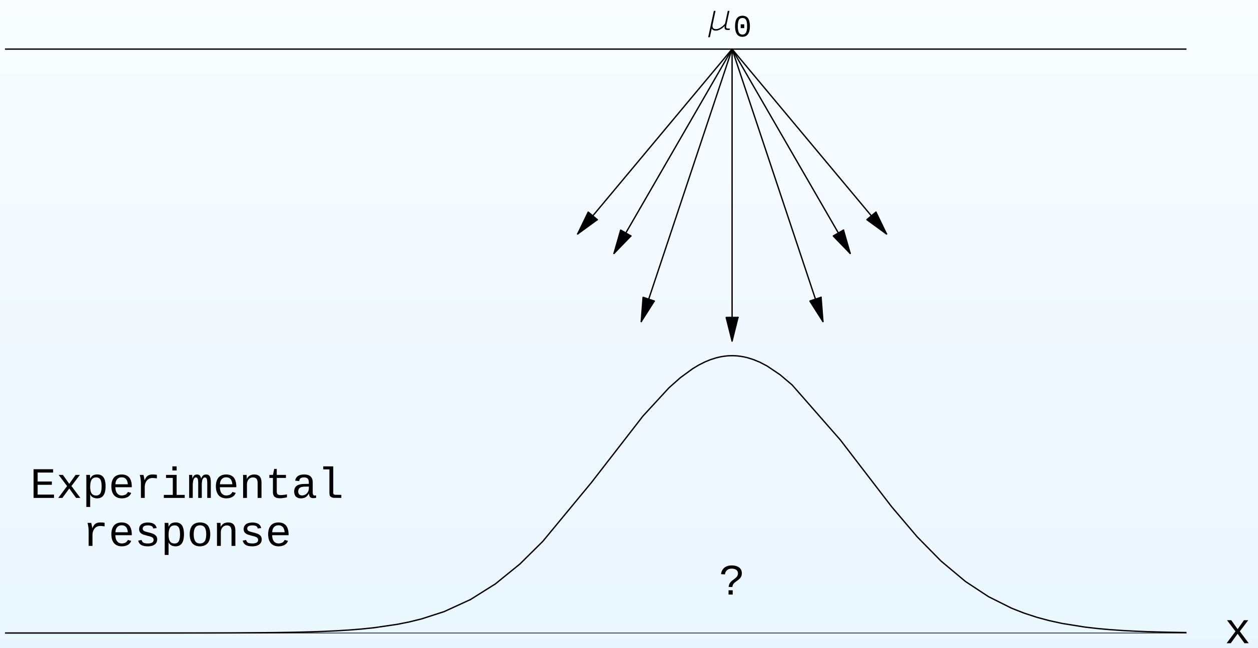
- Effect: number $x = 3$ extracted 'at random'
- Hypotheses: one of the following random generators:
  - H_1 Gaussian, with $\mu = 0$ and $\sigma = 1$
  - H_2 Gaussian, with $\mu = 3$ and $\sigma = 5$
  - H_3 Exponential, with $\tau = 2$

$\Rightarrow$ Which one to prefer?

Note: $\Rightarrow$ none of the hypotheses of this example can be excluded and, therefore, there is no way to reach a boolean conclusion. We can only state, somehow, our *rational preference*, based on the experimental result and our best knowledge of the behavior of each *model*.

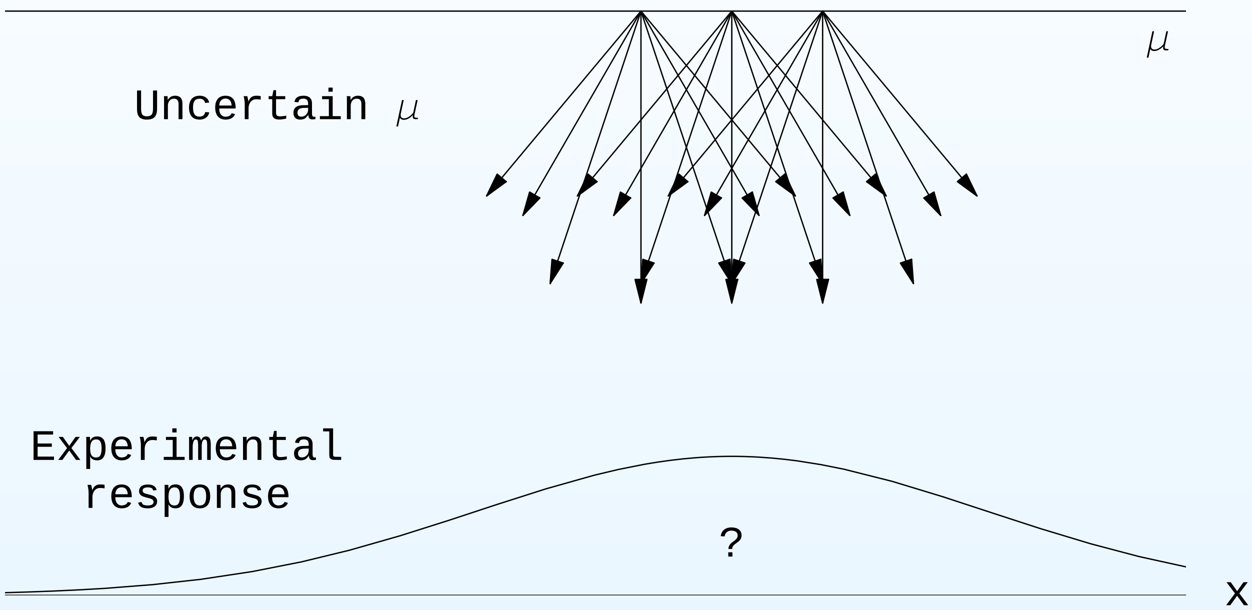
We shall come back to this example
$\longrightarrow$ Let's now move to 'measuring true values'

From ‘true value’ to observations



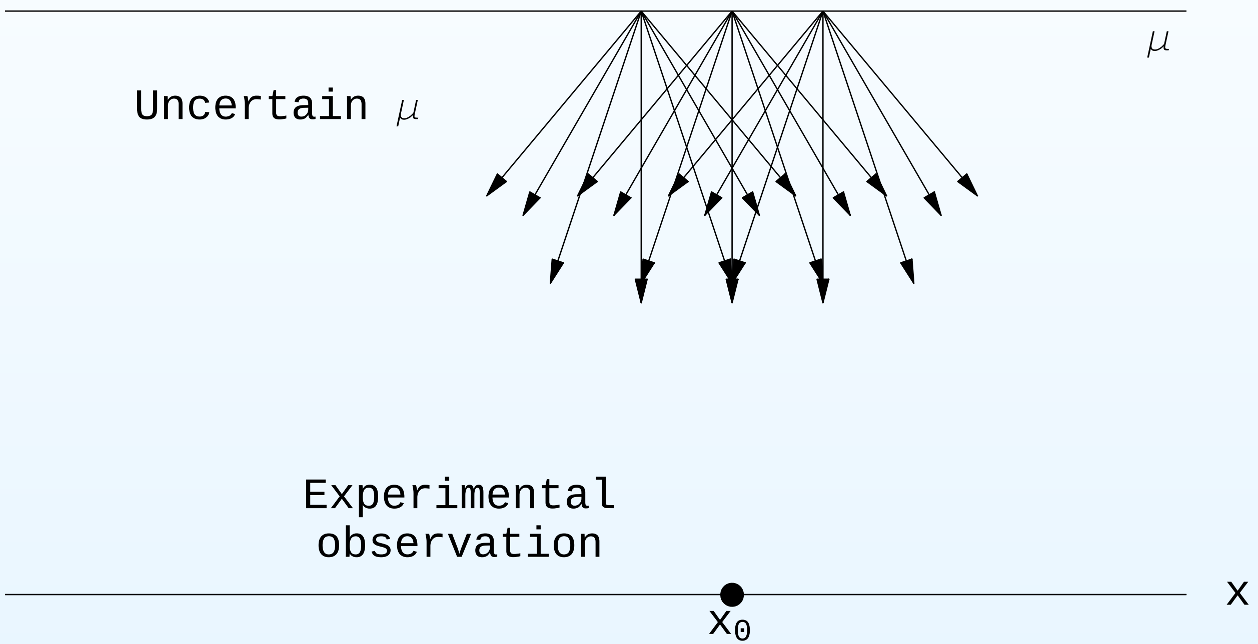
Given μ (exactly known) we are uncertain about x

From 'true value' to observations



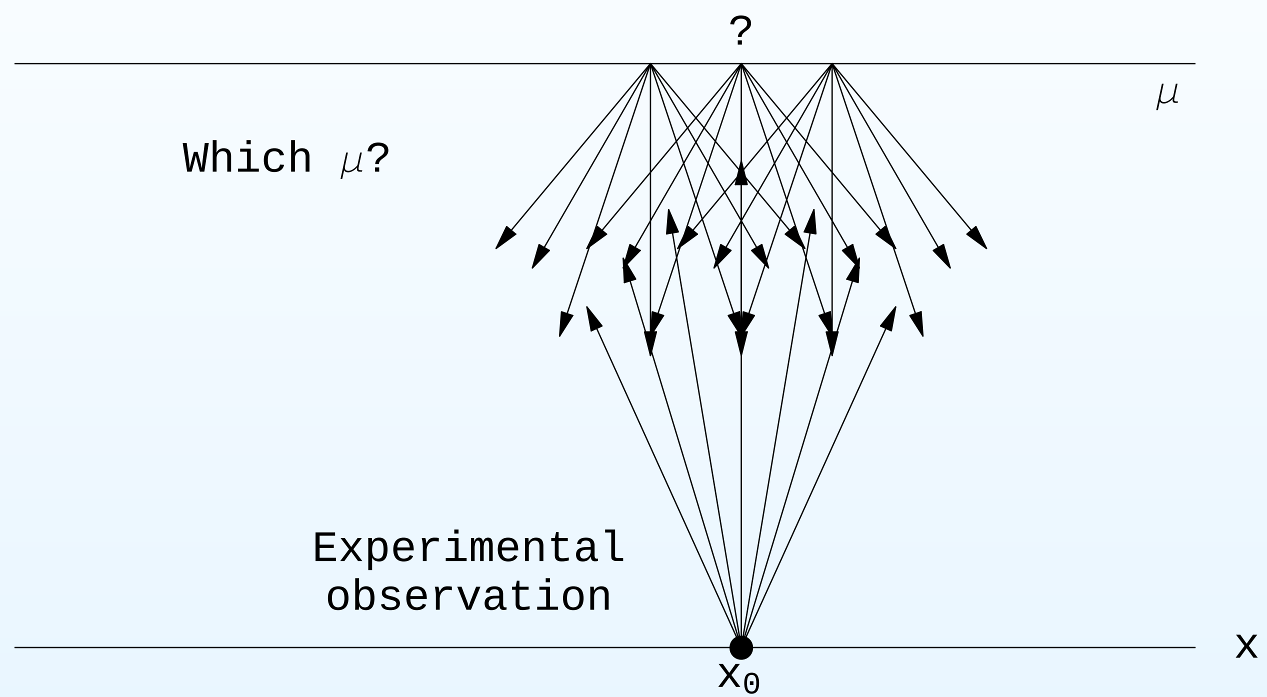
Uncertainty about μ makes us more uncertain about x

Inferring a true value



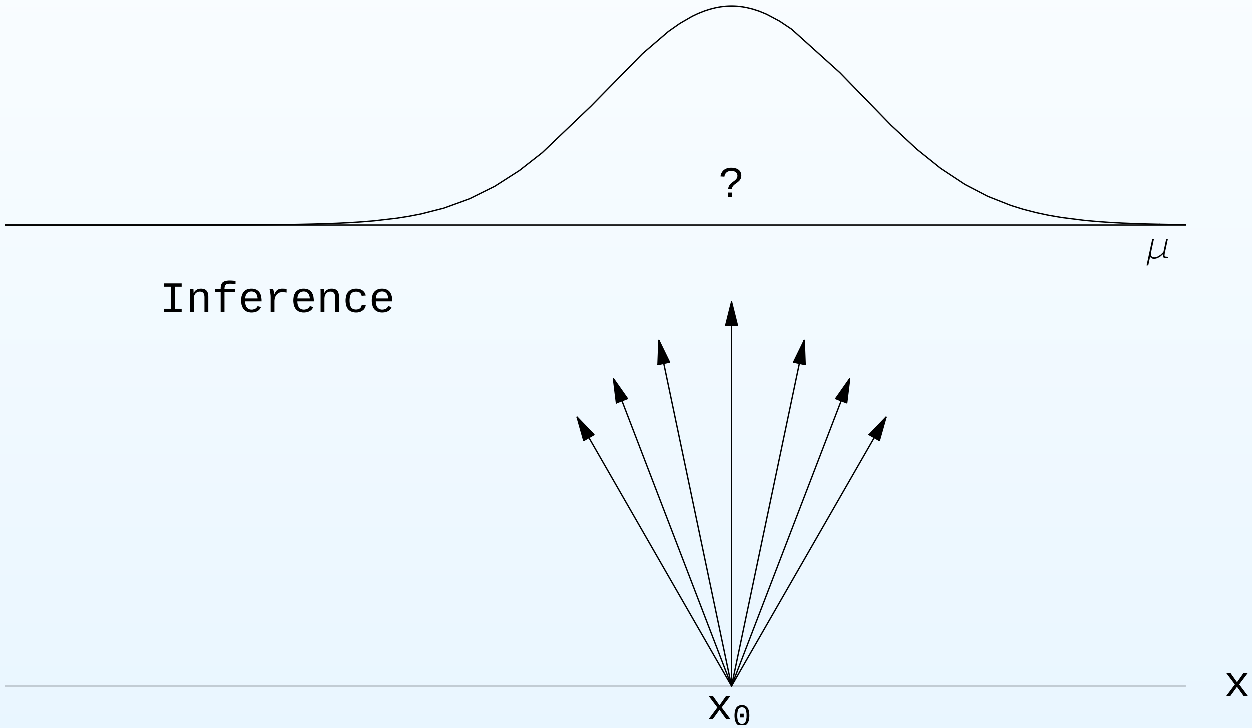
The observed data is <u>certain</u>: → 'true value' uncertain.

Inferring a true value



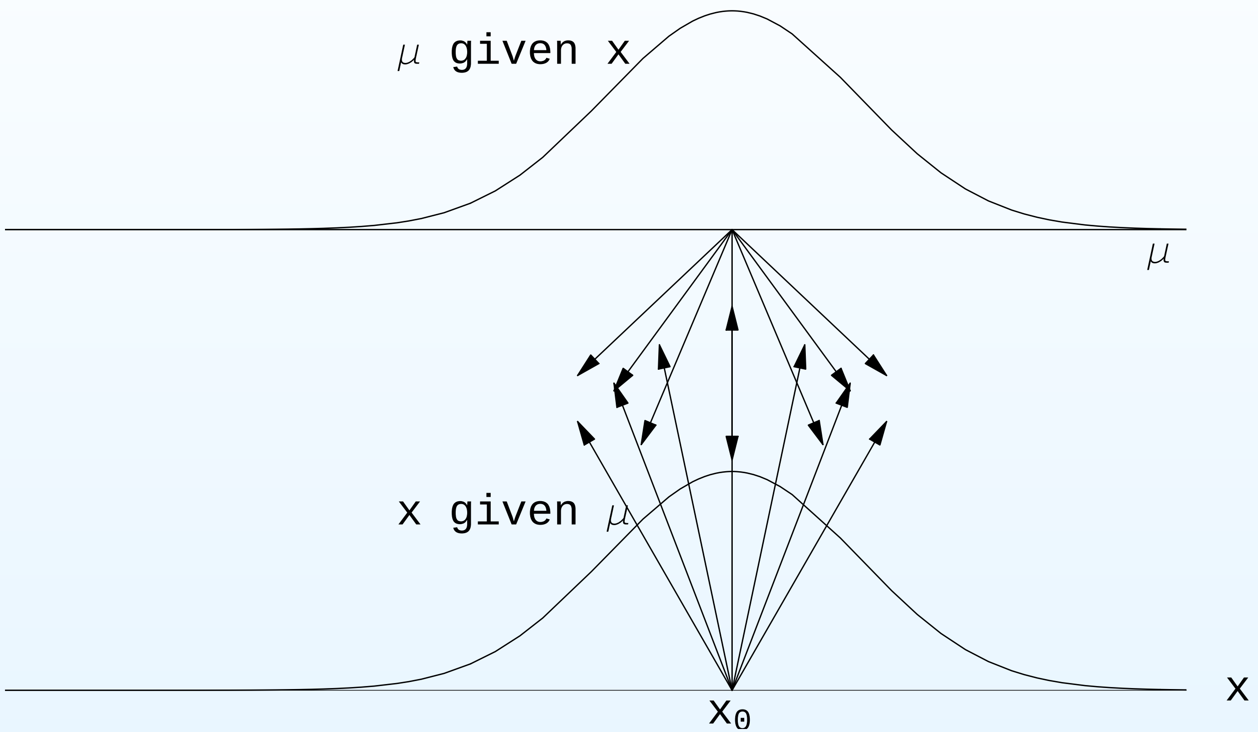
Where does the observed value of x comes from?

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

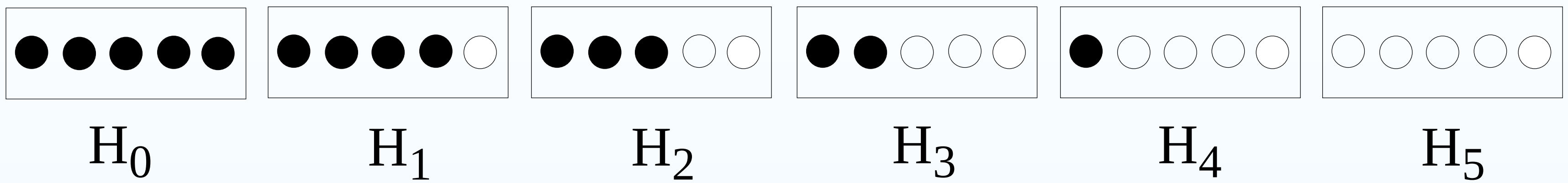
The human mind is used to live — and survive — in conditions of uncertainty and has developed mental categories to handle it.

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”.

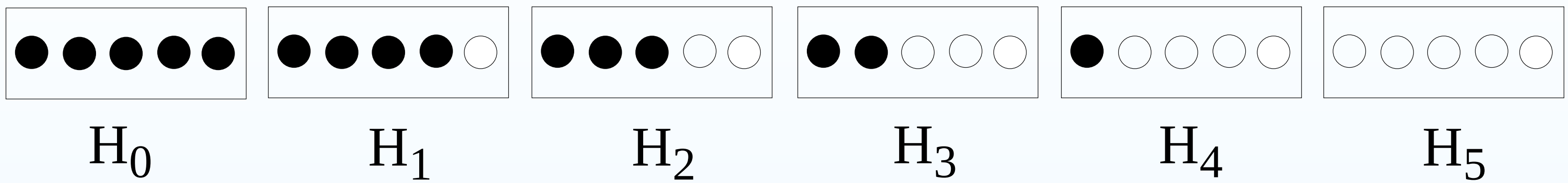
We can use similar expressions, all referring to the intuitive idea of probability.

The six box problem



Let us take randomly one of the boxes.

The six box problem



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:

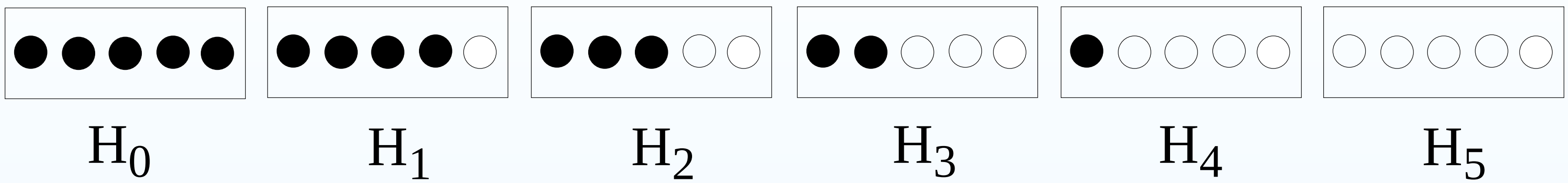
(a) Which box have we chosen, $H_0, H_1, \ldots, H_5$?

(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 certainty:

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

The six box problem



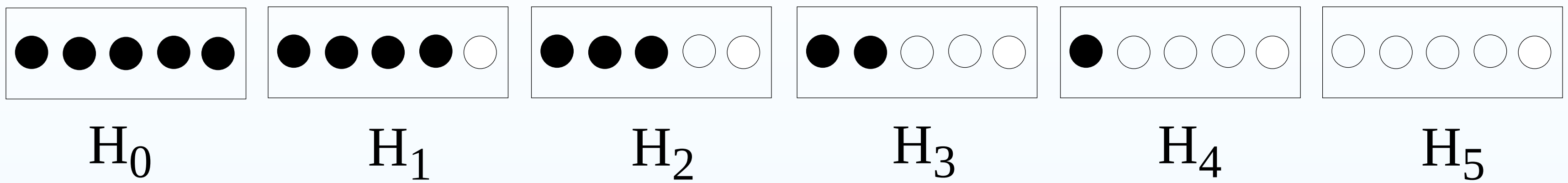
Let us take randomly one of the boxes.
We are in a state of uncertainty concerning several *events*, the most important of which correspond to the following questions:

(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?

- What happens after we have extracted one ball and looked its color?
  - Intuitively we now how to roughly change our opinion.
  - Can we do it quantitatively, in an objective way?

The six box problem



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?

- What happens after we have extracted one ball and looked its color?
  - Intuitively we now how to roughly change our opinion.
  - 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

- 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 and electron and read its properties, like we read the MAC address of a PC interface)

Doing 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 Science in conditions of uncertainty

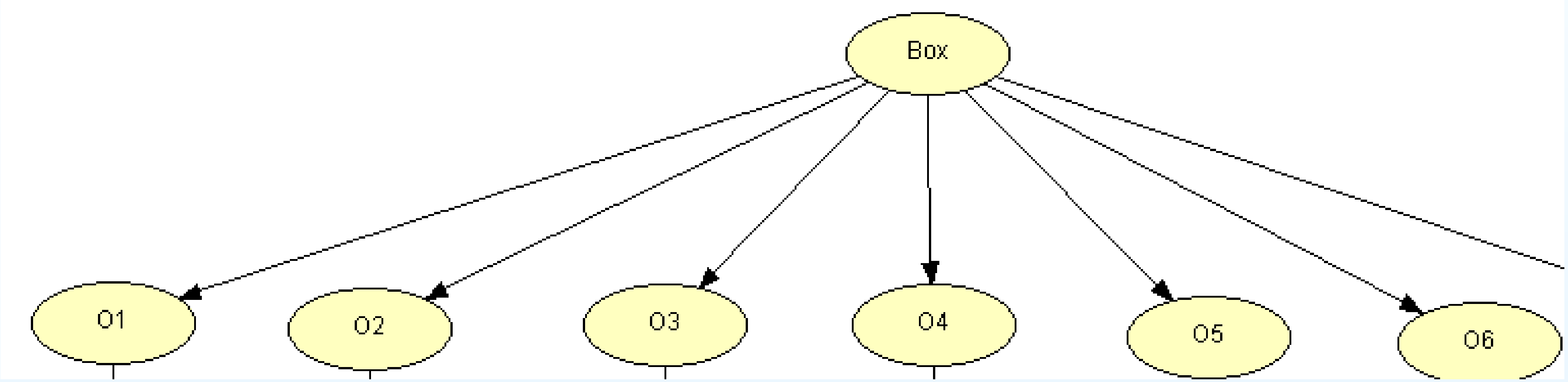
The constant status of uncertainty does not prevent us from doing Science (in the sense of Natural Science and not just Mathematics)

Indeed

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

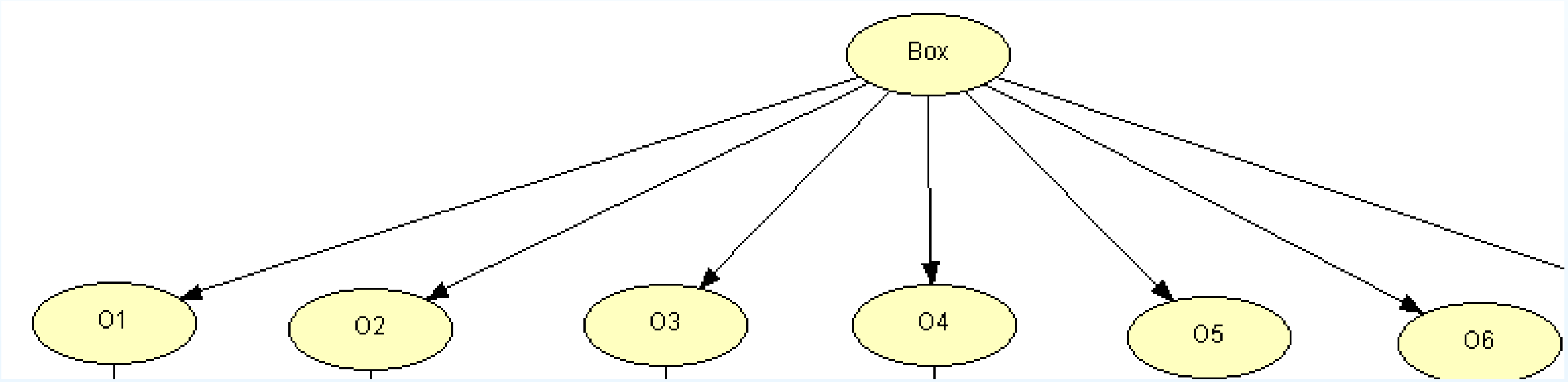
Cause-effect representation

box content $\longrightarrow$ observed color



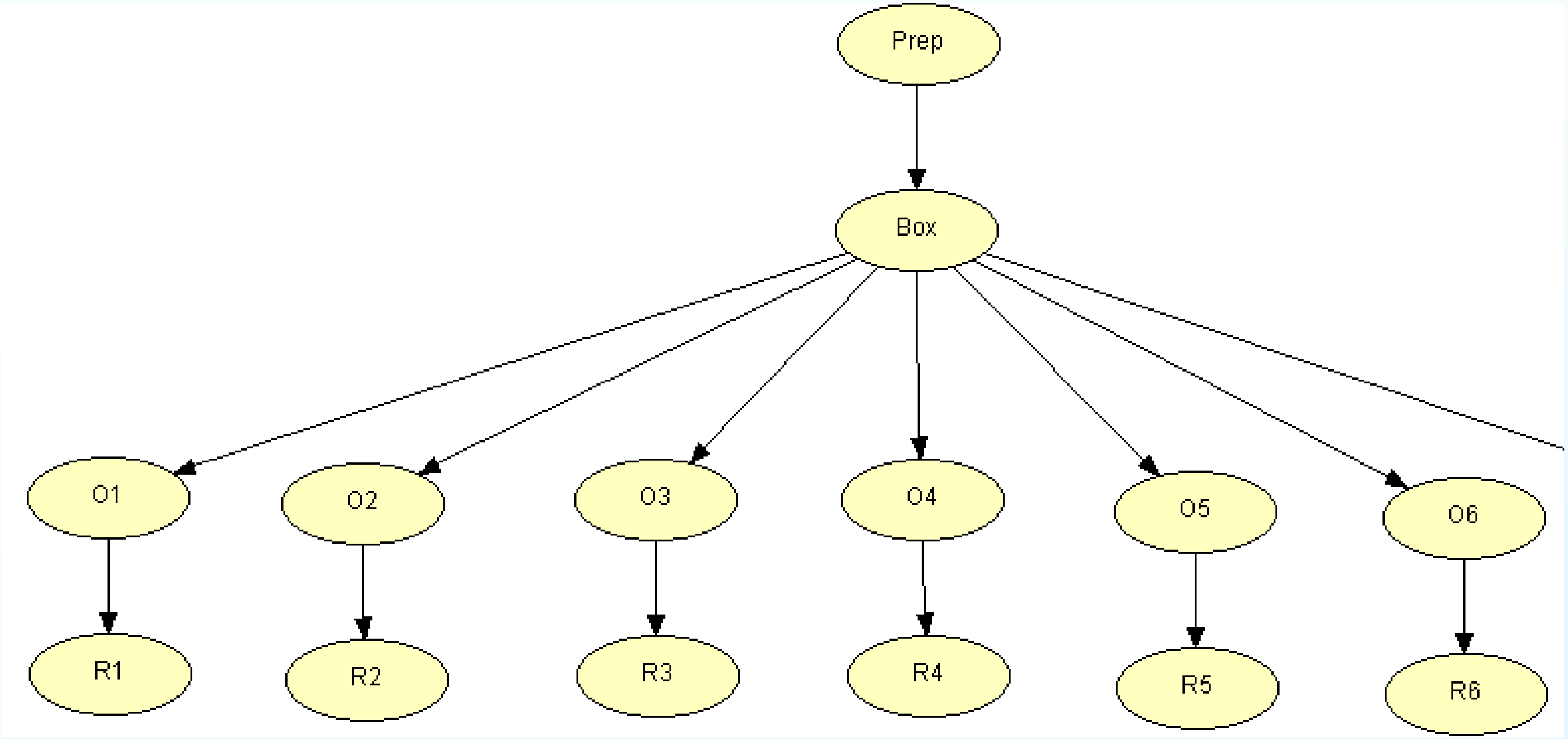
Cause-effect representation

box content $\longrightarrow$ observed color

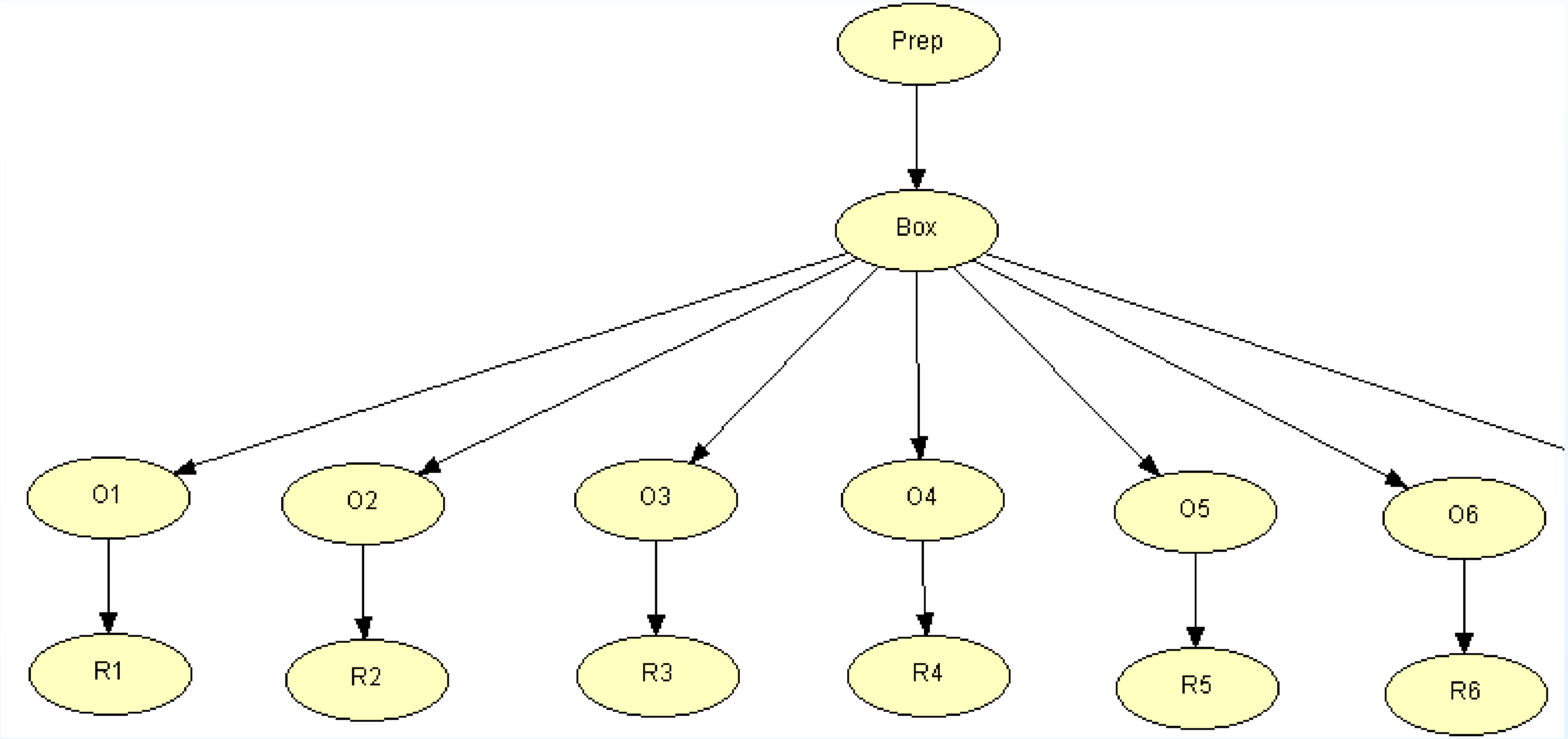


An effect might be the cause of another effect $\longrightarrow$

A network of causes and effects

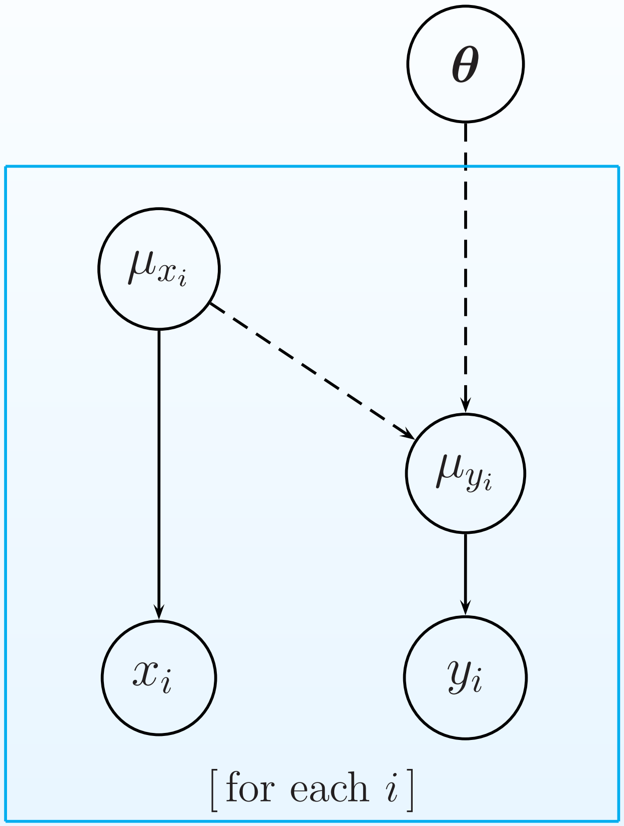


A network of causes and effects



and so on. . .　　$\Rightarrow$ Physics applications

A different way to view fit issues



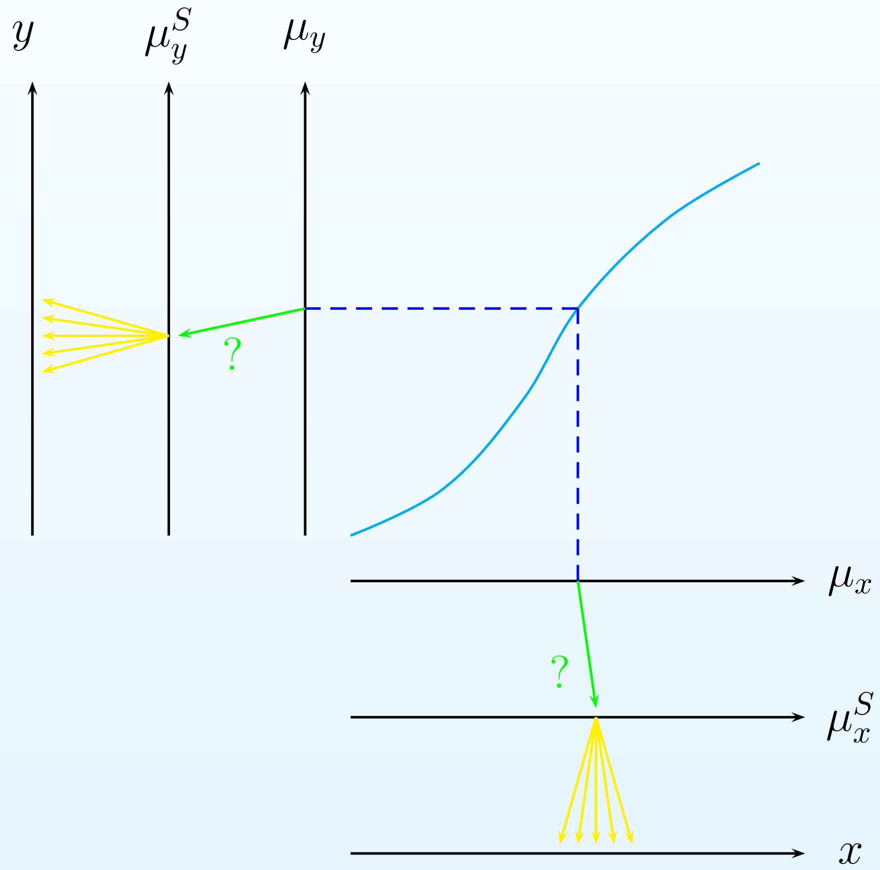
Determistic link μ_x's to μ_y's

Probabilistic links $\mu_x \longrightarrow x$, $\mu_y \longrightarrow y$

(errors on both axes!)

$\Rightarrow$ aim of fit: $\{\boldsymbol{x}, \boldsymbol{y}\} \longrightarrow \boldsymbol{\theta}$

A different way to view fit issues



How to quantify all that?

How to quantify all that?

- **Falsificationist approach**

  [and statistical variations over the theme].

How to quantify all that?

- **Falsificationist approach**

  [and statistical variations over the theme].

- **Probabilistic approach**

  [In the sense that probability theory is used throughly]

How to quantify all that?

- **Falsificationist approach**

  [and statistical variations over the theme].

- **Probabilistic approach**

  [In the sense that probability theory is used throughly]

  e . . .

How to quantify all that?

- **Falsificationist approach**

  [and statistical variations over the theme].

- **Probabilistic approach**

  [In the sense that probability theory is used throughly]

e . . .

- **. . . Fisichettume**

  [Le varie formulette di “calcolo e propagazione degli errori”]

$\Rightarrow$ **Segue su lucidi**: vedi pp. 13-26 Ref. [2]

Falsificationism

Usually associated to the name of Popper

and considered to be the *key to scientific progress*.

Falsificationism

Usually associated to the name of Popper

and considered to be the *key to scientific progress*.

$$\textbf{if } C_i \not\longrightarrow E, \textbf{ then } E_{\textbf{obs}} \not\longrightarrow C_i$$

$\Rightarrow$ Causes that cannot produce observed effects are ruled out ('falsified').

Falsificationism

Usually associated to the name of Popper

and considered to be the *key to scientific progress*.

$$\text{if } C_i \not\longrightarrow E, \text{ then } E_{\text{obs}} \not\longrightarrow C_i$$

$\Rightarrow$ Causes that cannot produce observed effects are ruled out ('falsified').

It seems OK, but it is naive for several aspects.

Falsificationism

Usually associated to the name of Popper

and considered to be the *key to scientific progress*.

$$\textbf{if } C_i \not\longrightarrow E, \textbf{ then } E_{\textbf{obs}} \not\longrightarrow C_i$$

$\Rightarrow$ Causes that cannot produce observed effects are ruled out ('falsified').

It seems OK, but it is naive for several aspects.

Let start realizing that the method is analogous with method of the proof by contradiction of classical, deductive logic.

Falsificationism

Usually associated to the name of Popper

and considered to be the *key to scientific progress*.

$$\text{if } C_i \not\longrightarrow E, \text{ then } E_{\text{obs}} \not\longrightarrow C_i$$

$\Rightarrow$ Causes that cannot produce observed effects are ruled out ('falsified').

It seems OK, but it is naive for several aspects.

Let start realizing that the method is analogous with method of the proof by contradiction of classical, deductive logic.

- Assume that a hypothesis is true
- Derive 'all' logical consequence
- If (at least) one of the consequences is known to be false, then the hypothesis is declared false.

Falsificationism? OK, but. . .

- What to do of all hypotheses that are not falsified? (Limbus? Get stuck?)

Falsificationism? OK, but...

- What to do of all hypotheses that are not falsified? (Limbus? Get stuck?)
- What to do is nothing of what can be observed is incompatible with the hypothesis (or with many hypotheses)?

  E.g. H_i being a Gaussian $f(x \,|\, \mu_i, \sigma_i)$

  $\Rightarrow$ Given any pair or parameters $\{\mu_i, \sigma_i\}$, <u>all values</u> of x between $-\infty$ and $+\infty$ <u>are possible</u>.

  $\Rightarrow$ Having observed any value of x, <u>none</u> of H_i can be, strictly speaking, <u>falsified</u>.

Falsificationism and statistics

. . . then, statisticians have invented the “hypothesis tests”

Falsificationism and statistics

... then, statisticians have invented the "hypothesis tests"

in which the impossible is replaced by the improbable!

Falsificationism and statistics

. . . then, statisticians have invented the “hypothesis tests”

in which the impossible is replaced by the improbable!

But from the impossible to the improbable there is not just a question of quantity, but a question of quality.

This mechanism, logically flawed, is particularly perverse, because deeply rooted in most people, due to education, but is not supported by logic.

$\Rightarrow$ Basically responsible of all fake claims of discoveries in the past decades.

[I am particularly worried about claims concerning our health, or the status of the planet, of which I have no control of the experimental data.]

In summary

A) **if** $C_i \not\longrightarrow E$, and **we observe** E

$\Rightarrow C_i$ is impossible ('false')

In summary

A) **if** $C_i \not\longrightarrow E$, and **we observe** E

$\Rightarrow C_i$ is impossible (‘false’)

B) **if** $C_i \xrightarrow[\text{small probability}]{} E$, and **we observe** E

$\Rightarrow C_i$ has small probability to be true

“most likely false”

In summary

A) **if** $C_i \not\longrightarrow E$, and **we observe** E OK

$\Rightarrow C_i$ is impossible ('false')

B) **if** $C_i \xrightarrow[\text{small probability}]{} E$, and **we observe** E

$\Rightarrow C_i$ has small probability to be true

"most likely false"

In summary

A) **if** $C_i \not\longrightarrow E$, and **we observe** E OK

$\Rightarrow C_i$ is impossible (‘false’)

B) **if** $C_i \xrightarrow[\text{small probability}]{\hspace{4em}} E$, and **we observe** E NO

$\Rightarrow C_i$ has small probability to be true

“most likely false”

Example 1

Playing lotto

H: “I play honestly at lotto, betting on a rare combination”
E: “I win”

$$H \xrightarrow[\text{“practically impossible”}]{\hspace{4cm}} E$$

Example 1

Playing lotto

H: “I play honestly at lotto, betting on a rare combination”

E: “I win”

$$\begin{tikzcd}
H \arrow[r, "\text{“practically impossible”}"'] & E \arrow[l, bend left=60, red]
\end{tikzcd}$$

“practically to exclude”

Example 1

Playing lotto

H: “I play honestly at lotto, betting on a rare combination”

E: “I win”

$$\begin{tikzcd} H \arrow[rr, "\text{“practically impossible”}"'] & & E \arrow[ll, bend left=60] \end{tikzcd}$$

“practically to exclude”

$\Rightarrow$ almost certainly I have cheated…

(or it is false that I won…)

Example 2

An Italian citizen is selected at random to undergo an AIDS test. Performance of clinical trial is not perfect, as customary.

Toy model:

$$\begin{aligned} P(\mathrm{Pos}\,|\,\mathrm{HIV}) &= 100\% \\ P(\mathrm{Pos}\,|\,\overline{\mathrm{HIV}}) &= 0.2\% \\ P(\mathrm{Neg}\,|\,\overline{\mathrm{HIV}}) &= 99.8\% \end{aligned}$$

H_1='HIV' (Infected) E_1 = Positive

H_2='$\overline{\mathrm{HIV}}$' (Healthy) E_2 = Negative

Example 2

An Italian citizen is selected at random to undergo an AIDS test. Performance of clinical trial is not perfect, as customary.

Toy model:

$$
\begin{aligned}
P(\text{Pos}\,|\,\text{HIV}) &= 100\% \\
P(\text{Pos}\,|\,\overline{\text{HIV}}) &= 0.2\% \\
P(\text{Neg}\,|\,\overline{\text{HIV}}) &= 99.8\%
\end{aligned}
$$

$$
\begin{tikzcd}
H_1=\text{'HIV' (Infected)} \arrow[r] & E_1 = \text{Positive} \\
H_2=\text{'}\overline{\text{HIV}}\text{' (Healthy)} \arrow[r] \arrow[ur] & E_2 = \text{Negative}
\end{tikzcd}
$$

Example 2

An Italian citizen is selected at random to undergo an AIDS test. Performance of clinical trial is not perfect, as customary.

Toy model:

$$\begin{aligned} P(\text{Pos}\,|\,\text{HIV}) &= 100\% \\ P(\text{Pos}\,|\,\overline{\text{HIV}}) &= 0.2\% \\ P(\text{Neg}\,|\,\overline{\text{HIV}}) &= 99.8\% \end{aligned}$$

$$\begin{tikzcd} H_1\text{='HIV' (Infected)} \arrow[r] & E_1 \text{ = Positive} \\ H_2\text{='}\overline{\text{HIV}}\text{' (Healthy)} \arrow[r] \arrow[ur] & E_2 \text{ = Negative} \end{tikzcd}$$

Result: $\Rightarrow$ <u>Positive</u>

Example 2

An Italian citizen is selected at random to undergo an AIDS test. Performance of clinical trial is not perfect, as customary.

Toy model:

$$
\begin{aligned}
P(\text{Pos}\,|\,\text{HIV}) &= 100\% \\
P(\text{Pos}\,|\,\overline{\text{HIV}}) &= 0.2\% \\
P(\text{Neg}\,|\,\overline{\text{HIV}}) &= 99.8\%
\end{aligned}
$$

? H_1='HIV' (Infected) $\longleftarrow$ E_1 = Positive

? H_2='$\overline{\text{HIV}}$' (Healthy) $\quad$ E_2 = Negative

Result: $\Rightarrow$ <u>Positive</u>

Infected or healthy?

Example 2

Being $P(\text{Pos}\,|\,\overline{\text{HIV}}) = 0.2\%$ and having observed 'Positive', can we say

- ”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”?

Example 2

Being $P(\text{Pos} \,|\, \overline{\text{HIV}}) = 0.2\%$ and having observed ‘Positive’,
can we say

- ”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”
- “There is only 0.2% probability that the person has no HIV” ?

Example 2

Being $P(\text{Pos} \,|\, \overline{\text{HIV}}) = 0.2\%$ and having observed 'Positive',
can we say

- ”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”
- “There is only 0.2% probability that the person has no HIV”
- “We are 99.8% confident that the person is infected”?

Example 2

Being $P(\text{Pos}\,|\,\overline{\text{HIV}}) = 0.2\%$ and having observed ‘Positive’, can we say

- ”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”
- “There is only 0.2% probability that the person has no HIV”
- “We are 99.8% confident that the person is infected”
- “The hypothesis H_1=Healthy is ruled out with 99.8% C.L.”

?

Example 2

Being $P(\text{Pos}\,|\,\overline{\text{HIV}}) = 0.2\%$ and having observed ‘Positive’,
can we say

- ~~”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”~~
- ~~“There is only 0.2% probability that the person has no HIV”~~
- ~~“We are 99.8% confident that the person is infected”~~
- ~~“The hypothesis H_1=Healthy is ruled out with 99.8% C.L.”~~

? NO

Instead, $P(\text{HIV}\,|\,\text{Pos, random Italian}) \approx 45\%$
(We will see in the sequel how to evaluate it correctly)

Example 2

Being $P(\text{Pos} \,|\, \overline{\text{HIV}}) = 0.2\%$ and having observed 'Positive', can we say

- ~~"It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive"~~
- ~~"There is only 0.2% probability that the person has no HIV"~~
- ~~"We are 99.8% confident that the person is infected"~~
- ~~"The hypothesis H_1=Healthy is ruled out with 99.8% C.L."~~

? NO

Instead, $P(\text{HIV} \,|\, \text{Pos, random Italian}) \approx 45\%$

$\Rightarrow$ Serious mistake! (not just 99.8% instead of 98.3% or so)

Example 2

Being $P(\text{Pos}\,|\,\overline{\text{HIV}}) = 0.2\%$ and having observed ‘Positive’,
can we say

- ~~”It is practically impossible that the person is healthy, since it was practically impossible that an healthy person would result positive”~~
- ~~“There is only 0.2% probability that the person has no HIV”~~
- ~~“We are 99.8% confident that the person is infected”~~
- ~~“The hypothesis H_1=Healthy is ruled out with 99.8% C.L.”~~

?

NO

Instead, $P(\text{HIV}\,|\,\text{Pos, random Italian}) \approx 45\%$

$\Rightarrow$ Serious mistake! (not just 99.8% instead of 98.3% or so)

... which might result into very bad decisions!

‘Standard’ statistical tests, p-values, etc

- This kind of logical mistake is quite common.
  “Si sbaglia da professionisti” (P. Conte)

‘Standard’ statistical tests, p-values, etc

- This kind of logical mistake is quite common.
  “Si sbaglia da professionisti” (P. Conte)
- Yes, statisticians have invented p-values (something like ‘probability of the tail(s)’ – I cannot enter into details) to overcome the problem that often the probability of any observation is always very small and the null hypotheses would always be rejected.

‘Standard’ statistical tests, p-values, etc

- This kind of logical mistake is quite common.
  “Si sbaglia da professionisti” (P. Conte)
- Yes, statisticians have invented p-values (something like ‘probability of the tail(s)’ – I cannot enter into details) to overcome the problem that often the probability of any observation is always very small and the null hypotheses would always be rejected.
  But
  - as far as logic is concerned, the situation is worsened (. . . although p-values ‘often, by chance work’).

‘Standard’ statistical tests, p-values, etc

- This kind of logical mistake is quite common.
  “Si sbaglia da professionisti” (P. Conte)
- Yes, statisticians have invented p-values (something like ‘probability of the tail(s)’ – I cannot enter into details) to overcome the problem that often the probability of any observation is always very small and the null hypotheses would always be rejected.
  But
  - as far as logic is concerned, the situation is worsened (. . . although p-values ‘often, by chance work’).
- Mistrust statistical tests, unless you know the details of what it has been done.
  $\longrightarrow$ You might take bad decisions!

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 ⇔ 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)

⇒ 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!

What to do?

$\Rightarrow$ Back to the past

What to do?

$\Longrightarrow$ **Back to the past**

But benefitting 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 benefitting 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!

$\Rightarrow$ Use consistently probability theory

What to do?

⇒ Back to the past

But benefitting of

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

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

  → no longer an excuse!

⇒ Use consistently probability theory

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

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

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

Standard textbook definitions

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

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

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

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

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

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

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

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

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

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.

BUT they cannot define the concept of probability!

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 assumption of equiprobability, when it applies).
- Rule B will result from a theorem (under well defined assumptions).

Probability

What is probability?

What is probability?

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

What is probability?

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

$\rightarrow$ how much we are confident that something is true

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

Probability

What <u>is</u> probability?

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

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

[Remark: ‘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…

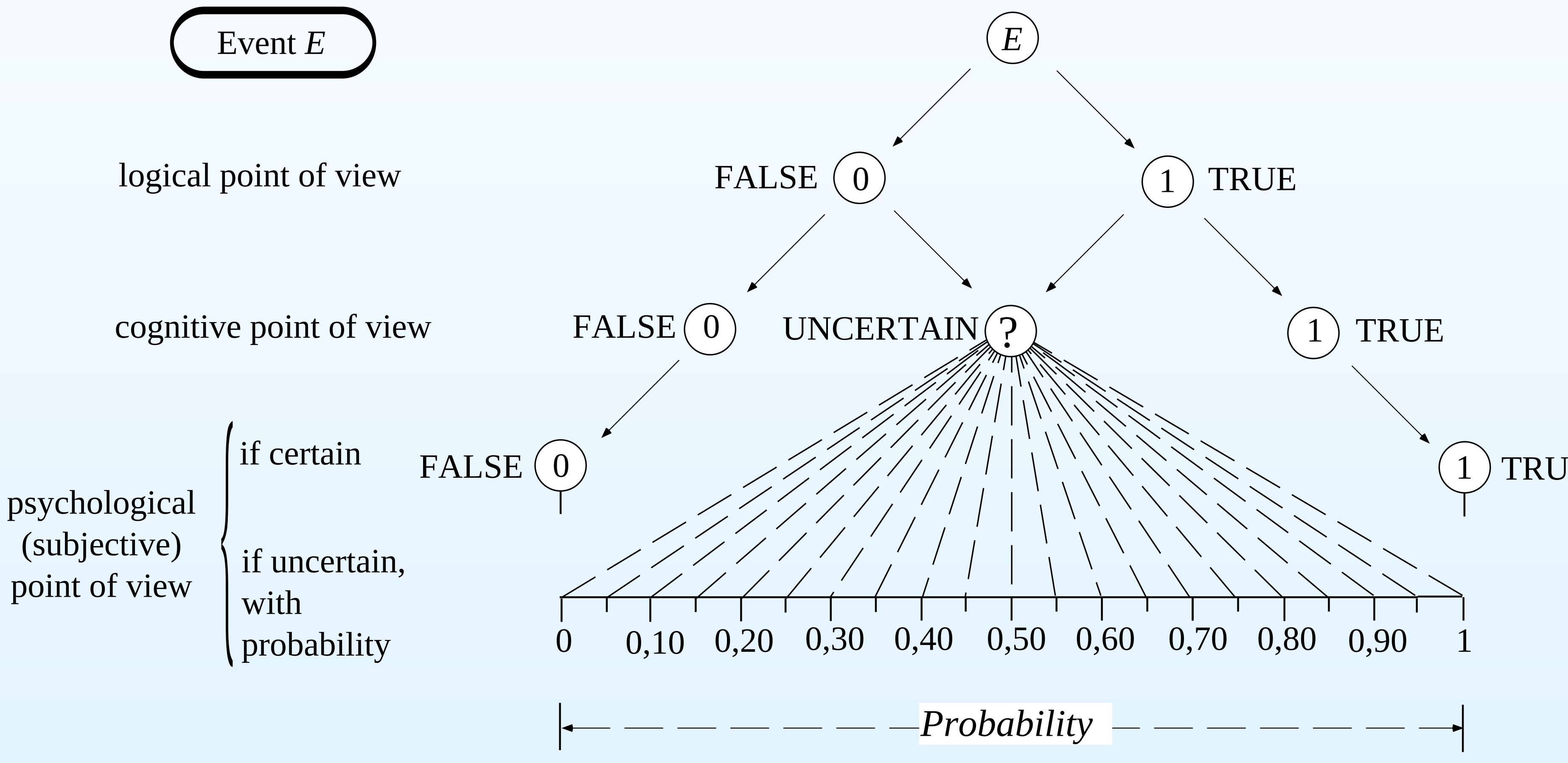
"Given the state of our knowledge about everything that could possible have any bearing on the coming true[1]*…, the numerical probability* p *of this event is to be a real number by the indication of which we try* in some cases *to setup a quantitative measure of the strength of our conjecture or anticipation,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

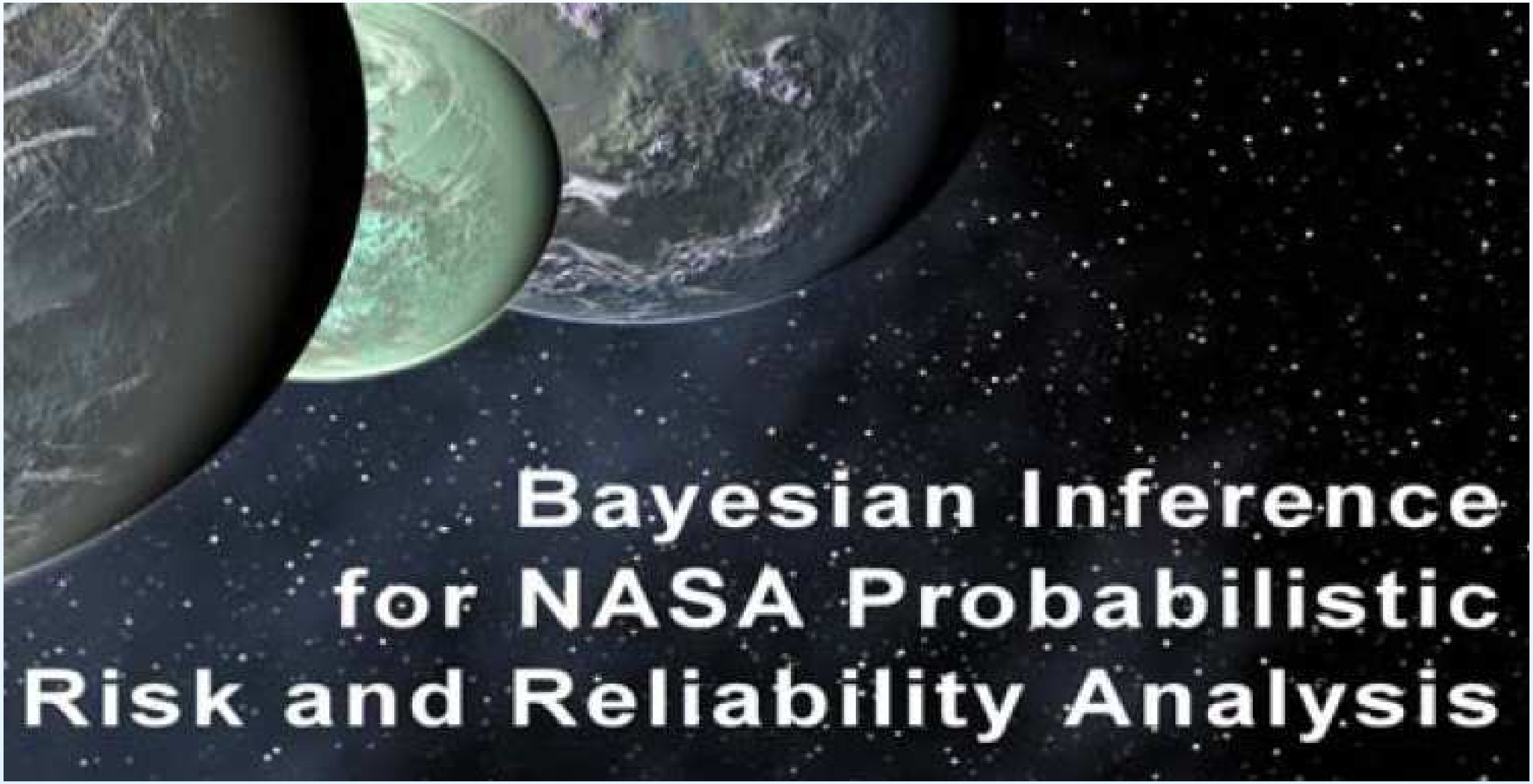
[1]*While in ordinary speech "to come true" usually refers to an event that is envisaged before it has happened, we use it here in the general sense, that the verbal description turns out to agree with actual facts.*

False, True and probable

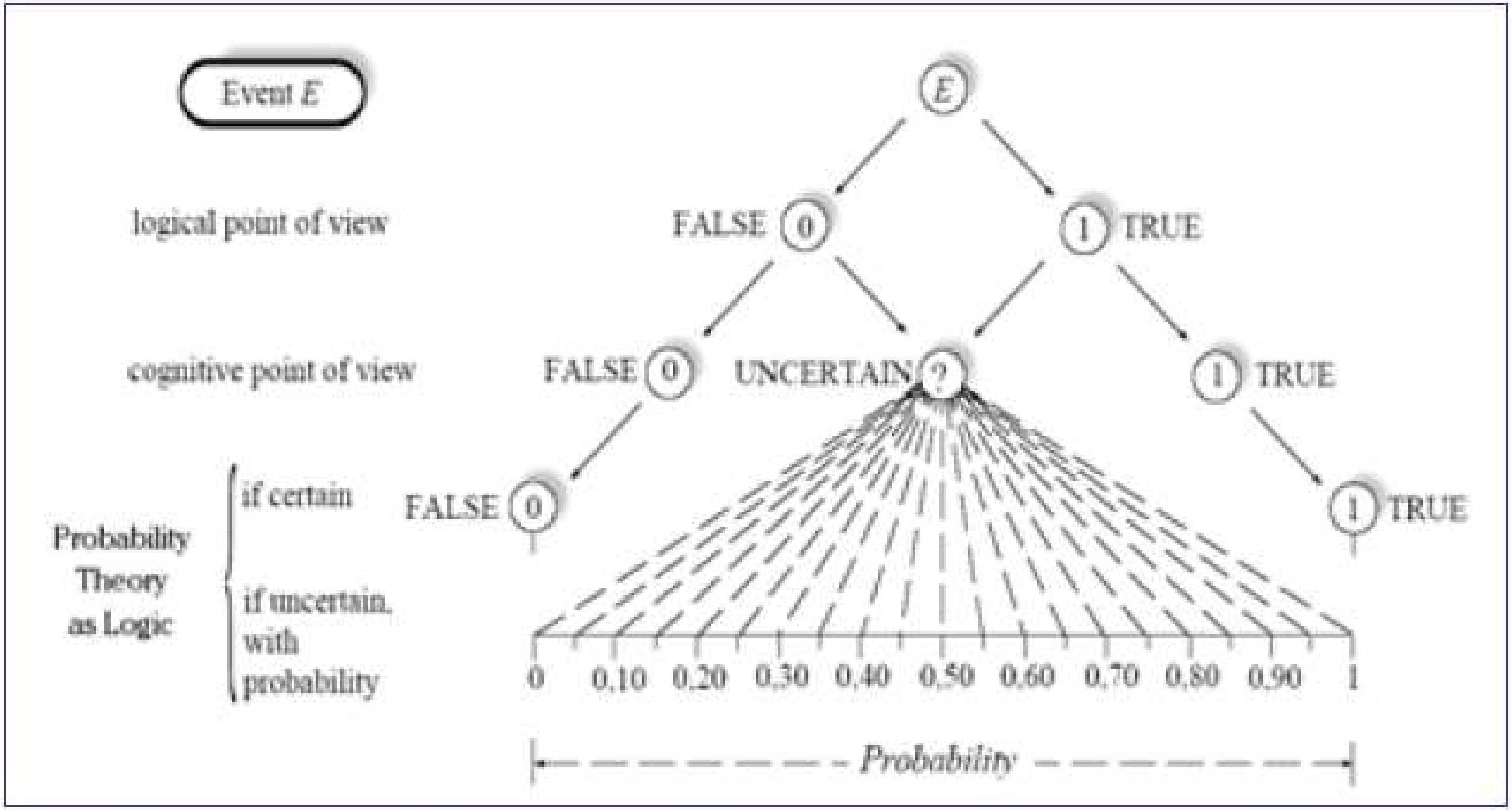


An helpful diagram

The previous diagram seems to help the understanding of the concept of probability



An helpful diagram



• Figure 2-1. Graphical abstraction of probability as a measure of information (adapted from "Probability and Measurement Uncertainty in Physics" by D'Agostini, [1995]).

(. . . but NASA guys are afraid of 'subjective', or 'psycological')