
Scoperte scientifiche annunciate a colpi di “sigma”

Giulio D'Agostini

`giulio.dagostini@roma1.infn.it`

Dipartimento di Fisica

Università di Roma La Sapienza

Abstract

La fisica di frontiera è caratterizzata da ricerche al limite della sensibilità strumentale e dal fatto che i fenomeni cercati possono essere 'mimati' da altri meno interessanti ('fondo'). Ne deriva che è impossibile arrivare a conclusioni logicamente *certe*, e si deve ricorrere a metodi statistici. Si sente allora parlare di evidenze sperimentali a 'n sigma'. Cosa significano? Il 2011 è stato un anno sotto questo aspetto esemplare. I neutrini superluminali di Opera erano dati a 10 sigma, mentre una nuova particella al Fermilab che avrebbe rivoluzionato la fisica a 3 sigma, più o meno come il bosone di Higgs al CERN.

Ad oggi soltanto l'Higgs è sopravvissuto, corroborato a 7-8 sigma (ma allora potrebbe ancora fare la stessa fine dei neutrini 'troppo veloci'?). Insomma, cosa rappresentano queste sigma e che informazione ci danno sulla probabilità che quanto annunciato sia veramente una scoperta e non l'ennesimo falso allarme?

2011: non solo Opera...

- Aprile, **CDF**: eccesso assolutamente inatteso a circa 150 GeV

$$\approx 3.2 \sigma$$

- Settembre, **Opera**: neutrini più veloci della luce

$$\approx 10 \sigma$$

- Dicembre, ATLAS e CMS a **LHC**: segnale compatibile con Higgs a circa 125 GeV:

$$\approx 3 \sigma$$

2011: non solo Opera...

- Aprile, **CDF**: eccesso assolutamente inatteso a circa 150 GeV

$$\approx 3.2 \sigma$$

- Settembre, **Opera**: neutrini più veloci della luce

$$\approx 10 \sigma$$

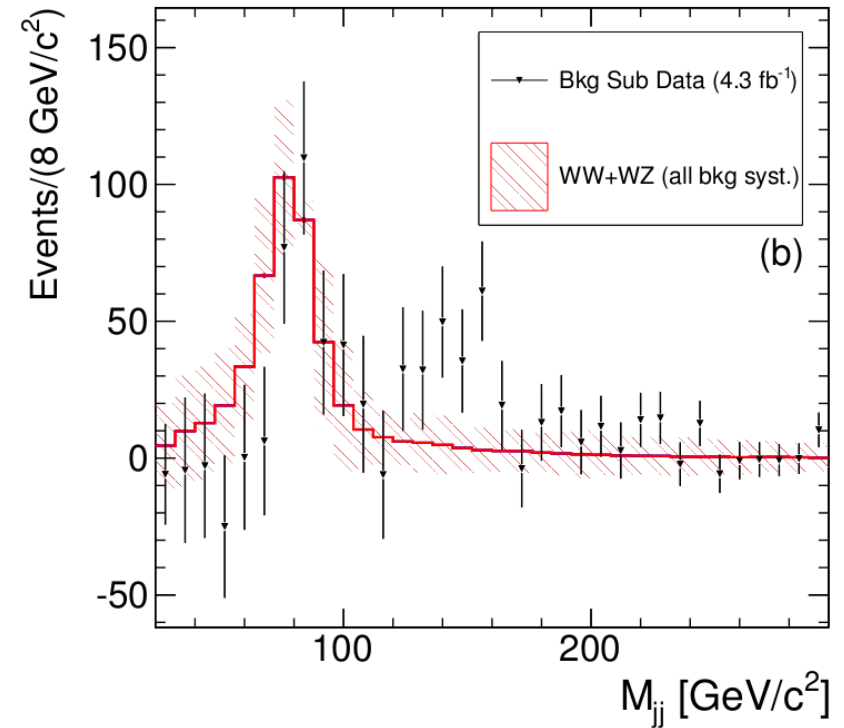
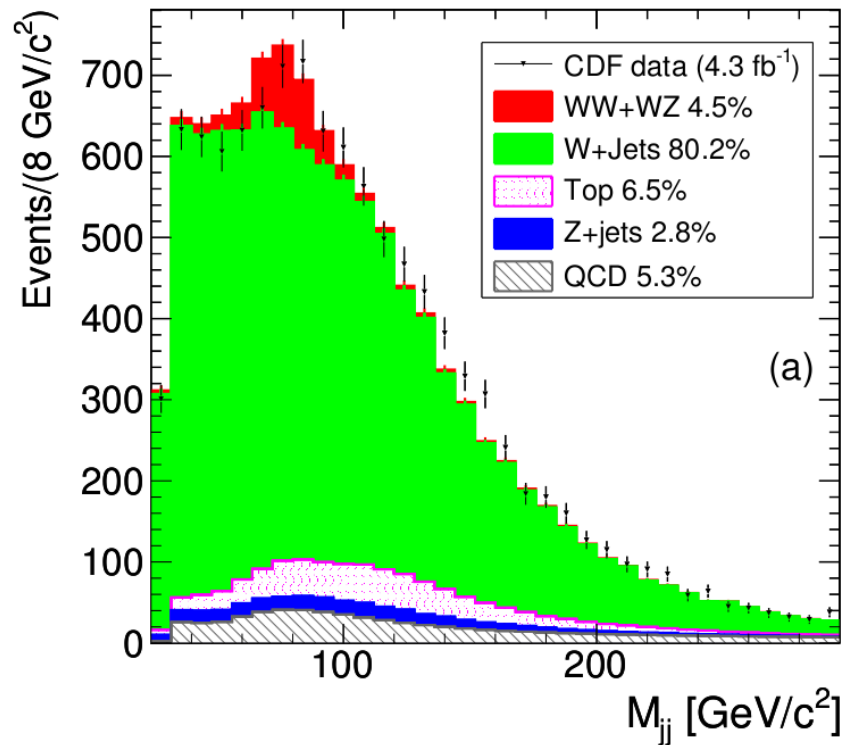
- Dicembre, ATLAS e CMS a **LHC**: segnale compatibile con Higgs a circa 125 GeV:

$$\approx 3 \sigma$$

Perché sostanziale **scetticismo per i primi due annunci**,
contro un cauto/deciso **ottimismo per il terzo?**

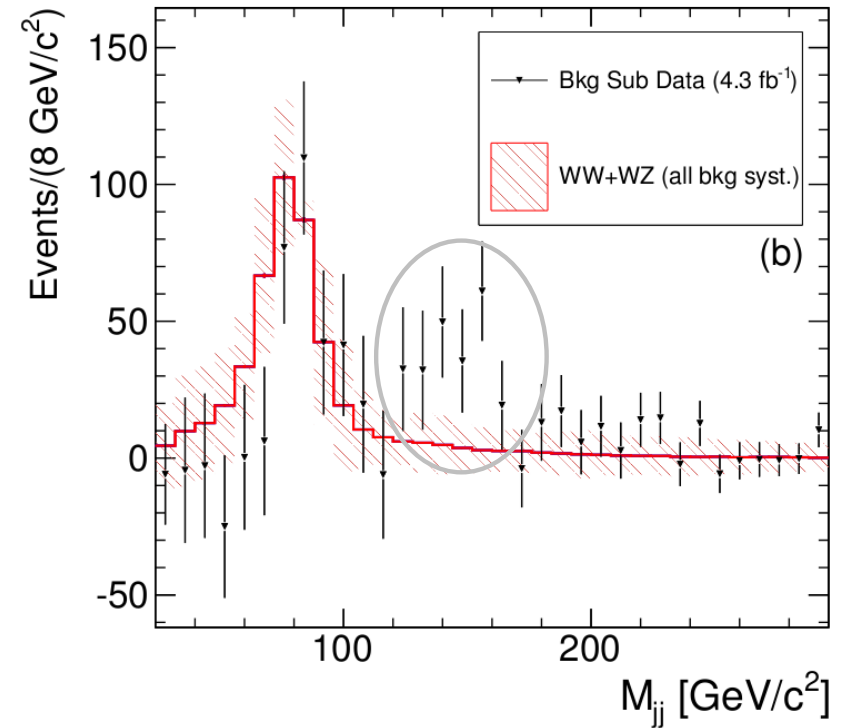
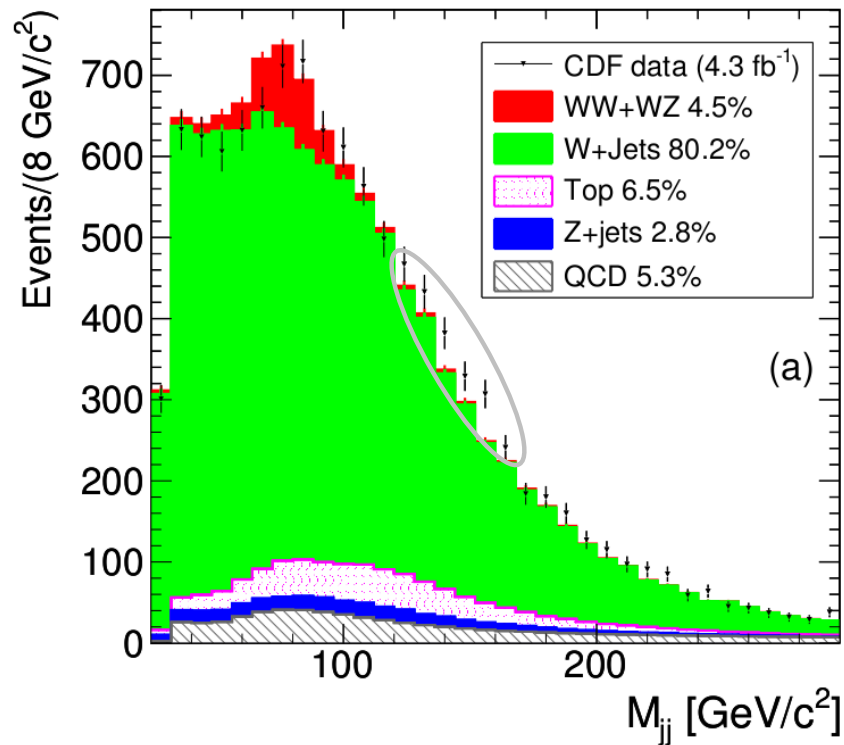
April 2011

CDF Collaboration at the Tevatron



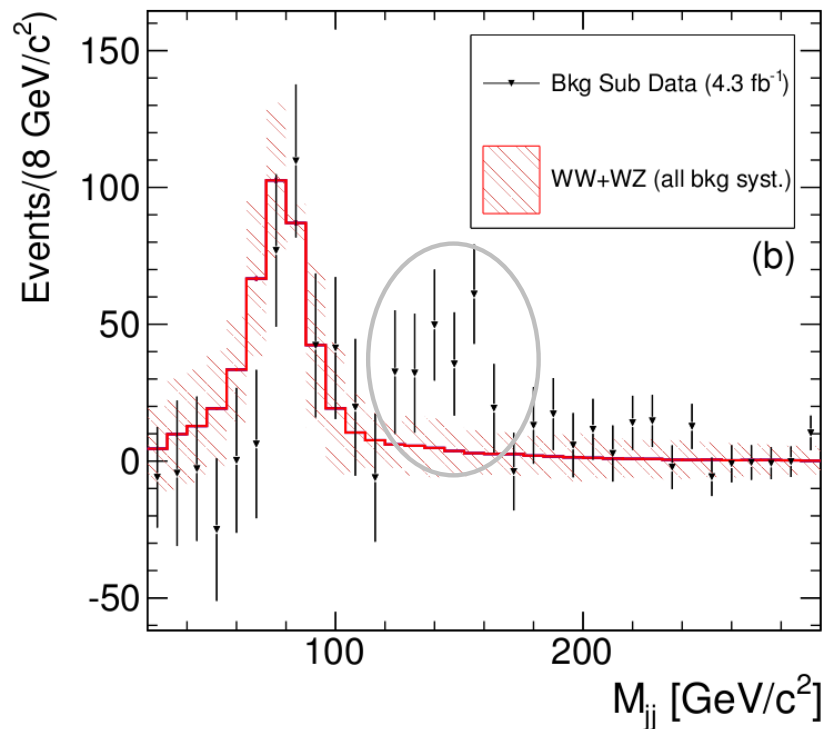
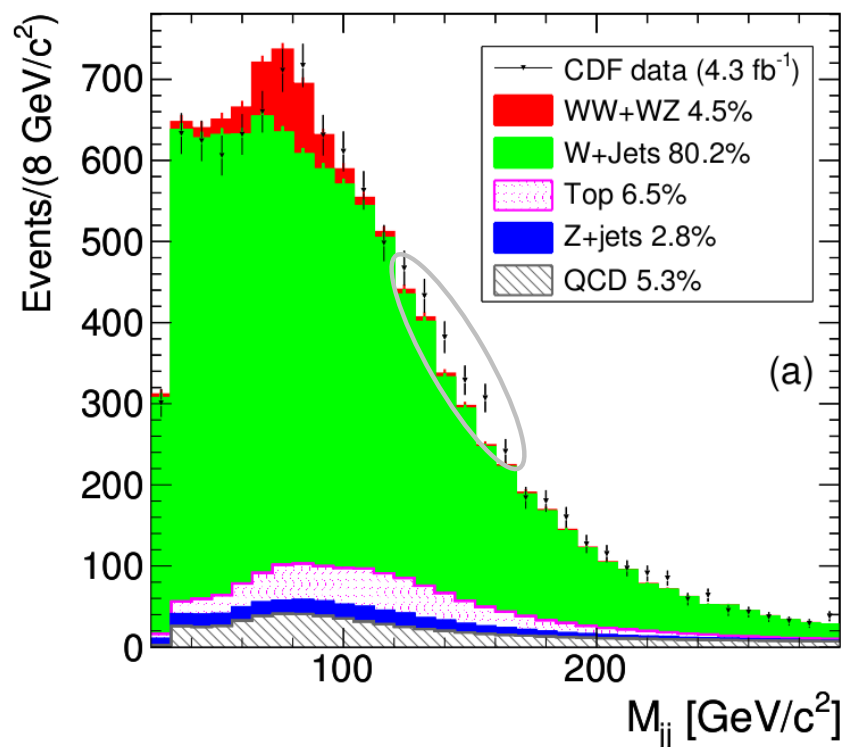
April 2011

CDF Collaboration at the Tevatron



April 2011

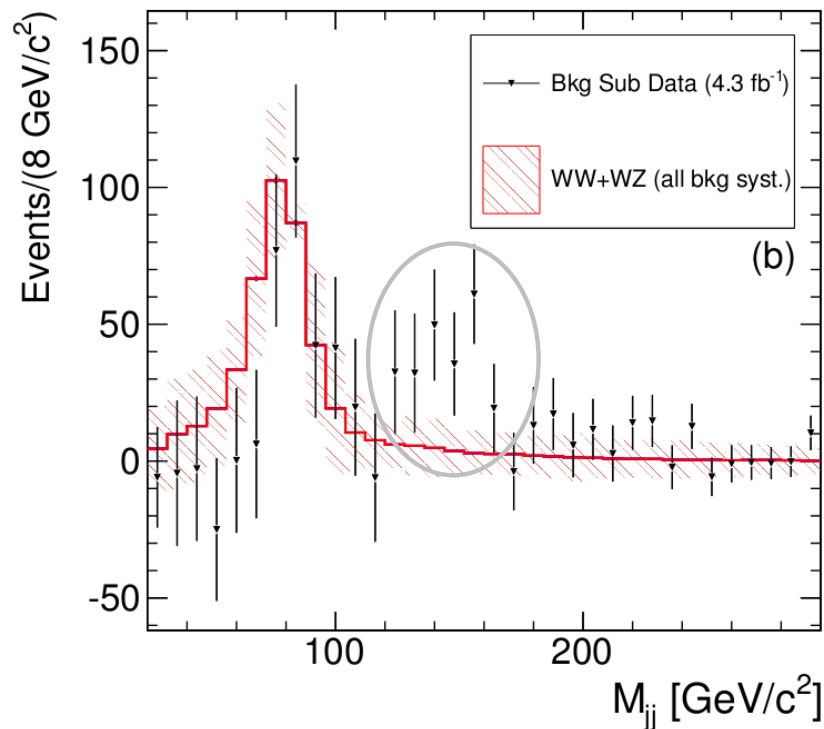
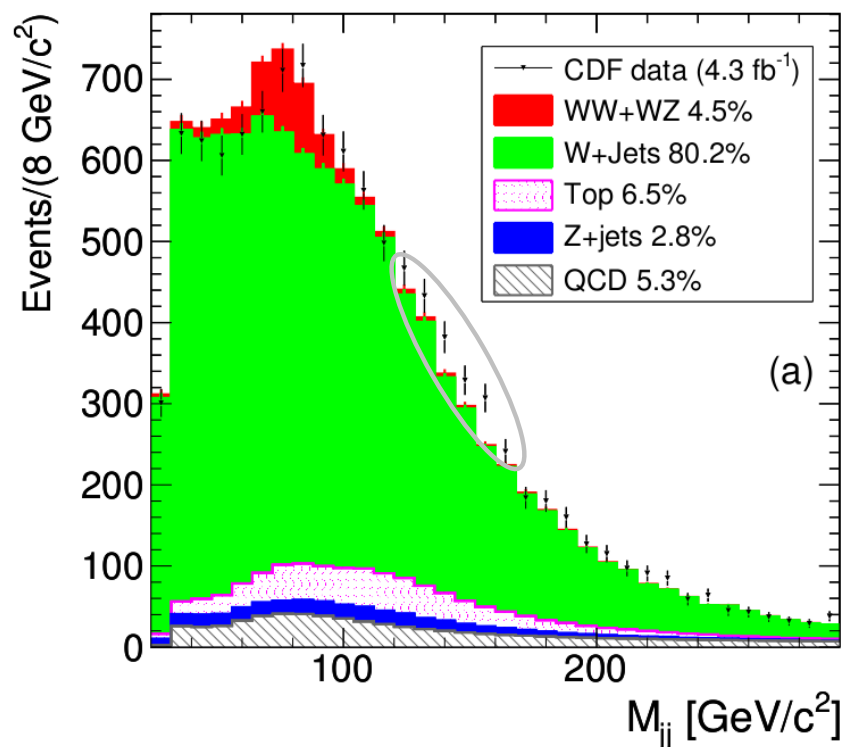
CDF Collaboration at the Tevatron



“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations”

April 2011

CDF Collaboration at the Tevatron

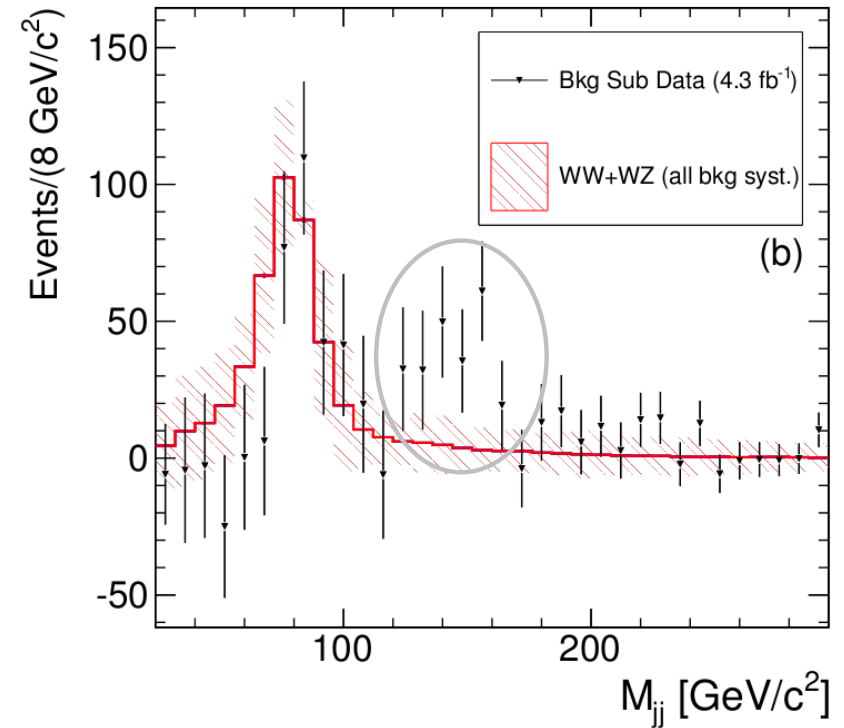
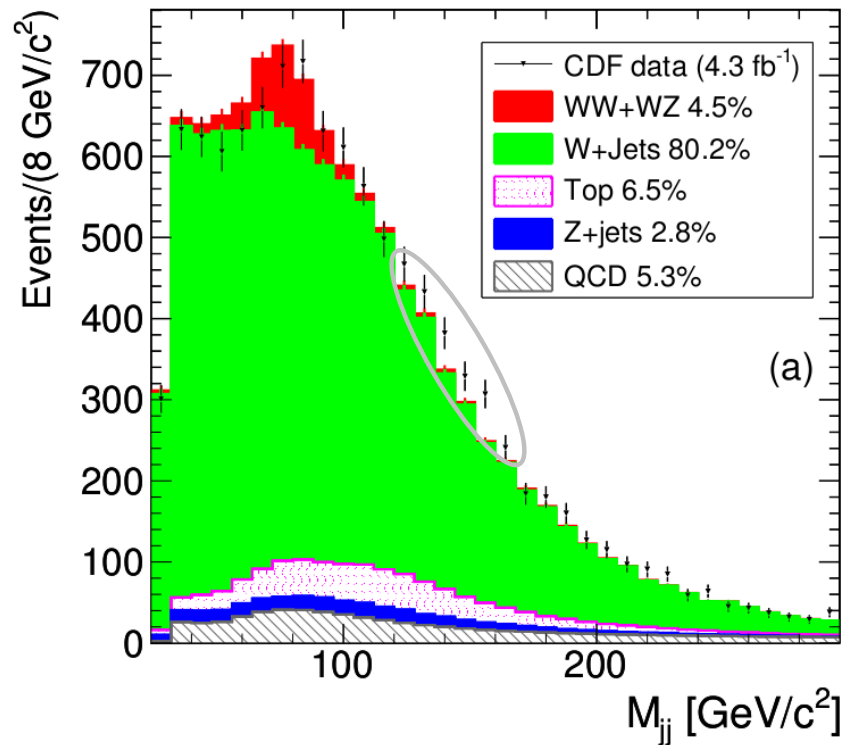


“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations”

3.2 σ !

April 2011

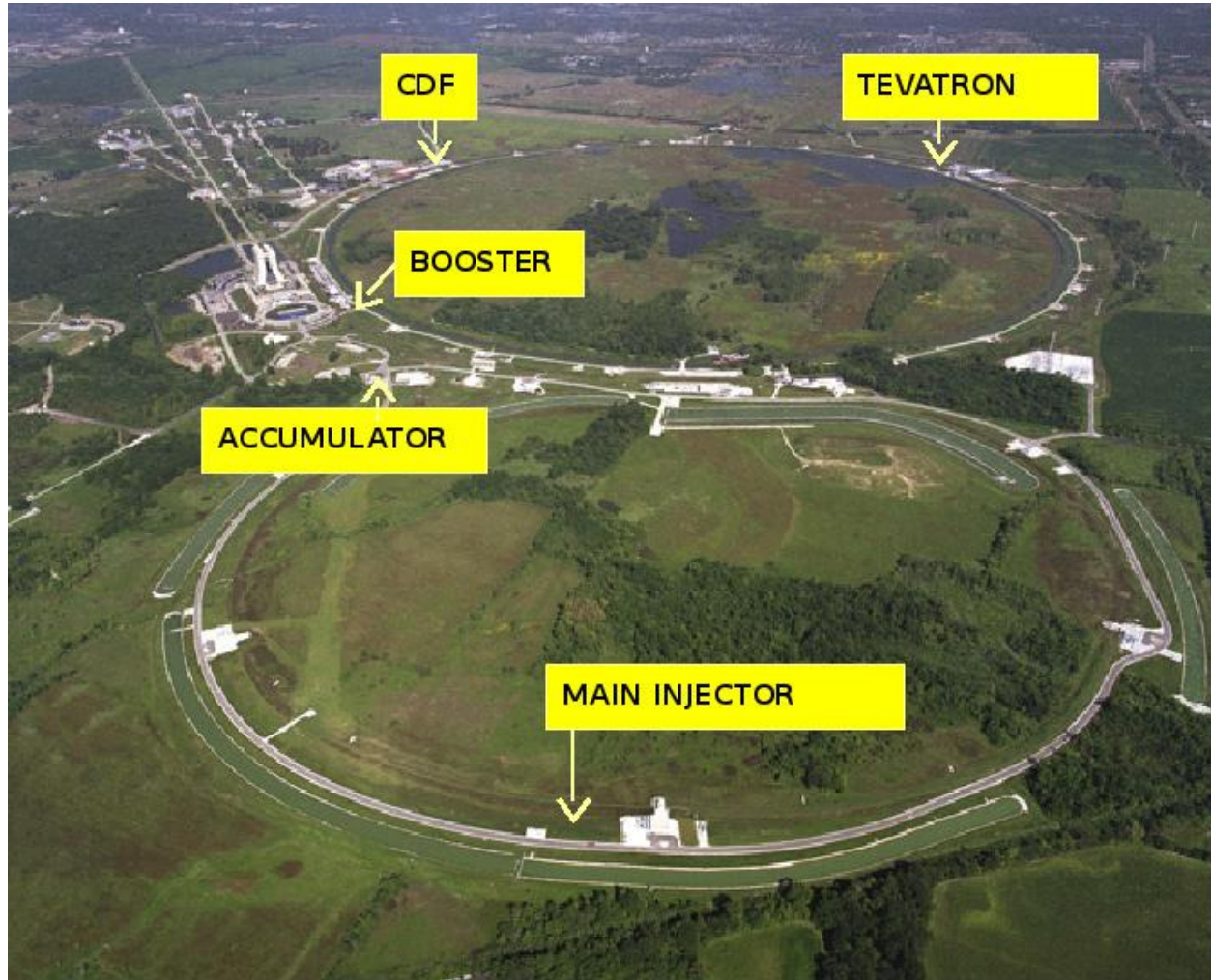
CDF Collaboration at the Tevatron



Che significa?

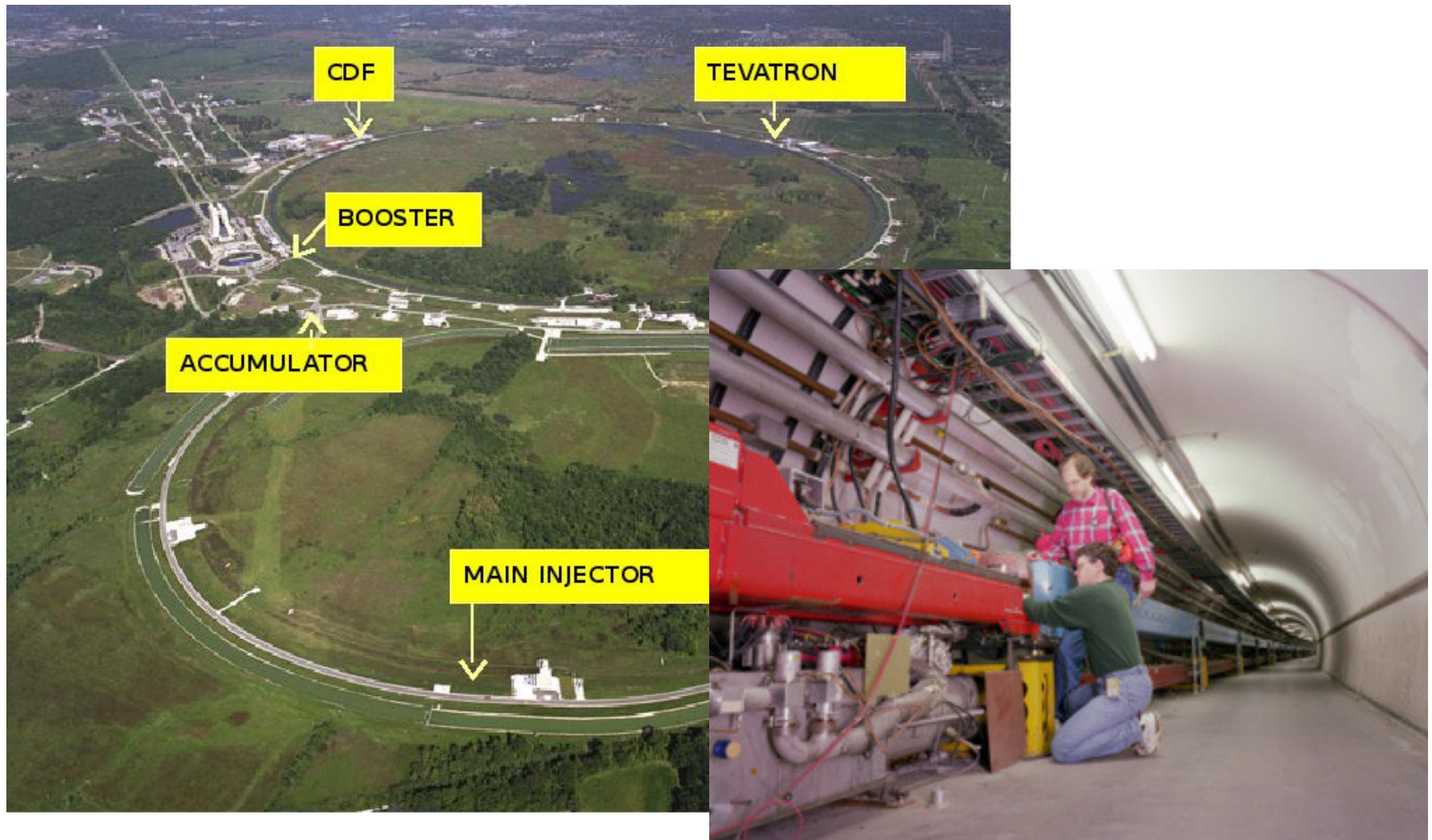
Tevatron and CDF

6.28 km, near Chicago



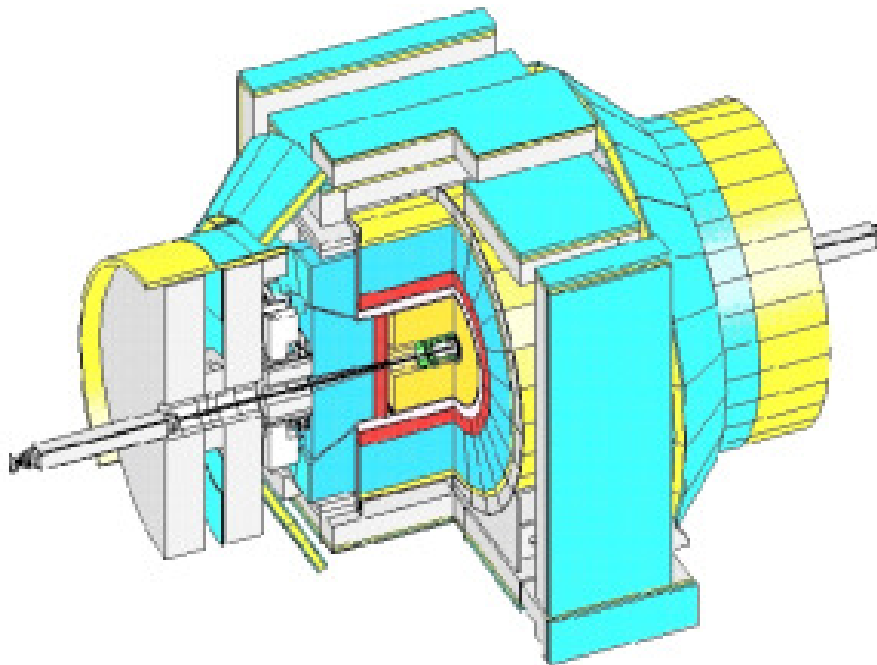
Tevatron and CDF

$$p \rightarrow \cdot \leftarrow \bar{p} \quad [\approx 1 \text{ TeV} + 1 \text{ TeV}]$$



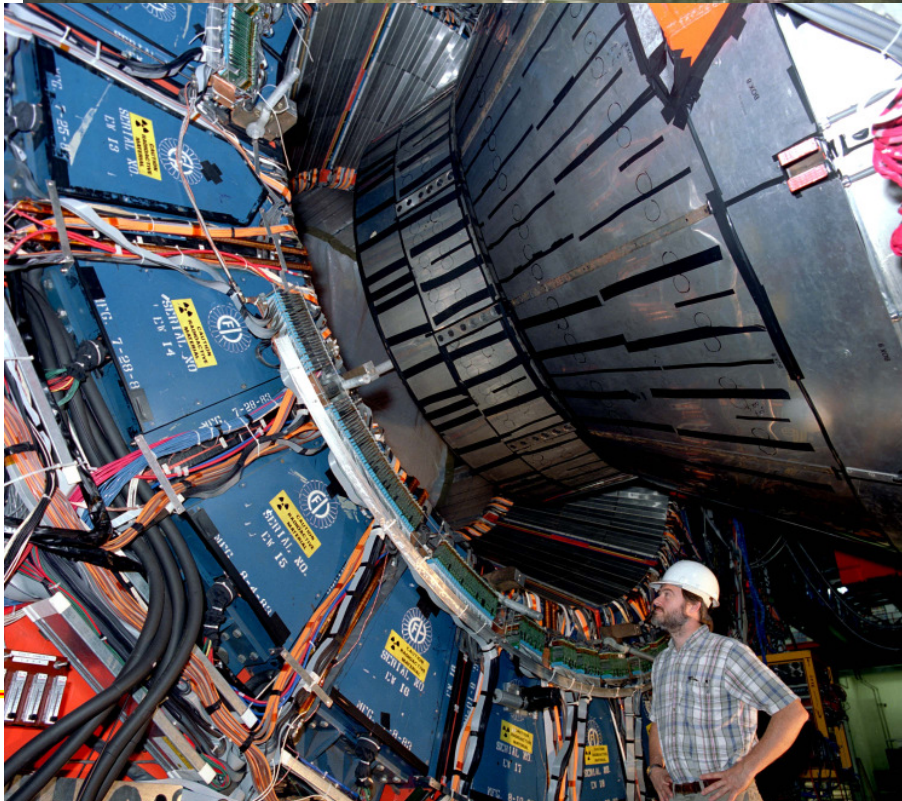
Tevatron and CDF

CDF: a multipurpose ('hermetic') detector



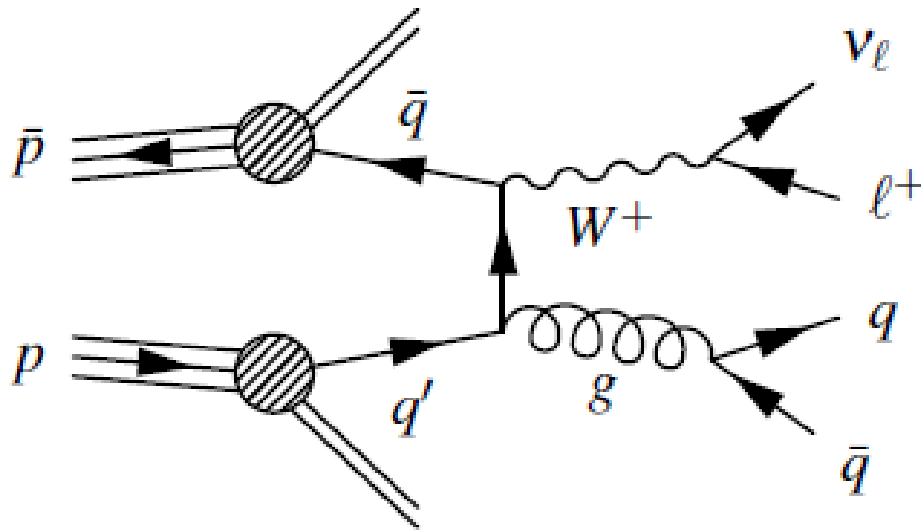
Tevatron and CDF

... a large, very sophisticated detector!



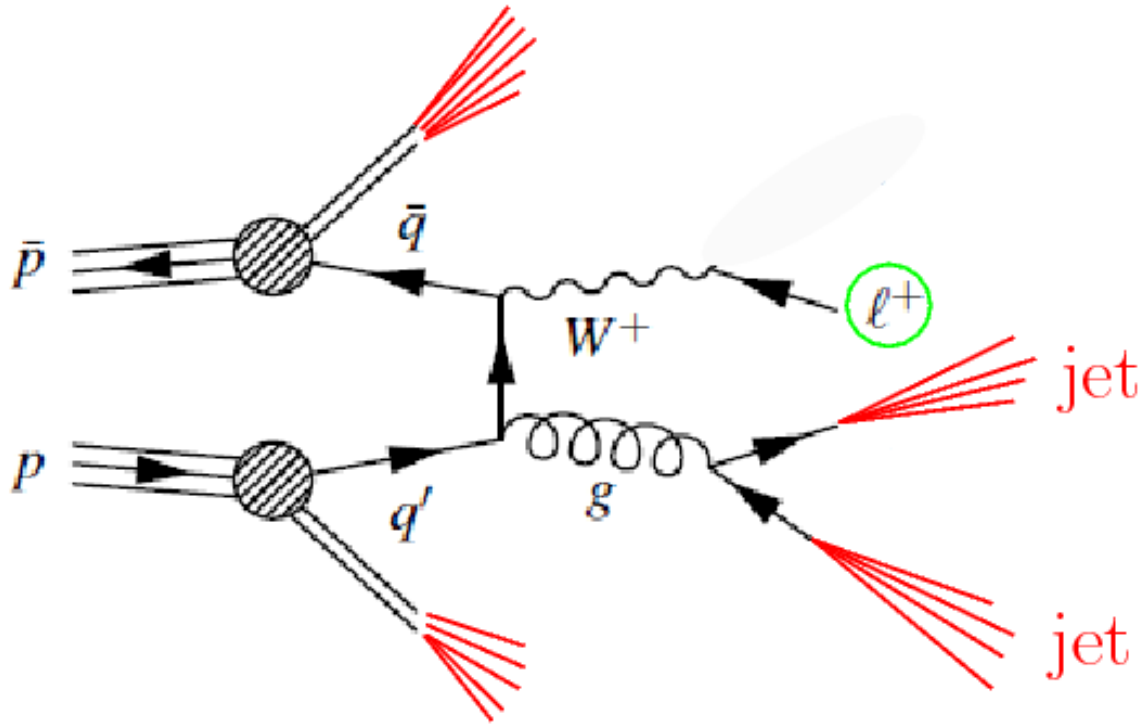
Jet-jet + W

$W + (q\bar{q})$ [+ 'remnants']



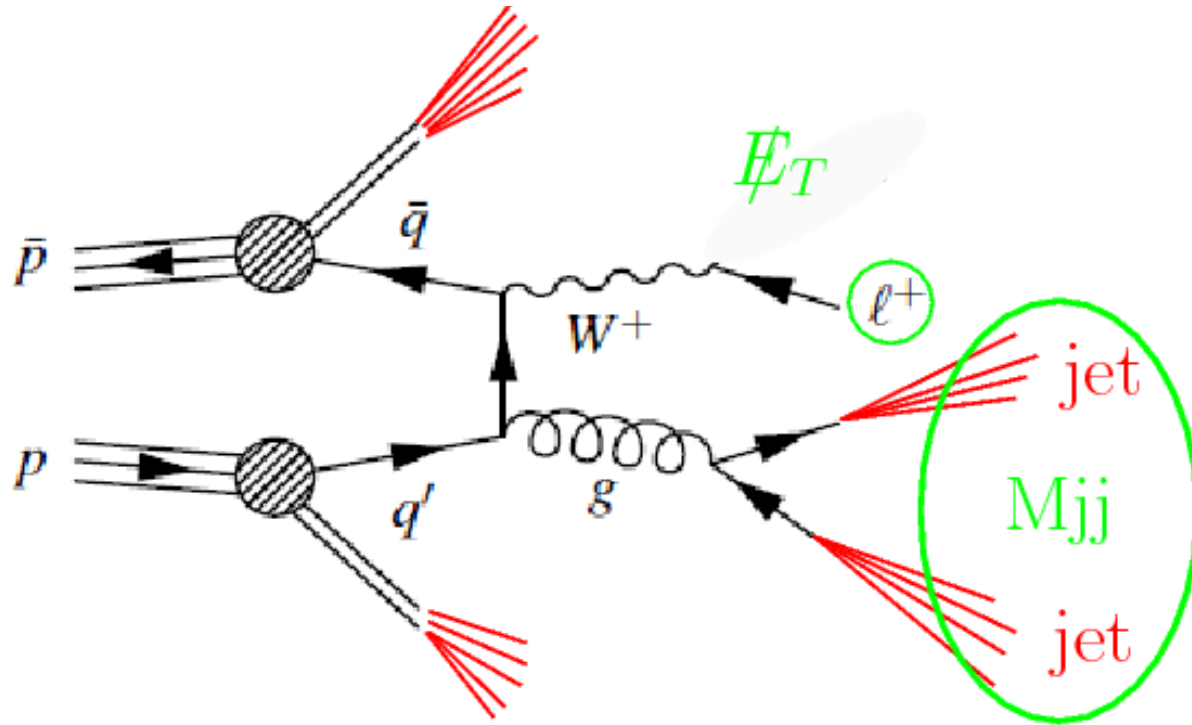
Jet-jet + W

$W + 2\text{jet}$ [+ molto altro]



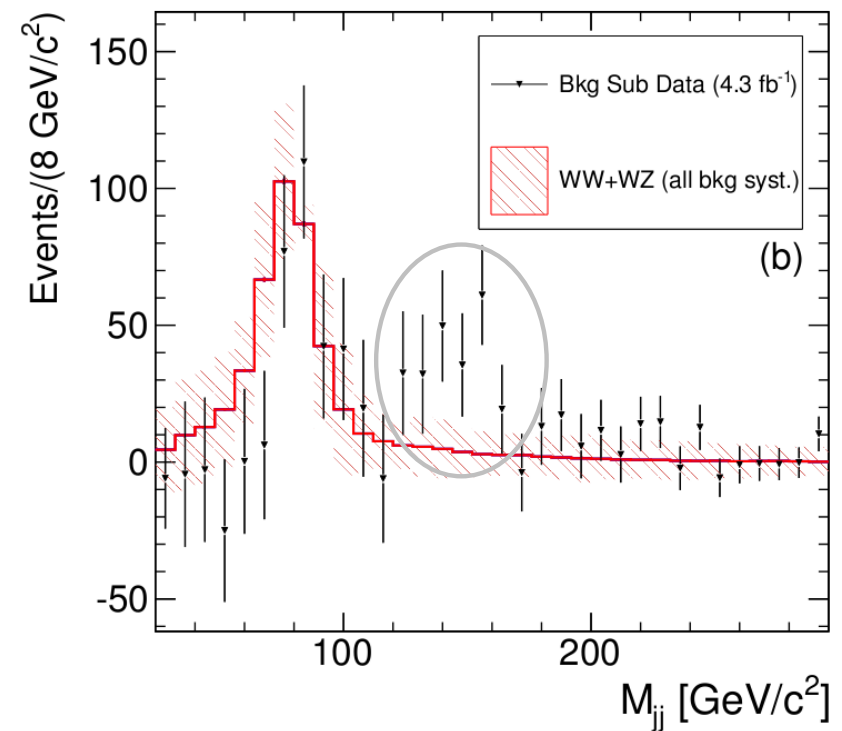
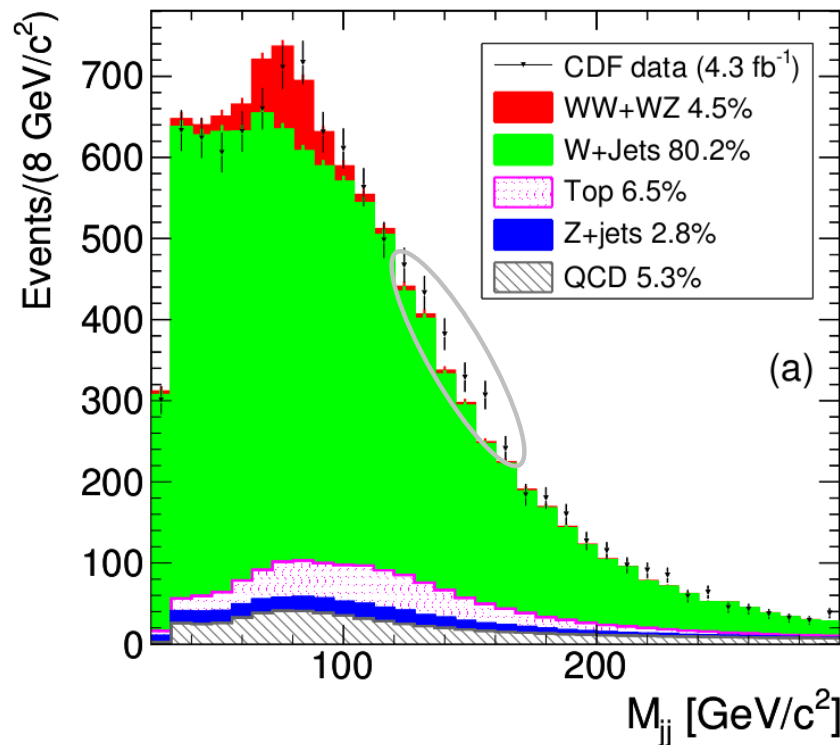
Jet-jet + W

$\Rightarrow M_{jj} + W + \dots$



The 'bump'!

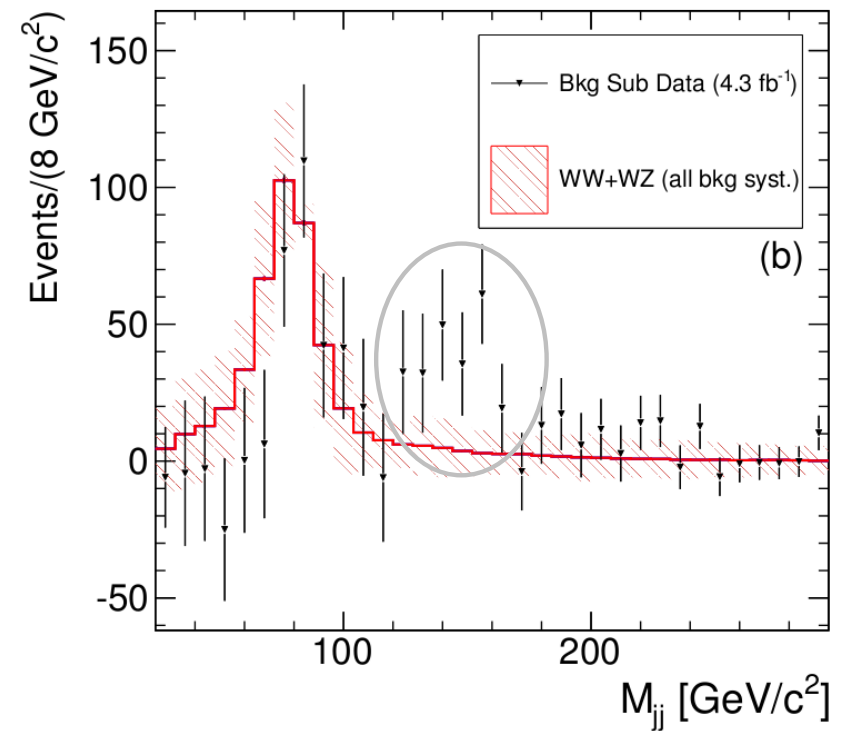
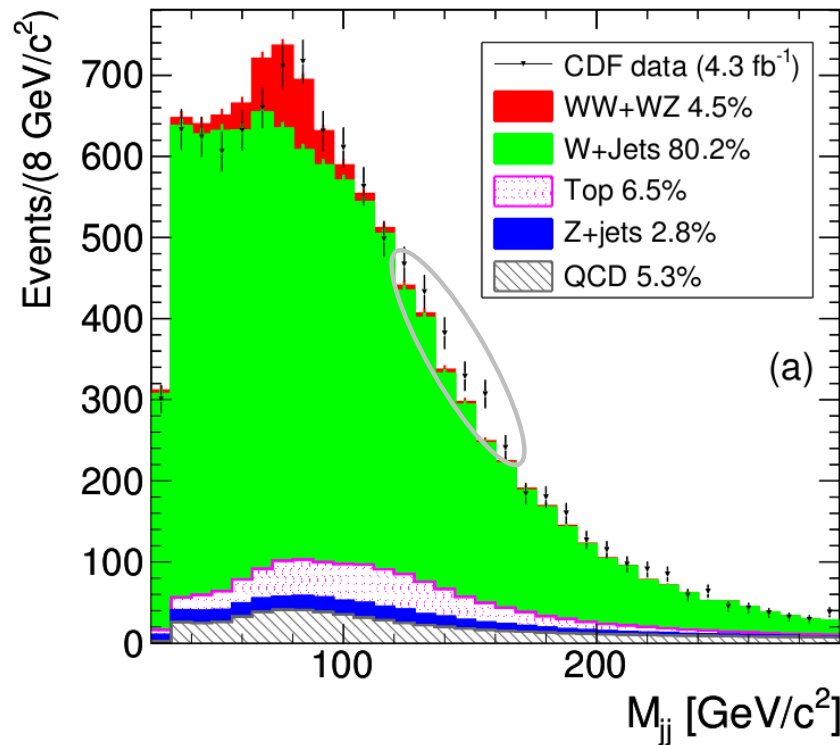
Invariant Mass Distribution of Jet Pairs Produced in Association with a W boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", (CDF, 4 aprile 2011)



“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations” [“3.2 σ ”]

The 'bump'!

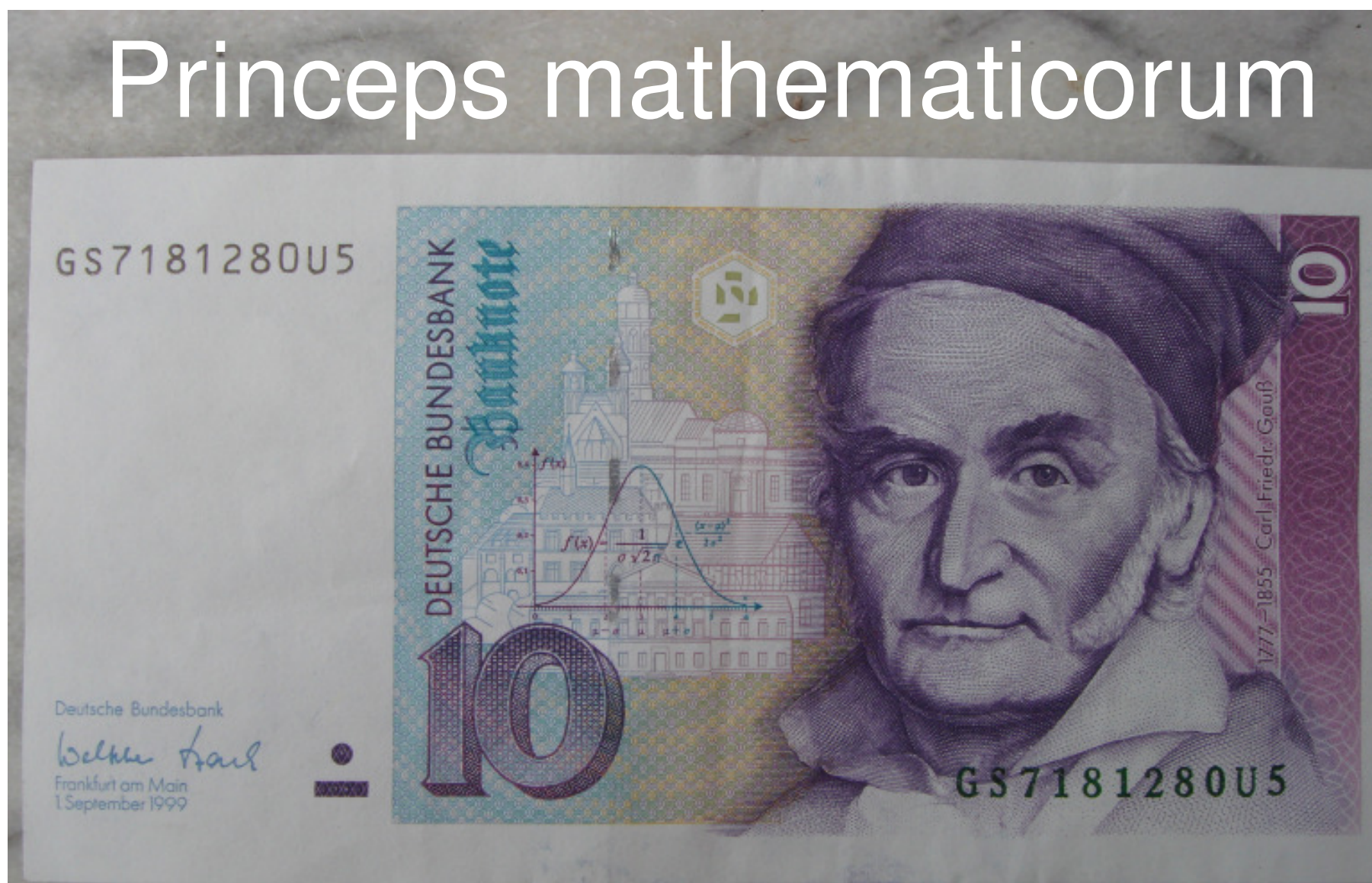
Invariant Mass Distribution of Jet Pairs Produced in Association with a W boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", (CDF, 4 aprile 2011)



Che significa?

Sigma e la gaussiana

Princeps mathematicorum



Sigma e la gaussiana



Sigma e probabilità [gaussiana!]

Se il numero aleatorio X è descritto da una gaussiana

$$P(-\sigma \leq X \leq +\sigma) = 68.3\%$$

$$P(-2\sigma \leq X \leq +2\sigma) = 95.4\%$$

$$P(-3\sigma \leq X \leq +3\sigma) = 99.73\%$$

$$1 - P(-3\sigma \leq X \leq +3\sigma) = 0.27\%$$

$$1 - P(-4\sigma \leq X \leq +4\sigma) = 6.3 \times 10^{-5}$$

$$\dots = \dots$$

$$1 - P(-10\sigma \leq X \leq +10\sigma) = 1.5 \times 10^{-23} \approx \frac{10}{N_A}$$

[1 preciso atomo in 1 g di ^{12}C !]

$$1 - P(-3.2\sigma \leq X \leq +3.2\sigma) = 1.4 \times 10^{-3}$$

$$P(X \geq +3.17\sigma) = 7.6 \times 10^{-4} \quad \checkmark$$

p-value, significatività e sigma

“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations” [3.2σ]

p-value, significatività e sigma

“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations” [“ 3.2σ ”]

Cominciare ad allacciare le cinture!

p-value, significatività e sigma

“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations” [“ 3.2σ ”]

Cominciare ad allacciare le cinture!

- Cos'è un **p-value**?
- In che senso ci fornisce una **'significance'**?

p-value, significatività e sigma

“we obtain a p-value of 7.6×10^{-4} , corresponding to a significance of 3.2 standard deviations” [“ 3.2σ ”]

Cominciare ad allacciare le cinture!

- Cos'è un **p-value**?
- In che senso ci fornisce una **'significance'**?

Insomma,

- 7.6×10^{-4} è una **probabilità**?
- **di cosa**?

April 2011: CDF sui media

Date: Fri, 8 Apr 2011 08:53:27

From: paolo.....

Subject: Nuova particella...

C'è da essere eccitati... noto che l'esperimento americano è coordinato da un italiano, e che dei 520 fisici che stanno lavorando al programma Cdf ben 70 sono italiani! Pensate che nazione saremmo se fossimo così bravi anche in politica!

<http://www.nibiru2012.it/forum/misteri/una-particella->

April 2011: CDF sui media

Date: Fri, 8 Apr 2011 08:10:38 +0100

From: alfredo.....

To: paolo....

Cc: giulio.dagostini@roma1.infn.it

Subject: Re: Nuova particella...

Purchè, italianamente, non sia una bufala per
prolungare la vita del Tevatron...

April 2011: CDF sui media

Date: Fri, 8 Apr 2011 08:10:38 +0100
From: alfredo.....
To: paolo....
Cc: giulio.dagostini@roma1.infn.it
Subject: Re: Nuova particella...

Purchè, italianamente, non sia una bufala per prolungare la vita del Tevatron...

Il quale dopo 5 minuti seguiva:

Date: Fri, 8 Apr 2011 08:15:50

....

sospetto aumentato dall'ultima frase dell'abstract
'We propose that the 3.2 excess at 150 GeV in the dijet mass spectrum of W+jets reported by CDF is the technipion T of low-scale technicolor. ...''

April 2011: CDF sui media

<http://www.nibiru2012.it/forum/misteri/una-particella-misteriosa>

Una Particella Misteriosa

07 Aprile 2011, 15:27:22 pm

USA: Una particella misteriosa perchè mai osservata finora è stata prodotta negli Stati Uniti, in un esperimento dell'acceleratore Tevatron del Fermilab di Chicago coordinato da un italiano, Giovanni Punzi, e potrebbe essere il primo passo nel terreno della nuova fisica.

April 2011: CDF sui media

<http://www.nibiru2012.it/forum/misteri/una-part>

Una Particella Misteriosa

07 Aprile 2011, 15:27:22 pm

...

Il segnale, descritto dal Fermilab in un articolo pubblicato online sul sito "arXiv", "non è fra quelli finora attesi", come il bosone di Higgs, ha commentato il presidente dell'Istituto Nazionale di Fisica Nucleare (Infn), Roberto Petronzio. Proprio per questo, ha aggiunto, "se si tratta di una particella completamente sconosciuta è molto probabile che si tratti di nuova fisica. Quello osservato dal Fermilab è sicuramente un effetto significativo".

April 2011: CDF sui media

<http://www.nibiru2012.it/forum/misteri/una-part>

Una Particella Misteriosa

07 Aprile 2011, 15:27:22 pm

...

Il segnale, descritto dal Fermilab in un articolo pubblicato online sul sito "arXiv", "non è fra quelli finora attesi", come il bosone di Higgs, ha commentato il presidente dell'Istituto Nazionale di Fisica Nucleare (Infn), Roberto Petronzio. Proprio per questo, ha aggiunto, "se si tratta di una particella completamente sconosciuta è molto probabile che si tratti di nuova fisica. Quello osservato dal Fermilab è sicuramente un effetto significativo".

'significativo': ??

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday, April 5:

“Physicists at the Fermi National Accelerator Laboratory are planning to announce Wednesday that they have found a suspicious bump in their data that could be evidence of a new elementary particle or even, some say, a new force of nature.

...

*The experimenters estimate that **there is a less than a quarter of 1 percent chance their bump is a statistical fluctuation**”*

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday, April 5:

“Physicists at the Fermi National Accelerator Laboratory are planning to announce Wednesday that they have found a suspicious bump in their data that could be evidence of a new elementary particle or even, some say, a new force of nature.

...

*The experimenters estimate that **there is a less than a quarter of 1 percent chance their bump is a statistical fluctuation**”*

$$P(\text{Statistical fluctuation}) \leq 0.25\%$$

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday, April 5:

“Physicists at the Fermi National Accelerator Laboratory are planning to announce Wednesday that they have found a suspicious bump in their data that could be evidence of a new elementary particle or even, some say, a new force of nature.

...

*The experimenters estimate that **there is a less than a quarter of 1 percent chance their bump is a statistical fluctuation**”*

$$P(\text{Statistical fluctuation}) \leq 0.25\%$$

$$P(\text{True Signal}) \geq 99.75\%!!$$

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday, April 5:

“Physicists at the Fermi National Accelerator Laboratory are planning to announce Wednesday that they have found a suspicious bump in their data that could be evidence of a new elementary particle or even, some say, a new force of nature.

...

*The experimenters estimate that **there is a less than a quarter of 1 percent chance their bump is a statistical fluctuation**”*

$$P(\text{Statistical fluctuation}) \leq 0.25\%$$

$$P(\text{True Signal}) \geq 99.75\%!!$$

Eureka!!

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday April 5:

“the most significant in physics in half a century”

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday April 5:

“the most significant in physics in half a century”

[Su come si passi poi da 7.6×10^{-4} a $< 2.5 \times 10^{-3}$
non chiedete a me (ma a questo livello è un dettaglio...)]

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday April 5:

“the most significant in physics in half a century”

Ben più importante è il fatto che un articolo dell'ArXiv di lunedì 4 sia ripreso dal NYT il giorno dopo!

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday April 5:

“the most significant in physics in half a century”

Ben più importante è il fatto che un articolo dell'ArXiv di lunedì 4 sia ripreso dal NYT il giorno dopo!

Chi credeva che si trattasse – al 99.75%! – di una scoperta?

- il giornalista che ha riportato la notizia?
- i responsabili di CDF e del Fermilab che lo hanno contattato?

Aprile 2011, the 'bump' explodes

The New York Times, Tuesday April 5:

“the most significant in physics in half a century”

Ben più importante è il fatto che un articolo dell'ArXiv di lunedì 4 sia ripreso dal NYT il giorno dopo!

Chi credeva che si trattasse – al 99.75%! – di una scoperta?

- il giornalista che ha riportato la notizia?
- i responsabili di CDF e del Fermilab che lo hanno contattato?

Dalla mia esperienza i giornalisti possono commettere imprecisioni, ma non si inventano le notizie (almeno quelle scientifiche...).

Aprile 2011, the 'bump' explodes

Fermilab Today, April 7:

“Wednesday afternoon, the CDF collaboration announced that it has evidence of a peak in a specific sample of its data. The peak is an excess of particle collision events that produce a W boson accompanied by two hadronic jets. This peak showed up in a mass region where we did not expect one.

...

Aprile 2011, the 'bump' explodes

Fermilab Today, April 7:

“Wednesday afternoon, the CDF collaboration announced that it has evidence of a peak in a specific sample of its data. The peak is an excess of particle collision events that produce a W boson accompanied by two hadronic jets. This peak showed up in a mass region where we did not expect one.

...

*The significance of this excess was determined to be 3.2 sigma, after accounting for the effect of systematic uncertainties. This means that **there is less than a 1 in 1375 chance that the effect is mimicked by a statistical fluctuation.**”*

Aprile 2011, the 'bump' explodes

Fermilab Today, April 7:

“Wednesday afternoon, the CDF collaboration announced that it has evidence of a peak in a specific sample of its data. The peak is an excess of particle collision events that produce a W boson accompanied by two hadronic jets. This peak showed up in a mass region where we did not expect one.

...

*The significance of this excess was determined to be 3.2 sigma, after accounting for the effect of systematic uncertainties. This means that **there is less than a 1 in 1375 chance that the effect is mimicked by a statistical fluctuation.**”*

$$1/1375 = 7.3 \times 10^{-4} \Rightarrow P(\text{No stat. fluct.}) = 99.93\% \quad !$$

Aprile 2011, the 'bump' explodes

Discovery News, April 7:

This is a big week for particle physicists, and even they will be having many sleepless nights over the coming months trying to grasp what it all means.

That's what happens when physicists come forward, with observational evidence, of what they believe represents something we've never seen before. Even bigger than that: something we never even expected to see.

...

Aprile 2011, the 'bump' explodes

Discovery News, April 7:

This is a big week for particle physicists, and even they will be having many sleepless nights over the coming months trying to grasp what it all means.

That's what happens when physicists come forward, with observational evidence, of what they believe represents something we've never seen before. Even bigger than that: something we never even expected to see.

...

*It is what is known as a "three-sigma event," and this refers to the statistical certainty of a given result. In this case, **this result has a 99.7 percent chance of being correct (and a 0.3 percent chance of being wrong).**"*

Aprile 2011, the 'bump' explodes

Discovery News, April 7:

This is a big week for particle physicists, and even they will be having many sleepless nights over the coming months trying to grasp what it all means.

That's what happens when physicists come forward, with observational evidence, of what they believe represents something we've never seen before. Even bigger than that: something we never even expected to see.

...

*It is what is known as a "three-sigma event," and this refers to the statistical certainty of a given result. In this case, **this result has a 99.7 percent chance of being correct (and a 0.3 percent chance of being wrong).**"*

Insomma, stiamo capendo bene, a parte il fatto che si passa dal 99.9% al 99.7%...

Aprile 2011, the 'bump' explodes

Jon Butterworth's blob on the Guardian, April 9:

"The last and greatest breakthrough from a fantastic machine, or a false alarm on the frontiers of physics?"

...

*If the histograms and data are exactly right, **the paper quotes a one-in-ten-thousand (0.0001) chance that this bump is a fluke.**"*

Aprile 2011, the 'bump' explodes

Jon Butterworth's blob on the Guardian, April 9:

"The last and greatest breakthrough from a fantastic machine, or a false alarm on the frontiers of physics?"

...

If the histograms and data are exactly right, the paper quotes a one-in-ten-thousand (0.0001) chance that this bump is a fluke." $\Rightarrow P(\text{Not Fluke}) = P(\text{"Genuine"}) = 99.99\%$

Aprile 2011, the 'bump' explodes

Jon Butterworth's blob on the Guardian, April 9:

"The last and greatest breakthrough from a fantastic machine, or a false alarm on the frontiers of physics?"

...

If the histograms and data are exactly right, the paper quotes a one-in-ten-thousand (0.0001) chance that this bump is a fluke." $\Rightarrow P(\text{Not Fluke}) = P(\text{"Genuine"}) = 99.99\%$

Ma, alla fine del pezzo:

1. "My money is on the false alarm at the moment,..."
2. "...but I would be very happy to lose it."
3. "And I reserve the right to change my mind rapidly as more data come in!"

Aprile 2011, the 'bump' explodes

Jon Butterworth's blob on the Guardian, April 9:

"The last and greatest breakthrough from a fantastic machine, or a false alarm on the frontiers of physics?"

...

If the histograms and data are exactly right, the paper quotes a one-in-ten-thousand (0.0001) chance that this bump is a fluke." $\Rightarrow P(\text{Not Fluke}) = P(\text{"Genuine"}) = 99.99\%$

Ma, alla fine del pezzo:

1. "My money is on the false alarm at the moment,..."
2. "...but I would be very happy to lose it."
3. "And I reserve the right to change my mind rapidly as more data come in!"

Assolutamente sensato! (A parte l'incongruenza iniziale)

“Significativo, ma non ci credo”...

Jon Butterworth's blob on the Guardian, April 9:

1. “My money is on the false alarm at the moment, . . .”

“Significativo, ma non ci credo”...

Jon Butterworth's blob on the Guardian, April 9:

1. “My money is on the false alarm at the moment, . . .”

“Non ci credo!”

“Significativo, ma non ci credo”...

Jon Butterworth's blob on the Guardian, April 9:

1. “My money is on the false alarm at the moment, . . .”

“Non ci credo!”

2. “. . . but I would be very happy to lose it.”

“Cosa credo” \neq “Cosa mi piacerebbe”

“Significativo, ma non ci credo”...

Jon Butterworth's blob on the Guardian, April 9:

1. “My money is on the false alarm at the moment, . . .”

“Non ci credo!”

2. “. . . but I would be very happy to lose it.”

“Cosa credo” \neq “Cosa mi piacerebbe”

3. “And I reserve the right to change my mind rapidly as more data come in!”

“Imparare dall'esperienza!”

⇒ Un fisico non è mai dogmatico

“Significativo, ma non ci credo”...

Jon Butterworth's blob on the Guardian, April 9:

1. “My money is on the false alarm at the moment, . . .”

“Non ci credo!”

2. “. . . but I would be very happy to lose it.”

“Cosa credo” \neq “Cosa mi piacerebbe”

3. “And I reserve the right to change my mind rapidly as more data come in!”

“Imparare dall'esperienza!”

⇒ Un fisico non è mai dogmatico

Ma come si devono cambiare razionalmente le proprie convinzioni scientifiche alla luce dei nuovi dati sperimentali?

“Significativo, ma non ci credo”...

E Jon Butterworth non era il solo a non crederci.
Anzi, **non ci credeva la stragrande maggioranza dei fisici.**

“Significativo, ma non ci credo”...

E Jon Butterworth non era il solo a non crederci.

Anzi, **non ci credeva la stragrande maggioranza dei fisici.**

⇒ Più o meno come nel caso, ben più noto,
dei neutrini di Opera... (**10σ** !)

“Significativo, ma non ci credo”...

E Jon Butterworth non era il solo a non crederci.

Anzi, **non ci credeva la stragrande maggioranza dei fisici.**

⇒ Più o meno come nel caso, ben più noto,
dei neutrini di Opera... (**10σ !**)

Ma allora, cosa significano queste
“significance”?

“Significativo, ma non ci credo”...

E Jon Butterworth non era il solo a non crederci.
Anzi, **non ci credeva la stragrande maggioranza dei fisici.**
⇒ Più o meno come nel caso, ben più noto,
dei neutrini di Opera... (**10σ** !)

Ma allora, cosa significano queste
“significance”?

“Paradosso di de Rujula”:

“If you disbelieve every result presented as having a 3 sigma – or “equivalently” a 99.7% chance – of being correct... You will turn out to be right 99.7% of the times.”

(Alvaro de Rujula, private communication)

Il cimitero della Fisica

THE CEMETERY OF PHYSICS
IS FULL OF WONDERFUL
EFFECTS...



...THAT VERY OFTEN LEAD
TO THEORETICAL, EXPERIMENTAL PROGRESS

Alvaro de Rujula

Testare una ipotesi

- Idea di base:
 - partiamo da un modello 'convenzionale'
[Modello Standard, 'teoria riconosciuta', etc:]
→ " H_0 " ("null hypothesis")

Testare una ipotesi

- Idea di base:
 - partiamo da un modello 'convenzionale'
[Modello Standard, 'teoria riconosciuta', etc:]
→ " H_0 " ("null hypothesis")
⇒ cercare violazioni di H_0

Testare una ipotesi

- Idea di base:
 - partiamo da un modello ‘convenzionale’
[Modello Standard, ‘teoria riconosciuta’, etc:]
→ “ H_0 ” (“null hypothesis”)
⇒ cercare violazioni di H_0
- Idealmente
→ ‘falsificare’ H_0

Testare una ipotesi

- Idea di base:
 - partiamo da un modello ‘convenzionale’
[Modello Standard, ‘teoria riconosciuta’, etc:]
→ “ H_0 ” (“null hypothesis”)
⇒ cercare violazioni di H_0
- Idealmente
→ ‘falsificare’ H_0
- In pratica:
 - ha senso?
 - come si fa?

Testare una ipotesi

- Idea di base:
 - partiamo da un modello 'convenzionale'
[Modello Standard, 'teoria riconosciuta', etc:]
→ " H_0 " ("null hypothesis")
⇒ cercare violazioni di H_0
- Idealmente
→ 'falsificare' H_0
- In pratica:
 - ha senso?
 - come si fa?

Rivediamo la prassi e cosa c'è dietro ⇒

Falsificazionismo

Usualmente associato a Popper
e considerato da molti ancor oggi la *chiave del
progresso scientifico*.

Falsificazionismo

Usualmente associato a Popper
e considerato da molti ancor oggi la *chiave del
progresso scientifico*.

se $C_i \not\rightarrow E_0$, **allora** $E_0^{(\text{mis})} \not\rightarrow C_i$

⇒ Cause che non possono produrre gli effetti osservati
sono scartate ('falsificate').

Falsificazionismo

Usualmente associato a Popper
e considerato da molti ancor oggi la *chiave del
progresso scientifico*.

se $C_i \not\rightarrow E_0$, **allora** $E_0^{(\text{mis})} \not\rightarrow C_i$

⇒ Cause che non possono produrre gli effetti osservati
sono scartate ('falsificate').

Sembra OK – 'ovvio'! – ma naïve per diversi aspetti.

Prova per assurdo ... 'estesa'

Regola della 'falsificazione': a cosa si ispira?

Prova per assurdo ... 'estesa'

Regola della 'falsificazione': a cosa si ispira?

Prova per assurdo della logica classica deduttiva:

- assumiamo un'ipotesi vera;
- deriviamo 'tutte' le conseguenze logiche;
- se (almeno) una delle conseguenze è manifestamente falsa l'ipotesi viene rigettata.

Prova per assurdo ... 'estesa'

Regola della 'falsificazione': a cosa si ispira?

Prova per assurdo della logica classica deduttiva:

- assumiamo un'ipotesi vera;
- deriviamo 'tutte' le conseguenze logiche;
- se (almeno) una delle conseguenze è manifestamente falsa l'ipotesi viene rigettata.

Falsificazionismo popperiano:

estende ragionamento alle scienze sperimentali

Prova per assurdo ... 'estesa'

Regola della 'falsificazione': a cosa si ispira?

Prova per assurdo della logica classica deduttiva:

- assumiamo un'ipotesi vera;
- deriviamo 'tutte' le conseguenze logiche;
- se (almeno) una delle conseguenze è manifestamente falsa l'ipotesi viene rigettata.

Falsificazionismo popperiano:

estende ragionamento alle scienze sperimentali

operazione legittima?

Falsificationismo? OK, ma...

Falsificationismo? OK, ma...

- Cosa fare di tutte le ipotesi che non sono falsificate? (Limbo? Come facciamo a progredire?)

Falsificationismo? OK, ma...

- Cosa fare di tutte le ipotesi che non sono falsificate? (Limbo? Come facciamo a progredire?)
- Cosa fare se niente di quello che può essere osservato è incompatibile con le ipotesi in gioco?

Falsificationismo? OK, ma...

- Cosa fare di tutte le ipotesi che non sono falsificate? (Limbo? Come facciamo a progredire?)
- Cosa fare se niente di quello che può essere osservato è incompatibile con le ipotesi in gioco?

Per es. se H_i sono gaussiane $f(x | \mu_i, \sigma_i)$

⇒ Data ogni coppia di parametri $\{\mu_i, \sigma_i\}$, tutti i valori di x fra $-\infty$ e $+\infty$ sono possibili $\forall H_i$.

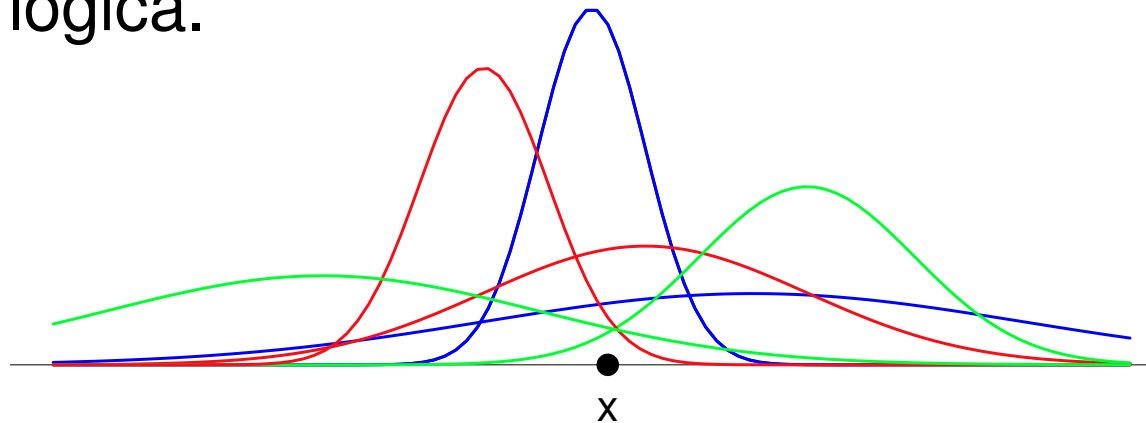
Falsificazionismo? OK, ma...

- Cosa fare di tutte le ipotesi che non sono falsificate? (Limbo? Come facciamo a progredire?)
- Cosa fare se niente di quello che può essere osservato è incompatibile con le ipotesi in gioco?

Per es. se H_i sono gaussiane $f(x | \mu_i, \sigma_i)$

⇒ Data ogni coppia di parametri $\{\mu_i, \sigma_i\}$, tutti i valori di x fra $-\infty$ e $+\infty$ sono possibili $\forall H_i$.

⇒ Avendo osservato un qualsiasi valore x , nessuna ipotesi H_i può essere falsificata, a stretto rigor di logica.



Falsificazionismo in azione...

Ovviamente, questo non significa che il falsificazionismo non funzioni mai

Falsificazionismo in azione...

Ovviamente, questo non significa che il falsificazionismo non funzioni mai

purché non siano coinvolti processi aleatori, o intrinseci ai processi fisici o legati agli 'errori' di misura.

Falsificazionismo in azione...

Ovviamente, questo non significa che il falsificazionismo non funzioni mai

purché non siano coinvolti processi aleatori, o intrinseci ai processi fisici o legati agli 'errori' di misura.

⇒ **praticamente mai nelle scienze sperimentali!**

Falsificazionismo in azione...

Ovviamente, questo non significa che il falsificazionismo non funzioni mai

purché non siano coinvolti processi aleatori, o intrinseci ai processi fisici o legati agli 'errori' di misura.

Sicuramente funziona contro se stesso:

- la scienza procede in pratica in ben altro modo:

Lo sviluppo naturale della scienza mostra che le ricerche sono effettuate nella direzione che sembra più credibile (e fruttuosa) ad un certo momento, un comportamento *"179 degrees or so out of phase from Popper's idea that we make progress by falsifying theories"*

(Wilczek,

<http://arxiv.org/abs/physics/0403115>)

Falsificazionismo in azione...

Ovviamente, questo non significa che il falsificazionismo non funzioni mai

purché non siano coinvolti processi aleatori, o intrinseci ai processi fisici o legati agli 'errori' di misura.

Sicuramente funziona contro se stesso:

⇒ a rigor di logica il falsificazionismo è da ritenersi... falsificato

Falsificationismo e statistica

... nel frattempo gli statistici hanno inventato i “test di ipotesi”

Falsificazionismo e statistica

... nel frattempo gli statistici hanno inventato i “test di ipotesi” nei quali l’impossibile è sostituito dall’improbabile!

Falsificazionismo e statistica

... nel frattempo gli statistici hanno inventato i “test di ipotesi” nei quali l'impossibile è sostituito dall'improbabile!

MA dall'impossibile all'improbabile
non è questione di quantità, bensì di qualità.

Falsificazionismo e statistica

... nel frattempo gli statistici hanno inventato i “test di ipotesi” nei quali l’impossibile è sostituito dall’improbabile!

MA dall’impossibile all’improbabile non è questione di quantità, bensì di qualità.

Questo meccanismo, *logicamente bacato*, è particolarmente pericoloso, perché fortemente radicato in molti, a tutti i livelli, a causa di prassi e insegnamento. (*)

⇒ Alla base di tutti gli annunci di false scoperte delle ultime decine di anni.

[Personalmente sono particolarmente preoccupato di tante affermazioni, basate su questi ragionamenti, a proposito di salute, stato del Pianeta, etc., molto più importanti dei nostri ‘giochini’.]

(*) “O è così o ti boccio!”...

Ricapitolando

A) **se** $C_i \not\rightarrow E$, e **osserviamo** E
 $\Rightarrow C_i$ è impossibile ('falsa')

Ricapitolando

A) **se** $C_i \not\rightarrow E$, e **osserviamo** E
 $\Rightarrow C_i$ è impossibile ('falsa')

B) **se** $C_i \xrightarrow{\text{bassa probabilità}} E$, e **osserviamo** E

$\Rightarrow C_i$ ha bassa probabilità di essere vera
ovvero “molto probabilmente falsa”

Ricapitolando

A) **se** $C_i \not\rightarrow E$, e osserviamo E
 $\Rightarrow C_i$ è impossibile ('falsa')

OK

B) **se** $C_i \xrightarrow{\text{bassa probabilità}} E$, e osserviamo E

$\Rightarrow C_i$ ha bassa probabilità di essere vera
ovvero "molto probabilmente falsa"

Ricapitolando

A) **se** $C_i \not\rightarrow E$, e osserviamo E
 $\Rightarrow C_i$ è impossibile ('falsa')

OK

B) ~~**se** $C_i \xrightarrow{\text{bassa probabilità}} E$, e osserviamo E
 $\Rightarrow C_i$ ha bassa probabilità di essere vera
ovvero "molto probabilmente falsa"~~

NO

Ma alla base dei test statistici!

Esempio

Italiano scelto a caso si sottopone al test dell'AIDS.
Test non perfetto, come succede in pratica.

Modello semplificato:

$$P(\text{Pos} \mid \text{HIV}) = 100\%$$

$$P(\text{Pos} \mid \overline{\text{HIV}}) = 0.2\%$$

$$P(\text{Neg} \mid \overline{\text{HIV}}) = 99.8\%$$

$H_1 = \text{'HIV'}$ (Infetto)

$E_1 = \text{Positivo}$

$H_2 = \overline{\text{'HIV'}}$ (Non infetto)

$E_2 = \text{Negativo}$

Esempio

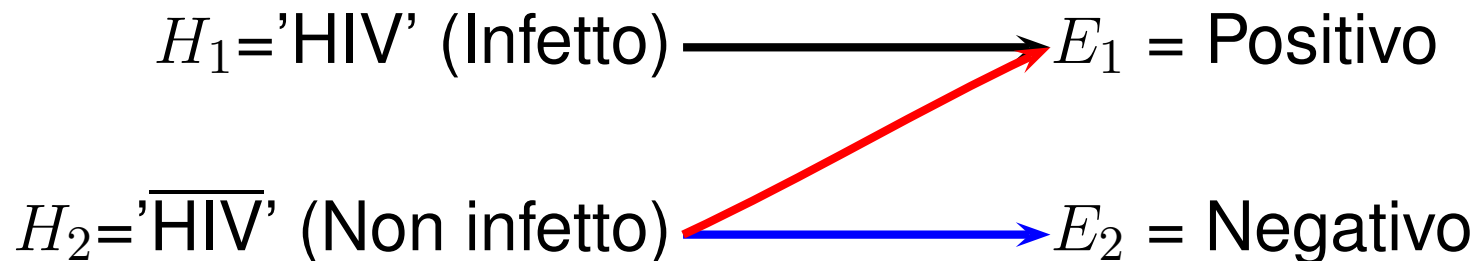
Italiano scelto a caso si sottopone al test dell'AIDS.
Test non perfetto, come succede in pratica.

Modello semplificato:

$$P(\text{Pos} \mid \text{HIV}) = 100\%$$

$$P(\text{Pos} \mid \overline{\text{HIV}}) = 0.2\%$$

$$P(\text{Neg} \mid \overline{\text{HIV}}) = 99.8\%$$



Esempio

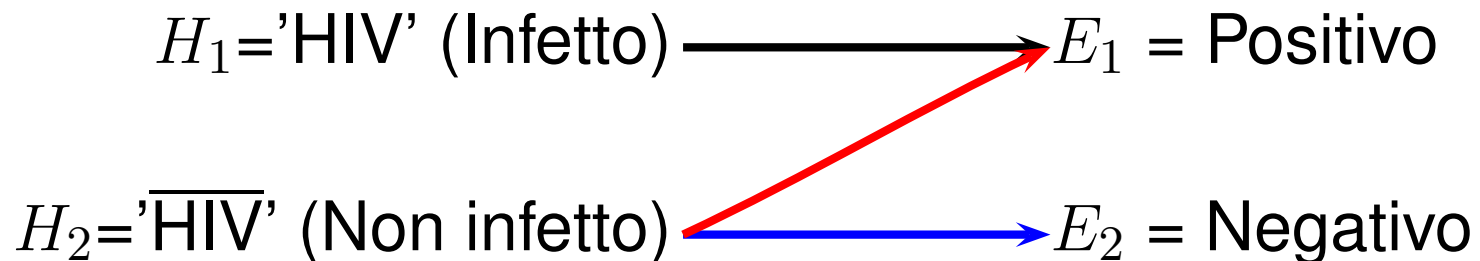
Italiano scelto a caso si sottopone al test dell'AIDS.
Test non perfetto, come succede in pratica.

Modello semplificato:

$$P(\text{Pos} \mid \text{HIV}) = 100\%$$

$$P(\text{Pos} \mid \overline{\text{HIV}}) = 0.2\%$$

$$P(\text{Neg} \mid \overline{\text{HIV}}) = 99.8\%$$



Risultato: \Rightarrow Positivo

Esempio

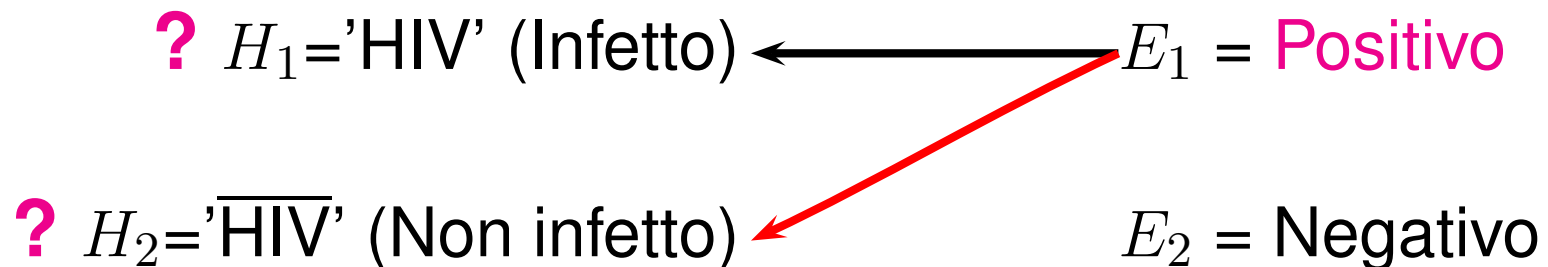
Italiano scelto a caso si sottopone al test dell'AIDS.
Test non perfetto, come succede in pratica.

Modello semplificato:

$$P(\text{Pos} \mid \text{HIV}) = 100\%$$

$$P(\text{Pos} \mid \overline{\text{HIV}}) = 0.2\%$$

$$P(\text{Neg} \mid \overline{\text{HIV}}) = 99.8\%$$



Risultato: \Rightarrow Positivo

Infetto o no?

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- "Praticamente impossibile che la persona sia infetta, in quanto era 'praticamente impossibile' che chi non ha il virus risulti positivo"?

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- "Praticamente impossibile che la persona sia infetta, in quanto era 'praticamente impossibile' che chi non ha il virus risulti positivo"
- "C'è solo lo 0.2% di probabilità che la persona non abbia l'HIV"?

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- "Praticamente impossibile che la persona sia infetta, in quanto era 'praticamente impossibile' che chi non ha il virus risulti positivo"
- "C'è solo lo 0.2% di probabilità che la persona non abbia l'HIV"
- "Siamo confidenti al 99.8% che la persona sia infetta"?

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- "Praticamente impossibile che la persona sia infetta, in quanto era 'praticamente impossibile' che chi non ha il virus risulti positivo"
- "C'è solo lo 0.2% di probabilità che la persona non abbia l'HIV"
- "Siamo confidenti al 99.8% che la persona sia infetta"
- "l'ipotesi $H_1 = \text{Non_HIV}$ è scartata al 99.8% C.L."

?

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- ~~”Praticamente impossibile che la persona sia infetta, in quanto era ‘praticamente impossibile’ che chi non ha il virus risulti positivo”~~
- ~~“C’è solo lo 0.2% di probabilità che la persona non abbia l’HIV”~~
- ~~“Siamo confidenti al 99.8% che la persona sia infetta”~~
- ~~“l’ipotesi $H_1 = \text{Non_HIV}$ è scartata al 99.8% C.L.”~~

?

NO

Invece, $P(\text{HIV} | \text{Pos, italiano a caso}) \approx 45\%$

(Lo lasciamo come esercizio – per la cui soluzione manca un dato!)

Cosa ne concludiamo?

Data $P(\text{Pos} | \overline{\text{HIV}}) = 0.2\%$ e avendo osservato 'Positivo', possiamo affermare:

- ~~● "Praticamente impossibile che la persona sia infetta, in quanto era 'praticamente impossibile' che chi non ha il virus risulti positivo"~~
- ~~● "C'è solo lo 0.2% di probabilità che la persona non abbia l'HIV"~~
- ~~● "Siamo confidenti al 99.8% che la persona sia infetta"~~
- ~~● "l'ipotesi $H_1 = \text{Non_HIV}$ è scartata al 99.8% C.L."~~

?

NO

Invece, $P(\text{HIV} | \text{Pos, italiano a caso}) \approx 45\%$

⇒ **Errore serio!** (non banalmente 99.8% invece di 98.3%)

$$P(A | B) \leftrightarrow P(B | A)$$

Attenzione al 'ribaltamento' arbitrario delle probabilità condizionate:

In generale $P(A | B) \neq P(B | A)$

$$P(A | B) \leftrightarrow P(B | A)$$

Attenzione al 'ribaltamento' arbitrario delle probabilità condizionate:

In generale $P(A | B) \neq P(B | A)$

- $P(\text{Positivo} | \overline{HIV}) \neq P(\overline{HIV} | \text{Positivo})$

$$P(A | B) \leftrightarrow P(B | A)$$

Attenzione al 'ribaltamento' arbitrario delle probabilità condizionate:

In generale $P(A | B) \neq P(B | A)$

- $P(\text{Positivo} | \overline{HIV}) \neq P(\overline{HIV} | \text{Positivo})$
- $P(\text{Vincere} | \text{Giocare}) \neq P(\text{Giocare} | \text{Vincere})$ [Lotto]

$$P(A | B) \leftrightarrow P(B | A)$$

Attenzione al 'ribaltamento' arbitrario delle probabilità condizionate:

In generale $P(A | B) \neq P(B | A)$

- $P(\text{Positivo} | \overline{HIV}) \neq P(\overline{HIV} | \text{Positivo})$
- $P(\text{Vincere} | \text{Giocare}) \neq P(\text{Giocare} | \text{Vincere})$ [Lotto]
- $P(\text{Pregnant} | \text{Woman}) \neq P(\text{Woman} | \text{Pregnant})$

$$P(A | B) \leftrightarrow P(B | A)$$

Attenzione al 'ribaltamento' arbitrario delle probabilità condizionate:

In generale $P(A | B) \neq P(B | A)$

- $P(\text{Positivo} | \overline{HIV}) \neq P(\overline{HIV} | \text{Positivo})$
- $P(\text{Vincere} | \text{Giocare}) \neq P(\text{Giocare} | \text{Vincere})$ [Lotto]
- $P(\text{Pregnant} | \text{Woman}) \neq P(\text{Woman} | \text{Pregnant})$

In particolare

- Una causa può produrre un certo effetto con bassissima probabilità, pur essendo tale causa la più probabile di quell'effetto, e talvolta la sola!

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Ad esempio, un generatore di numeri casuali gaussiano ($\mu = 3, \sigma = 1$) ci dà il numero $X = 3.1416$.

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Ad esempio, un generatore di numeri casuali gaussiano ($\mu = 3, \sigma = 1$) ci dà il numero $X = 3.1416$.

→ Qual'era la probabilità di dare proprio quel numero?:

$$\begin{aligned} P(X = 3.1416 | H_0) &= \int_{3.14155}^{3.14165} f_G(x | \mu, \sigma) dx \\ &\approx f_G(3.1416 | \mu, \sigma) \times \Delta x \\ &\approx f_G(3.1416 | \mu, \sigma) \times 0.0001 \\ &\approx 39 \times 10^{-6} \end{aligned}$$

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Ad esempio, un generatore di numeri casuali gaussiano ($\mu = 3, \sigma = 1$) ci dà il numero $X = 3.1416$.

→ Qual'è la probabilità che X venga da quel generatore?

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Ad esempio, un generatore di numeri casuali gaussiano ($\mu = 3, \sigma = 1$) ci dà il numero $X = 3.1416$.

- Qual'è la probabilità che X venga da quel generatore?
- Certamente **NON** $\approx 39 \times 10^{-6}$;

Eventi 'poco probabili'

Valori tipici nella prassi statistica per rifiutare un'ipotesi sono 5%, 1%, ... (vedi 'Test AIDS')

MA la stragante maggioranza degli eventi di interesse sono di bassissima probabilità

Ad esempio, un generatore di numeri casuali gaussiano ($\mu = 3, \sigma = 1$) ci dà il numero $X = 3.1416$.

→ Qual'è la probabilità che X venga da quel generatore?

- Certamente **NON** $\approx 39 \times 10^{-6}$;
- Anzi, è **esattamente 1**, in quanto H_0 è **unica causa che può produrre tale effetto**:

$$P(X = 3.1416 | H_0) \approx 39 \times 10^{-6}$$

$$P(H_0 | X = 3.1416) = 1.$$

Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

ma invece di ravvedersi, buttare tutto e cominciare finalmente a **leggere Laplace** (sì il 'nostro' Laplace!) ne inventa un'altra:

→ non bisogna basarsi sulla probabilità del numero ottenuto, ma sulla probabilità che uscisse o quel numero o un numero che si discosta più di esso dal valore atteso (la storia è un po' più lunga...):

$$P(X \geq 3.1416) = \int_{3.14155}^{+\infty} f_G(x | \mu, \sigma) dx \approx 44\%$$

Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

ma invece di ravvedersi, buttare tutto e cominciare finalmente a **leggere Laplace** (sì il 'nostro' Laplace!) ne inventa un'altra:

→ non bisogna basarsi sulla probabilità del numero ottenuto, ma sulla probabilità che uscisse o quel numero o un numero che si discosta più di esso dal valore atteso (la storia è un po' più lunga...):

$$P(X \geq 3.1416) [= P(X \geq x_{obs})] \Rightarrow \text{'p-value'}$$

Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

- ⇒ Come per magia **il risultato 'diventa' molto probabile!**
Di cosa ci preoccupavamo?
- ⇒ tutti contenti ...

Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

- ⇒ Come per magia **il risultato 'diventa' molto probabile!**
Di cosa ci preoccupavamo?
- ⇒ tutti contenti ... e co...

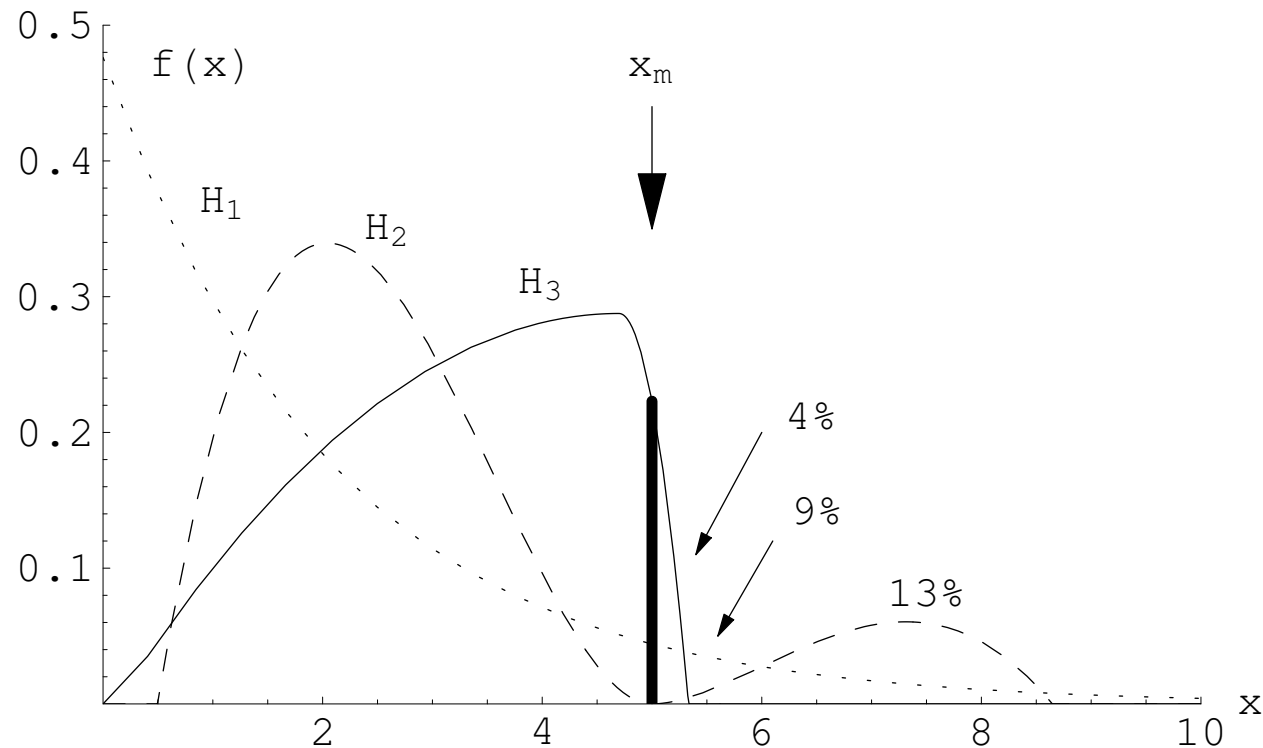
Probabilità di un'altra cosa...

A parte il fatto che il ragionamento basato sulla sola probabilità dell'effetto è bacato all'origine, il problema 'tecnico' degli **eventi di bassa probabilità che porterebbero a rifiutare qualsiasi ipotesi** induce lo statistico a riflettere sulla faccenda...

- ⇒ Come per magia il risultato 'diventa' molto probabile!
Di cosa ci preoccupavamo?
- ⇒ tutti contenti ... e co...
- ⇒ Dal punto di vista logico la situazione è peggiorata:
le nostre **conclusioni** non **dipendono** soltanto da quanto abbiamo osservato, ma anche **da eventi che non sono stati osservati!**

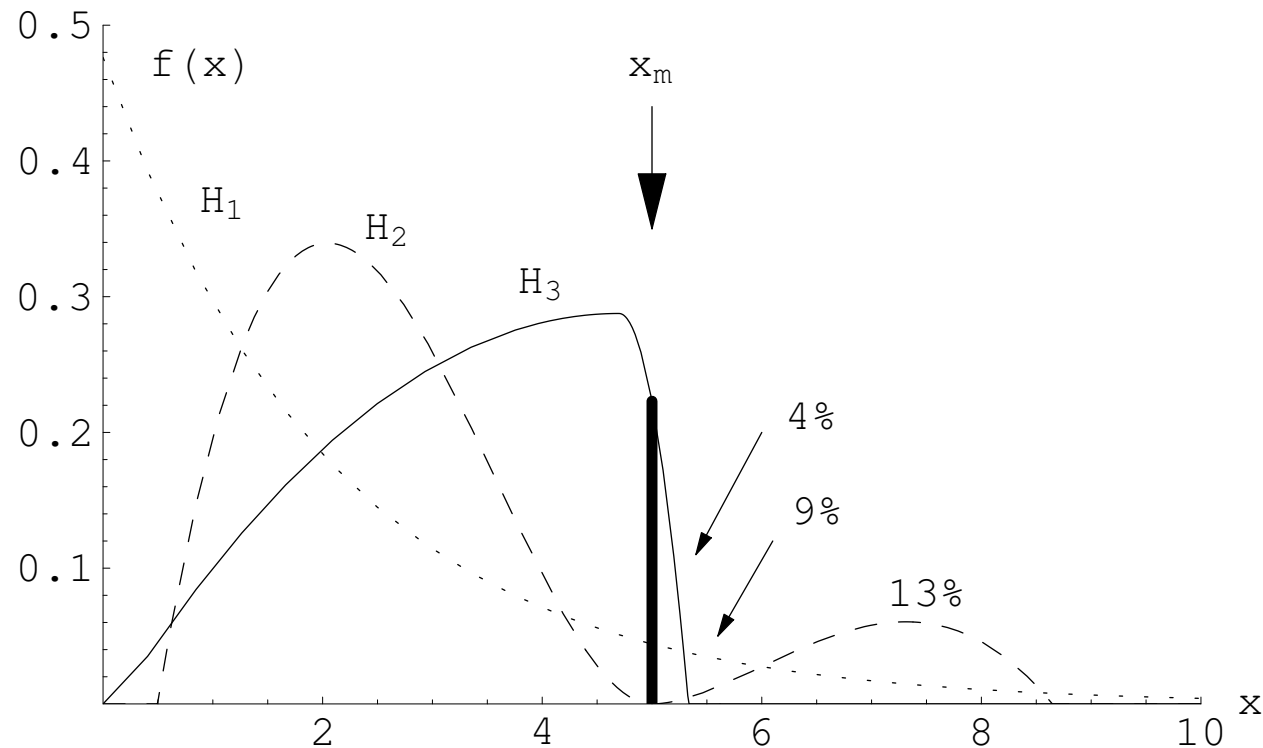
Tre ipotesi a confronto

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?



Tre ipotesi a confronto

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?

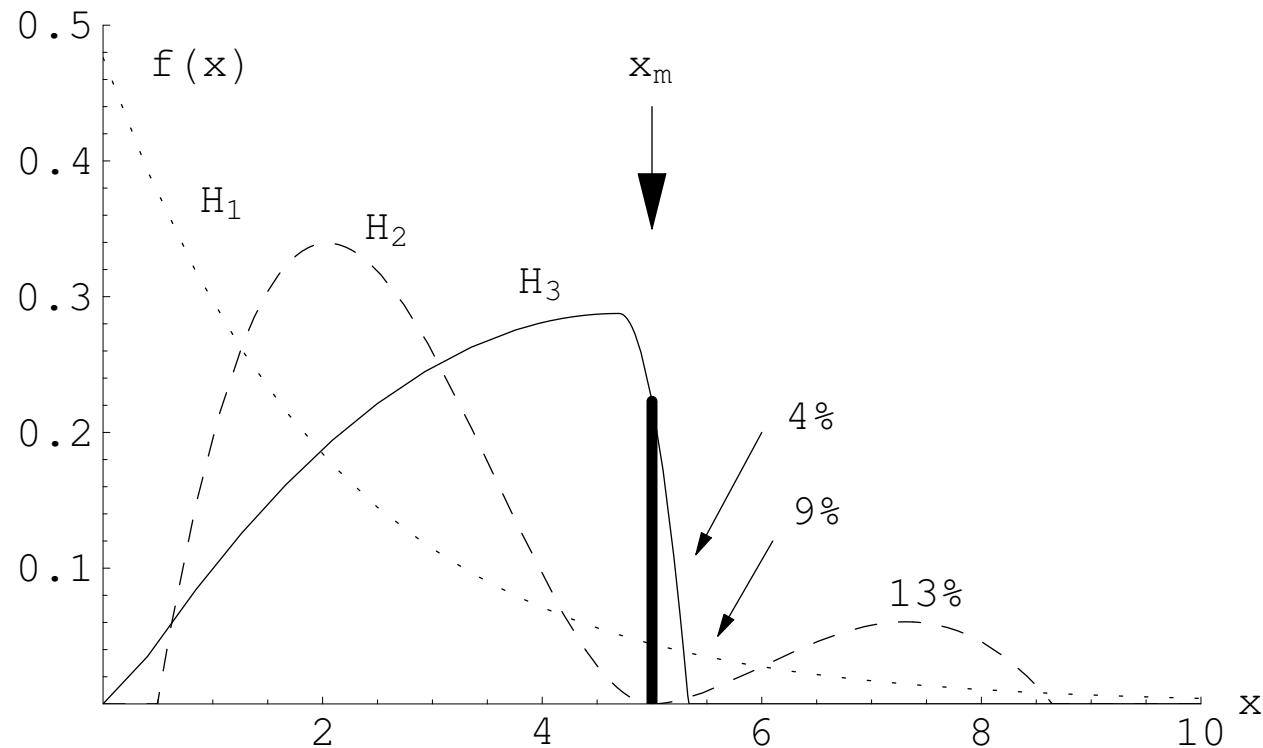


Anche se $P(x_m | H_i) \rightarrow 0$ (dipende dalla risoluzione)

$$P(x_m | H_3) > P(x_m | H_1) > P(x_m | H_2) = 0 \quad (!)$$

Tre ipotesi a confronto

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?

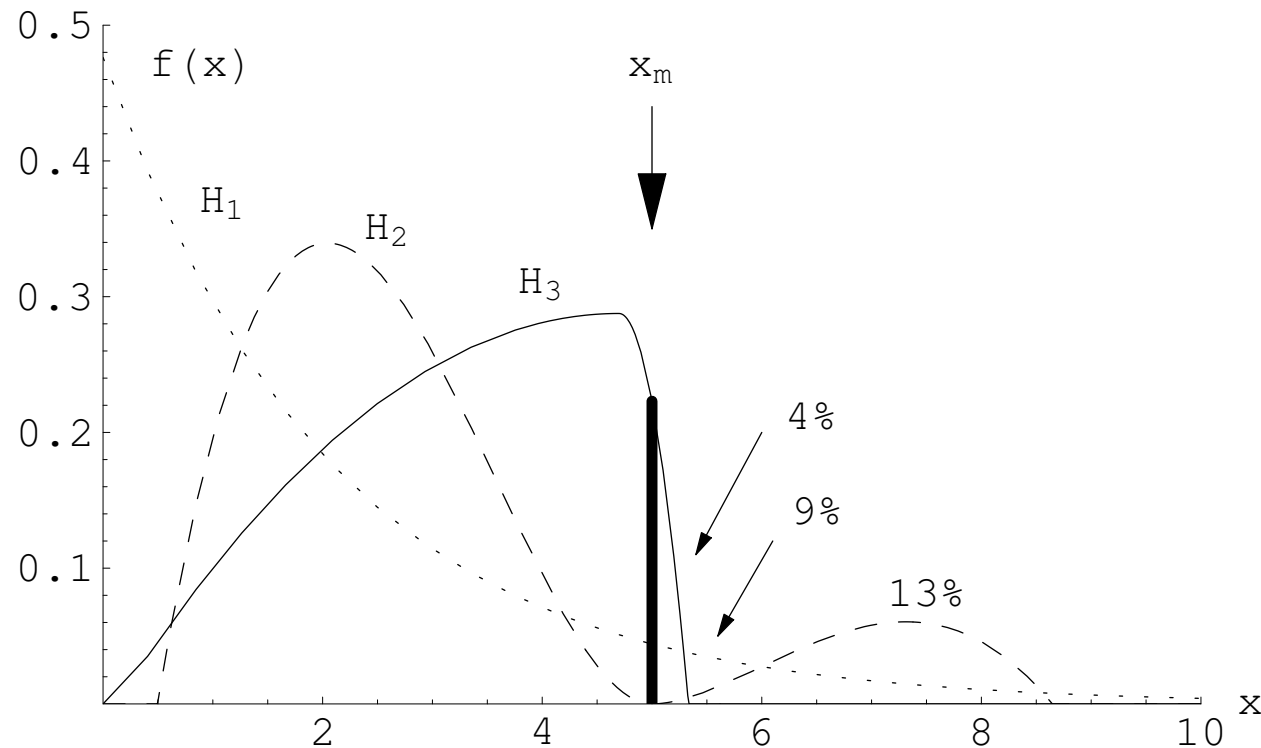


In particolare, l'ipotesi H_2 è (veramente) falsificata (in quanto impossibile), benché abbia il maggiore

'p-value'

Tre ipotesi a confronto

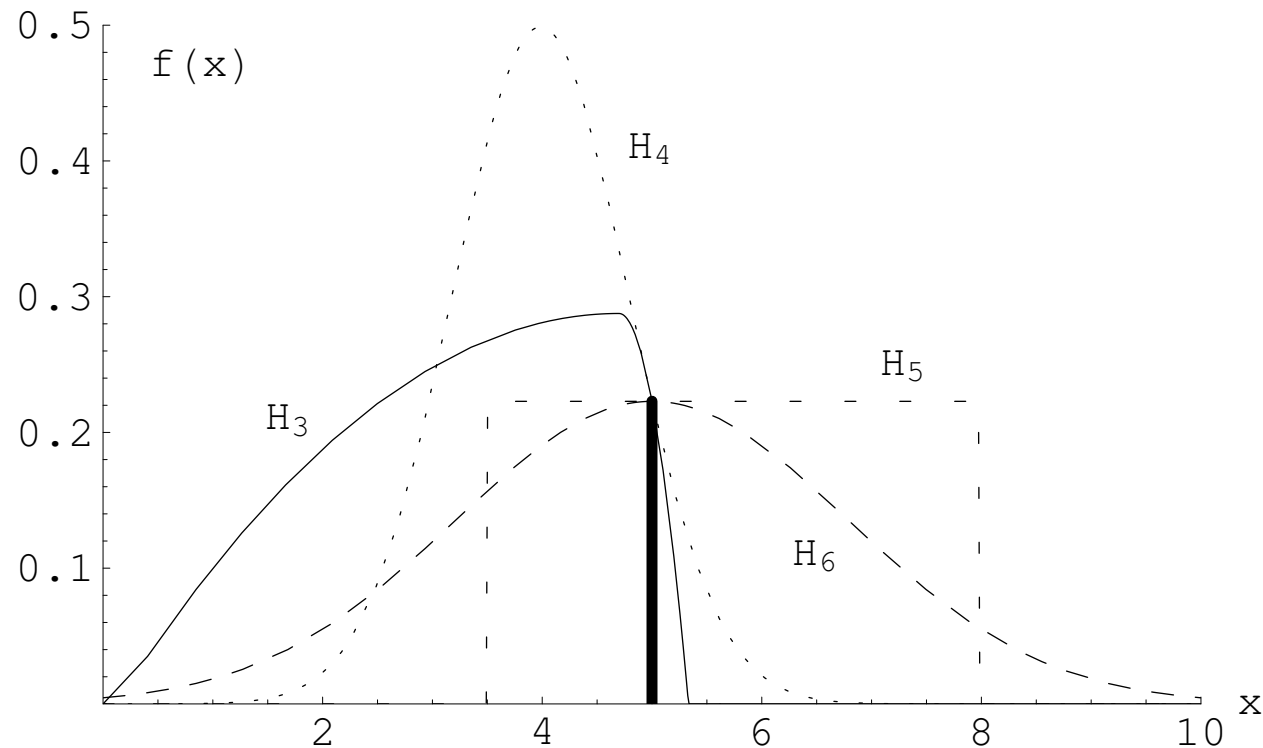
Quale ipotesi è favorita dall'osservazione sperimentale x_m ?



In particolare, l'ipotesi H_2 è (veramente) falsificata (in quanto impossibile), benché abbia la maggiore
'probabilità della coda'

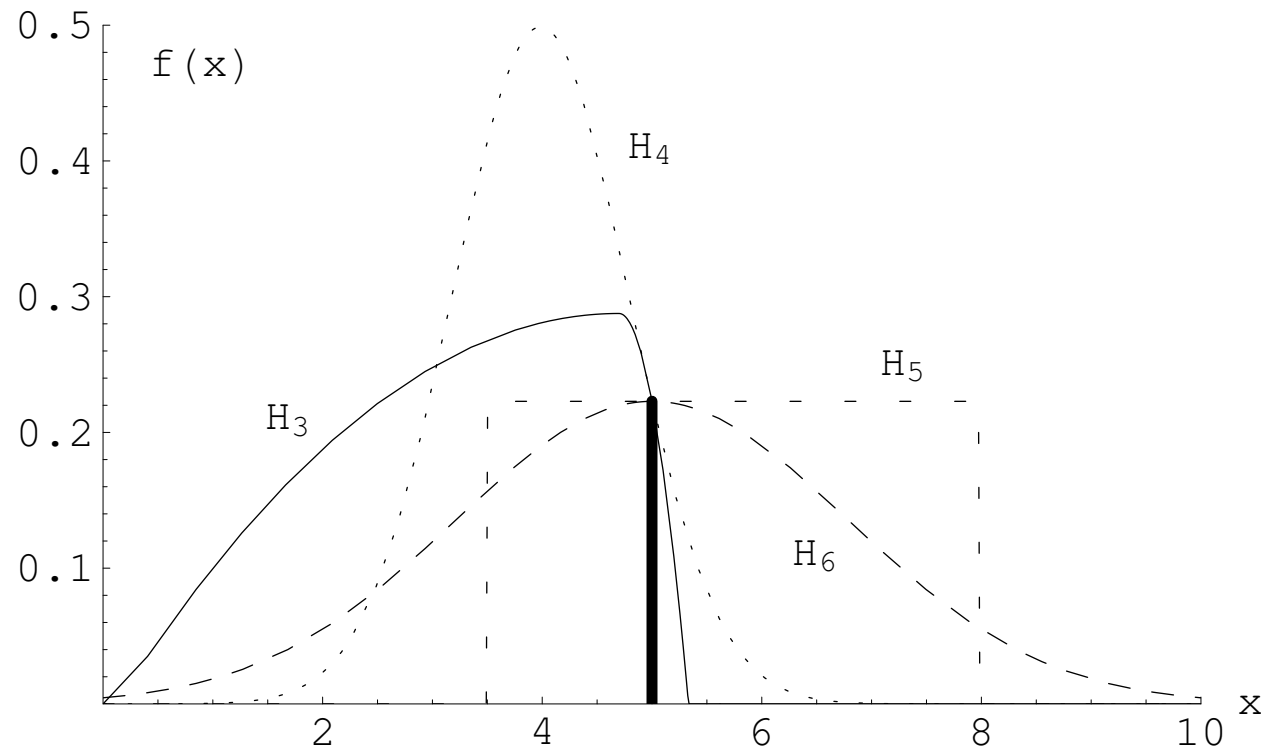
Un esperimento irrilevante

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?



Un esperimento irrilevante

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?

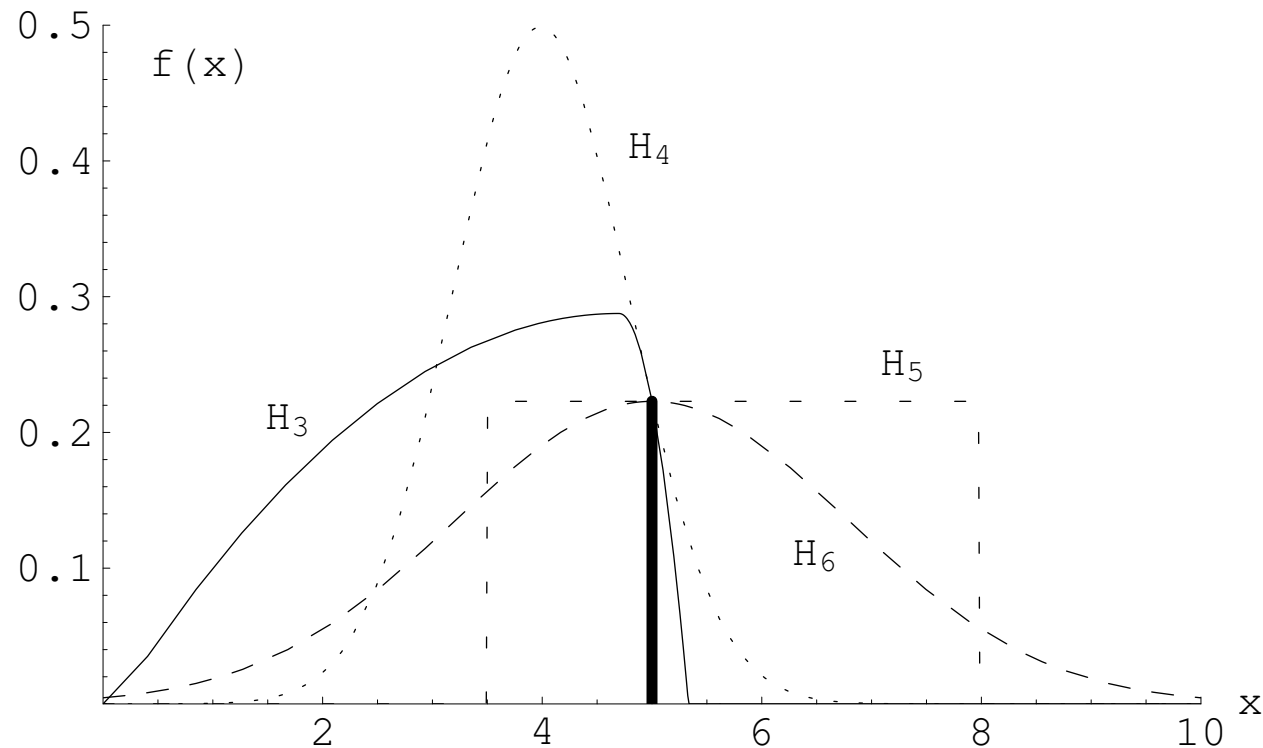


$$P(x_m | H_3) = P(x_m | H_4) = P(x_m | H_5) = P(x_m | H_6)$$

⇒ *Il risultato sperimentale è irrilevante!*

Un esperimento irrilevante

Quale ipotesi è favorita dall'osservazione sperimentale x_m ?



$$P(x_m | H_3) = P(x_m | H_4) = P(x_m | H_5) = P(x_m | H_6)$$

⇒ *Il risultato sperimentale è irrilevante!*

⇒ *... indipendentemente dal p-value!*

Uno, nessuno, 100000...

'p-value' = 'probabilità delle code'

Uno, nessuno, 100000...

'p-value' = 'probabilità delle code'

Coda di cosa?

Uno, nessuno, 100000...

'p-value' = 'probabilità delle code'

Coda di cosa?

→ la variabile di test (' θ ') è assolutamente arbitraria:

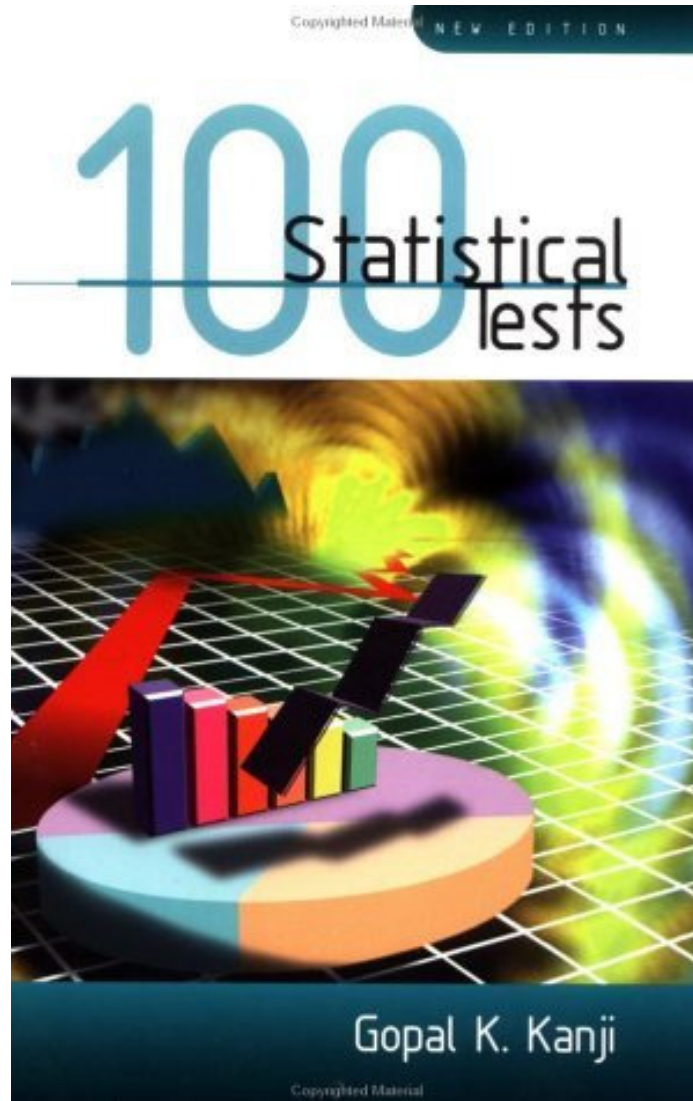
$$\theta = \theta(\mathbf{x})$$

$$\rightarrow f(\theta) \text{ [p.d.f]}$$

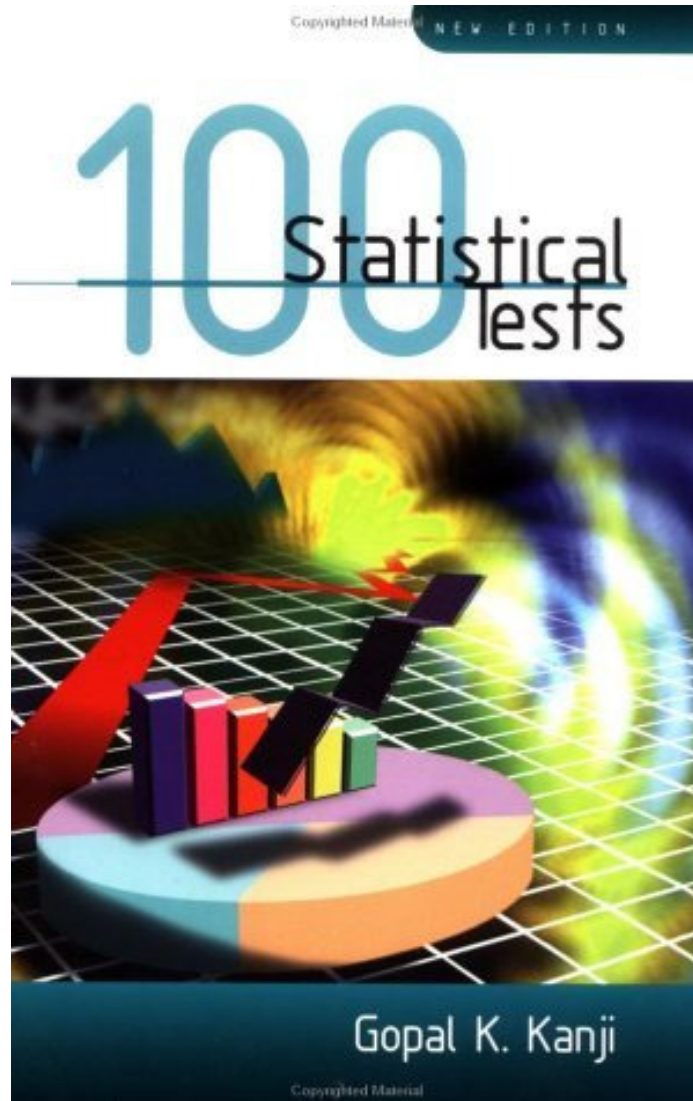
$$\text{Esperimento: } \rightarrow \theta_{mis} = \theta(\mathbf{x}_{mis})$$

$$\text{p-value} = P(\theta \geq \theta_{mis})$$

Uno, nessuno, 100000...



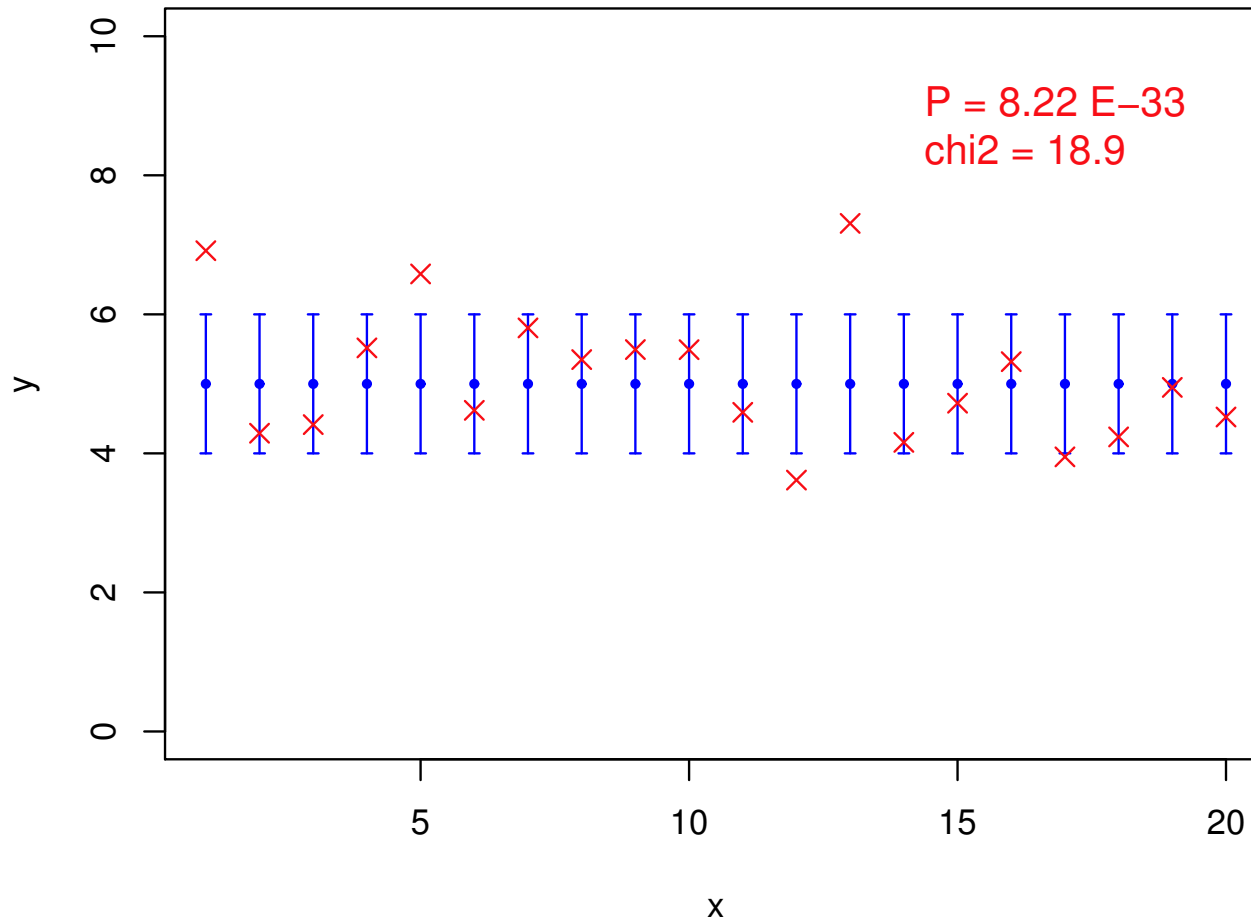
Uno, nessuno, 100000...



- lista non esaustiva,
 - con varianti, con arbitrarietà su quale scegliere.
- ⇒ si sceglie quello che fa più comodo:
- come andare in giro sperando di trovare “uno che ti dia ragione”...
- **‘Regola d’oro’** personale: “più esotico è il nome del test, meno credo al risultato”, perché penso che ne saranno stati scartati un certo numero che “gli davano torto”...

χ^2 ... il padre di tutti i p-value

Teoria ed esperimento:



Modello semplicissimo: $\forall x, y = y_0 = 5$ (u.a.) ed errore strumentale gaussiano con $\sigma = 1$.

Probabilità della configurazione

$P = 8.22 \times 10^{-33}$ è la probabilità della ‘configurazione’ di punti sperimentali:

- ottenuta moltiplicando la probabilità di ciascun punto (misure indipendenti):

$$P = \prod_i P_i$$

ove

$$P_i = \int_{y_{m_i} - \Delta y/2}^{y_{m_i} + \Delta y/2} f(y) dy$$

- come abbiamo visto, P_i dipende dalla ‘risoluzione’ Δy (‘quantizzazione’ dello strumento):

$$\rightarrow \text{usiamo } \Delta y = \frac{1}{10} \sigma$$

'Distanza' esperimento-teoria: χ^2

La costruzione del χ^2 è molto popolare
(i bei tempi di 'fisichetta!):

$$\chi^2 = \sum_i \left(\frac{y_{m_i} - y_{th_i}}{\sigma_i} \right)^2$$

$$\rightarrow \sum_i \left(\frac{y_{m_i} - y_0}{\sigma} \right)^2$$

$$\chi^2 \sim \Gamma(\nu/2, 1/2) \quad [\rightarrow \nu = 20]$$

$$\mathbf{E}[\chi^2] = \nu \quad [\rightarrow 20]$$

$$\mathbf{Var}[\chi^2] = 2\nu \quad [\rightarrow 40]$$

$$\mathbf{Std}[\chi^2] = \sqrt{2\nu} \quad [\rightarrow 6.3]$$

\Rightarrow

$$\boxed{\chi^2 = 20 \pm 6}$$

Le nostre aspettative sul χ^2

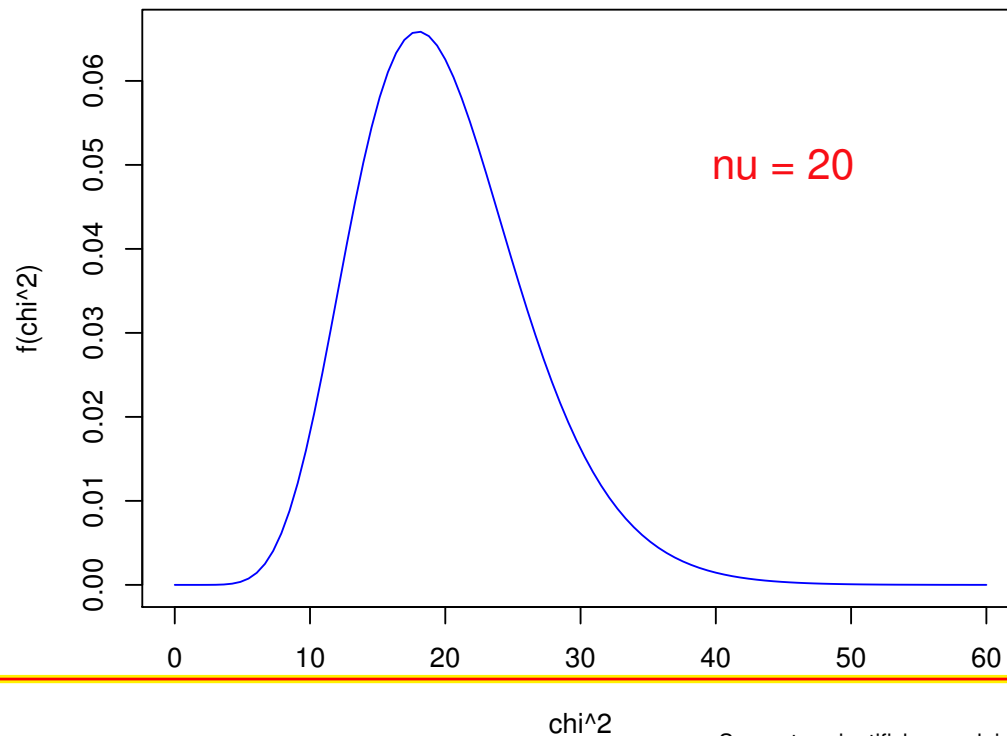
$$E[\chi^2] = \nu \quad [\rightarrow 20]$$

$$\text{Std}[\chi^2] = \sqrt{2\nu} \quad [\rightarrow 6.3]$$

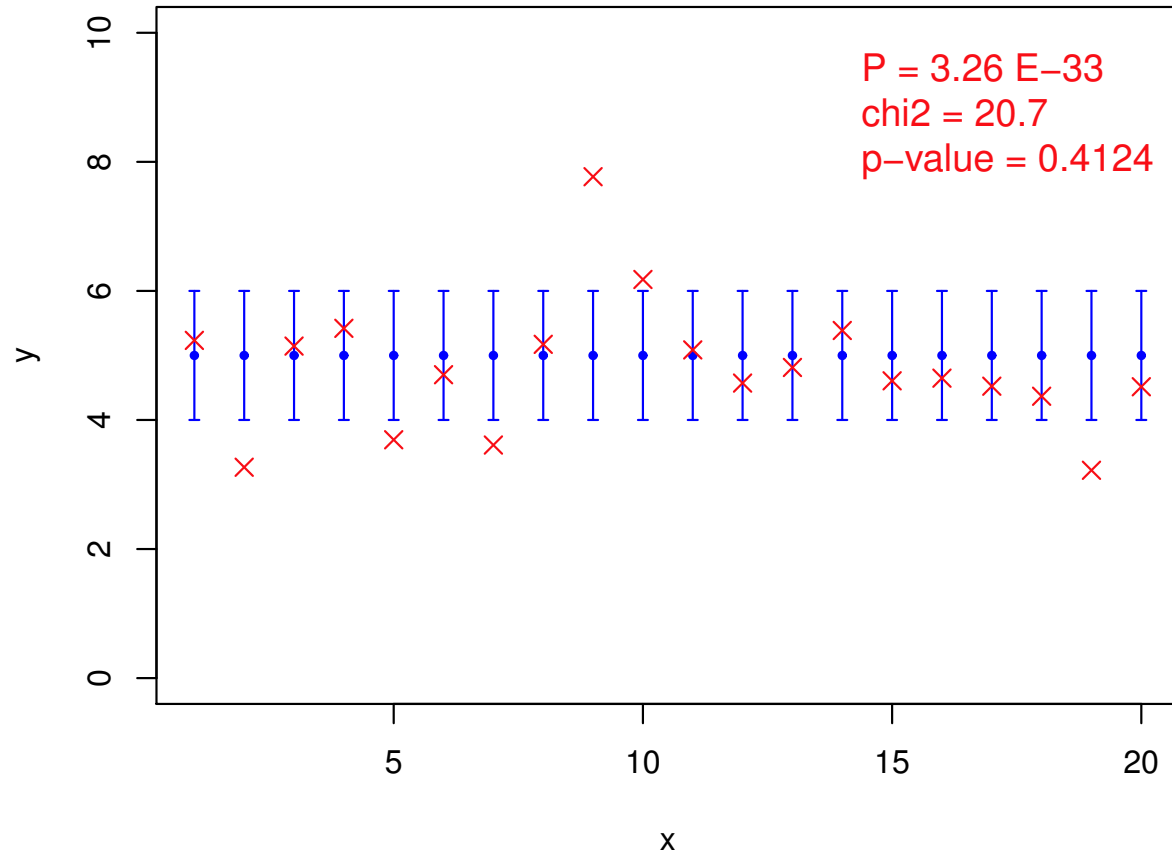
\Rightarrow

$$\chi^2 = 20 \pm 6$$

[moda: 18]



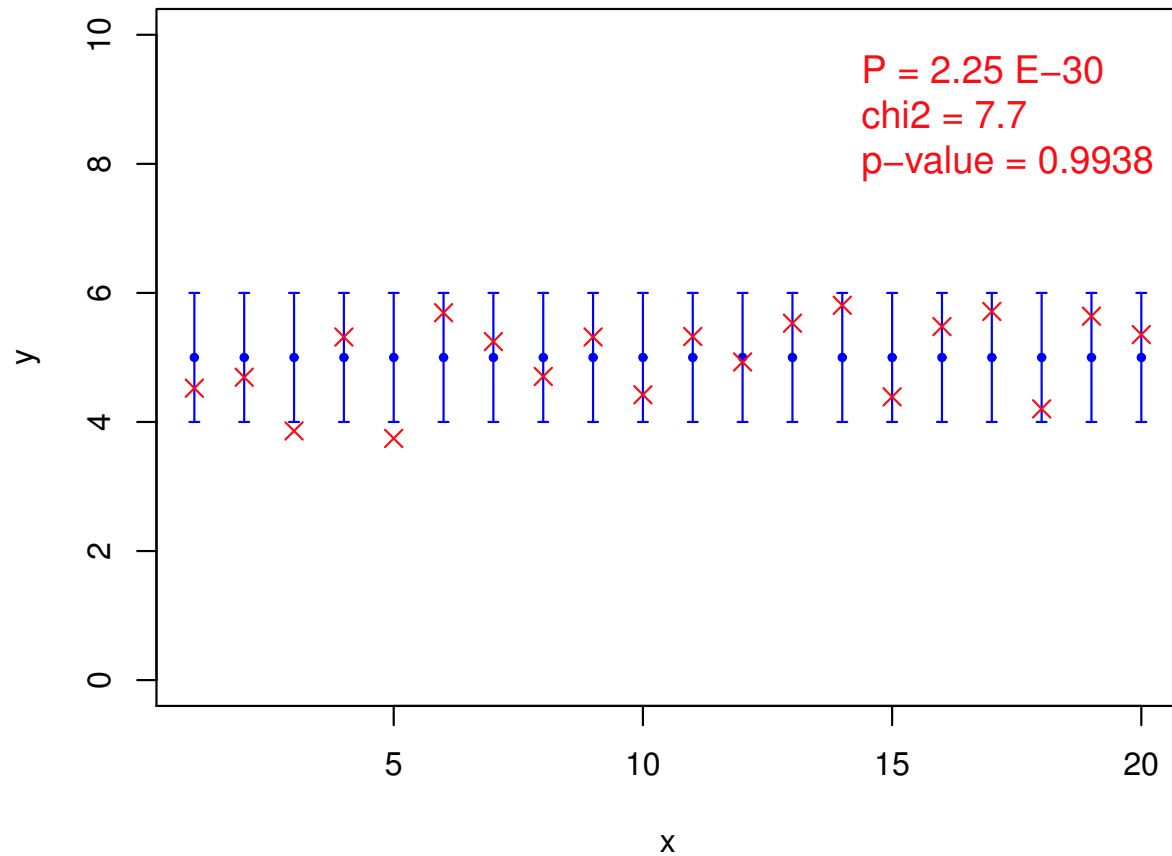
Alcuni esempi



Nella media.

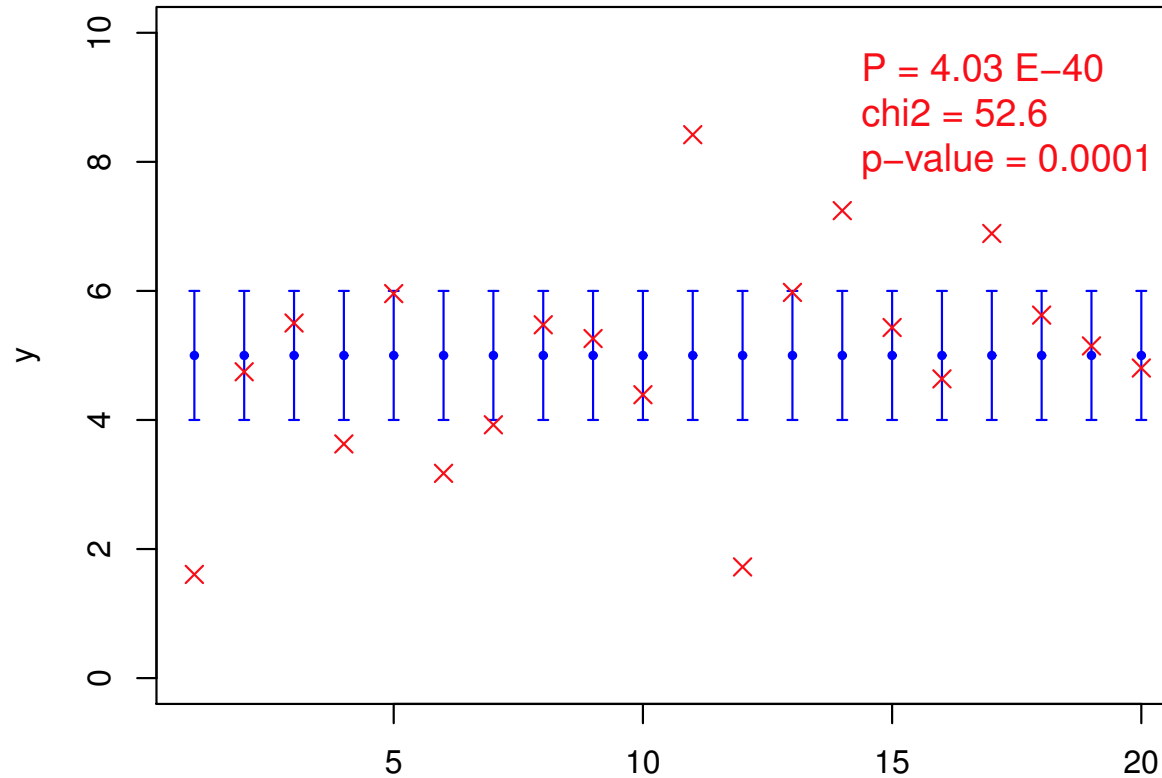
(Ma qualcuno comincerà a vederci qualche 'costellazione'...)

Alcuni esempi



Troppo buono?

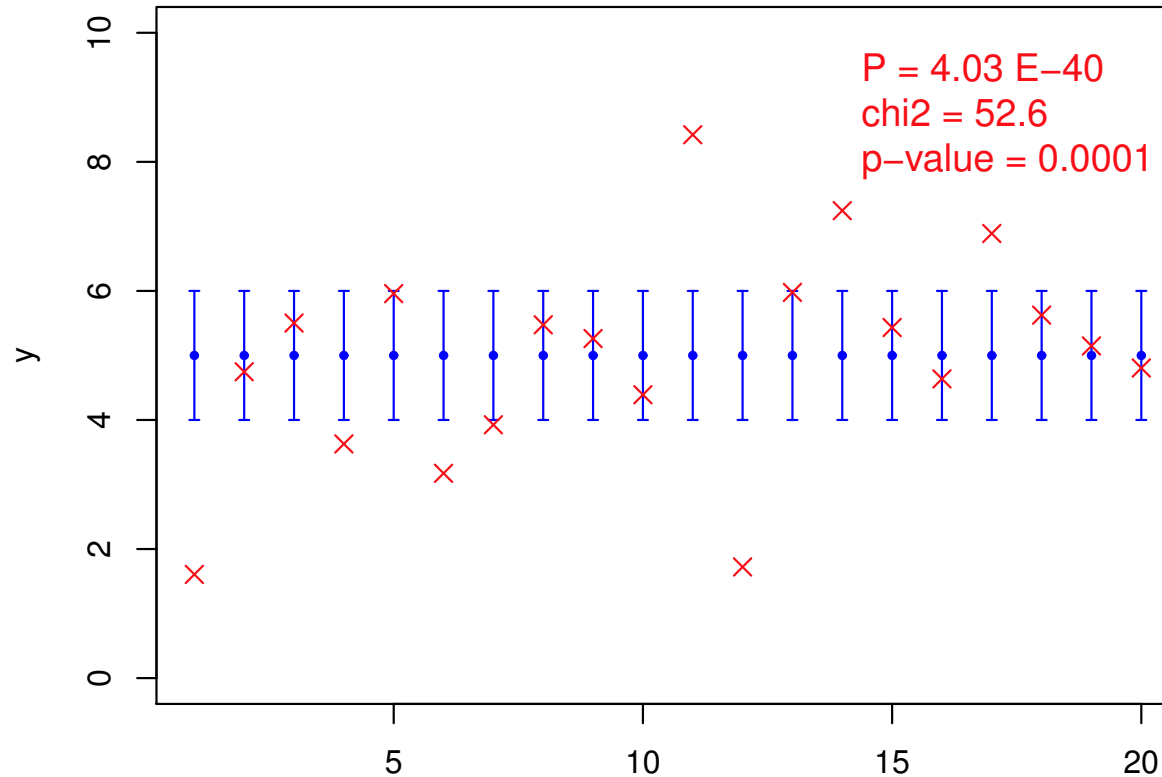
Alcuni esempi



$\chi^2 = 52.6$, con p-value = 0.93×10^{-4}

Al limite?

Alcuni esempi

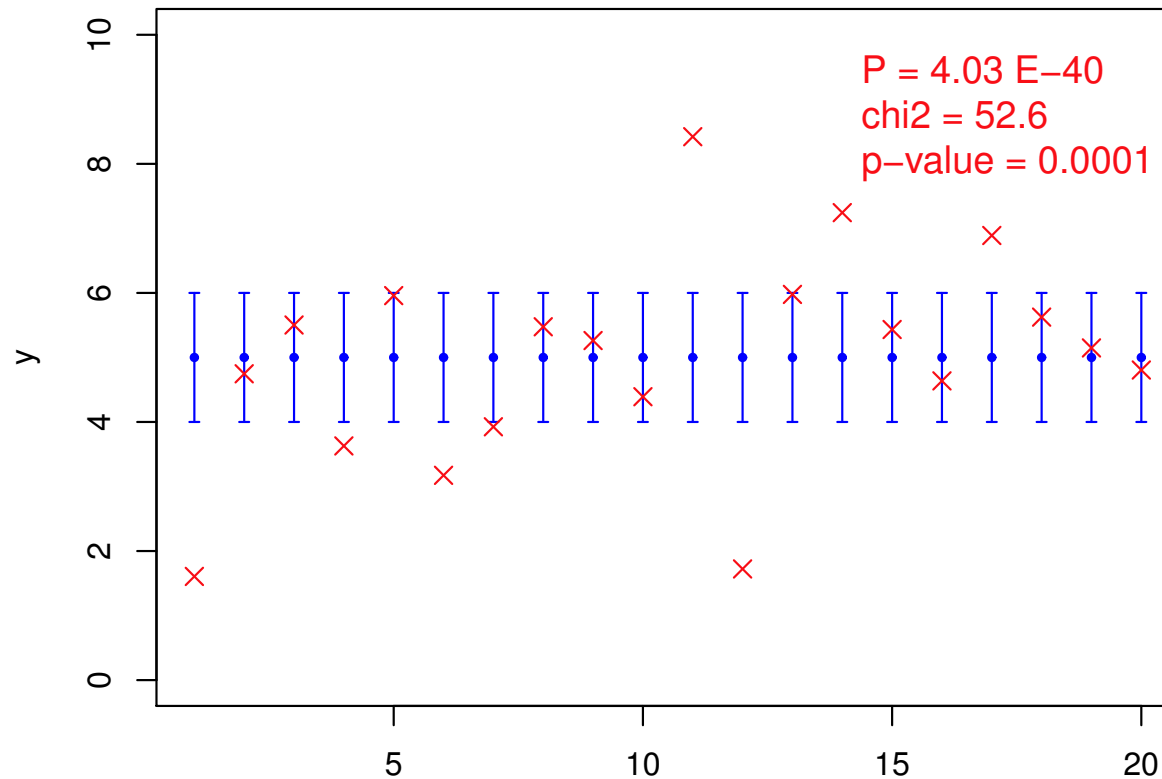


$\chi^2 = 52.6$, con p-value = 0.93×10^{-4}

Al limite? Uscito al 'primo colpo' (martedì 9 ottobre, 13:01)

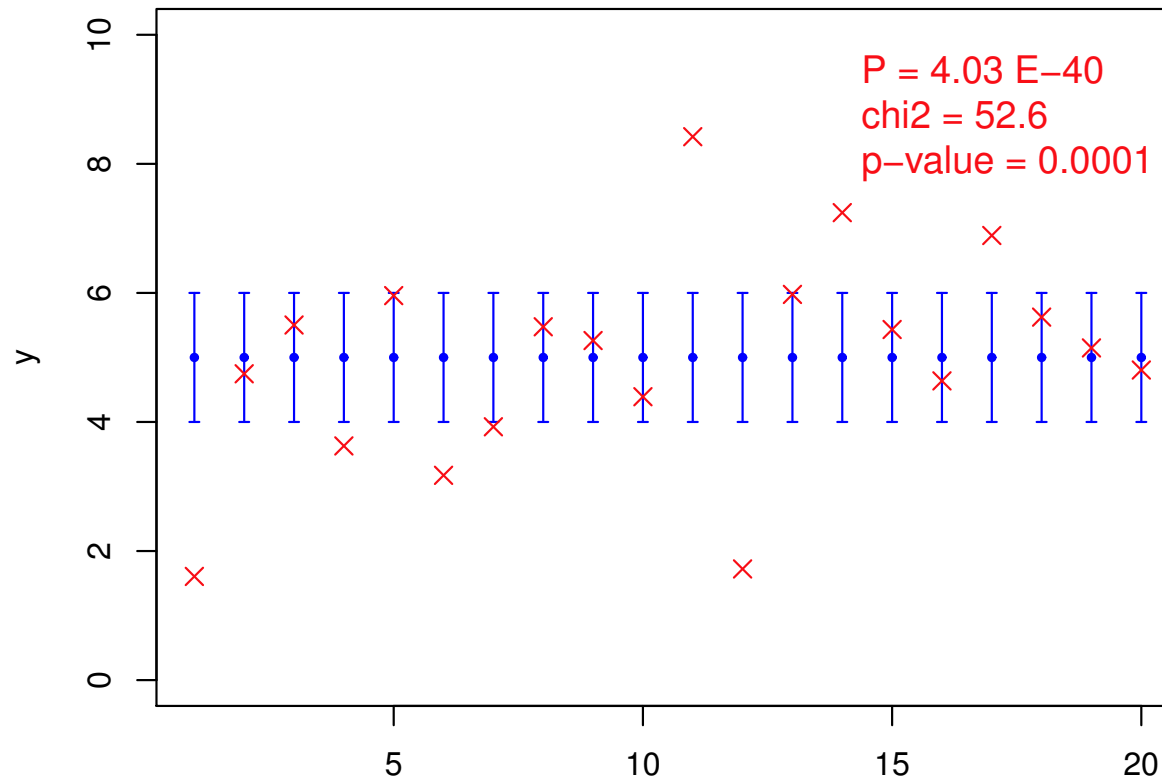
```
while (chi2.y() < 38) source ("chi2_1.R")
```

Alcuni esempi



Att: il χ^2_{mis} 52.6 è '5.1 sigma' dal suo valore atteso $[\frac{52.6-20}{\sqrt{40}}]$
ma il p-value è comunicato come "3.7 σ ", facendo riferimento alla probabilità della coda oltre 3.7 σ di una 'gaussiana equivalente'.
(come se non ci fosse già abbastanza confusione...)

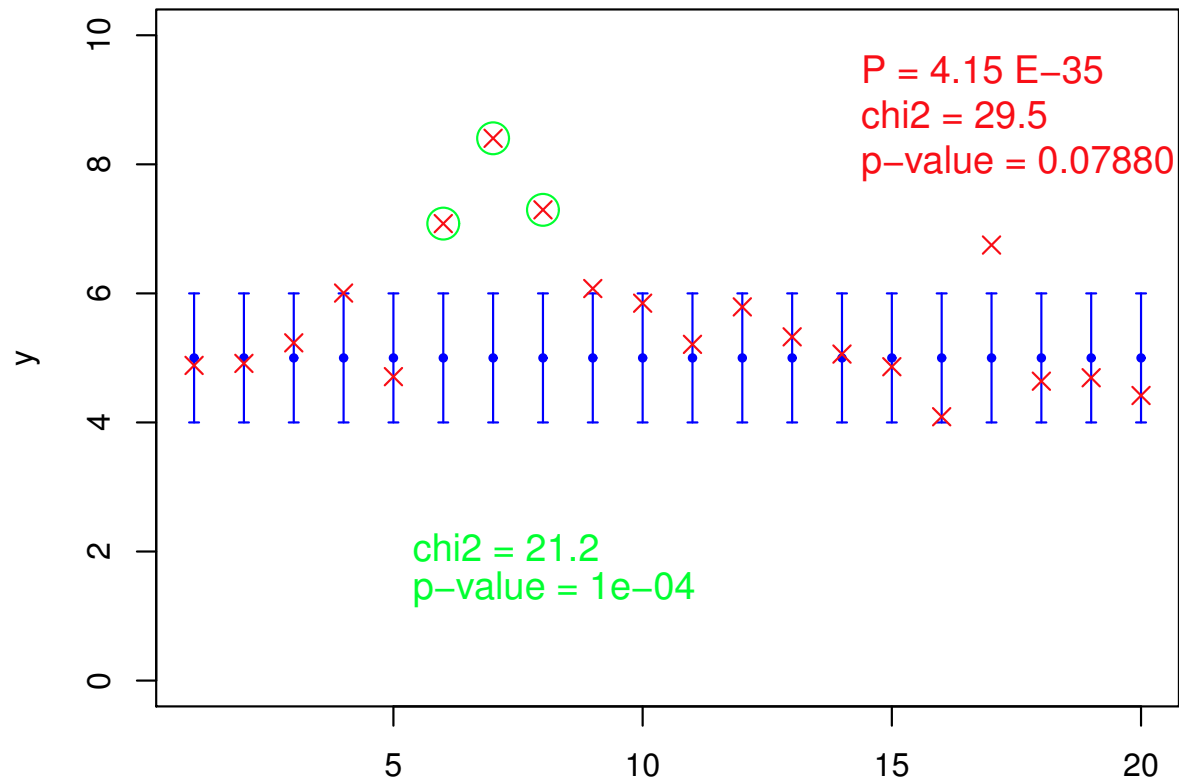
Alcuni esempi



Qualche simulazione in tempo reale \Rightarrow

L'arte del χ^2

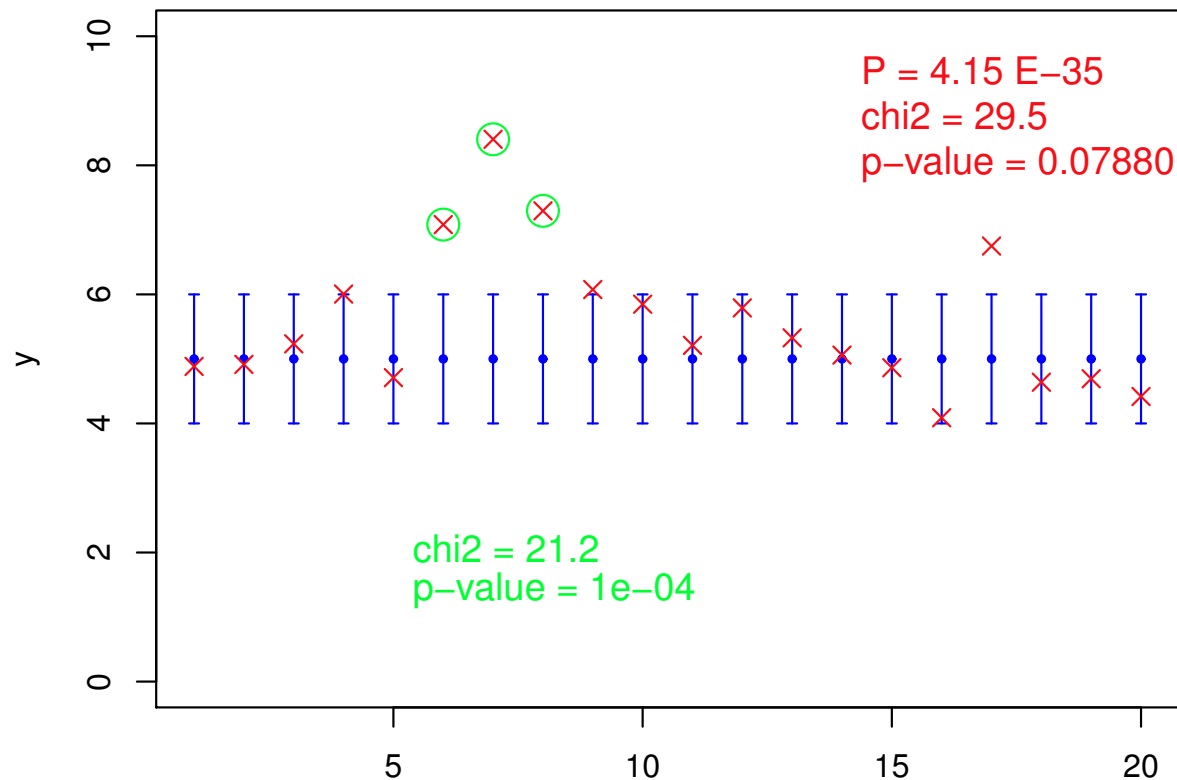
A volte il χ^2 non dà “quello che uno desidera”



E allora lo si calcola solo nella ‘zona sospetta’

L'arte del χ^2

A volte il χ^2 non dà “quello che uno desidera”



E allora lo si calcola solo nella ‘zona sospetta’

⇒ Se avessimo aggiunto i due punti laterali, il χ^2 sarebbe passato a **22.2**.

⇒ Ma con **5 punti** avremmo avuto un **p-value di 5×10^{-4}** .

p-value: cosa sono

p-value:

- Probabilità della(e) coda(e) di una ‘variabile di test’ (“statistica” – inglese ‘statistic’):

$$P(\theta \geq \theta_{mis}) = \int_{\theta_{mis}}^{\infty} f(\theta | H_0) d\theta$$

$$P[(\theta \geq \theta_{mis}) \cap (\theta \leq -\theta_{mis})] = 1 - \int_{-\theta_{mis}}^{\theta_{mis}} f(\theta | H_0) d\theta$$

- θ è funzione arbitraria dei dati,
- ... e spesso di un sottoinsieme dei dati.
- $f(\theta | H_0)$ è valutata in ‘qualche modo’, analiticamente, numericamente, o con metodi di Monte Carlo (simulazione mediante campionamento stocastico).

p-value: cosa non sono

- Cosa volevamo:
 - falsificare l'ipotesi H_0 :
 - ⇒ impossibile, dal punto di vista logico

p-value: cosa non sono

- Cosa volevamo:
 - falsificare l'ipotesi H_0 :
⇒ impossibile, dal punto di vista logico
- Ci accontentiamo allora di
 - rivedere il nostro grado di fiducia su H_0 alla luce dei dati sperimentali:

$$P(H_0 \mid \text{dati})$$

p-value: cosa non sono

- Cosa volevamo:
 - falsificare l'ipotesi H_0 :
⇒ impossibile, dal punto di vista logico
- Ci accontentiamo allora di
 - rivedere il nostro grado di fiducia su H_0 alla luce dei dati sperimentali:

$$P(H_0 \mid \text{dati})$$

⇒ **MA i p-value non ci danno questo:**

$$P(\theta \geq \theta_{mis} \mid H_0) \not\iff P(H_0 \mid \theta_{mis})$$

⇒ **Anche se erroneamente vengono confusi con questo!**

p-value: cosa non sono

- Cosa volevamo:
 - falsificare l'ipotesi H_0 :
⇒ impossibile, dal punto di vista logico
- Ci accontentiamo allora di
 - rivedere il nostro grado di fiducia su H_0 alla luce dei dati sperimentali:

$$P(H_0 \mid \text{dati})$$

⇒ MA i p-value non ci danno questo:

$$P(\theta \geq \theta_{mis} \mid H_0) \not\iff P(H_0 \mid \theta_{mis})$$

⇒ Anche se erroneamente vengono confusi con questo!

STRINGERE LE CINTURE!

Misunderstandings p-values

<http://en.wikipedia.org/wiki/P-value#Misunderstandings>

1. **The p-value is not the probability that the null hypothesis is true.** In fact, **frequentist statistics does not, and cannot, attach probabilities to hypotheses.** . . .
2. **The p-value is not the probability that a finding is “merely a fluke.”** . . .
3. **The p-value is not the probability of falsely rejecting the null hypothesis.**
 . . .
7. . . .

L'Higgs a 5 sigma!

4 Luglio 2012 (con 'rumors' nei giorni prima):

- “I dati confermano la soglia dei 5 sigma, **vale a dire** una probabilità di scoperta pari al 99,99994 per cento” spiega Gian Francesco Giudice, teorico del CERN (corriere.it, 3 luglio)

L'Higgs a 5 sigma!

4 Luglio 2012 (con 'rumors' nei giorni prima):

- “I dati confermano la soglia dei 5 sigma, **vale a dire** una probabilità di scoperta pari al 99,99994 per cento” spiega Gian Francesco Giudice, teorico del CERN (corriere.it, 3 luglio)
- “Ahead of the expected announcement, the journal Nature reported ‘pure elation’ Monday among physicists searching for the Higgs boson. *One team saw only “a 0.00006% chance of being wrong, the journal said.”* (USA Today, 2 July 2012).

L'Higgs a 5 sigma!

4 Luglio 2012 (con 'rumors' nei giorni prima):

- “I dati confermano la soglia dei 5 sigma, **vale a dire** una probabilità di scoperta pari al 99,99994 per cento” spiega Gian Francesco Giudice, teorico del CERN (corriere.it, 3 luglio)
- “Ahead of the expected announcement, the journal Nature reported ‘pure elation’ Monday among physicists searching for the Higgs boson. *One team saw only “a 0.00006% chance of being wrong, the journal said.”* (USA Today, 2 July 2012).
- Etc. etc.

<http://www.roma1.infn.it/~dagos/badmath/#added>

Sursum corda!

Ma allora non possiamo dire niente?

Sursum corda!

Ma allora non possiamo dire niente?

Possiamo, possiamo, a delle condizioni...

- Quando il gioco si fa probabilistico deve scendere in gioco la **teoria della probabilità**.

Sursum corda!

Ma allora non possiamo dire niente?

Possiamo, possiamo, a delle condizioni...

- Quando il gioco si fa probabilistico deve scendere in gioco la **teoria della probabilità**.

??

Ma non c'erano già gaussiane, chi-quadri, sigma, etc?

- Nel contesto statistico 'classico' dei test di ipotesi, manca (perché esplicitamente vietata) **la cosa fondamentale che ci serve** per il nostro gioco: la **probabilità delle ipotesi**.

- 'Mismatch' fra il nostro **modo naturale di pensare** e gli insegnamenti della **statistica**:

- $P(H_0 \mid \text{dati}) \longleftrightarrow P(\theta \geq \theta_{mis} \mid H_0)$

Sursum corda!

Ma allora non possiamo dire niente?

Possiamo, possiamo, a delle condizioni...

- Quando il gioco si fa probabilistico deve scendere in gioco la **teoria della probabilità**.

??

Ma non c'erano già gaussiane, chi-quadri, sigma, etc?

- Nel contesto statistico 'classico' dei test di ipotesi, manca (perché esplicitamente vietata) **la cosa fondamentale che ci serve** per il nostro gioco: **la probabilità delle ipotesi**.

- Basta 'buttare al mare' gli statistici del novecento (i 'frequentisti') e **ricominciare da 'gente seria'**, ripartendo da Laplace, insieme a Gauss, Bayes etc.,

Probabilità e scommessa

Recupero del concetto naturale di probabilità

- “quanto sono confidente in qualcosa”
- “quanto ci credo”

Probabilità e scommessa

Recupero del concetto naturale di probabilità

- “quanto sono confidente in qualcosa”
- “quanto ci credo”

“Mi gioco una palla che questo è un linfoma Hodgkin, due no ma una sì” (Nanni Moretti, Caro Diario)

Probabilità e scommessa

Recupero del concetto naturale di probabilità

- “quanto sono confidente in qualcosa”
- “quanto ci credo”

L'ordinaria pietra di paragone per vedere se qualche cosa, che uno afferma, sia una semplice persuasione, o almeno una convinzione soggettiva, cioè una ferma fede, è la scommessa. Spesso uno enuncia le sue proposizioni con una risolutezza così sicura e irriducibile da parere abbia deposto ogni tema di errore. Una scommessa lo fa adombrare. A volte si vede che egli possiede bensì una persuasione da poter essere apprezzata per un ducato, ma non per dieci. Infatti egli arrischia il primo ma, di fronte a dieci, comincia ad avvedersi di ciò che prima non avvertiva, essere cioè possibilissimo che si sia sbagliato. (Kant, Critica della Ragion Pura)

Probabilità e scommessa

Recupero del concetto naturale di probabilità

- “quanto sono confidente in qualcosa”
- “quanto ci credo”
- “più ci credo, più sono disposto a scommetterci”

Probabilità e scommessa

Recupero del concetto naturale di probabilità

- “quanto sono confidente in qualcosa”
- “quanto ci credo”
- “più ci credo, più sono disposto a scommetterci”
- “il mio grado di fiducia dipende dalle informazioni che io ho (nel cervello!)”
- “ritengo naturale – anzi sarei terrorizzato dal contrario!
– che altri cervelli abbiano diverse ‘informazioni’”
[*‘probabilità soggettiva’*]
- “razionalmente sono disposto a cambiare opinione”
- “... ma più una cosa sembra a priori inverosimile più ho bisogno di prove in supporto, fornite possibilmente da diverse persone nelle quali ripongo fiducia”

Il “teorema di Bayes”... di Laplace

[In termini di **Cause** (‘ipotesi’) ed **Effetti** (‘dati’)]

“Maggiore è la probabilità di un evento osservato, dato un certo numero di cause a cui l’evento può essere attribuito, maggiore è la probabilità di quella causa {dato quell’evento}.

$$P(C_i | E) \propto P(E | C_i)$$

Il “teorema di Bayes”... di Laplace

[In termini di **Cause** (‘ipotesi’) ed **Effetti** (‘dati’)]

“**Maggiore è la probabilità di un evento osservato**, dato un certo numero di cause a cui l’evento può essere attribuito, **maggiore è la probabilità di quella causa** {dato quell’evento}. La probabilità di una qualsiasi di tali cause {dato l’evento} è **così** una frazione il cui nominatore è la probabilità dell’evento data la causa, e il denominatore è la somma di simili probabilità, somma estesa a tutte le cause.

$$P(C_i | E) = \frac{P(E | C_i)}{\sum_j P(E | C_j)}$$

Il “teorema di Bayes”... di Laplace

[In termini di **Cause** (‘ipotesi’) ed **Effetti** (‘dati’)]

“Maggiore è la probabilità di un evento osservato, dato un certo numero di cause a cui l’evento può essere attribuito, maggiore è la probabilità di quella causa {dato quell’evento}. La probabilità di una qualsiasi di tali cause {dato l’evento} è **così** una frazione il cui nominatore è la probabilità dell’evento data la causa, e il denominatore è la somma di simili probabilità, somma estesa a tutte le cause. **Se le varie cause non sono ugualmente probabili *a priori***, è necessario, invece della probabilità dell’evento data ciascuna causa, usare il prodotto di tale probabilità e la ***possibilità della causa stessa.***”

$$P(C_i | E) = \frac{P(E | C_i) P(C_i)}{\sum_j P(E | C_j) P(C_j)}$$

Il “teorema di Bayes”... di Laplace

$$P(C_i | E) = \frac{P(E | C_i) P(C_i)}{\sum_j P(E | C_j) P(C_j)}$$

“Questo è il **principio fondamentale** (*) di quel ramo della teoria delle probabilità che consiste nel ragionare *a posteriori* **dagli eventi alle cause**”

(*) Nel “Saggio Filosofico” Laplace chiama ‘principi’ le regole fondamentali’.

Il “teorema di Bayes”... di Laplace

$$P(C_i | E) = \frac{P(E | C_i) P(C_i)}{\sum_j P(E | C_j) P(C_j)}$$

“Questo è il **principio fondamentale** (*) di quel ramo della teoria delle probabilità che consiste nel ragionare *a posteriori* **dagli eventi alle cause**”

(*) Nel “Saggio Filosofico” Laplace chiama ‘principi’ le regole fondamentali’.

Nota: il denominatore è un semplice fattore di normalizzazione.

$$\Rightarrow P(C_i | E) \propto P(E | C_i) P(C_i)$$

(Modo più conveniente per memorizzare il teorema)

Il “teorema di Bayes”... di Laplace

$$P(C_i | E) = \frac{P(E | C_i) P(C_i)}{\sum_j P(E | C_j) P(C_j)}$$

“Questo è il **principio fondamentale** (*) di quel ramo della teoria delle probabilità che consiste nel ragionare *a posteriori* **dagli eventi alle cause**”

(*) Nel “Saggio Filosofico” Laplace chiama ‘principi’ le regole fondamentali’.

Altra chiave di lettura (in termini di H_0 e H_1 e **dati**)

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

L'insegnamento di Laplace

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- 'Servirebbero' i dati e non le variabili di test ' θ ' (χ^2 etc);
[anche se in taluni casi esistono 'riassunti sufficienti']

L'insegnamento di Laplace

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- 'Servirebbero' i dati e non le variabili di test ' θ ' (χ^2 etc);
[anche se in taluni casi esistono 'riassunti sufficienti']
- Servono **almeno due ipotesi** alternative!

L'insegnamento di Laplace

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- 'Servirebbero' i dati e non le variabili di test ' θ ' (χ^2 etc);
[anche se in taluni casi esistono 'riassunti sufficienti']
- Servono **almeno due ipotesi** alternative!
- È di cruciale importanza quantificare la *ragionevolezza a priori* delle ipotesi alternative.

L'insegnamento di Laplace

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- ‘Servirebbero’ i dati e non le variabili di test ‘ θ ’ (χ^2 etc);
[anche se in taluni casi esistono ‘riassunti sufficienti’]
- Servono **almeno due ipotesi** alternative!
- È di cruciale importanza quantificare la *ragionevolezza a priori* delle ipotesi alternative.
- Se $P(\text{dati} | H_i) = 0$, ne segue $P(H_i | \text{dati}) = 0$:
⇒ la **falsificazione** (quella ‘seria’) è un **corollario del teorema** e non un ‘principio’.

L'insegnamento di Laplace

$$\frac{P(H_0 | \text{dati})}{P(H_1 | \text{dati})} = \frac{P(\text{dati} | H_0)}{P(\text{dati} | H_1)} \times \frac{P(H_0)}{P(H_1)}$$

- ‘Servirebbero’ i dati e non le variabili di test ‘ θ ’ (χ^2 etc); [anche se in taluni casi esistono ‘riassunti sufficienti’]
- Servono **almeno due ipotesi** alternative!
- È di cruciale importanza quantificare la *ragionevolezza a priori* delle ipotesi alternative.
- Se $P(\text{dati} | H_i) = 0$, ne segue $P(H_i | \text{dati}) = 0$:
⇒ la **falsificazione** (quella ‘seria’) è un **corollario del teorema** e non un ‘principio’.
- Non c’è **nessun problema** concettuale con il fatto che $P(\text{dati} | H_1) \rightarrow 0$ (ad es. 10^{-37}), purché rimanga finito il rapporto $P(\text{dati} | H_0)/P(\text{dati} | H_1)$.

Eppur funzionano!

Qualcuno obietterà che p-value e, in generale, ‘test di ipotesi’ *in genere* funzionano!

Eppur funzionano!

Qualcuno obietterà che p-value e, in generale, ‘test di ipotesi’ *in genere* funzionano!

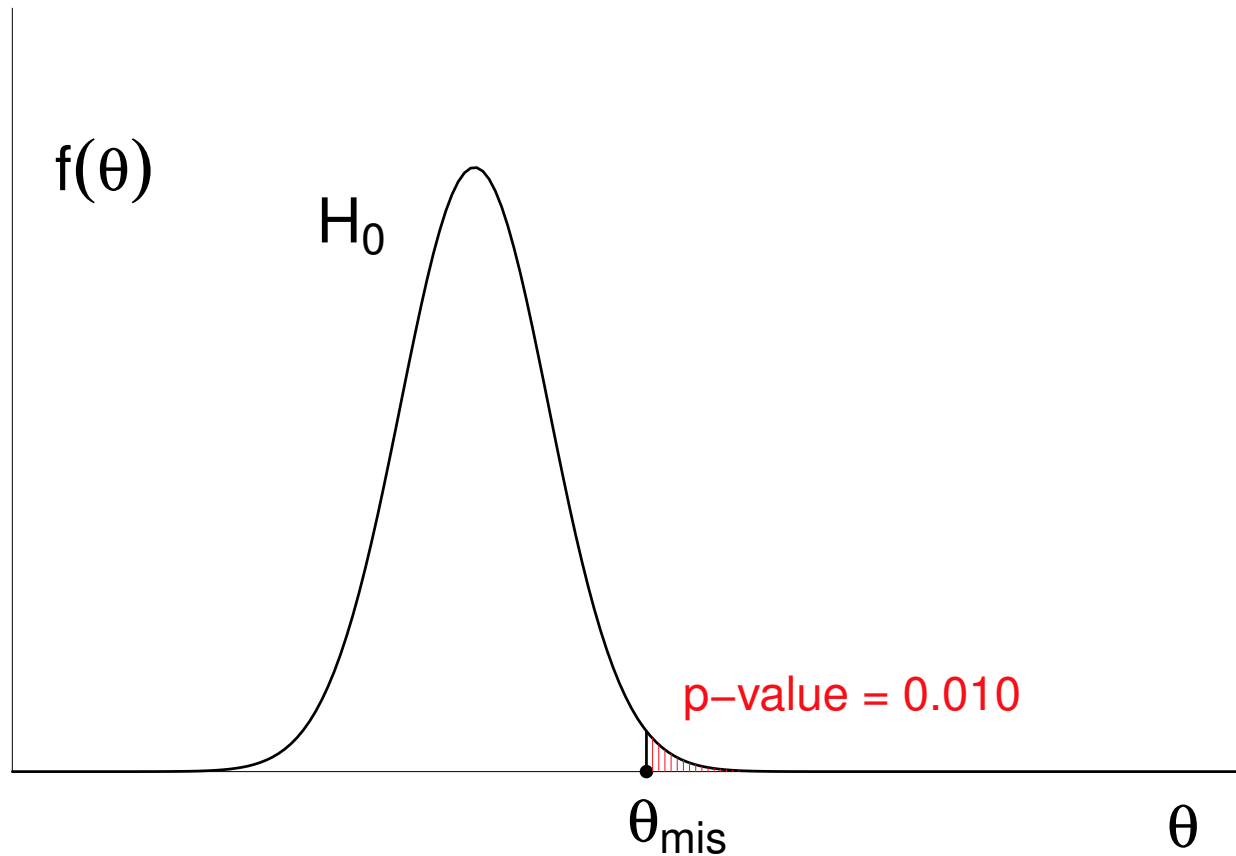
- **Certo!** Assolutamente d’accordo!
Così come *in genere funziona fare un sorpasso in curva* in una sperduta stradina di campagna!

Eppur funzionano!

Qualcuno obietterà che p-value e, in generale, ‘test di ipotesi’ *in genere* funzionano!

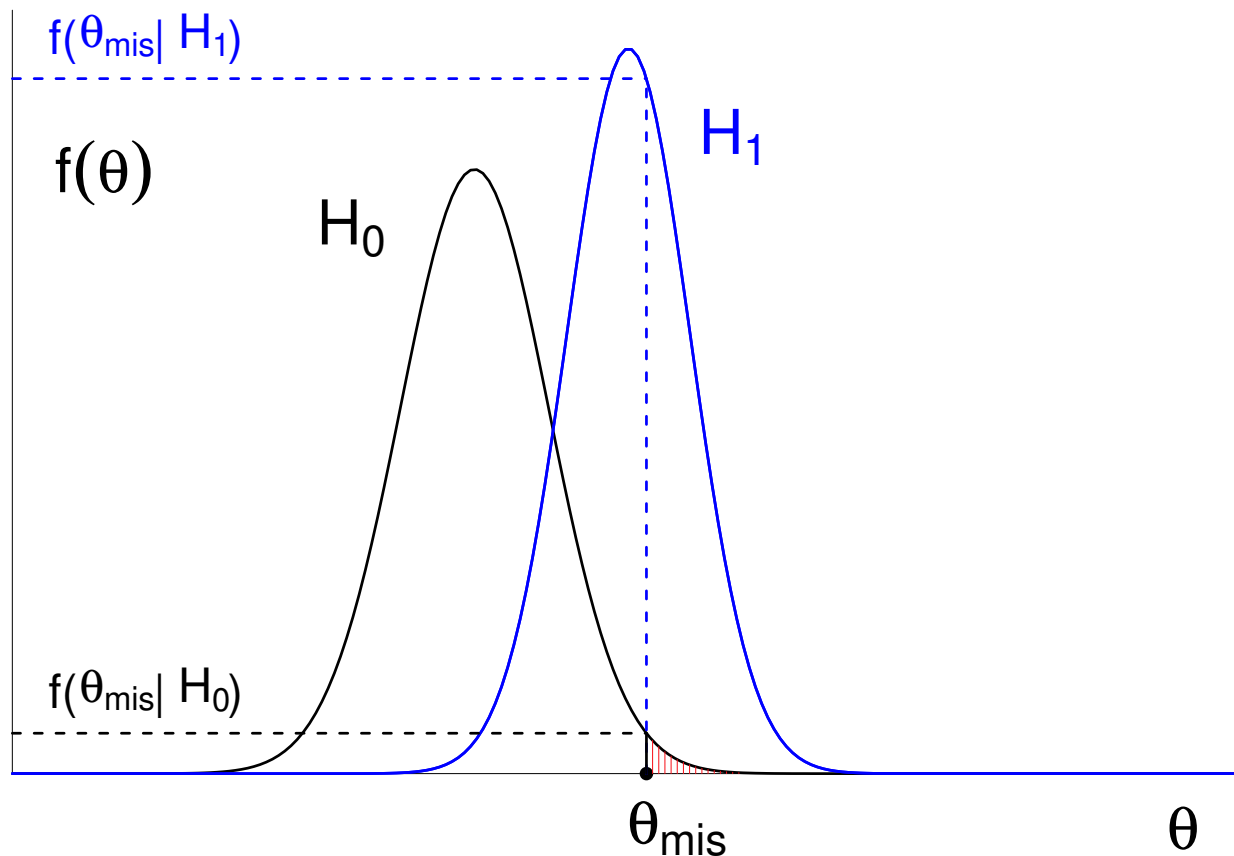
- Certo! Assolutamente d’accordo!
Così come *in genere funziona fare un sorpasso in curva* in una sperduta stradina di campagna!
- Ma adesso siamo anche di spiegare anche perché.

Eppur funzionano!



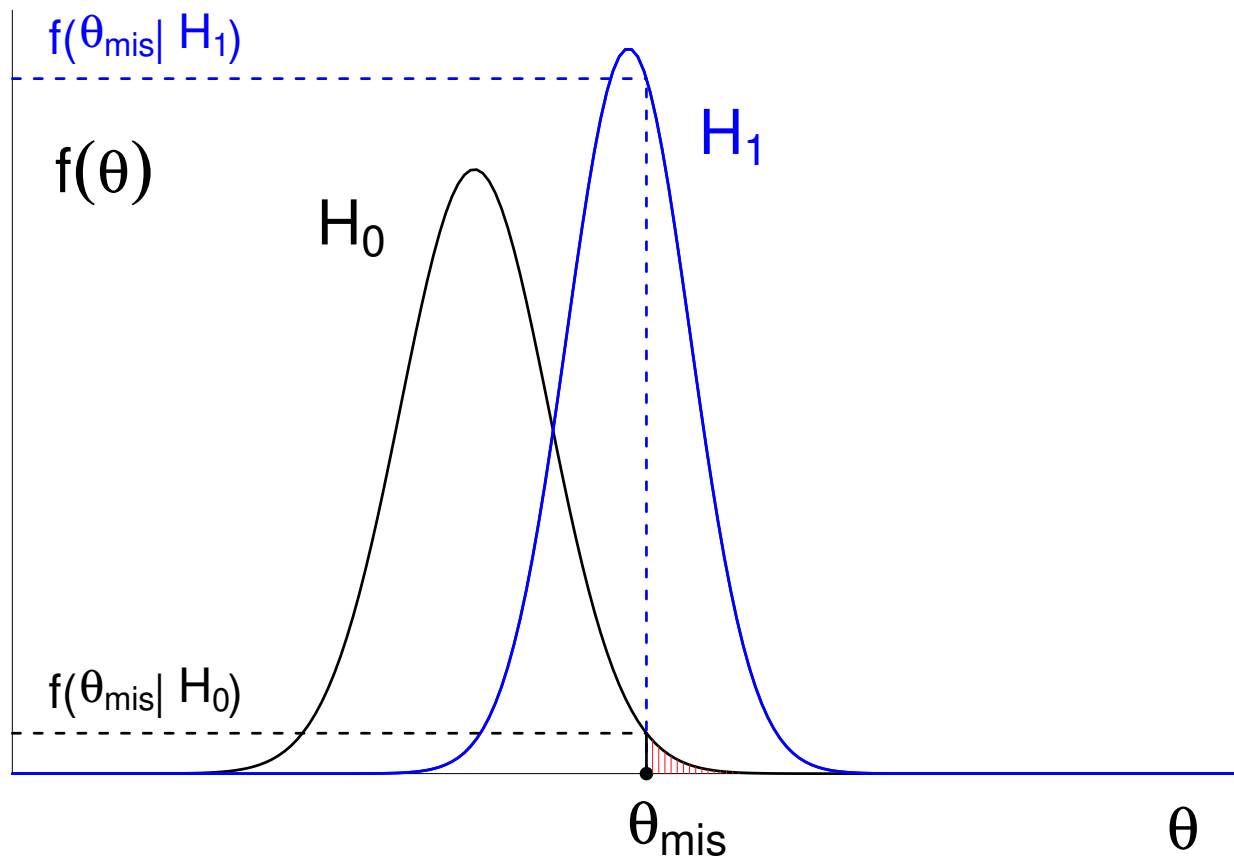
Perché mai l'osservazione di θ_{mis} dovrebbe diminuire la nostra fiducia in H_0 ?

Eppur funzionano!



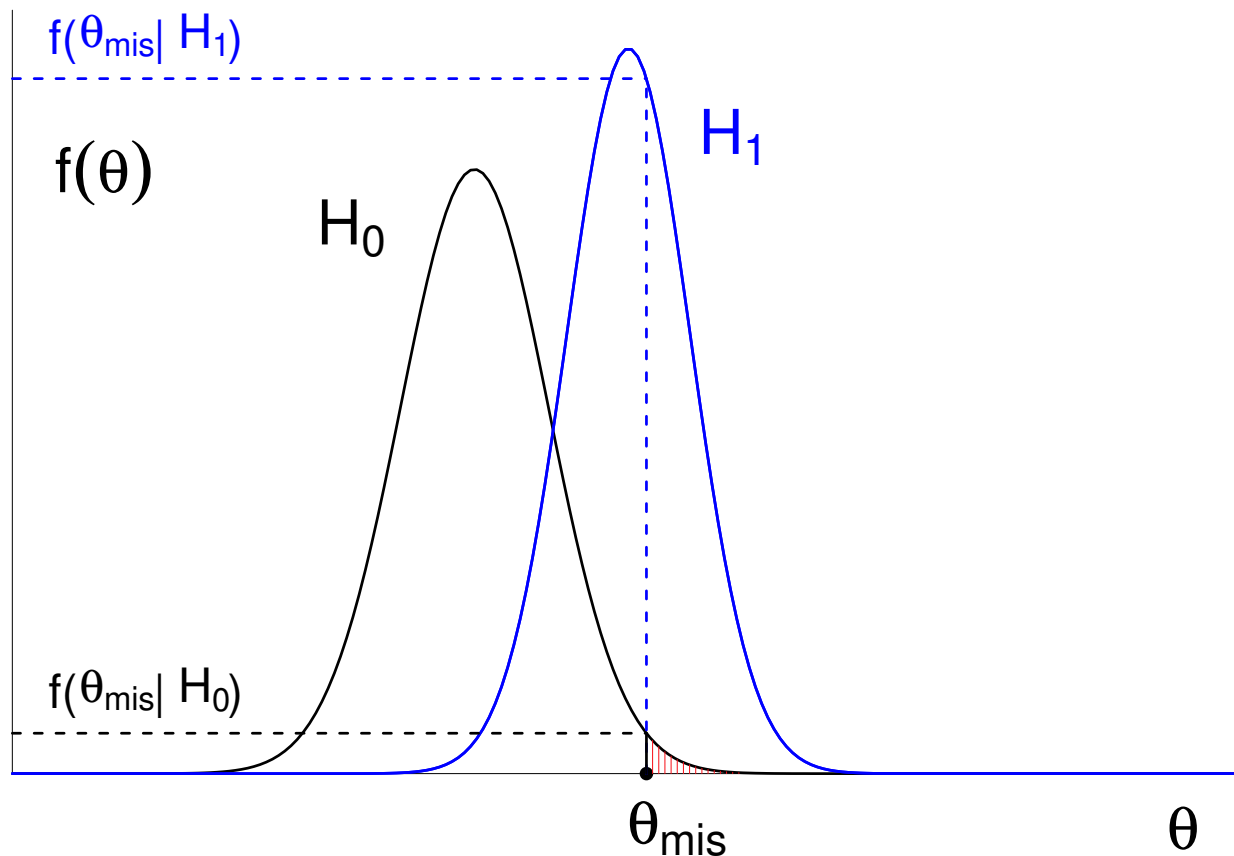
Perché *talvolta* possiamo dare una *possibilità concreta* ad una ipotesi alternativa H_1 , magari pur non essendo in grado di formularla esattamente.

Eppur funzionano!



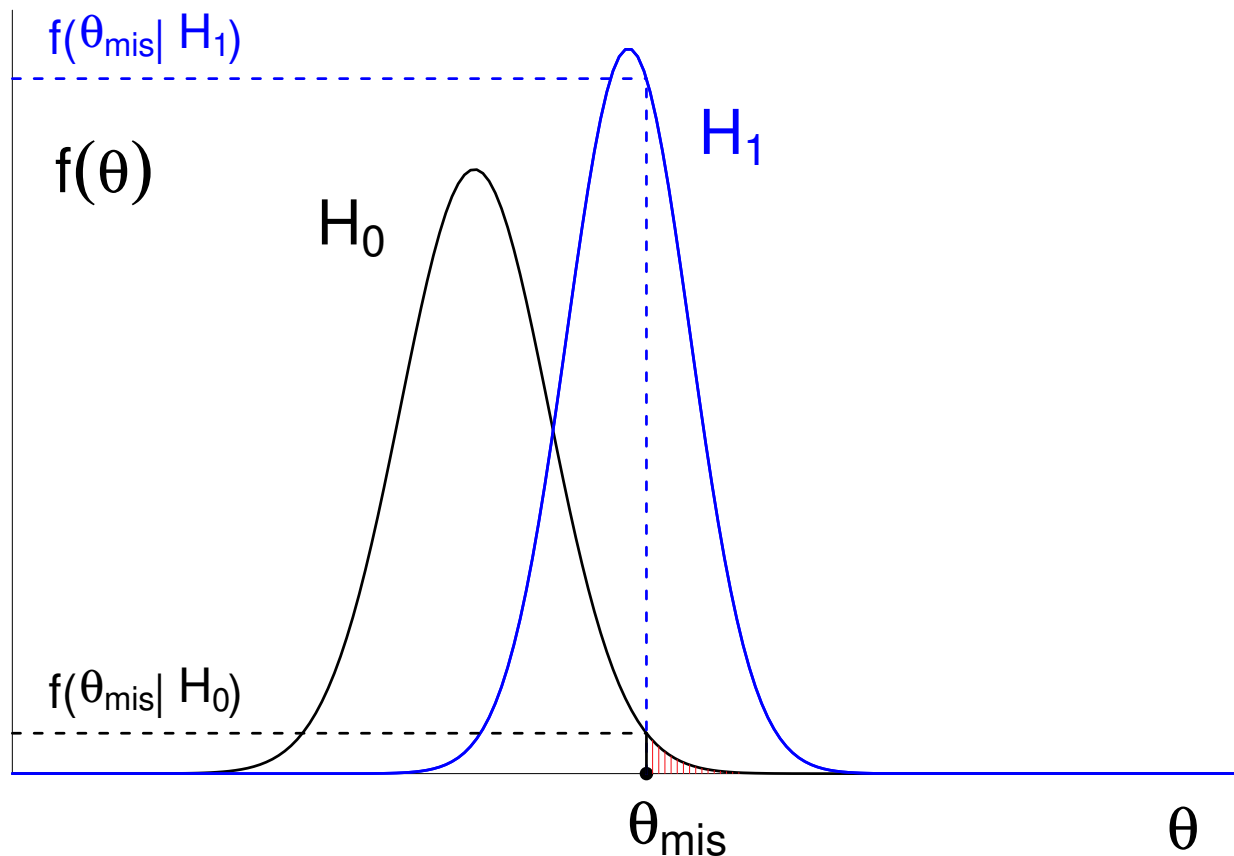
Ma quello che conta non è l'**area** a destra di θ_{mis} , bensì il rapporto fra $f(\theta_{mis} | H_1)$ e $f(\theta_{mis} | H_0)$!
 \Rightarrow a area 'piccola' corrisponde $f(\theta_{mis} | H_0)$ 'piccola'.

Eppur funzionano!



Ma se una ipotesi alternativa H_1 è impensabile non c'è 'piccolezza' dell'area che tenga!

Eppur funzionano!



Ma se una ipotesi alternativa H_1 è impensabile non c'è 'piccolezza' dell'area che tenga! *Fisico avvisato...!*

Annunci sensazionali e Fisica Sana

A questo punto è abbastanza chiaro come mai gli annunci di CDF e di Opera hanno lasciato i fisici increduli, con l'eccezione di qualche velleitario in cerca di protagonismo

Annunci sensazionali e Fisica Sana

A questo punto è abbastanza chiaro come mai gli annunci di CDF e di Opera hanno lasciato i fisici increduli, con l'eccezione di qualche velleitario in cerca di protagonismo (tanto sono finiti i tempi in cui si tagliavano le teste agli astronomi che sbagliavano la predizione delle eclissi...).

Annunci sensazionali e Fisica Sana

A questo punto è abbastanza chiaro come mai gli annunci di CDF e di Opera hanno lasciato i fisici increduli, con l'eccezione di qualche velleitario in cerca di protagonismo (tanto sono finiti i tempi in cui si tagliavano le teste agli astronomi che sbagliavano la predizione delle eclissi...).

Così come era abbastanza tranquillo che quello che si cominciava ad intravedere a dicembre dello scorso anno dovesse essere l'Higgs.

Annunci sensazionali e Fisica Sana

A questo punto è abbastanza chiaro come mai gli annunci di CDF e di Opera hanno lasciato i fisici increduli, con l'eccezione di qualche velleitario in cerca di protagonismo (tanto sono finiti i tempi in cui si tagliavano le teste agli astronomi che sbagliavano la predizione delle eclissi...).

Così come era abbastanza tranquillo che quello che si cominciava ad intravedere a dicembre dello scorso anno dovesse essere l'Higgs. (Anche perché in questo caso il 'botto' sarebbe stato non trovarlo!)

Annunci sensazionali e Fisica Sana

A questo punto è abbastanza chiaro come mai gli annunci di CDF e di Opera hanno lasciato i fisici increduli, con l'eccezione di qualche velleitario in cerca di protagonismo (tanto sono finiti i tempi in cui si tagliavano le teste agli astronomi che sbagliavano la predizione delle eclissi...).

Così come era abbastanza tranquillo che quello che si cominciava ad intravedere a dicembre dello scorso anno dovesse essere l'Higgs. (Anche perché in questo caso il 'botto' sarebbe stato non trovarlo!)

Non vi lasciate confondere da sigma e 'strane significance' che non dicono quanto veramente credere alla sensatezza del risultato.

Conclusioni

Philip Ball (Guardian, 23 dicembre 2011)

(<http://www.guardian.co.uk/commentisfree/2011/de>

“So D’Agostini recommends that, instead of heeding impressive-sounding statistics, we should ask what scientists actually believe. Better, we should find out if they had put money on it – and how much. After all, that is a tactic endorsed by none other than Kant.”

Conclusioni

Philip Ball (Guardian, 23 dicembre 2011)

(<http://www.guardian.co.uk/commentisfree/2011/de>

“So D’Agostini recommends that, instead of heeding impressive-sounding statistics, we should ask what scientists actually believe. Better, we should find out if they had put money on it – and how much. After all, that is a tactic endorsed by none other than Kant.”

Which is why I’m only being scientific when I say screw the sigmas: I’d place a tenner (but not a ton) on the Higgs, while offering to join Jim Al-Khalili in eating my shorts if neutrinos defy relativity.”

Conclusioni

Philip Ball (Guardian, 23 dicembre 2011)

(<http://www.guardian.co.uk/commentisfree/2011/de>

“So D’Agostini recommends that, instead of heeding impressive-sounding statistics, we should ask what scientists actually believe. Better, we should find out if they had put money on it – and how much. After all, that is a tactic endorsed by none other than Kant.”

Which is why I’m only being scientific when I say screw the sigmas: I’d place a tenner (but not a ton) on the Higgs, while offering to join Jim Al-Khalili in eating my shorts if neutrinos defy relativity.”

⇒ *Gli sono ‘andate bene’ entrambe le scommesse!*

Per approfondire

Vedi pagina dedicata alla presentazione:

<http://www.roma1.infn.it/%7Edagos/IdF2012/>

The End

FINE