Introduzione alle Esperienze di Laboratorio

La Fisica e La Statistica

Perche' e' necessario ricorrere ai metodi della statistica per interpretare correttamente i risultati di un esperimento ?

Rispondiamo eseguendo un esperimento (virtuale) :

Misura di π mediante un metodo "Monte Carlo"

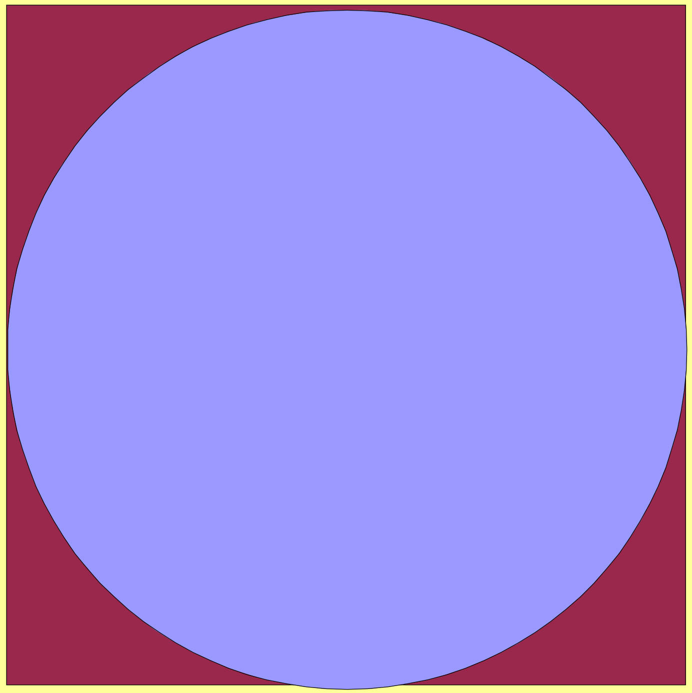
Il Principio

Consideriamo :

- un quadrato di lato $L=2R$
- il cerchio inscritto (raggio R)

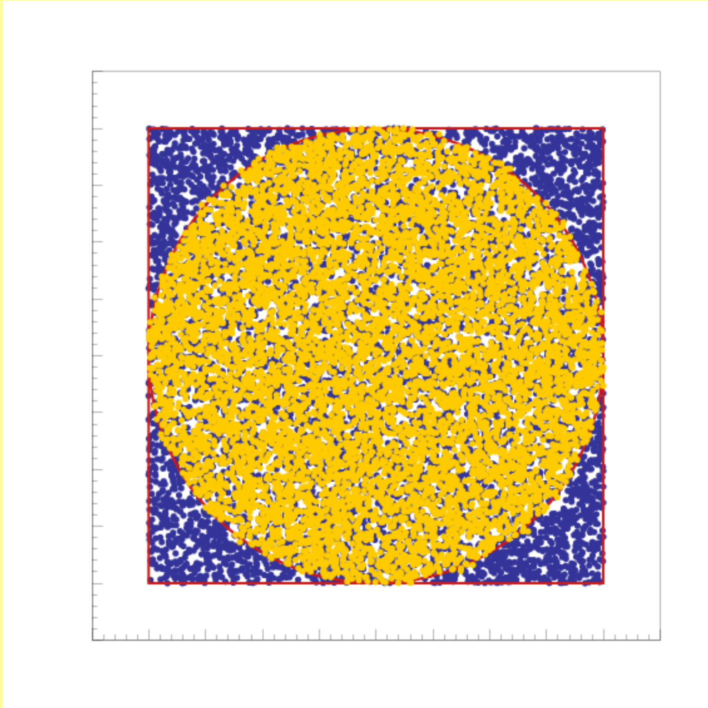
Il rapporto delle aree e':

$$r = \frac{area\,cerchio}{area\,quadrato} = \frac{\pi}{4}$$



Il Metodo

- Traccio le figure (sul suolo)
- Le ricopro lanciando a caso delle biglie (tutte uguali) su di esse
- Area quadrato $\propto$ **N** (biglie nel quadrato)
- Area cerchio $\propto$ **M** (biglie nel cerchio)



$$\frac{area\ cerchio}{area\ quadrato} = \frac{\pi}{4} = \frac{M}{N}$$

$$\pi = \frac{4M}{N}$$

La Misura

- La misura si ispira a quella compiuta dall'abate Buffon nel XVIII secolo: primo esempio di un metodo stocastico (Monte Carlo) per la misura di un parametro

- Nel seguito simuleremo una replica di questa misura utilizzando un semplice algoritmo :

  - **Itera $N = N+1$**
    - **genera un numero casuale x, $-1 < x < 1$**
    - **genera un numero casuale y, $-1 < y < 1$**
    - **se $(x^2+y^2) < 1$, itera $M = M+1$**
  - **Calcola $\pi = 4\ M\ /\ N$**

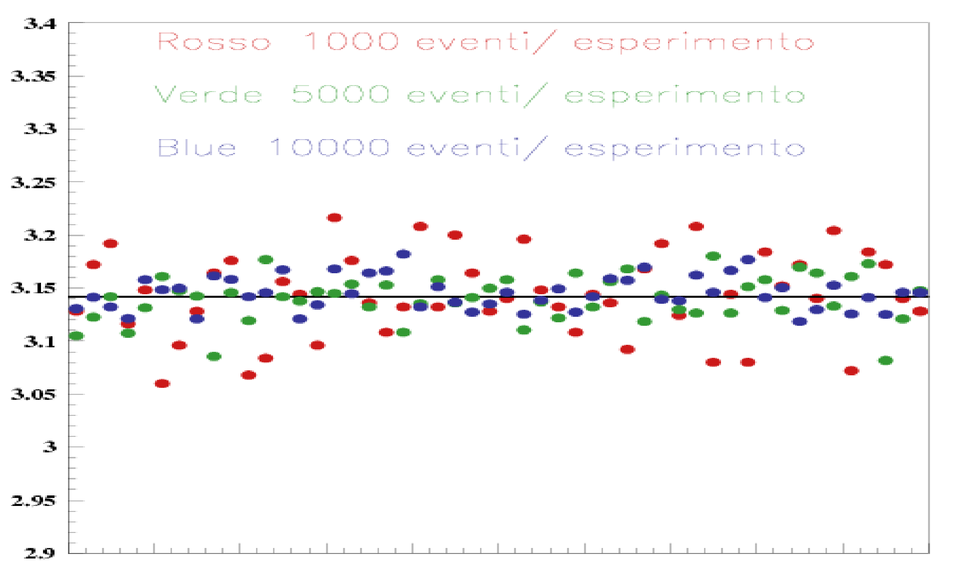
La simulazione

Risultati e Problemi

- La precisione migliora con il numero di lanci
- Quando "basta" ?
  - In un esperimento vero non conosciamo a priori il risultato !
- Qual e' l'errore della misura, inteso come la differenza tra il valore misurato e il valore "vero" (che non conosciamo) ?

Risposte

- Rispondiamo effettuando una successione (50) di prove (esperimenti) equivalenti e indipendenti, ognuna consistente del medesimo numero di lanci
- Rappresento i risultati mediante un grafico in cui in ascissa riporto l'ordine della prova e in ordinata il valore ottenuto (***ideogramma***)



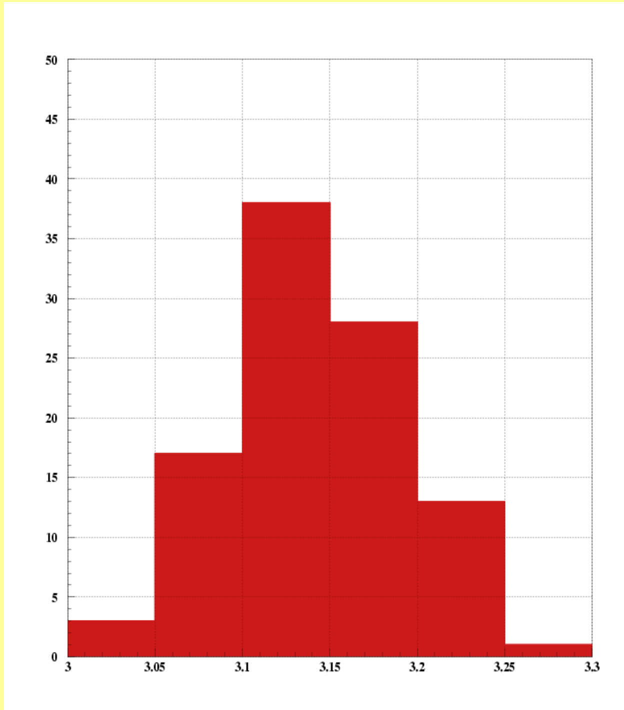
IDEOGRAMMA

Prime Conclusioni

- Aumentare il numero di lanci migliora le precisione
  - la dispersione tra diversi esperimenti diminusice se aumento il numero di lanci in ciascun esperimento
- Tuttavia tutti gli esperimenti sono, a priori, equivalenti
  - quale devo scegliere ?
  - esiste *piuttosto* una opportuna *combinazione* dei risultati dei vari (50) esperimenti che e' *intrisecamente* piu' precisa della singola prova ?
- Rispondo utilizzando a una diversa rappresentazione dei risultati: l'istogramma

L'istogramma

- Considero il dominio dei singoli risultati (e.g. 3. - 3.3)
- Lo suddivido in *N* intervalli consecutivi di uguale larghezza Δ (*e.g.* $\Delta = 0.05 \rightarrow N=6$)



- Sui ciascun intervallo a partire dal primo, itero questa procedura :
  - Conto quanti esperimenti, n_i, hanno dato un risultato compreso tra gli estremi dell' *i-esimo* intervallo
  - Traccio un rettangolo, di base uguale alla largezza Δ e altezza pari a n_i.

Un'altra simulazione

Eseguo cinque ripetizoni, ciascuna di 500 esperimenti

I. 1000 eventi / esperimento

II. 4000 eventi / esperimento

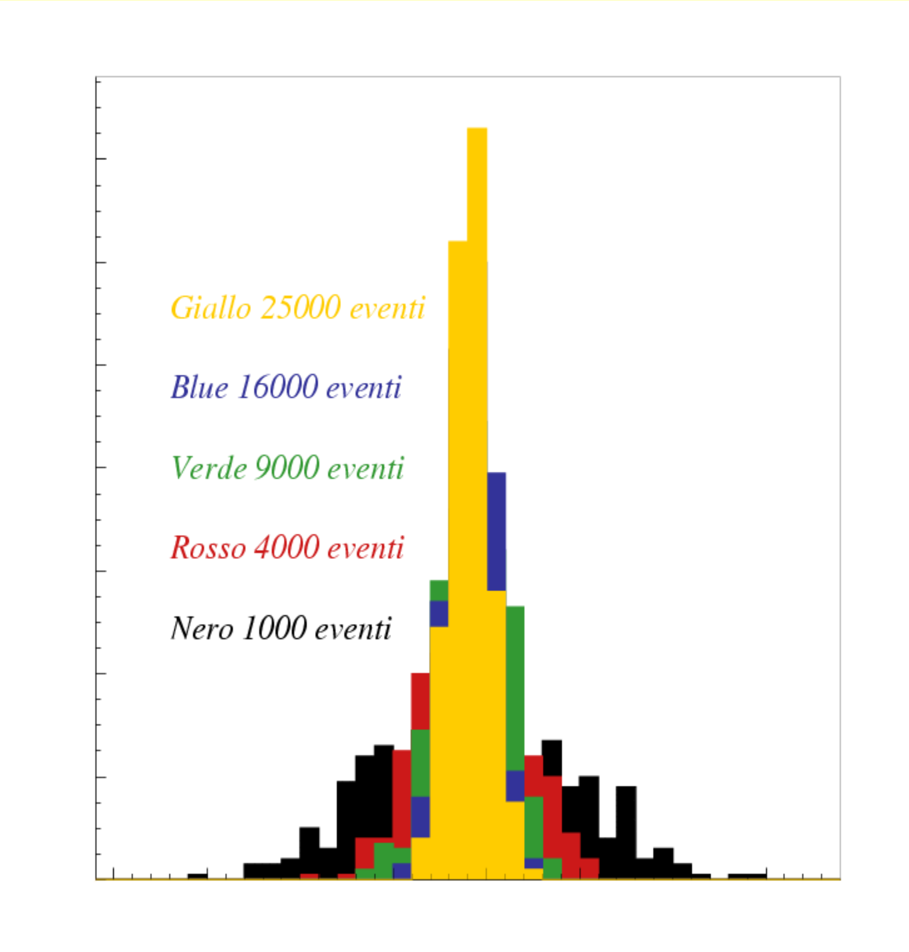
III. 9000

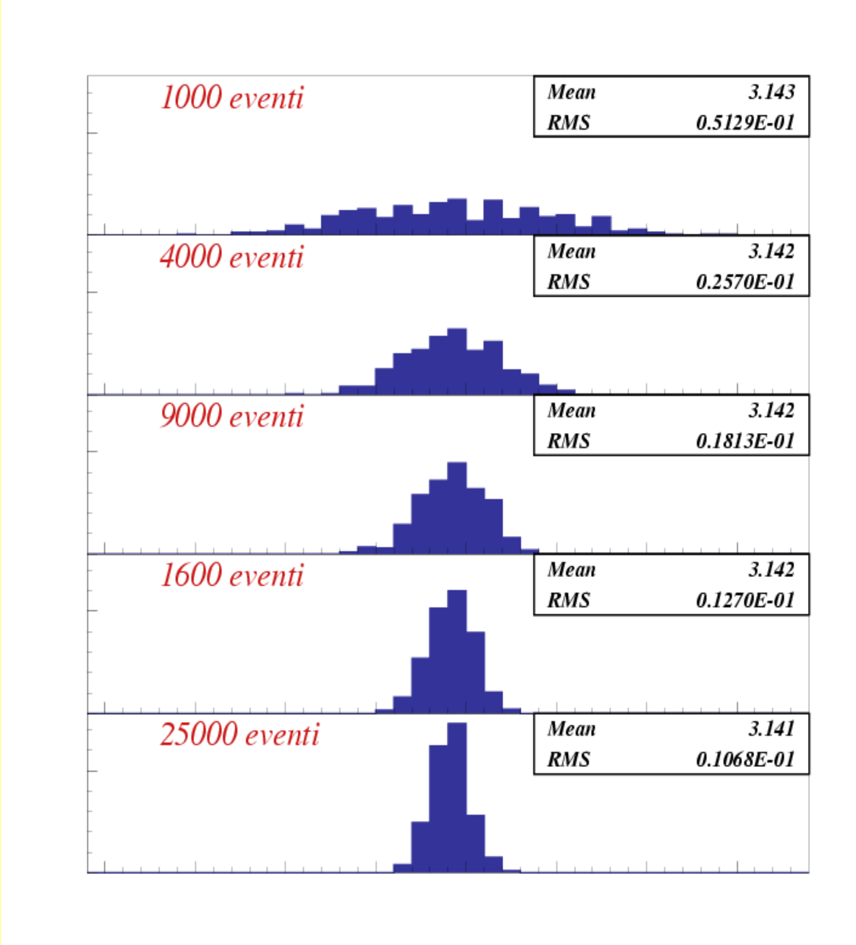
IV. 16000

V. 25000

e confronto, sovrapponendoli, gli istogrammi che ottengo per ciascuna

La Simulazione:
i risultati





Interpretazioni

- I risultati tendono a raggrupparsi attorno ad un valore (piu' probabile), che corrisponde alla loro *media aritmetica*

$$\bar{x} = \frac{1}{N} \Sigma_i x_i$$

- Aumentando il numero dei lanci, diminuisce la dispersione delle misure, che possiamo quantificare mediante la deviazione standard

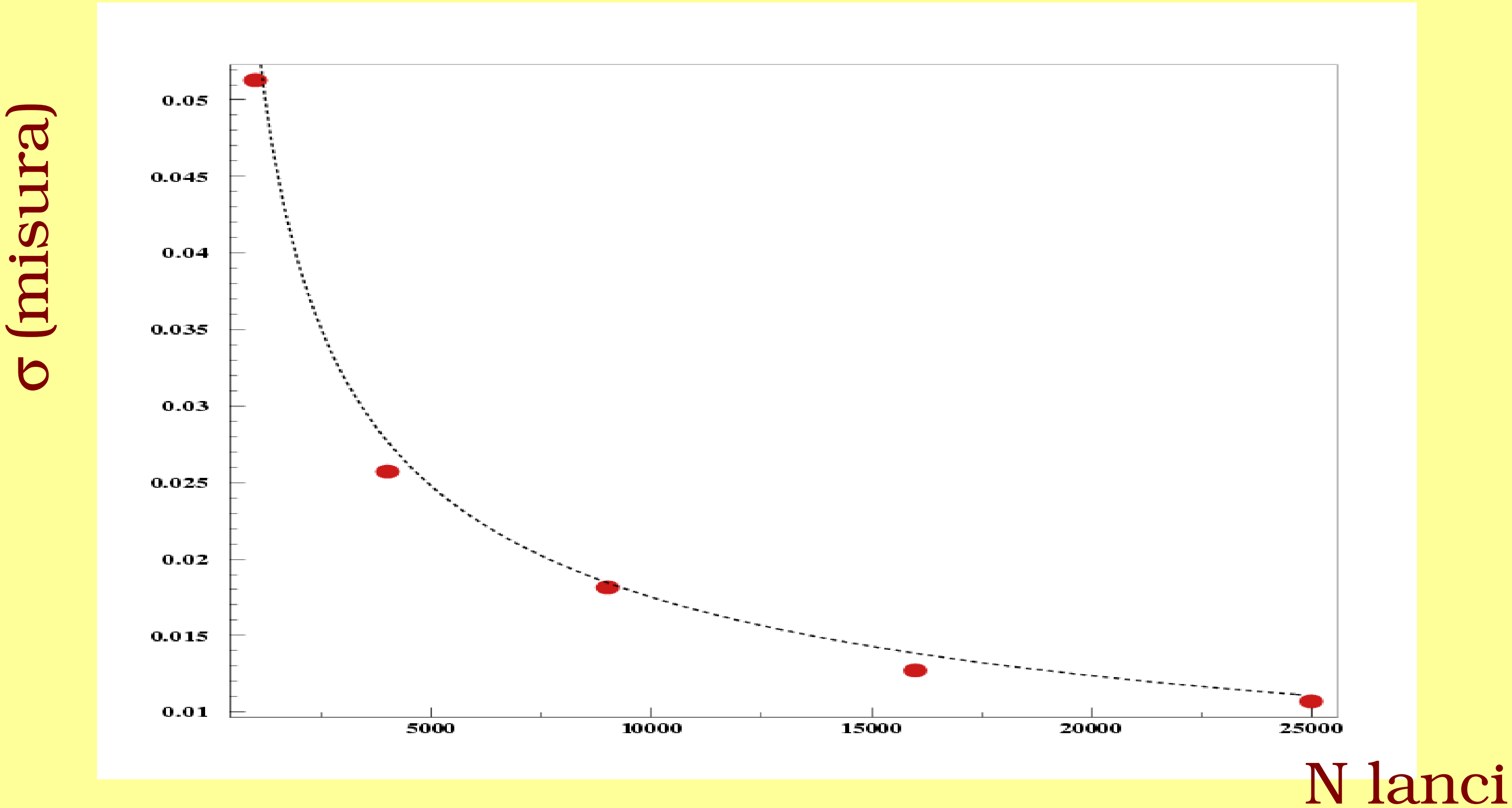
$$\sigma(x) = \sqrt{\frac{1}{N} \Sigma_i (x_i - \bar{x})^2} = \sqrt{\frac{1}{N} \Sigma_i z_i^2}$$

- Scarto iesimo $(x_i - \bar{x}) = z_i$

- Provare che: $\Sigma_i z_i = 0$

La deviazione standard

- E' una stima della precisione della misura
- Diminuisce in maniera proporzionale alla radice quadrata del numero di eventi (= lanci effettuati)



Un po' di teoria

Date N misure indipendenti di una variabile casuale:

- la media aritmetica e' la stima piu' precisa

$$\bar{x} = \frac{1}{N} \Sigma_i x_i$$

- l'errore della singola misura e' la deviazione standard

$$\sigma(x) = \sqrt{\frac{1}{N} \Sigma_i (x_i - \bar{x})^2}$$

- l'errore della media e' la deviazione standard, divisa per la radice del numero delle misure

$$\sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{N}}$$

Per dimezzare l'errore statistico e' necessario quadruplicare il numero delle misure !

Riassunto

- L'esito di una operazione di misura e' un processo stocastico, paragonabile all'estrazione di un numero casuale (il lotto, la roulette, i dadi)
- I risultati non si dispongono uniformemente, ma si *addensano* attorno ad un valore piu' probabile (la media aritmetica)
- Si puo' misurare la *dispersione* dei risultati (e quindi la *precisione* della misura) dalla deviazione standard (anche detta *scarto quadratico medio*): tanto piu' questa e' piccola , tanto migliore e' la misura.

Un po' di teoria (2)

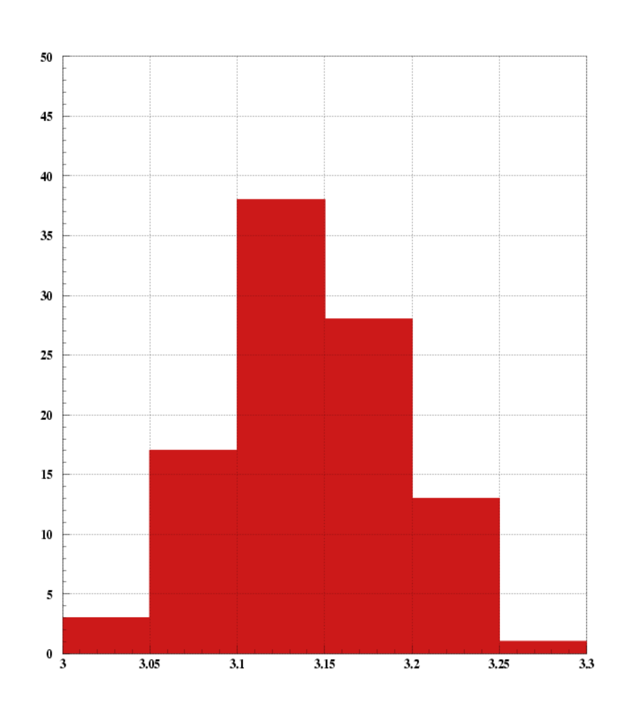
- I risultati delle misure di una variabile casuale non sono distribuiti a caso, ma secondo una funzione ben precisa (*la funzione Gaussiana*), la cui forma si puo' dedurre a partire da poche ipotesi elementari
- Qui non presento questa deduzione
- Schematizzo invece l'estrapolazione che porta dalla distribuzione (discreta) rappresentata tramite istogramma alla funzione (continua) Gaussiana

L'istogramma (ancora)

- compio N_e esperimenti
- definisco un intervallo (dominio), con limiti (X_{min} ; X_{max}) tali da contenere tutte i risultati
- suddivido il dominio in M intervalli di uguale dimensione

$$\Delta X = \frac{(X_{max} - X_{min})}{M}$$

- i-esimo intervallo (bin),
  - centro $x_i = X_{min} + i\,\Delta X$
  - n_i misure (quelle con risultati compresi entro gli estremi del bin)
- per definizione $\Sigma_i n_i = N_e$



$X_{min} = 3.0$, $X_{max} = 3.3$

$M = 6$, $\Delta X = 0.05$

Primo Passo: la Normalizzazione

Riscalo (normalizzo) l'istogramma:

- $n_i \rightarrow P_i = n_i/N_e = P(x_i)$
- $\Sigma_i n_i = N_e \Rightarrow \Sigma_i P_i = 1$

La rappresentazione in termini di P_i non dipende piu' dal numero di esperimenti effettuati

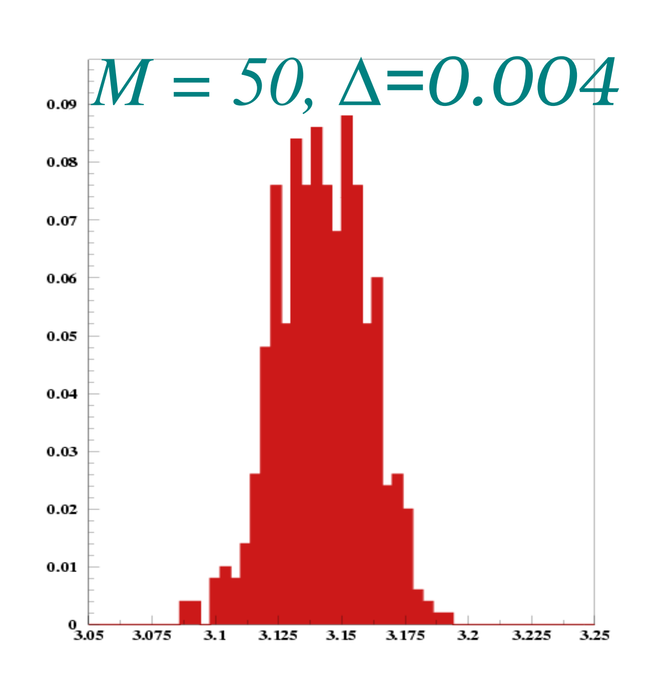
...verso il continuo...

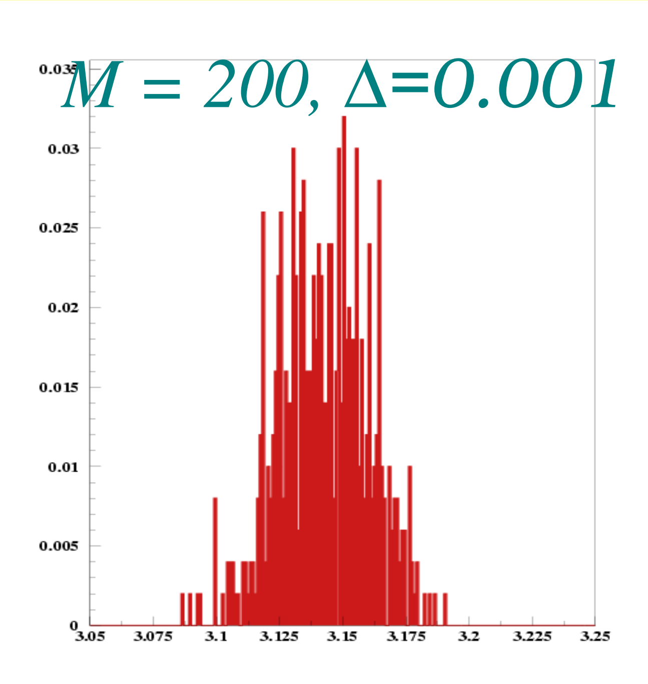
Se ho molti eventi posso aumentare il numero di bins:

$M \to M'$ ($>M$) ,

$\Delta X \to \Delta X'$ ($< \Delta X$)

Notare $M\Delta X = M'\Delta X' = X_{max} - X_{min}$





La Gaussiana

Con un numero infinito di eventi, nei limiti:

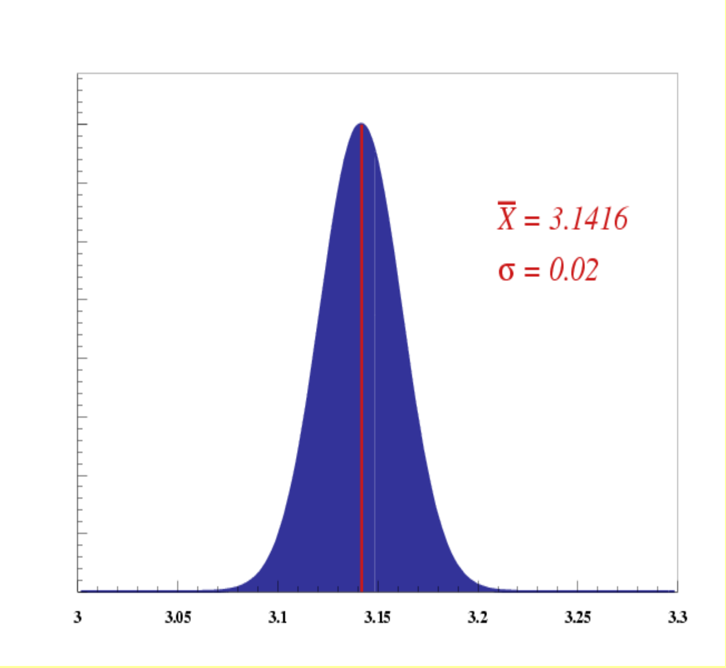
$M \rightarrow \infty,\ \Delta X \rightarrow dX,\ P(X_i) \rightarrow P(X)$

la distribuzione diventa continua, e viene chiamata *Densita' di Probabilita'* (*PDF*). La PDF per svariate misure indipendenti e ripetute della medesima grandezza e' rappresentata dalla funzione di Gauss:

$$P(x) = \frac{1}{\sqrt{2\pi}\,\mu} e^{\frac{-(x-x_0)^2}{2\mu^2}}$$

Proprieta' della Gaussiana

$$P(x) = \frac{1}{\sqrt{2\pi}\mu} e^{\frac{-(x-x_0)^2}{2\mu^2}}$$



- Simmetria

$$P(x-x_0) = P(x_0 - x)$$

- Normalizzazione:

$$\int_{-\infty}^{+\infty} P(x)\,dx = 1$$

- Interpretazione di μ

$$\int_{-\mu}^{+\mu} P(x)\,dx = 0.683$$

$$\int_{-3\mu}^{+3\mu} P(x)\,dx = 0.997$$

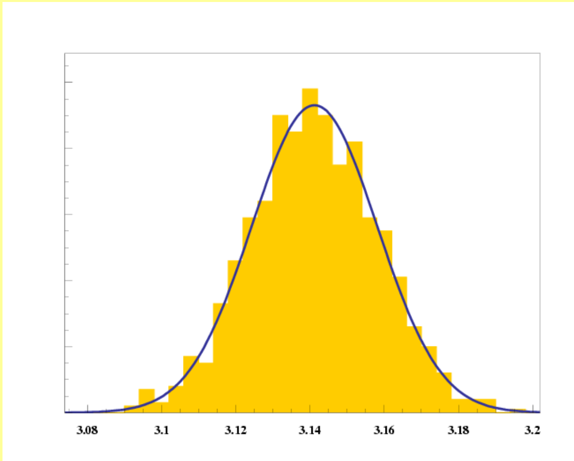
Tutti gli eventi compresi entro tre deviazioni standard dal valor medio

Il caso reale

- La Gaussiana e' il limite dell'istogramma osservato che si otterrebbe effettuando infiniti esperimenti, utilizzando infiniti bin ciascuno di dimensioni infinitesime
- In questo caso il valor medio x corrisponderebbe effettivamente al valore "vero"
- Possiamo pero' associare all'istogramma reale una funzione gaussiana e interpretare i risultati di conseguenza facendo le seguenti identificazioni

$$x_0 = \bar{x}$$

$$\mu = \sigma$$



Interpretazione

- Questa identificazione ci consente di interpretare gli esiti di una misura reale (con un numero finito di prove) dicendo che
  - la miglior stima del parametro e' la media aritmetica dei valori ottenuti
  - la probabilita' che il risultato differisca dal valor vero (incognito) meno dell'errore della media e' pari a ~ 68% (vedi pag 21)
  - la probabilita' che il risultato differisca dal valor medio piu' di tre volte l'errore e' pressoche' nulla (< 0.3%)

Applicazione

In laboratorio:

- effettuare le misure (200)
- produrre l'ideogramma
- calcolare la media aritmetica $\bar{x} = \frac{1}{N} \Sigma_i x_i$
- calcolare gli scarti dalla media $z_i = x_i - \bar{x}$
- calcolare la deviazione standard $\sigma = \left(\frac{1}{N} \Sigma_i z_i^2 \right)^{\frac{1}{2}}$
- calcolare l'errore della media $\sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{N}}$
- realizzare l'istogramma degli scarti:
  - intervallo: -4 σ ↔ 4 σ
  - larghezza del bin ~ ½ σ
- sovrapporre all'istogramma la corrispondente funzione gaussiana

Cautela

1. Normalizzo la funzione all'istogramma, imponendo la medesima area:

   A (istogramma) = $\Sigma_i (P_i \Delta x_i) = (\Sigma_i P_i) \Delta x = \Delta x$

   A (Gaussiana) = 1

   Funzione normalizzata $G(x) = P(x) \ \Delta x$

2. In ogni bin, calcolo il valore della funzione al centro del bin, $G(x_i)$, e lo riporto sul medesimo grafico su cui ho rappresentato l'istogramma

Osservazioni Finali

- L'errore statistico di una misura decresce con l'inverso del quadrato delle prove effettuate. Posso raggiungere una precisione infinita, oppure esistono dei fattori limitanti ?

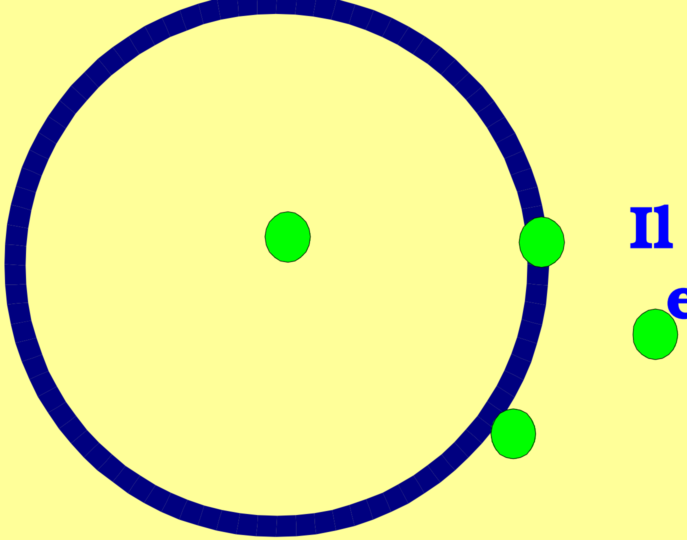
...ma soprattutto ...

Gli errori sistematici

Dovuti ad imperfezioni dell'apparato sperimentale usato per la misura o ad approssimazioni nei modelli teorici che la descrivono

Esempio (per p) :

- quanto e' quadrato il mio quadrato ?
- quanto e' tonda la circonferenza ?
- cose vuol dire che la biglia sta dentro o fuori ?



Il calcolo degli errori sistematici e' forse la fase piu' delicata per un esperimento moderno