STATISTICA DI CONTEGGIO

Serve per controllare il buon funzionamento della strumentazione

Si fanno tante misure in condizioni identiche e si verifica che le fluttuazioni siano come previsto

Serve per interpretare una singola misura sperimentale

Se ne stima accuratezza ed errore

Concetto di:

Media sperimentale $\overline{x}_e$ e media vera $\overline{x}$

Coincidono solo per N $\rightarrow \infty$

$$\overline{x}_e = \sum_{i=1}^{N} x_i / N$$

Funzione di distribuzione F(x)

$$\sum F(x) = 1 \qquad \sum x \cdot F(x) = \overline{x}_e$$

Residuo d_i e deviazione ε_i

$$d_i \equiv x_i - \overline{x}_e \qquad \varepsilon_i \equiv x_i - \overline{x} \qquad \sum_{i=1}^{N} d_i = 0$$

Varianza sperimentale s^2

$$s^2 \equiv \overline{\varepsilon^2} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \overline{x}_e\right)^2$$

$$s^2 = \sum_{x=0}^{\infty}\left(x - \overline{x}\right)^2 F(x) = \overline{x^2} - \left(\overline{x}\right)^2$$

MODELLI STATISTICI

Funzioni di distribuzione previste a priori in base al tipo di misura

Misura: contare il numero di successi per un dato numero n di tentativi binari (V/F)

La probabilità di successo p deve essere la stessa per tutti i tentativi

Esempio:
- argomento del test = osservazione di un nucleo radioattivo per un tempo t
- successo = il nucleo decade mentre e' sotto osservazione
- n. di tentativi = n. di nuclei sotto osservazione

Varianza prevista e deviazione standard

$$\sigma^2 \equiv \sum_{x=0}^{n} (x-\overline{x})^2 P(x) \qquad \sigma \equiv \sqrt{\sigma^2}$$

Binomiale

n e x devono essere degli interi. Occorre conoscere n e p

$$P(x) = \frac{n!}{(n-x)!\,x!} p^x (1-p)^{n-x}$$

$$\sum_{x=0}^{n} P(x) = 1 \qquad \overline{x} = \sum_{x=0}^{n} x \cdot P(x) = pn \qquad \sigma^2 \equiv \sum_{x=0}^{n} (x-\overline{x})^2 \cdot P(x) = pn(1-p)$$

Poisson

p deve essere oltre che costante anche piccola ($p \ll 1$). Si deve conoscere solo pn

$$P(x) = \frac{(\overline{x})^x e^{-\overline{x}}}{x!}$$

NON SIMMETRICA $\qquad x_{mp} \neq \overline{x}$

$$\sum_{x=0}^{n} P(x) = 1 \qquad \overline{x} = \sum_{x=0}^{n} x \cdot P(x) = pn \qquad \sigma^2 \equiv \sum_{x=0}^{n} (x-\overline{x})^2 \cdot P(x) = pn = \overline{x}$$

Gaussiana o Normale

$p \ll 1$ e $\overline{x} > 20$. Si deve conoscere solo $\overline{x}$

$$P(x) = \frac{1}{\sqrt{2\pi\overline{x}}} \exp\left(-\frac{(x-\overline{x})^2}{2\overline{x}} \right)$$

SIMMETRICA $\qquad x_{mp} = \overline{x}$

$$\sum_{x=0}^{n} P(x) = 1 \qquad \overline{x} = \sum_{x=0}^{n} x \cdot P(x) = pn \qquad \sigma^2 = \overline{x}$$

FORMA UNIVERSALE DI GAUSSIANA

Siccome P(x) dipende da x solo attraverso $|x - \bar{x}| \equiv \varepsilon$

possiamo riscrivere P(x) come:

$$G(\varepsilon) = \sqrt{\frac{2}{\pi \bar{x}}} \exp\left(-\frac{\varepsilon^2}{2\bar{x}}\right)$$

Funzione continua, solo positiva, con massimo in 0

Definiamo la "deviazione normalizzata" come: $t = \frac{\varepsilon}{\sigma} = \frac{\varepsilon}{\sqrt{\bar{x}}}$

Allora $$G(t) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{t^2}{2}\right)$$

$$\int_0^{t_0} G(t)dt \equiv f(t_0)$$

dà la probabilità che una misura singola con distribuzione gaussiana abbia una deviazione normalizzata t dal valor medio MINORE di t_0

Esempio

Siccome $$\int_0^1 G(t)dt = 0.683$$

possiamo dire che il 68% di tutte le misure con distribuzione gaussiana hanno una deviazione dal valor medio INFERIORE ad 1σ

$$f(t_0) = 1 - \alpha \qquad \varepsilon = \delta$$

I valori di f(t_0) si trovano tabulati

α	δ
0.3173	1σ
4.55×10^{-2}	2σ
2.7×10^{-3}	3σ
6.3×10^{-5}	4σ
5.7×10^{-7}	5σ
2.0×10^{-9}	6σ

α	δ
0.2	1.28σ
0.1	1.64σ
0.05	1.96σ
0.01	2.58σ
0.001	3.29σ
10^{-4}	3.89σ

STATISTICA DI POISSON MODIFICATA

Esempio CORRETTO di applicazione della statistica di Poisson $(\bar{x} = pn = \text{costante})$

misura di un decadimento radioattivo RARO per un tempo T << τ

n n. di nuclei in osservazione >> n. decad. osservati COSTANTE

$p = 1 - e^{-\lambda t}$ fissata ed uguale per tutti gli n COSTANTE

$p << 1$

Se il n. di decadimenti osservati durante la misura è comunque > 20
si può utilizzare la distribuzione gaussiana

Esempio SCORRETTO di applicazione della statistica di Poisson semplice

misura di un decadimento radioattivo per un tempo T CONFRONTABILE con τ

$p << 1$ MA $p = 1 - e^{-\lambda t}$ cambia durante la misura (dipende da t)

Se

$\mu(t) = \mu_0 e^{-\lambda t}$ rate di decadimento

Allora:

$$P(x) = \frac{(\mu_0 t_s)^x}{x!} e^{-\mu_0 t_s} \frac{1}{\mathrm{T}} \int_0^{\mathrm{T}} \exp\left[-x\lambda t + \mu_0 t_s \left(1 - e^{-\lambda t}\right)\right] dt$$

dove t_s è (tipicamente) il passo di campionamento della misura (T = nt_s), durante il quale si può considerare costante il valore di μ (t_s << τ)

CONTROLLO DELLA STRUMENTAZIONE

Occorre ripetere decine di volte una IDENTICA MISURA, verificando che tutte le condizioni sperimentali restino COSTANTI

Poi:

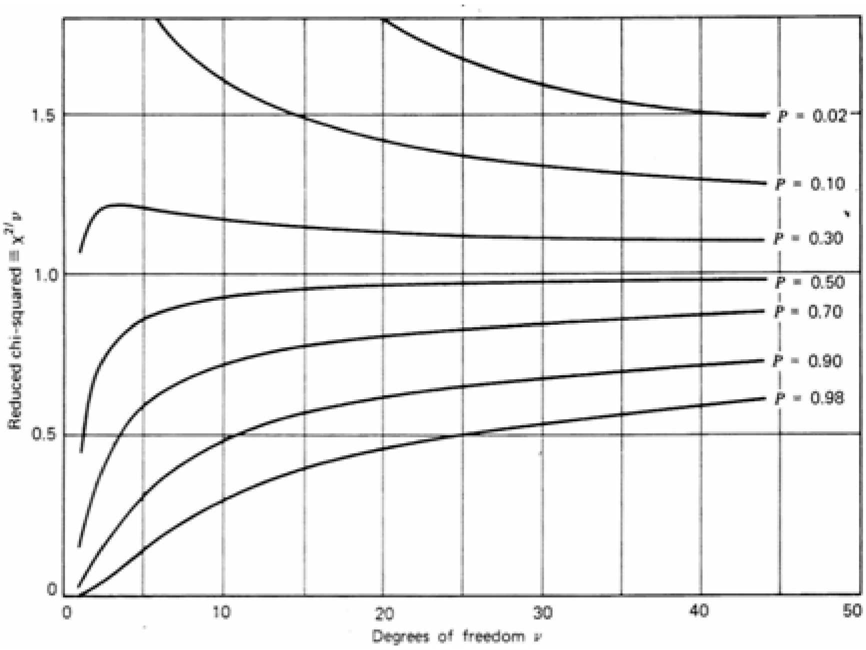
- estrarre $\bar{x}_e \quad s^2$
- calcolare $F(x)$
- porre $\bar{x}_e = \bar{x}$
- calcolare $P(x)$
- confrontare s^2 con σ^2 (devono essere simili)
- confrontare graficamente $F(x)$ con $P(x)$ (analisi qualitativa)
- calcolare $\chi^2 \equiv \frac{1}{\bar{x}_e}\sum_i \left(x_i - \bar{x}_e\right)^2 = \frac{(N-1)s^2}{\bar{x}_e}$

Il χ^2 normalizzato $\left(\equiv \frac{\chi^2}{\mathrm{d.o.f}} = \frac{\chi^2}{N\text{-}1} = \frac{s^2}{\bar{x}_e}\right)$

deve essere circa 1 $\left(= \frac{\sigma^2}{\bar{x}}\right)$, e comunque compreso tra 0.5 e 1.5

Nelle tabelle si trova la probabilità corrispondente ad una certa coppia (d.o.f. , χ^2) di avere un χ^2 calcolato dalla poissoniana MAGGIORE del χ^2 sperimentale.

$$\chi^2_{\mathrm{calc}} \equiv \frac{(N-1)\sigma^2}{\bar{x}} = N-1$$



Tale probabilità deve essere 0.5

Se e' molto minore di 0.5, vuol dire che le fluttuazioni sperimentali sono eccessivamente elevate

Se e' molto maggiore di 0.5, vuol dire che le fluttuazioni sperimentali sono stranamente piccole

STATISTICA CON UNA SINGOLA MISURA

★ Vogliamo poter estrarre un s^2 SENSATO da una sola misura che ci consenta di sapere cosa ci dobbiamo aspettare se ripetiamo la misura

Singola misura: $\hat{x}$ devo forzare $\bar{x} = \hat{x}$

Statistica di Poisson: $\sigma^2 \equiv \bar{x}$ devo forzare $s^2 = \sigma^2 = \hat{x}$
(o di Gauss)

Da una sola misura dico che mi aspetto una distribuzione con $\bar{x} = \hat{x}$ $\sigma = \sqrt{\hat{x}}$

Nel caso GAUSSIANA ($\bar{x}$ grande)

$$\bar{x} \in \left[\hat{x} \pm \sqrt{\hat{x}}\right] \qquad \text{68\% C. L.}$$

DEVIAZIONE STANDARD PERCENTUALE $\sigma_{\%} \equiv \dfrac{\sigma}{x} = \dfrac{\sqrt{x}}{x} = \dfrac{1}{\sqrt{x}}$

Es: $x = 100 \quad \sigma = 10 \quad \sigma_{\%} = 10\%$

Per avere $\sigma_{\%} = 1\%$ devo avere $x = 1000($

L'errore percentuale migliora con la √ dell'aumento della statistica

ATTENZIONE

Se in un grafico si riportano i punti sperimentali con la barra d'errore tipica ± 1 σ
il fit dei punti deve passare AL MEGLIO per ≈ 2/3 delle barre di errore.
Se passa su un numero molto maggiore, vuol dire che si è SOVRASTIMATO l'errore

PROPAGAZIONE DEGLI ERRORI

ATTENZIONE

$\sigma = \sqrt{x}$ sse x rappresenta un CONTEGGIO DI SUCCESSI

Non è invece esatto se x è:

- il tasso di conteggi
- la somma o la differenza di conteggi
- le medie di conteggi indipendenti
- ogni altra quantità DERIVATA

In tutti questi altri casi occorre calcolare σ secondo la propagazione degli errori

ALCUNE FORMULE

$$\sigma_u^2 = \left(\frac{\partial u}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial u}{\partial y}\right)^2 \sigma_y^2 + \left(\frac{\partial u}{\partial z}\right)^2 \sigma_z^2 + \ldots$$

$$\sigma_{x\pm y} = \sqrt{\sigma_x^2 + \sigma_y^2} \qquad \sigma_{Ax} = A\sigma_x \qquad \sigma_{x/B} = \frac{\sigma_x}{B}$$

se $u = x \cdot y$ o $u = x/y$ $\qquad \left(\frac{\sigma_u}{u}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2$

$$\sigma_{\bar{x}} = \sqrt{\frac{\bar{x}}{N}}$$

Se voglio il best value di una combinazione di misure indipendenti con errori diversi

$\langle x \rangle \equiv \dfrac{\sum_{i=1}^{N} a_i x_i}{\sum_{i=1}^{N} a_i}$ con $a_i = \dfrac{1}{\sigma_{x_i}^2}\left(\sum_{j=1}^{N} \dfrac{1}{\sigma_{x_j}^2}\right)^{-1}$ e $\dfrac{1}{\sigma_{\langle x \rangle}^2} = \sum_{i=1}^{N} \dfrac{1}{\sigma_{x_i}^2}$

OTTIMIZZAZIONE DEI CONTEGGI

Ipotesi

S tasso di conteggi di sorgente

B tasso di conteggi di fondo

$t_{1/2} >> t_{mis}$ B stabile nel tempo

Allora $S = \frac{N_{S+B}}{T_{S+B}} - \frac{N_B}{T_B}$

$$\sigma_s^2 = \left[\left(\frac{\sigma_{N_{S+B}}}{T_{S+B}}\right)^2 + \left(\frac{\sigma_{N_B}}{T_B}\right)^2\right] = \frac{N_{S+B}}{T_{S+B}^2} + \frac{N_B}{T_B^2} = \frac{S+B}{T_{S+B}} + \frac{B}{T_B} \qquad (1)$$

Ottimizzazione del tempo complessivo $T = T_{S+B} + T_B$ a disposizione

Differenziando la (1): $2\sigma_s \mathrm{d}\sigma_s = -\frac{S+B}{T_{S+B}^2}\mathrm{d}T_{S+B} - \frac{B}{T_B^2}\mathrm{d}T_B$

Affinche' σ_S sia minima pongo $d\sigma_S = 0$ e ottengo:

$\frac{S+B}{T_{S+B}^2}\mathrm{d}T_{S+B} = -\frac{B}{T_B^2}\mathrm{d}T_B$ ma $\mathrm{d}T_{S+B} + \mathrm{d}T_B = 0$ perché T è richiesto costante

$\frac{T_{S+B}^2}{T_B^2} = -\frac{\mathrm{d}T_{S+B}}{\mathrm{d}T_B}\frac{S+B}{B}$ quindi $-\frac{\mathrm{d}T_{S+B}}{\mathrm{d}T_B} = 1$

Pertanto:

$\left.\frac{T_{S+B}}{T_B}\right|_{opt} = \sqrt{\frac{S+B}{B}}$ Il rivelatore deve essere ottimizzato in modo tale da massimizzare T^{-1} con T tale da determinare il tasso S con una data accuratezza statistica

Se $S << B$, si ha l'ottimo per $T_{S+B} = T_B$

Caso $S >> B$ $\frac{1}{T} \approx \varepsilon_{\%}^2 S$ Occorre massimizzare S (efficienza)

Caso $S << B$ $\frac{1}{T} \approx \varepsilon_{\%}^2 \frac{S^2}{4B}$ Occorre massimizzare $S^2/4B$ (configurazione)

LIMITI DI RIVELABILITA'

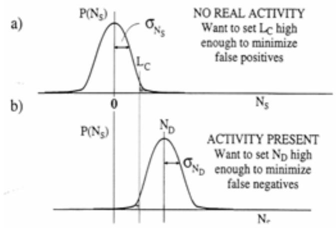
Definizione di "Minimum Detectable Amount" (MDA) Caso con $S << B$

$N_S = N_T - N_B$ $\quad N_T\ , N_B > 30$ $\quad$ Gaussiana con $\sigma^2_{NS} = \sigma^2_{NT} + \sigma^2_{NB}$

Senza fluttuazioni statistiche ne' variazioni strumentali

$N_S > 0$ c'e' contaminazione $\quad N_S = 0$ non c'e' contaminazione

Occorre fissare un *Livello critico* L_c che ci permetta di scartare con buona efficienza i falsi positivi, limitando i rischi di falsi negativi



Caso 1. Non c'e' reale contaminazione

$\overline{N}_T = \overline{N}_B\ ,\ \overline{N}_S = 0 \qquad \sigma_{N_T} = \sigma_{N_B} = \sqrt{N_B}\ ,\ \sigma_{N_S} = \sqrt{2N_B}$ (fluttuazioni solo statistiche)

Allora scegliendo $L_C = 1.645\ \sigma_{N_S} = 2.326\ \sigma_{N_B}$ si limitano i falsi positivi a non piu' del 5%

Caso 2. C'e' una reale contaminazione (conto semplificato)

Dato L_C come sopra, se fosse $\overline{N}_S = L_C$ avremmo il 50% di falsi negativi

Definiamo la MDA per il ns rivelatore quella attivita' che dara' luogo ad un massimo del 5% di falsi negativi

Il minimo n. di conteggi N_D da una sorgente in studio che garantisca un tasso di falsi negativi non superiore al 5% in un rivelatore che sta operando secondo un livello critico L_C tale da assicurare un tasso di falsi positivi non superiore al 5%, puo' essere cosi' calcolato:

Se le fluttuazioni sono solo statistiche e $N_D << N_B$, in prima approssimazione si ha:

$$\sigma_{N_D} = \sqrt{2N_B + N_D} \approx \sqrt{2N_B} \qquad N_D = L_C + 1.645\ \sigma_{N_D} \approx 4.653\ \sigma_{N_B} \left(+\ 2.706\right)$$

calcolo meno approssimato

Allora $\quad MDA = \dfrac{N_D}{f\varepsilon T}$

f radiation yield

ε efficienza assoluta

T tempo di misura

DISTRIBUZIONE DEGLI INTERVALLI TEMPORALI

Ipotesi Processo random con probabilità di avvenire COSTANTE NEL TEMPO $dp = r\,dt$

Il sistema non ha memoria

Esempio Segnali di un rivelatore di rad. posto ad adeguata distanza da una sorgente radioatt. a vita lunga

I decadimenti radioattivi sono un processo poissoniano

Intervalli tra eventi successivi (primo evento a t=0)

$$N = rt$$

numero medio di eventi in un tempo t

$$P(n) = \frac{(rt)^n e^{-rt}}{n!}$$

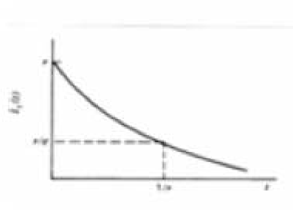
distribuzione poissoniana di n eventi

$$P(0) = \frac{(rt)^0 e^{-rt}}{0!} = e^{-rt}$$

distribuzione poissoniana di 0 eventi

$$I_1(t)dt = P(0) \times r\,dt = re^{-rt}dt$$

prob. di avere il primo nuovo evento tra t e t+dt



$$\bar{t} = \frac{\int_0^\infty t I_1(t)dt}{\int_0^\infty I_1(t)dt} = \frac{1}{r}$$

Vale anche prendendo l'istante iniziale in modo casuale (Time-to-Next_Event)

Intervalli di tempo per N eventi successivi

E' utile nel caso di utilizzo di uno "scaler" o buffer durante l'acquisizione

$$I_N(t)dt = P(N-1) \times r\,dt = \frac{(rt)^{N-1} e^{-rt}}{(N-1)!} r\,dt$$

$$\bar{t} = \frac{\int_0^\infty t I_N(t)dt}{\int_0^\infty I_N(t)dt} = \frac{N}{r} \qquad t_{mp} = \frac{N-1}{r}$$