

UNIVERSITA' DEGLI STUDI DI FIRENZE
Facoltà di Scienze M.F.N.
Corso di Laurea in Fisica

Prof. Roberto Falciani

Prof. Andrea Stefanini

**Appunti aggiuntivi al corso di
ESPERIMENTAZIONI I**

Analisi statistica degli errori accidentali

Anno Accademico 2002-2003

Versione del 17 Gennaio 2003

1 Trattazione dei dati

Supponiamo di voler misurare una grandezza x e, dopo aver cercato di identificare le varie sorgenti di errore sistematico, di aver ridotto il loro contributo all'incertezza di misura a livelli trascurabili rispetto al contributo dovuto alle sorgenti di errore casuale.

Le rimanenti sorgenti di incertezza saranno quindi esclusivamente di tipo accidentale e dovremmo essere capaci di rilevarne il contributo ripetendo la misura parecchie volte.

Il nostro primo problema è quello di escogitare metodi adeguati per registrare un gran numero di valori misurati e per metterne in evidenza le caratteristiche. Supponiamo ad es. di aver realizzato 10 misure della lunghezza di una sbarretta metallica, utilizzando un calibro decimale, e di aver ottenuto i seguenti valori x_i (in cm):

11.35, 11.33, 11.34, 11.35, 11.36, 11.34, 11.34, 11.36, 11.37, 11.36

Per meglio trattare i nostri dati potremmo riordinarli in ordine crescente ed associare ad ogni valore misurato il numero delle volte in cui lo si è ottenuto. Indichiamo con x_k i diversi valori misurati e con n_k il numero delle volte che il valore x_k è stato ottenuto. Potremo quindi costruire la tabella sotto riportata.

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
x_k (cm)	11.33	11.34	11.35	11.36	11.37
n_k	1	3	2	3	1

Sembra ragionevole determinare la miglior stima della grandezza misurata tramite la media \bar{x} dei 10 valori misurati, data da

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} = \frac{\sum_{i=1}^{10} x_i}{10} = 11.35 \text{ cm} \quad (1)$$

La media può anche essere espressa nella forma

$$\bar{x} = \frac{\sum_{k=1}^5 x_k n_k}{N}$$

dove la sommatoria è estesa ai diversi valori di x misurati.

La sommatoria $\sum_{k=1}^5 x_k n_k$ viene anche detta **somma pesata** in quanto ciascun valore x_k viene pesato tramite il numero di volte di occorrenza n_k . Notiamo che la frazione

$$F_k = \frac{n_k}{N}$$

rappresenta la “frequenza” con cui si è ottenuto il risultato x_k nelle nostre N misure. Potremo ora scrivere

$$\bar{x} = \sum_{k=1}^5 x_k F_k$$

e quindi la media è proprio la media pesata di tutti i diversi valori x_k misurati. Notiamo inoltre che

$$\sum_{k=1}^5 F_k = 1$$

ovvero gli F_k costituiscono un **insieme normalizzato** e la relazione precedente viene detta **condizione di normalizzazione**.

La distribuzione delle nostre misure può essere evidenziata graficamente in un **istogramma a barre**, ovvero in un grafico in cui in ascissa sono riportati i diversi valori x_k misurati ed in ordinata le corrispondenti frazioni F_k (vedi Fig.1a); in tale rappresentazione la distribuzione dei risultati è manifestata dall'andamento delle altezze delle barre verticali.

La sola media \bar{x} non è sufficiente a caratterizzare la distribuzione delle nostre misure. Infatti una distribuzione del tipo di quella mostrata in Fig.1b è caratterizzata dalla stessa media di quella di Fig.1a ma è chiaro che le due distribuzioni sono alquanto diverse tra

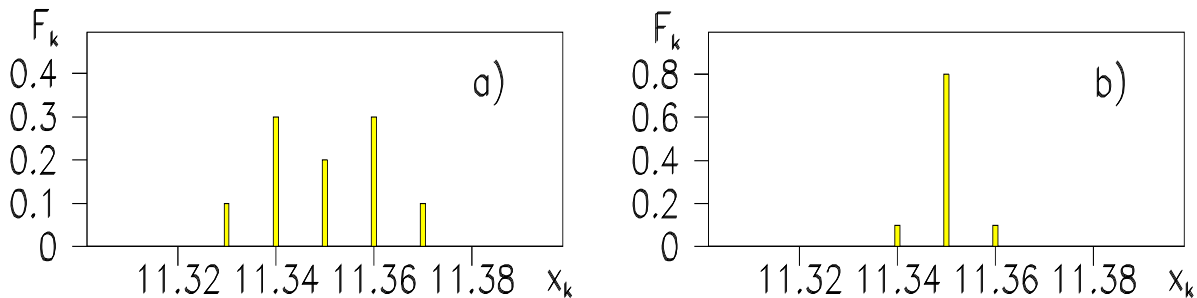


Figura 1:

loro. Dovremo quindi introdurre un nuovo parametro che permetta di caratterizzare come le misure x_i si distribuiscono intorno al valor medio, cioè di stimare la loro incertezza “media” .

Consideriamo innanzitutto la differenza $x_i - \bar{x} = d_i$, detta **deviazione**. Tanto minori in valore assoluto saranno le deviazioni tanto più precise saranno le misure. Le deviazioni assumeranno valori positivi o negativi come conseguenza della presenza di sorgenti di errore casuale; potremo pensare di stimare l'incertezza nelle nostre misure facendo una media delle deviazioni, ottenendo

$$\bar{d} = \sum_{i=1}^N \frac{(x_i - \bar{x})}{N} = \sum_{i=1}^N \frac{x_i}{N} - \sum_{i=1}^N \frac{\bar{x}}{N} = \bar{x} - \frac{\bar{x} N}{N} = 0 \text{ cm}$$

Tale risultato è conseguenza diretta della definizione (1) della media ed è indipendente dai valori delle x_i . La deviazione media non può quindi dare nessuna informazione sulla distribuzione delle nostre misure.

Potremo allora prendere in considerazione i quadrati delle deviazioni, che sono sempre positivi, e mediare su tali valori. Se poi prendiamo la radice quadrata del risultato

otteniamo una grandezza avente le stesse dimensioni di x che viene chiamata **deviazione standard** ed è indicata da σ_x

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N}} = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (2)$$

Per le due distribuzioni di Fig.1 la deviazione standard è

$$\text{caso a) } \sigma_x = 0.012 \text{ cm}, \quad \text{caso b) } \sigma_x = 0.004 \text{ cm}$$

Si vede quindi che la distribuzione del caso b), in cui le misure sono più concentrate nell'intorno del valore medio, è caratterizzata da una σ_x minore.

Una trattazione più rigorosa (come fatta nella statistica matematica) suggerisce di sostituire N con $N - 1$ nell'espressione della deviazione standard ottenendo così

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad (3)$$

Questa sostituzione è essenzialmente dovuta al fatto che degli N **gradi di libertà** contenuti nelle N misure indipendenti di x uno è stato utilizzato per determinare il valor medio \bar{x} ; ne restano quindi $N - 1$ a disposizione per la stima della deviazione standard. D'altra parte la sostituzione ci permette di superare un'incongruenza della (2). Infatti supponiamo di aver eseguito una sola misura di x e di aver ottenuto il valore x_1 ; evidentemente sarà $\bar{x} = x_1$. Se calcoliamo la deviazione standard con la (2) otteniamo $\sigma_x = 0$ ovvero un risultato assurdo. Valutando invece la deviazione standard con la (3) otteniamo un risultato indeterminato, in quanto rapporto tra due valori nulli, che è coerente con il fatto che avendo a disposizione una sola misura è impossibile determinare due parametri indipendenti tra loro, quali la media e la deviazione standard.

Discuteremo successivamente in maniera più rigorosa la scelta della deviazione standard come parametro che caratterizza l'incertezza delle nostre misure.

Le misure delle grandezze fisiche sono in generale caratterizzate da uno spettro continuo di valori possibili. Quindi se avessimo realizzato le precedenti misure utilizzando un

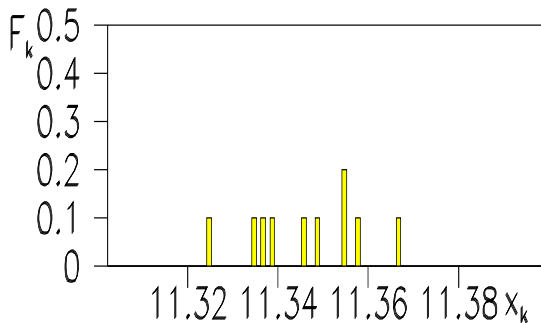


Figura 2:

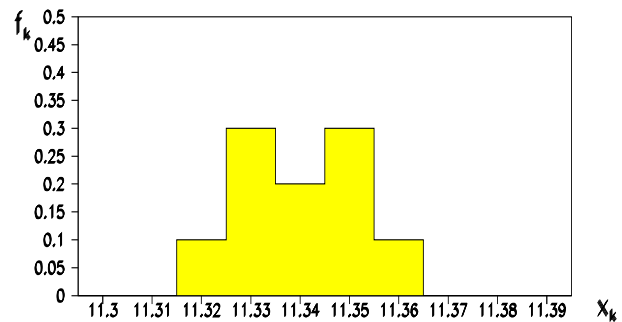


Figura 3:

calibro con errore di sensibilità 0.001 cm anziché 0.01 cm le nostre 10 misure si sarebbero

distribuite nell'intervallo tra 11.3 e 11.4 *cm* dando luogo ad un istogramma a barre del tipo di quello riportato in Fig.2, dove l'altezza delle barre non è più sufficiente a caratterizzare la distribuzione. In questo caso risulta più utile dividere la regione su cui si distribuiscono i valori misurati in **intervalli** o **classi** e contare quanti valori cadono in ciascun intervallo. I risultati così ottenuti possono essere riportati in grafico (vedi Fig.3) ottenendo il cosiddetto **istogramma ad intervalli**, in cui la frazione di misure che cadono in ciascun intervallo è data dall'area del rettangolo disegnato su ogni intervallo, ovvero

$$f_k \Delta_k = \text{frazione delle misure nell'intervallo } k\text{-esimo}$$

In via di principio i Δ_k potrebbero essere anche diversi tra loro, in quanto la caratteristica importante della distribuzione dei dati è l'area di ogni rettangolo, ma risulta più semplice sceglierli uguali per una migliore interpretazione a prima vista.

All'aumentare del numero delle misure l'istogramma ad intervalli assume via via una forma più regolare e potremo anche ridurre la larghezza degli intervalli in modo

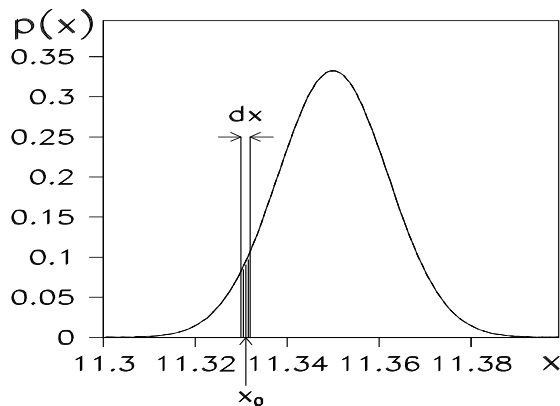


Figura 4:

la striscia ombreggiata in figura. E' per tale motivo che la curva limite $p(x)$ viene anche detta **densità di probabilità**.

Possiamo a questo punto considerare il nostro istogramma come la miglior stima sperimentale della distribuzione limite per cui $p(x_i) dx$ è la **probabilità che una singola misura di x dia un risultato compreso tra x_i e $x_i + dx$** .

Se si vuole la probabilità che una singola misura cada nell'intervallo (x_1, x_2) basta determinare l'integrale definito

$$\int_{x_1}^{x_2} p(x) dx$$

Dalla definizione della funzione limite segue inoltre che

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

ovvero la funzione limite è normalizzata.

Nel caso che nella misura di x , dopo aver ridotto le incertezze sistematiche a livelli trascurabili rispetto alle corrispondenti incertezze accidentali, siano presenti molti piccoli errori accidentali, la distribuzione limite assume una forma simmetrica a campana (vedi Appendice A), centrata sul valore vero di x , che indicheremo con X , ovvero su quel valore a cui si tenderebbe potendo ridurre l'incertezza di misura.

La funzione matematica che descrive tale curva è chiamata **distribuzione normale** o **funzione di Gauss** ed una rappresentazione di tale funzione è data da

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}}$$

ove σ è il cosiddetto **parametro di larghezza**. In Appendice A sono riportate le principali caratteristiche della distribuzione normale.

2 Miglior stima del valore vero e del parametro di larghezza della funzione di Gauss

Scopo della misura di una grandezza fisica è quello di dare una stima dei parametri che caratterizzano la corrispondente funzione limite (X e σ) mediante un numero limitato di misure. In pratica noi misuriamo N volte la grandezza x ottenendo i risultati

$$x_1, x_2, x_3, \dots, x_N$$

Per quanto visto nel par. precedente sulla funzione limite, sappiamo che la probabilità di ottenere una misura compresa tra x_1 e x_1+dx , con dx larghezza infinitesima dell'intervallo, è data da:

$$P(x_1, x_1 + dx) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_1-X)^2}{2\sigma^2}} dx \propto \frac{1}{\sigma} e^{-\frac{(x_1-X)^2}{2\sigma^2}}$$

Analogamente per quanto riguarda x_2 abbiamo

$$P(x_2, x_2 + dx) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_2-X)^2}{2\sigma^2}} dx \propto \frac{1}{\sigma} e^{-\frac{(x_2-X)^2}{2\sigma^2}}$$

La probabilità di ottenere la serie di misure x_1, \dots, x_N è data, per la regola della probabilità composta, dal prodotto delle singole probabilità, ovvero

$$\begin{aligned} P(x_1, \dots, x_N) &= P(x_1, x_1 + dx) P(x_2, x_2 + dx) \dots P(x_N, x_N + dx) \\ &\propto \frac{1}{\sigma^N} e^{-\frac{\sum_{i=1}^N (x_i - X)^2}{2\sigma^2}} \end{aligned} \quad (4)$$

D'altra parte il **principio di massima verosimiglianza** afferma che “avendo realizzato una serie di N misure della grandezza x ed avendo ottenuto i valori $x_1, x_2, x_3, \dots, x_N$ la probabilità di ottenere tale serie di valori deve essere massima” e quindi le miglior stime

di X e σ saranno quelle che rendono massima la (4). Per quanto riguarda X , che appare solo nell'esponente della (4), la probabilità sarà massima quando

$$\frac{\sum_{i=1}^N (x_i - X)^2}{2\sigma^2} = \text{minimo} \quad \Rightarrow \quad \frac{\partial \sum_{i=1}^N (x_i - X)^2}{\partial X} = 0 \quad \Rightarrow \quad X = \frac{\sum_{i=1}^N x_i}{N} = \bar{x} \quad (5)$$

Si vede quindi che la media è proprio la miglior stima del valore vero della nostra grandezza.

Analogamente potremo operare per determinare la miglior stima della larghezza della funzione limite, ottenendo

$$\frac{\partial P(x_1, \dots, x_N)}{\partial \sigma} = 0 \quad \Rightarrow \quad \sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - X)^2}{N}} \quad (6)$$

Tale valore di σ massimizza la probabilità $P(x_1, \dots, x_N)$ ma dipende ancora da X che è incognita. Potremo sostituire a X la media \bar{x} che sappiamo minimizza $\sum_{i=1}^N (x_i - X)^2$. Con tale sostituzione sottostimiamo quindi il numeratore della (6) e si può dimostrare che per compensare tale sottostima basta sostituire N con $N - 1$ nella (6). Si ottiene quindi

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} = \sigma_x \quad (7)$$

Sulla base di quanto ora ottenuto e delle proprietà della funzione di Gauss, potremo quindi affermare che nell'intervallo $[\bar{x} - \sigma_x, \bar{x} + \sigma_x]$ cadono il 68% delle misure, ovvero che avremo una probabilità del 68% che una singola misura cada nell'intervallo $[\bar{x} - \sigma_x, \bar{x} + \sigma_x]$.

3 Criterio di Chauvenet

Talvolta accade che in una serie di dati misurati una o più misure sembrano essere in disaccordo con le altre. Supponiamo ad esempio di aver realizzato con un calibro decimale 15 misure della larghezza x di una sbarretta e di aver ottenuto i seguenti risultati (in mm)

10.1, 10.0, 10.2, 10.3, 10.2, 10.1, 10.0, 10.2, 10.1, 10.2, 10.1, 10.3, 11.0, 11.5, 11.6

E' evidente che le ultime tre misure si discostano dalle altre ma abbiamo bisogno di un criterio oggettivo per giustificare un loro eventuale scarto. Infatti per quanto visto sulla distribuzione normale dei dati è certo possibile che una misura differisca significativamente dalle altre ma sappiamo d'altra parte che a tale possibilità è associata una probabilità molto piccola.

Ci sono varie linee di pensiero su come comportarsi in queste situazioni, anche tenendo conto del fatto che alcune volte nella storia della fisica *anomalie* di questo tipo hanno portato a sensazionali scoperte. Il comportamento più corretto è quello di ripetere molte volte la misura in modo da identificare se l'*anomalia* è legata ad un fatto fisico reale.

Qualora questo non sia possibile, si può procedere come segue:

- a) si determinano valor medio e deviazione standard delle misure (nel nostro caso $\bar{x} \simeq 10.4 \text{ mm}$, $\sigma_x \simeq 0.5 \text{ mm}$);

b) si determina la deviazione da \bar{x} per la misura che maggiormente si discosta da esso ($11.6 \text{ mm} \rightarrow d = 1.2 \text{ mm}$ ovvero $2.4 \sigma_x$). Supponendo che le misure siano distribuite normalmente con migliori stime del valore vero e del parametro di larghezza date nel punto a), possiamo determinare la probabilità di ottenere misure che differiscano dal valore vero almeno di questa quantità

$$\begin{aligned} P(\text{al di fuori di } 2.4 \sigma_x) &= 1 - P(\text{entro } 2.4 \sigma_x) = \\ &= 1 - 0.984 = 0.016 \simeq \frac{1}{63} \end{aligned}$$

In altre parole dovremmo aspettarci che una misura ogni 63 differisca dal valore medio di almeno $2.4 \sigma_x$. Se avessimo fatto 63 misure sarebbe quindi normale aspettarci la presenza del valore 11.6 mm . Ma noi abbiamo fatto solo 15 misure e quindi il numero di misure con tale deviazione dovrebbe essere 0.24.

Il **criterio di Chauvenet** stabilisce che se il numero atteso di misure con deviazione rispetto al valore medio maggiore o uguale a quello della misura sospetta è minore di 0.5 la misura sospetta può essere scartata.

Nel nostro caso saremo quindi giustificati nello scartare il valore 11.6 mm .

Potremo ripetere la procedura del punto b) anche per le altre due misure sospette ottenendo

$$\begin{aligned} x = 11.5 \text{ mm} \quad d = 2.2 \sigma_x \quad P(\text{al di fuori di } 2.0 \sigma_x) &= 1 - 0.972 = 0.028 \simeq \frac{1}{36} \\ x = 11.0 \text{ mm} \quad d = 1.2 \sigma_x \quad P(\text{al di fuori di } 1.2 \sigma_x) &= 1 - 0.77 = 0.23 \simeq \frac{1}{4} \end{aligned}$$

Sulla base del criterio di Chauvenet potremo quindi scartare anche il valore 11.5 mm mentre il valore 11.0 mm risulta pienamente giustificato.

Una volta operato lo scarto dei due valori sospetti determineremo le nuove migliori stime del valore vero e del parametro di larghezza ottenendo $\bar{x} = 10.3 \text{ mm}$, $\sigma_x \simeq 0.3 \text{ mm}$. Con i nuovi valori così ottenuti la terza misura sospetta (quella a 11.0 mm , corrispondente ora a $2.3 \sigma_x$) sarebbe scartabile sulla base del criterio di Chauvenet. La maggior parte degli esperti di statistica concordano però sul fatto che il criterio di Chauvenet non possa essere applicato più di una volta e che anche in sede di prima applicazione del criterio il numero di valori scartati debba essere trascurabile rispetto al numero totale delle misure.

Per completezza citiamo una seconda linea di pensiero più moderata che richiede di individuare i dati sospetti mediante il criterio di Chauvenet e di riportare il risultato finale della misura in due modi diversi, con e senza i dati sospetti, dichiarando esplicitamente la procedura seguita per ottenerli.

4 Propagazione degli errori nel caso di incertezze accidentali ed indipendenti tra loro

Per ricavare le regole di propagazione dell'incertezza dalle grandezze misurate direttamente a quelle misurate indirettamente utilizzeremo un metodo induttivo, ricavando l'espressione dell'incertezza in alcuni casi semplici e da questi ottenendo poi la formula generale.

Prendiamo quindi in esame alcuni casi semplici di misure indirette di una grandezza z ottenuta, nota la relazione funzionale $z = f(x, y, \dots)$, dalle grandezze x, y, \dots distribuite normalmente intorno ai corrispondenti valori veri X, Y, \dots con deviazioni standard $\sigma_x, \sigma_y, \dots$:

a) $z = x + A$ con A costante nota

La probabilità di ottenere un certo valore x_i è data da

$$P_x(x_i) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{(x_i - X)^2}{2\sigma_x^2}} dx$$

La probabilità di avere il corrispondente valore $z_i = x_i + A$ di z è quindi data

$$P_z(z_i) = P_x(z_i - A) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{[(z_i - A) - X]^2}{2\sigma_x^2}} dx = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{[z_i - (X + A)]^2}{2\sigma_x^2}} dz$$

Da tale relazione si vede subito che i valori di z sono normalmente distribuiti intorno al valore vero $Z = X + A$ con larghezza σ_x .

b) $z = B x$ con B costante nota

In modo del tutto analogo a quanto visto nel caso a) avremo che la probabilità di ottenere un certo valore $z_i = B x_i$ è data da

$$P_z(z_i) = P_x(z_i/B) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{[(z_i/B) - X]^2}{2\sigma_x^2}} dx = \frac{1}{\sqrt{2\pi} B \sigma_x} e^{-\frac{[z_i - (BX)]^2}{2 B^2 \sigma_x^2}} dz$$

I valori di z sono quindi distribuiti normalmente intorno al valore vero $Z = BX$ con larghezza $B\sigma_x$.

c) $z = x + y$

Per le due grandezze misurate direttamente avremo

$$P_x(x_i) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{(x_i - X)^2}{2\sigma_x^2}} dx \quad \text{e} \quad P_y(y_i) = \frac{1}{\sqrt{2\pi} \sigma_y} e^{-\frac{(y_i - Y)^2}{2\sigma_y^2}} dy$$

La probabilità di ottenere un certo valore $z_i = x_i + y_i$ è data da

$$\begin{aligned} P_z(z_i) &= P_x(x_i) P_y(y_i) = \frac{1}{2\pi \sigma_x \sigma_y} e^{-\frac{(x_i - X)^2}{2\sigma_x^2}} e^{-\frac{(y_i - Y)^2}{2\sigma_y^2}} dx dy = \\ &= \frac{1}{2\pi \sigma_x \sigma_y} e^{-\frac{1}{2} \left[\frac{(x_i - X)^2}{\sigma_x^2} + \frac{(y_i - Y)^2}{\sigma_y^2} \right]} dx dy \end{aligned}$$

Possiamo ora utilizzare l'identità

$$\begin{aligned} \frac{(x_i-X)^2}{\sigma_x^2} + \frac{(y_i-Y)^2}{\sigma_y^2} &= \frac{(x_i-X+y_i-Y)^2}{\sigma_x^2 + \sigma_y^2} + \frac{[\sigma_y^2 (x_i-X) - \sigma_x^2 (y_i-Y)]^2}{\sigma_x^2 \sigma_y^2 (\sigma_x^2 + \sigma_y^2)} = \\ &= \frac{(x_i-X+y_i-Y)^2}{\sigma_x^2 + \sigma_y^2} + \alpha^2 \end{aligned}$$

ottenendo

$$\begin{aligned} P_z(z_i) &= \frac{1}{2\pi \sigma_x \sigma_y} e^{-\frac{1}{2} \left[\frac{(x_i-X+y_i-Y)^2}{\sigma_x^2 + \sigma_y^2} \right]} e^{-\frac{\alpha^2}{2}} dx dy = \\ &= \frac{1}{2\pi \sigma_x \sigma_y} e^{-\frac{1}{2} \left[\frac{(z_i-X-Y)^2}{\sigma_x^2 + \sigma_y^2} \right]} e^{-\frac{\alpha^2}{2}} dx dy \end{aligned} \quad (8)$$

Ma la probabilità sopra riportata è quella di ottenere il valore z_i con una particolare coppia di valori x_i, y_i . Se vogliamo la probabilità di ottenere un valore z_i indipendentemente dalla particolare coppia scelta dovremo sommare su tutti i contributi a P_z dovuti alle coppie di valori x_j, y_j tali che $z_i = x_j + y_j$. Ognuna di queste coppie fornisce un contributo del tipo di quello riportato in (8) in cui il primo termine esponenziale è sempre lo stesso mentre il secondo differisce a seconda della coppia x_j, y_j . La somma (integrazione) di tutti questi termini dà quindi un risultato del tipo

$$P_z(z_i) \propto \frac{1}{\sqrt{2\pi} \sqrt{\sigma_x^2 + \sigma_y^2}} e^{-\frac{1}{2} \left[\frac{(z_i-X-Y)^2}{\sigma_x^2 + \sigma_y^2} \right]} \quad (9)$$

ottenuto tenendo conto che l'integrazione dell'esponenziale in α^2 dà un contributo pari a $\sqrt{2\pi \sigma_x^2 \sigma_y^2 / (\sigma_x^2 + \sigma_y^2)}$.

La (9) evidenzia il risultato che i valori di z sono distribuiti normalmente intorno al valore vero $X+Y$ e con parametro di larghezza $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$. Si vede quindi che le incertezze misurate direttamente si sommano "in quadratura" dando un risultato inferiore rispetto a quello che si otterrebbe se si sommassero linearmente. Tale risultato è conseguenza del fatto che è estremamente improbabile una fluttuazione statistica concorde in ambedue le grandezze misurate direttamente.

d) caso generale $z = z(x, y)$

Per trattare il caso generale aggiungeremo alle ipotesi introdotte nei casi precedenti quella che le incertezze relative sulle grandezze misurate direttamente σ_x/\bar{x} e σ_y/\bar{y} siano molto minori di 1. In tali condizioni potremo sviluppare la funzione $z(x, y)$ in serie e, fermandoci al 1° ordine, avremo

$$z_i = z(x_i, y_i) = z(X, Y) + \left(\frac{\partial z}{\partial x} \right)_{X,Y} (x_i - X) + \left(\frac{\partial z}{\partial y} \right)_{X,Y} (y_i - Y)$$

Tale relazione esprime i valori della grandezza misurata indirettamente z_i come somma di tre termini:

$x_{i1}, x_{i2}, \dots, x_{iN}$, $i = 1 \dots M$ i corrispondenti valori misurati) che sono distribuite normalmente, dimenticando per il momento che i vari x_{ij} non sono altro che le misure di un'unica grandezza fisica. Con tale assunzione e per quanto visto nel par.4 casi b) e c), la variabile \bar{x} è quindi distribuita normalmente intorno ad un valore vero $X_{\bar{x}}$ dato da

$$X_{\bar{x}} = \frac{X_1 + X_2 + \dots + X_N}{N} \quad (11)$$

e parametro di larghezza

$$\sigma_{\bar{x}} = \sqrt{\left(\frac{\partial \bar{x}}{\partial x_1}\right)^2 \sigma_1^2 + \dots + \left(\frac{\partial \bar{x}}{\partial x_N}\right)^2 \sigma_N^2} = \sqrt{\frac{1}{N^2} \sigma_1^2 + \dots + \frac{1}{N^2} \sigma_N^2} \quad (12)$$

Se a questo punto ricordiamo che le variabili x_1, x_2, \dots, x_N non sono distinte tra loro ma sono solo il risultato di una particolare suddivisione delle misure di una stessa variabile x , abbiamo immediatamente che $X_1 = X_2 = \dots = X_N = X$ ed inoltre $\sigma_1 = \sigma_2 = \dots = \sigma_N = \sigma_x$; operando tali sostituzioni nelle (11) e (12) si ricava infine

$$X_{\bar{x}} = X \quad \text{e} \quad \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}} \quad (13)$$

ovvero la media è distribuita normalmente intorno al valore vero X con deviazione standard $\sigma_{\bar{x}}$, detta **deviazione standard della media**, che viene quindi a rappresentare l'incertezza con la quale il valor medio misurato stima statisticamente il valor vero. Il significato di $\sigma_{\bar{x}}$ è quindi il seguente: se realizziamo nuovamente una serie di N misure e determiniamo il nuovo valore medio abbiamo una probabilità del 68% che esso si trovi entro una distanza pari a $\sigma_{\bar{x}}$ dal valore vero X .

Concludiamo con qualche commento sulla (13). Ricordiamo innanzitutto che la deviazione standard σ_x rappresenta l'incertezza nelle singole misure x_{ij} e non cambierebbe apprezzabilmente aumentando il numero delle misure. D'altra parte la deviazione standard della media $\sigma_{\bar{x}}$ diminuisce all'aumentare del numero di misure come conseguenza del fatto che meglio riusciamo ad individuare la forma della distribuzione limite ed in particolare il suo massimo. Il fattore \sqrt{N} che appare nella (13) ci dice però che per migliorare la precisione di un fattore 10 dovremo aumentare N di un fattore 100!

Vogliamo infine mettere in guardia dalla facile considerazione che "basta avere pazienza e, aumentando il numero di misure, si riesce sempre ad ottenere la precisione voluta con qualsiasi strumento si operi". Per capire meglio come tale affermazione sia illusoria, supponiamo di aver misurato 100 volte una lunghezza con un calibro decimale e di aver ottenuto sempre lo stesso valore 1.35 cm. Tenendo conto dell'errore di sensibilità dello strumento (0.01 cm) l'unica affermazione che potremo fare è che le 100 misure sono distribuite nell'intervallo tra 1.345 e 1.355 cm ma non abbiamo nessuna informazione su come le misure sono distribuite all'interno dell'intervallo. Ci troviamo quindi nella condizione di costruire un istogramma ad intervalli di larghezza Δ pari all'errore di sensibilità in cui si avranno conteggi in un solo intervallo. Si potrebbe determinare la deviazione standard di tale distribuzione, calcolandola rispetto alla mediana dell'intervallo, e otterremmo che essa è pari a $\Delta/12$ (nel nostro caso $0.01 \text{ cm}/12 \simeq 8 \cdot 10^{-4} \text{ cm}$) che costituisce quindi un

limite inferiore all'incertezza ottenibile con lo strumento a disposizione.

Vogliamo infine ricordare che una delle ipotesi di partenza di tutta la trattazione statistica è quella di aver ridotto a livello trascurabile (rispetto all'errore di sensibilità dello strumento usato) il contributo all'incertezza dovuto agli errori sistematici. Tale ipotesi non è più verificata quando, aumentando il numero delle misure, si riesca a spingere l'incertezza di misura molto al di sotto dell'errore di sensibilità dello strumento. In tal caso rischieremo di realizzare una misura estremamente precisa ma affetta da errori sistematici, che non sappiamo valutare, ben più grossi dell'incertezza sperimentalmente ottenuta.

6 Covarianza nella propagazione degli errori

Supponiamo di misurare indirettamente una grandezza fisica $z = z(x, y)$ misurando direttamente le grandezze x e y . Siano \bar{x} , \bar{y} e σ_x , σ_y le migliori stime del valore vero e dell'incertezza di misura ricavate da una serie di N misure $(x_1, y_1, x_2, y_2, \dots, x_N, y_N)$. In corrispondenza delle N coppie x_i, y_i potremo determinare N valori di z_i e successivamente determinarne la media \bar{z} , che assumiamo come migliore stima del valore vero di z , e la deviazione standard σ_z , nostra miglior stima dell'incertezza sui valori z_i .

Assumeremo nuovamente che le incertezze relative nelle misure dirette siano $\ll 1$; in tali ipotesi potremo scrivere, trascurando termini di ordine superiore al 1°,

$$z_i = z(x_i, y_i) = z(\bar{x}, \bar{y}) + \left(\frac{\partial z}{\partial x}\right)_{\bar{x}, \bar{y}} (x_i - \bar{x}) + \left(\frac{\partial z}{\partial y}\right)_{\bar{x}, \bar{y}} (y_i - \bar{y})$$

Utilizzando tale espressione per determinare la media avremo

$$\bar{z} = \frac{\sum_{i=1}^N z_i}{N} = \frac{1}{N} \sum_{i=1}^N \left[z(\bar{x}, \bar{y}) + \left(\frac{\partial z}{\partial x}\right)_{\bar{x}, \bar{y}} (x_i - \bar{x}) + \left(\frac{\partial z}{\partial y}\right)_{\bar{x}, \bar{y}} (y_i - \bar{y}) \right]$$

Tale valore medio è la somma di tre termini di cui solo il primo dà un contributo diverso da zero una volta che venga applicata l'operazione di sommatoria. Otteniamo quindi

$$\bar{z} = z(\bar{x}, \bar{y})$$

ovvero per determinare \bar{z} basta valutare $z(x, y)$ per $x = \bar{x}$ e $y = \bar{y}$.

Determiniamo ora la deviazione standard di z usando le espressioni precedenti; avremo

$$\begin{aligned} \sigma_z^2 &= \frac{1}{N-1} \sum_{i=1}^N (z_i - \bar{z})^2 = \frac{1}{N-1} \sum_{i=1}^N \left[\left(\frac{\partial z}{\partial x}\right)_{\bar{x}, \bar{y}} (x_i - \bar{x}) + \left(\frac{\partial z}{\partial y}\right)_{\bar{x}, \bar{y}} (y_i - \bar{y}) \right]^2 = \\ &= \left(\frac{\partial z}{\partial x}\right)_{\bar{x}, \bar{y}}^2 \sigma_x^2 + \left(\frac{\partial z}{\partial y}\right)_{\bar{x}, \bar{y}}^2 \sigma_y^2 + 2 \left(\frac{\partial z}{\partial x}\right)_{\bar{x}, \bar{y}} \left(\frac{\partial z}{\partial y}\right)_{\bar{x}, \bar{y}} \sigma_{xy} \end{aligned} \quad (14)$$

dove abbiamo indicato con σ_{xy} , chiamata **covarianza**, l'espressione

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

L'espressione (14) per σ_z^2 è indipendente dal fatto che le grandezze x e y siano indipendenti tra loro e che siano o meno distribuite normalmente.

Se le misure di x e y sono indipendenti tra loro e distribuite normalmente avremo che la covarianza σ_{xy} dovrebbe tendere a zero al crescere del numero delle misure in quanto ambedue i termini $(x_i - \bar{x})$ e $(y_i - \bar{y})$ possono assumere, con uguale probabilità, sia valori positivi che negativi. Nel caso limite in cui σ_{xy} tenda a zero otterremo

$$\sigma_z^2 = \left(\frac{\partial z}{\partial x} \right)_{\bar{x}, \bar{y}}^2 \sigma_x^2 + \left(\frac{\partial z}{\partial y} \right)_{\bar{x}, \bar{y}}^2 \sigma_y^2 \quad (15)$$

ovvero il risultato già ricavato nel par.4.

Se le misure non sono indipendenti tra loro la covarianza rimane diversa da zero: in tal caso diremo che le incertezze su x e su y sono **correlate** e la varianza su z sarà sempre maggiore del valore (15). Potremo cercare un limite superiore per la varianza che sia valido in tutti i casi. Per questo, ricordando che $|\sigma_{xy}| \leq \sigma_x \sigma_y$, nota come disuguaglianza di Schwarz, si ha

$$\sigma_z^2 \leq \left[\left| \frac{\partial z}{\partial x} \right|_{\bar{x}, \bar{y}} \sigma_x + \left| \frac{\partial z}{\partial y} \right|_{\bar{x}, \bar{y}} \sigma_y \right]^2$$

da cui segue

$$\sigma_z \leq \left| \frac{\partial z}{\partial x} \right|_{\bar{x}, \bar{y}} \sigma_x + \left| \frac{\partial z}{\partial y} \right|_{\bar{x}, \bar{y}} \sigma_y$$

Si vede quindi che il massimo valore per la deviazione standard sulla misura di z è quello che si ottiene sommando linearmente i contributi delle incertezze misurate direttamente.

7 Livello di confidenza

Accade spesso di dover confrontare tra loro due misure di una grandezza z per verificare l'accordo o meno tra i valori misurati. Nell'ipotesi che le due misure siano affette solo da incertezze accidentali, potremo rendere quantitativo questo confronto.

Siano infatti

$$z_1 \pm \sigma_1 \quad z_2 \pm \sigma_2$$

i valori misurati, con σ_1 e σ_2 le rispettive deviazioni standard della media. La deviazione τ tra le due misure è anch'essa distribuita normalmente e le migliori stime del suo valore vero e della sua incertezza sono date da

$$\tau_M = |z_1 - z_2| \quad \text{e} \quad \sigma_\tau = \sqrt{\left(\frac{\partial \tau}{\partial z_1} \right)^2 \sigma_1^2 + \left(\frac{\partial \tau}{\partial z_2} \right)^2 \sigma_2^2} = \sqrt{\sigma_1^2 + \sigma_2^2}$$

In generale τ_M sarà diverso da zero (valore corrispondente all'accordo perfetto tra le due misure, ma statisticamente poco probabile) e potremo determinare la sua distanza, in

unità σ_τ , dallo zero. Essa è data da $t = |\tau_M/\sigma_\tau|$ e la probabilità di ottenere un valore di τ che disti da τ_M quanto o più dello zero è

$$P(\text{al di fuori di } t\sigma_\tau) = 1 - P(\text{entro } t\sigma_\tau) = 1 - \int_{\tau_M - t\sigma_\tau}^{\tau_M + t\sigma_\tau} p(\tau) d\tau$$

dove $p(\tau)$ è la densità di probabilità gaussiana associata alla deviazione τ . Se tale probabilità è grande, la deviazione misurata è statisticamente accettabile e le due misure sono **consistenti** tra loro. Se viceversa la probabilità è molto piccola, la deviazione non è statisticamente accettabile e le misure risultano **inconsistenti** tra loro.

Il limite di accettabilità o meno dipende dal livello al di sotto del quale riteniamo che una deviazione sia irragionevolmente improbabile: tale limite, detto **livello di confidenza**, è convenzionalmente posto al 5%. Avremo quindi che una deviazione misurata pari a $2\sigma_\tau$ ($P(\text{entro } 2\sigma_\tau) = 95.4\% \rightarrow P(\text{al di fuori di } 2\sigma_\tau) = 4.6\%$) non è accettabile al livello di confidenza del 5% e quindi le due misure risultano inconsistenti tra loro.

8 Medie pesate

Un problema frequente è quello di combinare più misure di una certa grandezza al fine di ottenere un unico risultato finale. Supponiamo ad esempio di avere a disposizione le seguenti 3 misure della grandezza x

$$x_1 \pm \sigma_1 \quad x_2 \pm \sigma_2 \quad x_3 \pm \sigma_3$$

dove σ_1 , σ_2 e σ_3 sono le migliori stime delle incertezze di misura, valutate come deviazioni standard della singola misura o della media.

Come primo passo dovremo valutare la deviazione $\tau_{ij} = x_i - x_j$ per ciascuna coppia di valori misurati e confrontarla con le relative incertezze: se dovesse risultare che $\tau_{ij} \gg \sigma_i, \sigma_j$ dovremo sospettare la presenza di errori sistematici in una o ambedue le misure in esame e diremo che le misure sono **inconsistenti** tra loro. In tal caso non potremo ricavare un unico risultato finale delle nostre misure.

Nel caso che tutte le misure siano **consistenti** tra loro, potremmo pensare di determinare il risultato finale mediando i valori misurati. Ma tale procedura non tiene conto dell'eventuale differenza nella precisione delle tre misure: ci aspetteremmo che una misura intrinsecamente più precisa dovesse avere un'importanza maggiore sul risultato finale.

Sfruttando il fatto, insito nella scelta della miglior stima dell'incertezza di misura, che le nostre misure sono distribuite normalmente intorno al valore vero X , potremo affermare che, sulla base della prima misura, la probabilità di ottenere un valore di x tra x_1 e $x_1 + dx$ è data da

$$P_x(x_1) \propto \frac{1}{\sigma_1} e^{-\frac{(x_1 - X)^2}{2\sigma_1^2}}$$

Analogamente per le altre due misure avremo

$$P_x(x_2) \propto \frac{1}{\sigma_2} e^{-\frac{(x_2-X)^2}{2\sigma_2^2}} \quad P_x(x_3) \propto \frac{1}{\sigma_3} e^{-\frac{(x_3-X)^2}{2\sigma_3^2}}$$

La probabilità di ottenere la terna di misure a nostra disposizione è quindi data da

$$P_x(x_1, x_2, x_3) = P_x(x_1) P_x(x_2) P_x(x_3) \propto \frac{1}{\sigma_1 \sigma_2 \sigma_3} e^{-\frac{\chi^2}{2}} \quad (16)$$

dove abbiamo indicato

$$\chi^2 = \left(\frac{x_1 - X}{\sigma_1}\right)^2 + \left(\frac{x_2 - X}{\sigma_2}\right)^2 + \left(\frac{x_3 - X}{\sigma_3}\right)^2$$

Il principio di massima verosimiglianza asserisce che la nostra miglior stima di X è quel valore per il quale le effettive misure x_1, x_2, x_3 hanno la massima probabilità. La nostra miglior stima di X , che indicheremo con X_M , è quindi quella per cui la (16) è massima, ovvero per cui χ^2 è minimo. Annullando la derivata di χ^2 rispetto a X abbiamo quindi

$$\frac{\partial \chi^2}{\partial X} = 0 \quad \Rightarrow \quad X_M = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2} + \frac{x_3}{\sigma_3^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2}}$$

Introducendo i pesi

$$w_1 = \frac{1}{\sigma_1^2} \quad w_2 = \frac{1}{\sigma_2^2} \quad w_3 = \frac{1}{\sigma_3^2}$$

potremo scrivere

$$X_M = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3}{w_1 + w_2 + w_3}$$

La grandezza X_M viene detta **media pesata**; pesando le varie misure tramite il reciproco della varianza si dà maggior peso alla misura più precisa.

Utilizzando la propagazione dell'incertezza (10) potremo determinare l'incertezza su X_M ottenendo

$$\sigma_{X_M} = \sqrt{\sum_i \left(\frac{\partial X_M}{\partial x_i}\right)^2 \sigma_i^2} = \frac{1}{\sqrt{\sum_i \frac{1}{\sigma_i^2}}} = \frac{1}{\sqrt{\sum_i w_i}}$$

9 Metodo dei minimi quadrati

Uno dei più comuni ed interessanti problemi sperimentali è quello che consiste nella misura di una serie di valori di due grandezze fisiche per determinare la relazione funzionale che lega le due grandezze.

Il caso più semplice è quello in cui la relazione funzionale è lineare del tipo

$$y = a + b x$$

dove a e b sono due costanti incognite. Per determinarle si misurano N valori di x (x_1, x_2, \dots, x_N) ed i corrispondenti valori di y (y_1, y_2, \dots, y_N); se le nostre misure non fossero affette da incertezze ciascuno dei punti (x_i, y_i) dovrebbe giacere esattamente sulla retta $y = a + b x$ ed i valori di a e b sarebbero immediatamente determinati.

In realtà le incertezze sperimentali fanno sì che i dati sperimentali fluttuino intorno

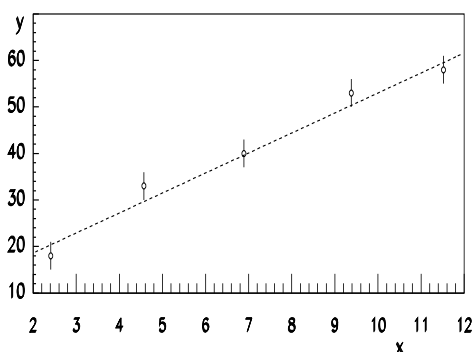


Figura 5:

alla retta ed al massimo possiamo aspettarci che la distanza tra i dati sperimentali e la retta sia confrontabile con le incertezze (vedi Fig.5). In tali condizioni potremo cercare di determinare la migliore stima di a e b , e della loro incertezza, compatibilmente con i dati sperimentali. Tale procedura sarà esposta nel successivo paragrafo e potremo successivamente verificare se la nostra serie di valori misurati è effettivamente consistente con la relazione lineare trovata.

a) Determinazione delle costanti a e b

Ipotizziamo innanzitutto che le misure delle due grandezze x e y siano affette esclusivamente da incertezze accidentali (ovvero di aver ridotto il contributo all'incertezza da parte delle sorgenti sistematiche a livello trascurabile rispetto al contributo delle sorgenti accidentali). Supponiamo inoltre, per semplicità di trattazione, che l'incertezza relativa nella misura di una delle due grandezze, ad es. x , sia molto minore di quella dell'altra. Se conoscessimo le costanti a e b , per ogni x_i misurato potremmo determinare il corrispondente valore vero Y_i della grandezza y per mezzo della relazione

$$Y_i = a + b x_i$$

Esso costituisce infatti la miglior stima del valore vero della grandezza y associato a x_i . La probabilità di trovare un valore di y compreso tra y_i e $y_i + dy$ è allora data da

$$P_y(y_i) = \frac{1}{\sqrt{2\pi}\sigma_{y_i}} e^{-\frac{(y_i - Y_i)^2}{2\sigma_{y_i}^2}} dy = \frac{1}{\sqrt{2\pi}\sigma_{y_i}} e^{-\frac{(y_i - a - bx_i)^2}{2\sigma_{y_i}^2}} dy$$

Se le N misure sono indipendenti tra loro la probabilità di ottenere la serie di valori y_1, y_2, \dots, y_N in corrispondenza di x_1, x_2, \dots, x_N è data da

$$\begin{aligned} P_y(y_1, \dots, y_N) &= P_y(y_1) P_y(y_2) \dots P_y(y_N) \propto \\ &\propto \frac{1}{\sigma_{y_1} \sigma_{y_2} \dots \sigma_{y_N}} e^{-\frac{(y_1 - Y_1)^2}{2\sigma_{y_1}^2}} e^{-\frac{(y_2 - Y_2)^2}{2\sigma_{y_2}^2}} \dots e^{-\frac{(y_N - Y_N)^2}{2\sigma_{y_N}^2}} \\ &\propto \frac{1}{\sigma_{y_1} \sigma_{y_2} \dots \sigma_{y_N}} e^{-\frac{\chi^2}{2}} \end{aligned}$$

dove abbiamo definito χ^2 tramite la relazione

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - a - bx_i)^2}{\sigma_{y_i}^2} \quad (17)$$

Le migliori stime di a e b sono quei valori per i quali $P_y(y_1, \dots, y_N)$ è massima ovvero per i quali il χ^2 è minimo. Cerchiamo di giustificare con parole diverse tale minimizzazione. Se nella (17) non ci fossero i $\sigma_{y_i}^2$ a denominatore, minimizzare χ^2 significherebbe minimizzare la distanza tra i nostri dati sperimentali e la retta che rappresenta la relazione tra y e x . La presenza dei termini $\sigma_{y_i}^2$ fa sì che le distanze corrispondenti a misure più precise abbiano un peso maggiore nella determinazione della retta.

Tale criterio di minimizzazione viene chiamato **criterio del minimo di χ^2** . Ne esamineremo un caso particolare: considereremo infatti solo il caso in cui le incertezze sulle singole misure di y sono tutte uguali tra loro. Supporremo quindi che

$$\sigma_1 = \sigma_2 = \dots = \sigma_N = \sigma_y$$

In tal caso il metodo prende il nome di **metodo dei minimi quadrati**.

In tali ipotesi sarà quindi

$$\chi^2 = \sum_{i=1}^N \frac{(y_i - a - bx_i)^2}{\sigma_y^2}$$

e la minimizzazione del χ^2 corrisponde alla minimizzazione della distanza dei punti sperimentali dalla retta. Per trovare i valori di a e b che minimizzano il χ^2 dovremmo porre uguali a zero le derivate di χ^2 rispetto ad a e b , ovvero

$$\frac{\partial \chi^2}{\partial a} = 0 \quad \Rightarrow \quad \sum_{i=1}^N y_i - Na - b \sum_{i=1}^N x_i = 0$$

$$\frac{\partial \chi^2}{\partial b} = 0 \quad \Rightarrow \quad \sum_{i=1}^N x_i y_i - a \sum_{i=1}^N x_i - b \sum_{i=1}^N x_i^2 = 0$$

e quindi potremo ricavare a e b risolvendo il sistema di equazioni

$$\begin{cases} Na + b \sum_{i=1}^N x_i = \sum_{i=1}^N y_i \\ a \sum_{i=1}^N x_i + b \sum_{i=1}^N x_i^2 = \sum_{i=1}^N x_i y_i \end{cases}$$

la cui soluzione è

$$a = \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}$$

$$b = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}$$

La retta così ottenuta è chiamata **retta dei minimi quadrati** o **linea di regressione di y su x** . E' chiaro che le costanti a e b così determinate saranno affette da incertezza: vedremo successivamente come determinarla.

b) Verifica della linearità

La serie di valori misurati (x_1, x_2, \dots, x_N ed i corrispondenti y_1, y_2, \dots, y_N) non ci permette di determinare l'incertezza su y a meno che per un dato valore x_i non ripetiamo più volte la misura del corrispondente y_i determinandone quindi la σ_{y_i} . Tuttavia, avendo ipotizzato che le y_i siano distribuite normalmente intorno al valore vero $a + bx_i$, le deviazioni $y_i - a - bx_i$ sono distribuite anch'esse normalmente intorno al valore vero 0 e con lo stesso parametro di larghezza σ_y . Ciò suggerisce che una buona stima per σ_y dovrebbe essere data dalla radice quadrata della media delle deviazioni quadratiche dalla retta ovvero

$$\sigma_y^2 = \frac{\sum_{i=1}^N (y_i - Y_i)^2}{N} = \frac{\sum_{i=1}^N (y_i - a - bx_i)^2}{N}$$

In effetti la miglior stima non è la precedente ma bensì

$$\sigma_y^2 = \frac{\sum_{i=1}^N (y_i - a - bx_i)^2}{N - 2} \quad (18)$$

Come già visto nel par.1 per la definizione della deviazione standard, nel corso di un calcolo statistico si definisce **numero di gradi di libertà** il numero di misure indipendenti meno il numero di parametri calcolati dalle misure. Nel caso del metodo dei minimi quadrati i parametri calcolati sono due (a e b) e da ciò segue il fattore $N - 2$.

Inoltre, se avessimo misurato solo due coppie di valori x_1, y_1 e x_2, y_2 , potremmo sempre trovare una retta che passa esattamente per entrambi i punti ma non potremmo dire niente sull'incertezza di misura e quindi sull'affidabilità della retta. In tal caso la (18) darebbe un risultato indeterminato in quanto rapporto di due quantità nulle.

Se abbiamo già una stima indipendente della nostra incertezza sulla grandezza y (la σ_{y_i} ottenuta dalle misure ripetute di y_i per un dato valore di x_i), dovremo aspettarci che essa sia confrontabile con la σ_y ottenuta dalla (18). Si possono presentare 3 casi:

- $\sigma_{y_i} \gg \sigma_y$
l'incertezza misurata sperimentalmente sovrastima la fluttuazione dei dati sperimentali rispetto all'andamento lineare determinato. In questo caso la stima sperimentale dell'incertezza σ_{y_i} non è quantitativamente rappresentativa delle fluttuazioni casuali su y e l'andamento lineare risulta giustificato sperimentalmente;

- $\sigma_{y_i} \simeq \sigma_y$
le due incertezze sono confrontabili e quindi anche l'andamento lineare è ben verificato sperimentalmente;
- $\sigma_{y_i} \ll \sigma_y$
la fluttuazione dei dati sperimentali rispetto all'andamento lineare è molto maggiore di quanto ci si potrebbe attendere sulla base dell'incertezza misurata sperimentalmente: l'andamento lineare determinato è alquanto dubbio.

Per completezza ricordiamo che, qualora non fosse disponibile una stima sperimentale dell'incertezza di misura sulla grandezza y (σ_{y_i}), è possibile determinare dai dati sperimentali un parametro, detto **coefficiente di correlazione lineare**, dato da

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

dove \bar{x} e \bar{y} sono i valori medi delle misure di x e y , il cui valore può essere utilizzato per giustificare l'andamento lineare tra le due variabili. Rimandiamo ai testi di statistica per maggiori dettagli su tale parametro.

c) Incertezza sui parametri a e b

Una volta determinata la fluttuazione dei dati sperimentali rispetto all'andamento lineare potremo ricavare le incertezze sui parametri a e b utilizzando la formula di propagazione dell'incertezza (10) ed ottenendo

$$\sigma_a = \sqrt{\sum_{i=1}^N \left(\frac{\partial a}{\partial y_i} \right)^2} \sigma_y^2$$

$$\sigma_b = \sqrt{\sum_{i=1}^N \left(\frac{\partial b}{\partial y_i} \right)^2} \sigma_y^2$$

Sviluppando i calcoli (non semplici!) si ottiene

$$\sigma_a = \sigma_y \sqrt{\frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}$$

$$\sigma_b = \sigma_y \sqrt{\frac{N}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}}$$

A APPENDICE A

CENNI SULLE DISTRIBUZIONI STATISTICHE

A.1 Introduzione

Scopo di questa appendice è quello di giustificare la rappresentazione funzionale della distribuzione gaussiana utilizzata per descrivere la distribuzione dei dati nel caso di una misura affetta esclusivamente da errori accidentali.

Come primo passo dovremo introdurre i concetti di “permutazione” e “combinazione” di n oggetti. Una permutazione riguarda l'ordine di n oggetti diversi tra loro, che indicheremo con a_1, a_2, \dots, a_n .

Se ad esempio gli n oggetti sono le sette note musicali, una sequenza di essi del tipo

$$a_1 \ a_5 \ a_7 \ a_3 \ a_2 \ a_4 \ a_6$$

costituisce una melodia. Ci si può chiedere quante melodie possono essere costruite prendendo una sequenza delle sette note in modo che ogni nota appaia solo una volta.

Se non ci poniamo nessuna regola di composizione musicale abbiamo 7 possibili scelte per la prima nota, 6 per la seconda, 5 per la terza e così via. Il numero di differenti melodie è quindi

$$7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040 = 7!$$

Ognuna di queste è detta una permutazione delle sette note date. Più in generale, nel caso di n oggetti, avremo che il numero delle permutazioni è dato da $n!$.

Supponiamo ora che tre delle nostre sette note coincidano tra loro (ad es. $a_1 = a_2 = a_3$). La nostra melodia è quindi costituita da solo 5 note diverse tra loro, 4 delle quali appaiono solo una volta e 1 di esse appare 3 volte. Delle 7! melodie iniziali quelle costituite da

$$a_1 \ a_4 \ a_2 \ a_5 \ a_6 \ a_7 \ a_3 \quad \text{e} \quad a_2 \ a_4 \ a_1 \ a_5 \ a_6 \ a_7 \ a_3$$

sono la stessa melodia. Quindi per ogni disposizione degli altri oggetti (a_4, a_5, a_6, a_7) avremo 3! melodie, ottenute permutando tra loro a_1, a_2 e a_3 , uguali tra loro. Ciò riduce il numero delle melodie diverse tra loro a

$$\frac{7!}{3!} = \frac{5040}{6} = 840$$

Supponiamo infine che le altre 4 note coincidano tra loro, così che la nostra melodia sia costituita da solo due note diverse, una che appare 3 volte e l'altra 4 volte. La permutazione delle seconde 4 note non cambia la melodia, una volta fissata la disposizione delle altre 3. Quindi per ottenere il numero delle diverse melodie dovremo ancora dividere per 4! ottenendo così:

$$\frac{7!}{4! \cdot 3!} = \frac{5040}{24 \cdot 6} = 35$$

In altre parole potremo dire che abbiamo a disposizione 7 posizioni in tre delle quali poniamo ogni volta una data nota mentre nelle rimanenti poniamo una seconda nota. Il numero di tali combinazioni di 7 posizioni prese 3 alla volta è denotato con $C_{7,3}$ o $\binom{7}{3}$ ed è dato da

$$C_{7,3} = \binom{7}{3} = \frac{7!}{3!4!}$$

Se ci fossimo riferiti alla seconda nota avremmo dovuto determinare le combinazioni di 7 posizioni prese 4 alla volta, ottenendo

$$C_{7,4} = \binom{7}{4} = \frac{7!}{4!3!}$$

A.2 Richiami di calcolo combinatorio

Richiamiamo alcune nozioni di calcolo combinatorio che possono essere approfondite in testi di Analisi Matematica, anche a livello liceale.

Dati n oggetti (o elementi) distinti a_1, a_2, \dots, a_n ed indicato con k un numero intero positivo non superiore ad n ($0 < k \leq n$), chiameremo **disposizioni semplici** di questi n oggetti a k a k (o di classe k) tutti i gruppi di k oggetti distinti che si possono formare con gli n oggetti dati, intendendo due gruppi distinti se differiscono o per un elemento o per l'ordine degli elementi. Il loro numero $D_{n,k}$ sarà dato da

$$D_{n,k} = n (n - 1) (n - 2) \dots (n - k + 1) \quad k \text{ fattori}$$

Ponendo per convenzione $D_{n,0} = 1$ qualunque sia l'intero $n > 0$, avremo che $D_{n,1} = n$. Se $k = n$ tutti gli n oggetti vengono considerati e 2 gruppi si differenzieranno solo per l'ordine secondo cui gli n oggetti sono disposti. In tal caso il numero delle disposizioni sarà dato da

$$D_{n,n} = n! = P_n$$

e chiameremo P_n il numero di **permutazioni** degli n elementi dati.

Chiamiamo poi **combinazioni semplici** di classe k degli n oggetti $C_{n,k}$ il numero dei gruppi di k oggetti che si possono formare con gli n oggetti dati, considerando distinti due gruppi che differiscono per almeno un elemento. Due gruppi di k oggetti, contenenti gli stessi elementi in ordine diverso, rappresentano una stessa combinazione semplice. Si può facilmente vedere che $C_{n,1} = n$ e che $C_{n,n} = 1$.

Supponiamo di aver determinato tutte le $C_{n,k}$ combinazioni semplici; da ogni combinazione, permutando i k elementi, si otterranno P_k disposizioni distinte, cioè

$$\begin{aligned} D_{n,k} &= C_{n,k} \cdot P_k \\ C_{n,k} &= \frac{D_{n,k}}{P_k} = \frac{n (n-1) (n-2) \dots (n-k+1)}{k!} = \\ &= \frac{n (n-1) (n-2) \dots (n-k+1) (n-k)!}{k! (n-k)!} = \frac{n!}{(n-k)! k!} \end{aligned}$$

$C_{n,k}$ si chiama anche **coefficiente binomiale di classe k e di grado n** e può essere indicato con la notazione $\binom{n}{k}$. Si può vedere che

$$\binom{n}{k} = \binom{n}{n-k} \quad , \quad \binom{n}{1} = \binom{n}{n-1} = n \quad , \quad \binom{n}{n} = 1 \quad (19)$$

e che si assume $\binom{n}{0} = 1$ per convenzione.

I coefficienti binomiali $\binom{n}{k}$ sono di particolare utilità per lo sviluppo dei binomi. Infatti

$$(a + b)^n = a^n + na^{(n-1)}b + \frac{n(n-1)}{2}a^{(n-2)}b^2 + \frac{n(n-1)(n-2)}{3!}a^{(n-3)}b^3 + \dots + \frac{n(n-1)}{2}a^2b^{(n-2)} + nab^{(n-1)} + b^n = \sum_{k=0}^n \binom{n}{k} a^k b^{(n-k)}$$

A.3 Distribuzione binomiale

Supponiamo ora di far coincidere le n posizioni prese in esame precedentemente con le n prove di un dato esperimento, ad esempio il lancio di un dado. Ogni prova può avere vari risultati ovvero presentare una qualunque faccia da 1 a 6. Denotiamo con p la probabilità di ottenere un dato risultato (ad esempio quella di ottenere 1) e con $q = 1 - p$ quella di ottenere un risultato diverso da quello prescelto (q è evidentemente il complementare a 1 di p in quanto la somma di p e q dà la probabilità di ottenere un qualsiasi risultato che deve essere necessariamente uguale a 1, ovvero il 100%). Nel caso del lancio del dado avremo, se non ci sono trucchi, $p = \frac{1}{6}$ e $q = \frac{5}{6}$.

La probabilità di ottenere r volte il valore prescelto in n prove è data dalla cosiddetta distribuzione binomiale

$$P(r, n) = b_{n,p}(r) = \binom{n}{r} p^r q^{(n-r)} \quad (20)$$

in quanto la probabilità di ottenere r volte il risultato prescelto in n prove con una data sequenza è $p^r q^{(n-r)}$ (la probabilità della simultanea occorrenza di eventi indipendenti [probabilità composta] è data dal prodotto delle probabilità dei singoli eventi) e, se non fissiamo la sequenza, abbiamo $\binom{n}{r}$ combinazioni possibili.

Se, ad esempio, vogliamo ottenere il risultato 1 per 3 volte in 5 lanci del dado avremo che la probabilità di avere 1 nei primi 3 lanci ed un valore diverso da 1 negli altri due è data da

$$\left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^2 = 3.2 \cdot 10^{-3}$$

Se poi vogliamo la probabilità di ottenere 1 per 3 volte su 5 lanci, indipendentemente dal lancio in cui avviene, avremo

$$\binom{5}{3} \cdot 3.2 \cdot 10^{-3} = 10 \cdot 3.2 \cdot 10^{-3} = 3.2 \cdot 10^{-2}$$

Potremo costruirci la distribuzione di probabilità, fissato il numero di prove, di ottenere 0,1,2,3,4 o 5 volte la faccia 1 nei 5 lanci del dado. Otterremo la distribuzione riportata in Fig.6a) (in Fig.6b) e 6c) sono mostrate le distribuzioni corrispondenti ai casi di 25 e 125 lanci).

Si vede innanzitutto che la distribuzione così ottenuta è normalizzata ovvero è a somma unitaria.

Si potrebbe dimostrare, ma non lo faremo, che, se ripetessimo le nostre n prove molte

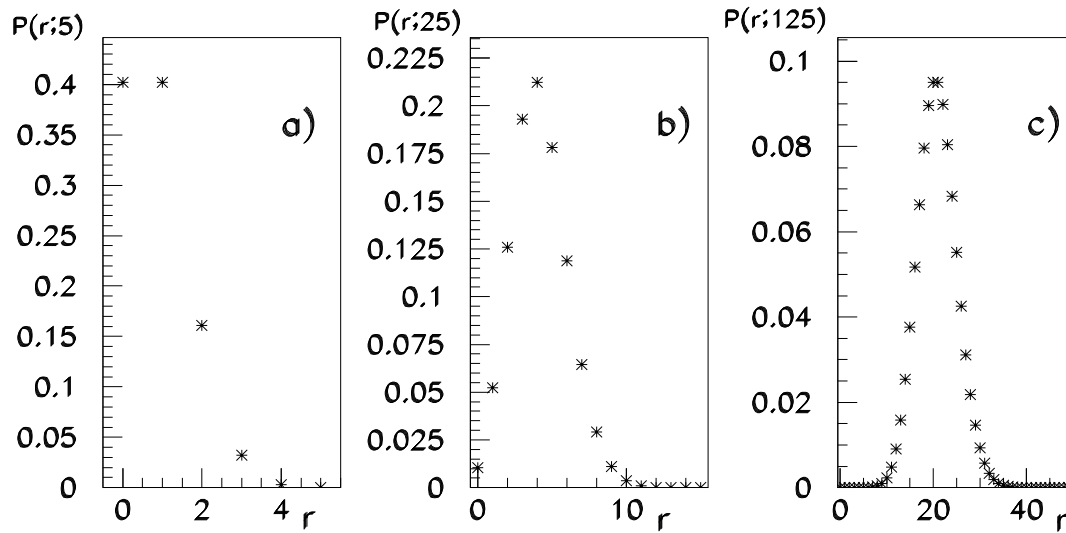


Figura 6: Distribuzioni binomiali per 5, 25 e 125 lanci del dado

volte, il numero medio di volte in cui si ottiene il risultato prescelto (la faccia 1 del dado) è dato da

$$\bar{r} = \sum_{r=0}^n r \cdot b_{n,p}(r) = n \cdot p$$

(nell'esempio $\bar{r} = 5 \cdot \frac{1}{6} = 0.83$), pari cioè al prodotto del numero delle misure per la probabilità di ottenere il risultato prescelto.

Per la varianza otterremo

$$\sigma_r^2 = \sum_{r=0}^n (r - \bar{r})^2 \cdot b_{n,p}(r) = np(1 - p)$$

da cui si ricava per la deviazione standard

$$\sigma_r = \sqrt{np(1 - p)}$$

(nell'esempio $\sigma_r = 0.83$ ovvero lo stesso valore ottenuto per \bar{r} ; questo non è un risultato generale ma è solo conseguenza del fatto che il numero delle prove coincide con il rapporto tra q e p).

In generale la distribuzione binomiale non è simmetrica (vedasi il caso riportato in Fig.6). Solo nel caso in cui $p = q = \frac{1}{2}$ la distribuzione diventa simmetrica in quanto, ricordando le (19) e (20), si ha

$$P(r, n) = P(n - r, n).$$

Una distribuzione di questo tipo la si ottiene considerando il lancio di una moneta (ad esempio da 2 Euro). La probabilità di ottenere europa in un lancio è data da $p = \frac{1}{2}$, quella di ottenere dante è $q = \frac{1}{2}$. Applicando la relazione (20) si può costruire la distribuzione di probabilità riportata in Fig.7.

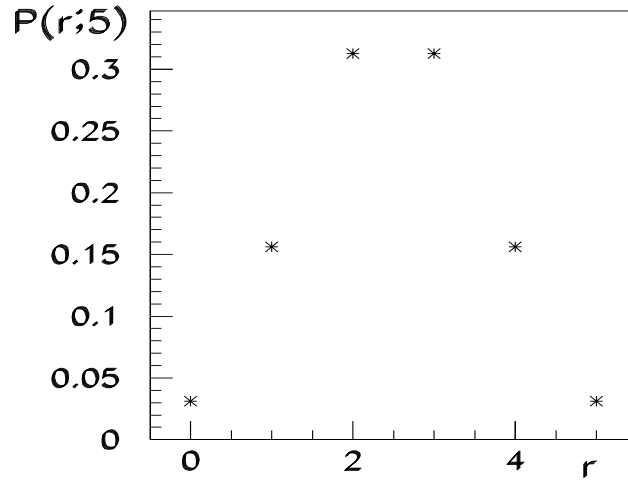


Figura 7: Distribuzione binomiale per 5 lanci di una moneta

Si verifica immediatamente che

$$\bar{r} = n \cdot p = 5 \cdot \frac{1}{2} = 2.5$$

e

$$\sigma_r = \sqrt{np(1-p)} = \sqrt{5 \cdot \frac{1}{2} \cdot \frac{1}{2}} = \sqrt{\frac{5}{4}} \simeq 1.12$$

A.4 Distribuzione di Gauss degli errori casuali

Supponiamo che nella nostra misura siano presenti n sorgenti di errore, ognuna delle quali può contribuire con un valore $+e_c$ oppure $-e_c$ nel modificare il valore vero X della grandezza fisica x che stiamo misurando. Le probabilità di modificare in positivo (p^+) o in negativo (p^-) il valore vero siano uguali tra loro e pari a $\frac{1}{2}$.

Possiamo costruire un diagramma del tipo di quello in Fig.8.

Le misure finali possono variare tra $X - ne_c$ e $X + ne_c$ a passi di $2e_c$. Ogni particolare combinazione degli errori, o cammino del diagramma da X ad uno dei possibili valori finali, ha una probabilità che può essere rappresentata come un prodotto

$$p^+ p^- p^+ p^+ \dots p^+ p^-$$

Se sono presenti n fonti di errore sono possibili n scelte. Se di queste un numero r corrispondono ad una probabilità p^+ e $n - r$ corrispondono ad una p^- , avremo che l'errore finale di misura sarà dato da

$$r \cdot e_c + (n - r)(-e_c) = (2r - n) \cdot e_c$$

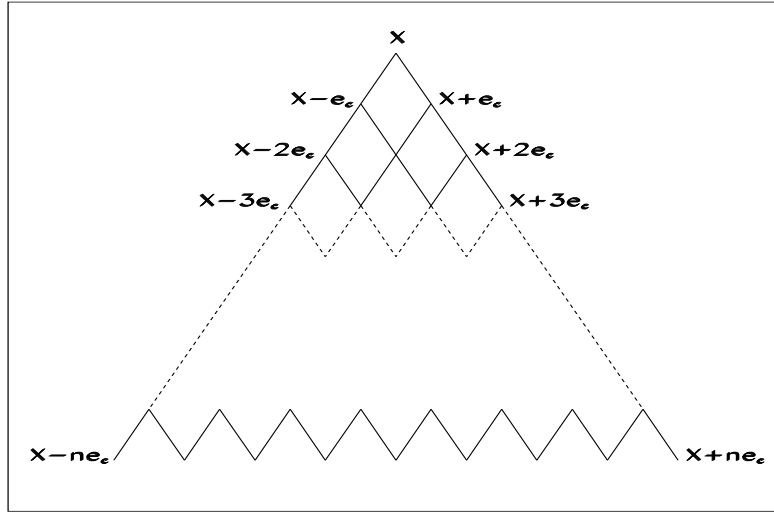


Figura 8: Diagramma delle sorgenti di errore

Per ogni errore $(2r - n) \cdot e_c$, e quindi per la corrispondente probabilità $(p^+)^r (p^-)^{(n-r)}$, ci sono molti modi diversi di scrivere l'ordine dei singoli fattori. Tali modi possibili corrispondono al numero di combinazioni di n presi a r ogni volta. Avremo quindi che la probabilità di avere un valore dell'errore di misura pari a $(2r - n) \cdot e_c$ è data da

$$P(2r - n) = \binom{n}{r} (p^+)^r (p^-)^{(n-r)}$$

Ma abbiamo detto che $p^+ = p^- = \frac{1}{2}$ e quindi otteniamo

$$P(2r - n) = \binom{n}{r} \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{(n-r)} = \binom{n}{r} \left(\frac{1}{2}\right)^n = \frac{n!}{r!(n-r)!} \left(\frac{1}{2}\right)^n$$

Chiamando a questo punto $m = 2r - n$, abbiamo che $r = \frac{1}{2}(n + m)$ e che la probabilità di avere un errore $E = (2r - n)e_c = me_c$, ovvero di avere una misura $x = X + E$, è data da

$$P(m, n) = \frac{n!}{2^n \left(\frac{n+m}{2}\right)! \left(\frac{n-m}{2}\right)!} \quad (21)$$

E' chiaro che in questo caso la distribuzione di probabilità sarà simmetrica e centrata al valore $\bar{r} = \frac{n}{2}$ ovvero centrata a $m = 0$. In Fig.9 sono riportati i grafici ottenibili per $n = 5$, $n = 25$ e $n = 125$.

Si può notare che quando n cresce la distribuzione si avvicina sempre più ad una curva continua e simmetrica del tipo di quella vista nel caso degli errori accidentali. Da notare che m può prendere solo valori pari o dispari a seconda che n sia pari o dispari (il termine $2r$ nella definizione di m è sempre pari) e che E può variare solo con passo pari a $2e_c$.

Vediamo ora cosa succede quando n e m crescono indefinitamente. Possiamo allora utilizzare la formula approssimata di Stirling per il fattoriale di un numero j molto

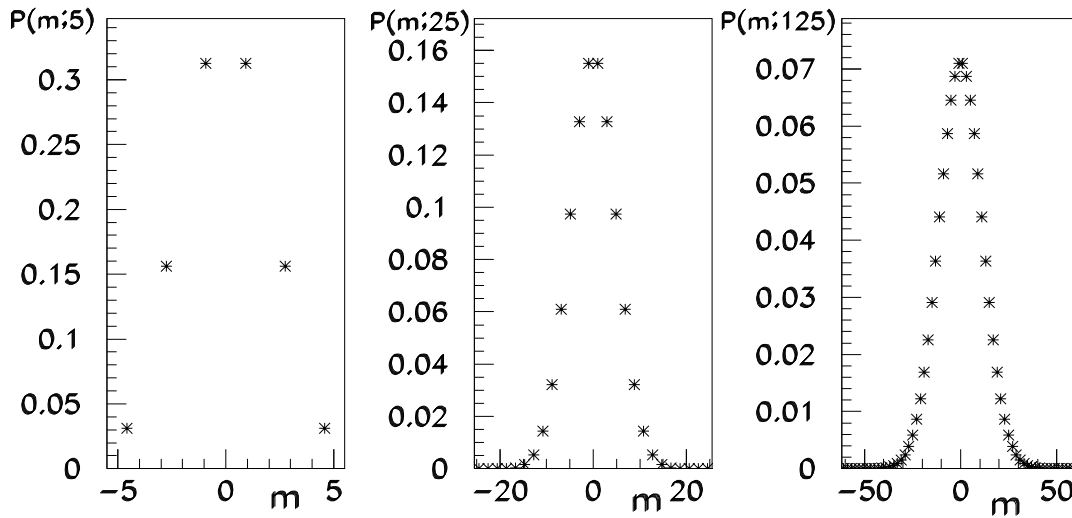


Figura 9: Distribuzioni di Gauss con 5, 25 e 125 sorgenti accidentali di errore

maggiore di 1

$$\ln(j!) \simeq \frac{1}{2} \ln(2\pi j) + j(\ln j - 1)$$

Prendendo il ln di ambedue i membri della (21) ed introducendo la formula di Stirling si ottiene

$$\begin{aligned} \ln P(m, n) &= \ln(n!) - \ln\left(\frac{n+m}{2}\right)! - \ln\left(\frac{n-m}{2}\right)! - n \ln 2 = \dots\dots\dots = \\ &= \frac{1}{2} \ln\left(\frac{2}{\pi n}\right) - \frac{1}{2}(n+1) \ln\left(1 - \frac{m^2}{n^2}\right) - \frac{1}{2}m \ln\left(\frac{1 + \frac{m}{n}}{1 - \frac{m}{n}}\right) \end{aligned}$$

Se a questo punto facciamo l'ipotesi (che verificheremo in seguito) che n cresca più velocemente di m , potremo sostituire ai ln a secondo membro i corrispondenti sviluppi al primo ordine, ovvero

$$(n+1) \ln\left(1 - \frac{m^2}{n^2}\right) = (n+1) \left(-\frac{m^2}{n^2} + \dots\right) \simeq -\frac{m^2}{n}$$

e

$$m \ln\left(\frac{1 + \frac{m}{n}}{1 - \frac{m}{n}}\right) = m \left(\frac{m}{n} + \dots - \left(-\frac{m}{n} + \dots\right)\right) \simeq 2\frac{m^2}{n}$$

ottenendo infine

$$\ln P(m, n) \simeq \frac{1}{2} \ln \frac{2}{\pi n} - \frac{m^2}{2n}$$

Ma sappiamo che m è legato all'errore di misura E dalla relazione $E = me_c$ e quindi

$$\ln P(m, n) \simeq \frac{1}{2} \ln \frac{2}{\pi n} - \frac{E^2}{2ne_c^2}$$

da cui

$$P(m, n) \simeq \left(\frac{2}{\pi n}\right)^{\frac{1}{2}} e^{-\frac{E^2}{2ne_c^2}} = \frac{2e_c}{(2\pi ne_c^2)^{\frac{1}{2}}} e^{-\frac{E^2}{2ne_c^2}}$$

Se definiamo una funzione di distribuzione dell'errore tramite la relazione

$$P(E, ne_c^2) = \frac{1}{(2\pi ne_c^2)^{\frac{1}{2}}} e^{-\frac{E^2}{2ne_c^2}}$$

abbiamo

$$P(m, n) \simeq P(E, ne_c^2)E'$$

dove $E' = 2e_c$ è l'incremento di errore da un valore possibile al successivo. Quindi, a parte il fattore costante E' , $P(E, ne_c^2)$ è la curva continua che approssima la distribuzione binomiale. E' allora chiaro che se $P(E, ne_c^2)$ deve rimanere una funzione finita e diversa da zero, il fattore ne_c^2 dovrà mantenersi finito e quindi al tendere di e_c a zero n tenderà all'infinito come $\frac{1}{e_c^2}$. Ma anche $E = me_c$ deve mantenersi finito quando e_c tende a zero: per ottenere tale risultato m dovrà tendere ad infinito come $\frac{1}{e_c}$, il che giustifica l'ipotesi fatta inizialmente che m tenda all'infinito più lentamente di n .

Al tendere di e_c a zero possiamo considerare che esso sia diventato così piccolo che molti passi di valore $2e_c$ si trovino nell'intervallo di errore da E a $E + dE$, anche se scegliamo dE così piccolo che $P(E, ne_c^2)$ non cambi apprezzabilmente in questo intervallo. Quindi la somma dei termini della distribuzione binomiale $P(m, n)$ per cui i corrispondenti errori stanno nell'intervallo

$$E \leq me_c \leq E + dE$$

sarà

$$\sum P(m, n) \simeq \sum P(E, ne_c^2)E' \simeq P(E, ne_c^2)dE = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{E^2}{2\sigma^2}} dE$$

Questa è la probabilità di misurare la grandezza x con un errore nell'intervallo $(E, E + dE)$ e può essere riscritta nella forma (ricordando che X rappresenta il valore vero della nostra grandezza)

$$p(x)dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}} dx$$

ovvero l'espressione funzionale utilizzata per la distribuzione di Gauss.

A.5 Caratteristiche della distribuzione normale

La funzione di Gauss, o distribuzione normale, può essere rappresentata nella forma

$$p(x) = K e^{-\frac{(x-X)^2}{2\sigma^2}}$$

ove K è il **fattore di normalizzazione** e σ è una grandezza positiva denominata **parametro di larghezza**. Imponendo la condizione di normalizzazione è possibile ricavare K ; infatti

$$\int_{-\infty}^{+\infty} p(x) dx = \int_{-\infty}^{+\infty} K e^{-\frac{(x-X)^2}{2\sigma^2}} dx = 1 \quad \rightarrow \quad K = \frac{1}{\sqrt{2\pi} \sigma}$$

e quindi la funzione di Gauss diventa

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-X)^2}{2\sigma^2}}$$

Si vede quindi che la funzione di Gauss dipende essenzialmente da due parametri (X, σ): per evidenziare tale aspetto indicheremo d'ora in poi la $p(x)$ con la notazione $p_{X,\sigma}(x)$. Si può inoltre immediatamente notare che $p_{X,\sigma}(x) \geq 0$ per ogni x ed inoltre

$$\begin{aligned} x \rightarrow \pm\infty &\Rightarrow p_{X,\sigma}(x) \rightarrow 0 \\ p'_{X,\sigma}(x) = 0 &\Rightarrow x = X &\Rightarrow p_{X,\sigma}(X) = \frac{1}{\sqrt{2\pi} \sigma} \\ &&\Rightarrow p''_{X,\sigma}(X) < 0 \end{aligned}$$

e quindi la funzione limite ha un massimo in $x = X$.

Per caratterizzare ulteriormente l'andamento della curva determiniamo i valori di x per cui essa assume valori pari alla metà del valore massimo; basterà imporre

$$p_{X,\sigma}(x_{1/2}) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x_{1/2}-X)^2}{2\sigma^2}} = \frac{1}{2} \frac{1}{\sqrt{2\pi} \sigma} \quad \rightarrow \quad (x_{1/2} - X) = \pm 1.18 \sigma$$

e quindi la larghezza a mezza altezza della curva è data da 2.36σ . Ciò giustifica il nome parametro di larghezza dato a σ : maggiore è σ più larga è la distribuzione. Il fatto che σ appaia anche nel fattore di normalizzazione a denominatore implica che ad una curva più stretta compete un valore massimo maggiore, come c'era da aspettarsi dovendosi mantenere unitario l'integrale della curva.

Se a questo punto volessimo caratterizzare la funzione di Gauss tramite due parametri come abbiamo fatto per la distribuzione ad intervalli dovremmo introdurre per una funzione continua i concetti analoghi a quelli della media e della deviazione standard. Definiremo quindi il **valore di aspettazione** della grandezza x tramite la relazione

$$\bar{x} = \int_{-\infty}^{+\infty} x p_{X,\sigma}(x) dx$$

mentre potremo ricavare la deviazione standard estraendo la radice quadrata della **varianza** definita da

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x - X)^2 p_{X,\sigma}(x) dx$$

Determinando valore di aspettazione e varianza per la funzione di Gauss si ottiene

$$\bar{x} = X, \quad \sigma_x^2 = \sigma^2$$

La probabilità $P(X, \sigma)$ che una singola misura dia un risultato tra $X - \sigma$ e $X + \sigma$ è data da

$$P(X, \sigma) = \int_{X-\sigma}^{X+\sigma} p_{X,\sigma}(x) dx = \frac{1}{\sqrt{2\pi} \sigma} \int_{X-\sigma}^{X+\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}} dx$$

o, passando al caso generale, la probabilità $P(X, t\sigma)$ tra $X - t\sigma$ e $X + t\sigma$, dove t è un parametro positivo variabile, sarà data da

$$P(X, t\sigma) = \int_{X-t\sigma}^{X+t\sigma} p_{X,\sigma}(x) dx = \frac{1}{\sqrt{2\pi} \sigma} \int_{X-t\sigma}^{X+t\sigma} e^{-\frac{(x-X)^2}{2\sigma^2}} dx$$

Il problema può essere semplificato operando un cambiamento di variabile da x a $z = (x - X)/\sigma$, che equivale a traslare la funzione di Gauss studiata in modo da centrarla intorno all'origine e a prendere σ come unità di misura delle deviazioni rispetto al valore vero X . Si ottiene quindi

$$P(X, t\sigma) = \int_{X-t\sigma}^{X+t\sigma} p_{X,\sigma}(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{z^2}{2}} dz \quad (22)$$

Questo cambiamento di variabile ci permette quindi di ricondurre il problema dello studio di una qualunque funzione di Gauss al caso di una funzione di Gauss di "riferimento" (avente $X = 0$ e $\sigma = 1$).

L'integrale riportato nella (22) è chiamato **funzione degli errori** ed indicato con **erf(t)**;

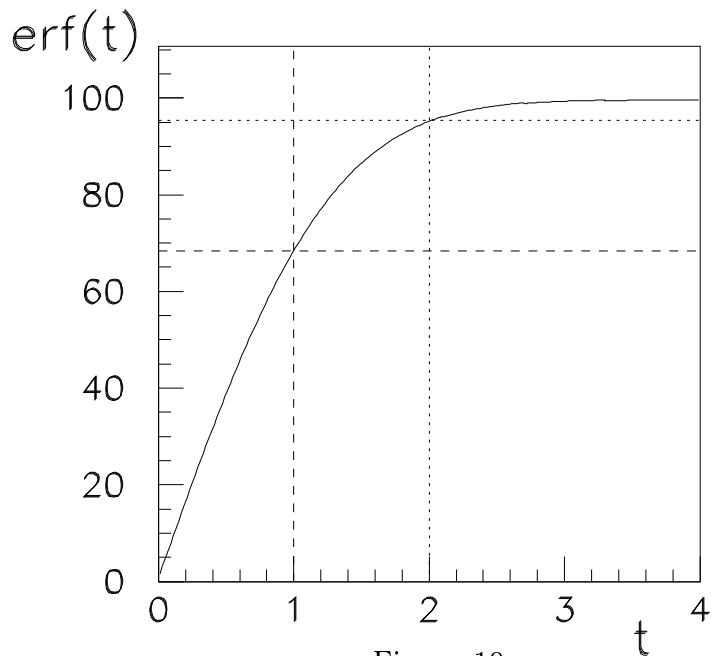


Figura 10:

esso non può essere valutato analiticamente e viene calcolato numericamente. In tabella e in Fig.10 sono riportati i valori di $erf(t)$ (in %) calcolati per vari valori di t .

t	0.00	0.02	0.04	0.06	0.08
0.0	0.00	1.60	3.19	4.78	6.38
0.1	7.97	9.55	11.13	12.71	14.28
0.2	15.85	17.41	18.97	20.51	22.05
0.3	23.58	25.10	26.61	28.12	29.61
0.4	31.08	32.55	34.01	35.45	36.88
0.5	38.29	39.69	41.08	42.25	43.81
0.6	45.15	46.47	47.78	49.07	50.35
0.7	51.61	52.85	54.07	55.27	56.46
0.8	57.63	58.78	59.91	61.02	62.11
0.9	63.19	64.24	65.28	66.29	67.29
1.0	68.27	69.23	70.17	71.09	71.99
1.1	72.87	73.73	74.57	75.40	76.20
1.2	76.99	77.75	78.50	79.23	79.95
1.3	80.64	81.32	81.98	82.62	83.24
1.4	83.85	84.44	85.01	85.57	86.11
1.5	86.64	87.15	87.64	88.12	88.59
1.6	89.04	89.48	89.90	90.31	90.70
1.7	91.09	91.46	91.81	92.16	92.49
1.8	92.81	93.12	93.42	93.71	93.99
1.9	94.26	94.51	94.76	95.00	95.23
2.0	95.45	95.66	95.86	96.06	96.25
2.1	96.43	96.60	96.76	96.92	97.07
2.2	97.22	97.36	97.49	97.62	97.74
2.3	97.86	97.97	98.07	98.17	98.27
2.4	98.36	98.45	98.53	98.61	98.69
2.5	98.76	98.83	98.89	98.95	99.01
2.6	99.07	99.12	99.17	99.22	99.26
2.7	99.31	99.35	99.39	99.42	99.46
2.8	99.49	99.52	99.55	99.58	99.60
2.9	99.93	99.65	99.67	99.69	99.71

A.6 Distribuzione di Poisson

In alcuni casi la precisione dell'apparato di misura non ha praticamente nessun effetto sulla misura, poichè la misura stessa mostra fluttuazioni di natura casuale.

Un tipico esempio riguarda il numero di particelle di polvere, o piccoli meteoriti, che urtano un satellite che viaggia nello spazio. Sebbene si possa parlare di densità media delle particelle, esse non saranno tutte spaziate tra loro di una stessa distanza. Talvolta due particelle saranno più vicine della media, talvolta più lontane. Di conseguenza durante ogni chilometro di cammino del satellite il numero di particelle che lo urtano varierà.

Abbiamo quindi una quantità fisica (il numero di urti) che ha una distribuzione di errore intrinseca.

Per descrivere tale effetto utilizzeremo ancora la distribuzione binomiale dandone una diversa interpretazione. Cercheremo di descrivere il processo utilizzando un modello in cui si considera la situazione del satellite ogni 10 metri del cammino e rappresenteremo ciò con il diagramma mostrato in Fig.11.

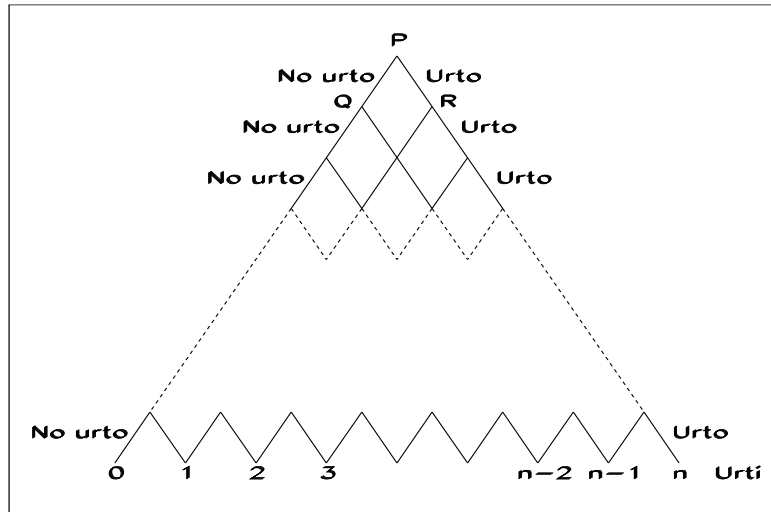


Figura 11: Diagramma degli urti con il satellite

Se non si ha nessuna collisione durante i primi 10 m il punto P si muove verso sinistra (se l'urto è avvenuto P si sposta verso destra). Q e R rappresentano la condizioni del satellite all'inizio dei secondi 10 m di percorso. Dopo 100 passi di questo tipo è stato coperto 1 km da parte del satellite e si raggiunge la linea finale del diagramma. Ogni posizione finale può essere in generale raggiunta da una varietà di cammini. In generale un viaggio del satellite che sia diviso in n stadi, dei quali r comportano urti, sarà descritto da una distribuzione binomiale del tipo

$$P(r; n, p) = \binom{n}{r} p^r q^{n-r} = \binom{n}{r} p^r (1 - p)^{n-r}$$

dove p è la probabilità di urto in un cammino di 10 m . Mentre nel caso della distribuzione di Gauss le due possibili scelte del diagramma erano equiprobabili, in questo caso, essendo bassa la densità delle particelle di polvere, la probabilità che avvenga un urto è chiaramente molto minore di quella di non avere urti. Avremo quindi:

$$p \ll 1 - p \quad \text{o} \quad p \ll 1$$

In quanto detto precedentemente non abbiamo però considerato la possibilità di avere due o più urti durante ognuno degli stadi. E' d'altra parte vero che, se la probabilità di avere un urto è piccola, quella di averne due o più è ancora più piccola (dell'ordine di p^2 , p^3 , ecc.). Ciò nonostante tale probabilità è finita e potremo ridurla ad un livello tale da poterla trascurare rendendo ciascuno stadio sempre più piccolo, ovvero n sempre

più grande. Al diminuire della lunghezza dello stadio diminuirà anche la probabilità di avere un urto in esso e ci si può aspettare che, una volta che la lunghezza sia diventata sufficientemente piccola, la probabilità p sia proporzionale alla lunghezza stessa. Ma la lunghezza di uno stadio è inversamente proporzionale a n e quindi il prodotto np sarà una costante che indicheremo con

$$\mu = np \tag{23}$$

Useremo ancora la formula di Stirling per dare un'espressione approssimata di $P(r; n, p)$; la distribuzione $P(r; n, p)$ sarà quindi ottenuta dalla distribuzione binomiale per $n \rightarrow \infty$ e corrispondentemente $p \rightarrow 0$ in modo che μ resti finito.

Con procedura analoga a quella sviluppata per la distribuzione di Gauss e tenendo conto della (23) si arriva all'espressione:

$$\ln P(r; n, p) \xrightarrow{n \rightarrow \infty} \ln P(r; \mu) = -\mu + r \ln \mu - \ln r!$$

da cui si ottiene

$$P(r; \mu) = \frac{e^{-\mu} \mu^r}{r!}$$

che è nota come distribuzione di Poisson.

Vediamo alcune proprietà della distribuzione di Poisson.

- la distribuzione è normalizzata in quanto

$$\sum_{r=0}^{\infty} \frac{e^{-\mu} \mu^r}{r!} = e^{-\mu} \sum_{r=0}^{\infty} \frac{\mu^r}{r!} = e^{-\mu} e^{\mu} = 1$$

- il numero medio di conteggi atteso, se ripetiamo il nostro esperimento di conteggio degli urti più volte, è dato da

$$\bar{r} = \sum_{r=0}^{\infty} r P(r; \mu) = \sum_{r=0}^{\infty} r \frac{e^{-\mu} \mu^r}{r!} = \mu$$

- la varianza della distribuzione è data da

$$\sigma^2 = \sum_{r=0}^{\infty} (r - \mu)^2 P(r; \mu) = \dots\dots\dots = \mu$$

Quindi la distribuzione di Poisson con conteggio medio μ ha deviazione standard $\sqrt{\mu}$. In altre parole, se facciamo una misura di un numero di eventi in un qualche intervallo di lunghezza (o di tempo) ed otteniamo il valore ν , allora il nostro risultato finale per il conteggio medio atteso in quell'intervallo è

$$\nu \pm \sqrt{\nu}$$

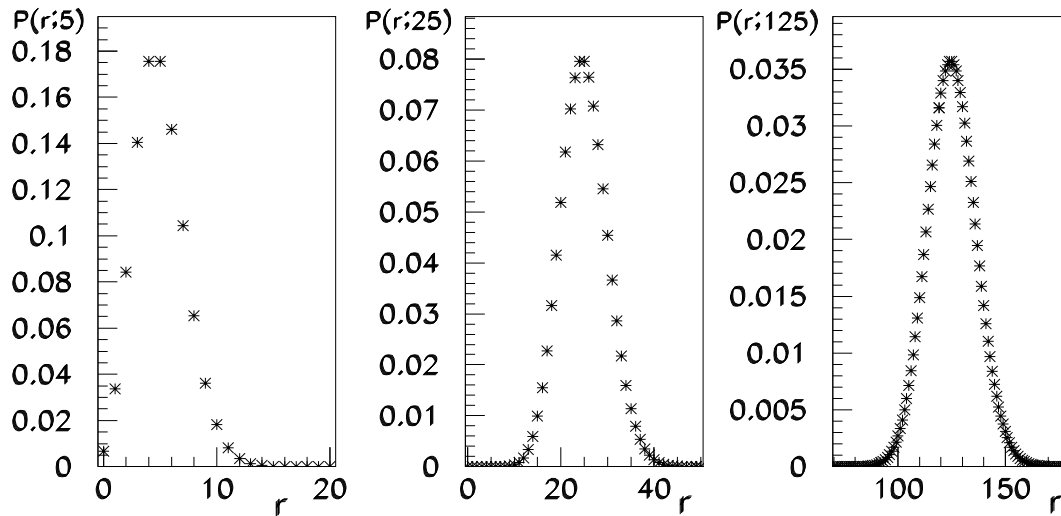


Figura 12: Distribuzioni di Poisson per $\mu = 5, 25$ e 125

La distribuzione di Poisson differisce dalla distribuzione di Gauss in 3 importanti aspetti:

- la distribuzione di Poisson si riferisce ad una variabile che assume valori discreti, r , e che può assumere valori positivi o nulli, mentre la distribuzione di Gauss si riferisce ad una variabile continua;
- la distribuzione di Poisson non è simmetrica: quando μ è piccolo l'asimmetria è molto accentuata e si attenua man mano che μ aumenta (vedi Fig.12);
- la distribuzione di Poisson è completamente determinata da un unico parametro μ mentre la distribuzione di Gauss è determinata una volta noti X e σ .

In Fig.12 sono mostrate le distribuzioni di Poisson che si ottengono nei casi in cui il numero medio dei conteggi atteso μ nell'intervallo assume i valori 5, 25 e 125.