

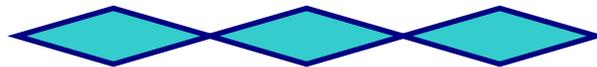
# ELABORAZIONE STATISTICA DI DATI SPERIMENTALI

Finora abbiamo accennato ad alcuni elementi di calcolo delle probabilità ed abbiamo visto alcune distribuzioni "teoriche" nel senso che si riferivano ad entità (eventi dello spazio S) che seguivano regole ben precise (probabilità note, variabili continue, numero infinito di valutazioni delle variabili aleatorie). In qualche caso si è accennato ad alcuni legami con la realtà.

La determinazione delle caratteristiche di tali eventi aleatori è argomento della statistica che, basandosi su un numero limitato di informazioni (misure) cerca di stimare i parametri delle distribuzioni di probabilità. Il nostro interesse è dovuto al fatto che tali metodologie consentono di dare valutazioni oggettive (in termini probabilistici) delle misure delle grandezze fisiche che risultano sempre soggette al caso.

Introduciamo alcune definizioni che pongono in correlazione il mondo delle probabilità con quello della statistica: presa la variabile aleatoria X possiamo immaginare di eseguirne infinite determinazioni (individui<sup>[1]</sup>) che rappresentano una popolazione statistica (popolazione madre) che segue la distribuzione di probabilità della variabile aleatoria X. Un sottoinsieme di N individui (con N finito<sup>[2]</sup>) di tale popolazione prende il nome di campione di grandezza N.

Lo scopo della statistica è quello di stimare i parametri della distribuzione di probabilità che regola l'intera popolazione dalla quale è stato estratto un campione di misure (**inferenza statistica**) e valutare in termini probabilistici in che misura i risultati di tale indagine, eseguita necessariamente su un campione finito, possano approssimare i valori veri (**stima delle incertezze**).



## INFERENZA STATISTICA

### VEROSIMIGLIANZA

Consideriamo una popolazione individuata dalla v.a. X che segue una distribuzione  $f(x)$  [da ora in poi indicheremo con  $f(x)$  sia distribuzioni di probabilità discrete che densità di probabilità continue] con media pari a  $m$  e varianza pari a  $\sigma^2$ .

Per costruire un campione si esegue una prima misura  $X_1$  (si estrae cioè un individuo dalla popolazione) e si associa ad essa il suo valore numerico  $x_1$ ; quindi si determina il valore della seconda misura  $X_2 = x_2$  e così fino a  $X_N = x_N$ . Il campione sarà quindi descritto dagli N valori assunti dagli N individui:  $[x_1, x_2, \dots, x_N]$  e su questi sarà definita la densità di probabilità congiunta

$$f(X_1=x_1 \text{ e } X_2=x_2 \text{ e } \dots \text{ e } X_N=x_N) = f(x_1, x_2, \dots, x_N).$$

<sup>1</sup> nel nostro caso gli individui sono risultati di misure; la popolazione madre è costituita dagli infiniti risultati possibili; il campione è costituito dal particolare insieme di risultati di una serie (finita) di misure

<sup>2</sup> la bontà delle conclusioni che possono trarsi dall'analisi di un campione dipende fortemente dalle modalità di estrazione di un campione dalla popolazione madre: si pensi ai diversi risultati di un'inchiesta relativa all'abolizione degli esami se l'indagine venisse svolta intervistando solo studenti o solo docenti.

Ovviamente se si eseguisse una nuova serie di misure si otterrebbe un diverso campione costituito da  $[x_1', x_2', \dots, x_N']$  a causa del carattere intrinsecamente aleatorio dell'operazione di estrazione di un individuo e quindi una diversa  $f(x_1', x_2', \dots, x_N')$ . Questa probabilità è detta verosimiglianza: indica quanto è verosimile che si ottenga il campione effettivamente ottenuto.

La motivazione di molte delle formule statistiche incontrate in questo corso si basa sul principio della massima verosimiglianza: la probabilità di ottenere i risultati effettivamente ottenuti deve essere elevata (la massima, in questo metodo) altrimenti sarebbero stati ottenuti altri risultati.

Senza entrare in dettagli limitiamoci a considerare due esempi per intuire la potenza del metodo della massima verosimiglianza:

**STIMA DELLA PROBABILITÀ IN UN PROCESSO BINOMIALE**

Supponiamo di aver ottenuto  $k$  successi su  $N$  prove in un esperimento che si pensa sia descritto da una binomiale. Vogliamo stimare la probabilità  $p$  associata al singolo evento.

La probabilità associata ai  $k$  successi ottenuti è  $P_{N,p}(k) = \binom{N}{k} p^k (1-p)^{N-k}$ ;

per trovare il massimo deriviamo rispetto a  $p$ :

$$\frac{\partial P_{N,p}(k)}{\partial p} = \binom{N}{k} [k p^{k-1} (1-p)^{N-k} + p^k (-N+k)(1-p)^{N-k-1}] = \binom{N}{k} p^{k-1} (1-p)^{N-k-1} [k(1-p) + p(-N+k)]$$

e uguagliando a 0 si ottiene  $k - N p = 0$  cioè  $p = \frac{k}{N} = \frac{\text{numero di successi ottenuti}}{\text{numero delle prove effettuate}}$

Si vede quindi che, in questo caso, la migliore stima per la probabilità non è nient'altro che la frequenza relativa dell'evento considerato.

**STIMA DEL VALORE ATTESO DI UNA GAUSSIANA A PARTIRE DA N MISURE**

Se nell'eseguire le misure non è stato alterato il sistema in esame<sup>[3]</sup> possiamo ritenere che le  $N$  determinazioni siano indipendenti e quindi la funzione di verosimiglianza  $f(x_1, x_2, \dots, x_N)$ , che è una probabilità congiunta, diventa pari al prodotto delle probabilità e quindi può essere riscritta come  $f(x_1)f(x_2) \dots f(x_N)$ . Nel caso di una distribuzione gaussiana la verosimiglianza vale:

$$\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_1-m)^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_2-m)^2}{2\sigma^2}} \times \dots \times \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_N-m)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^N e^{-\sum_{i=1,N} \frac{(x_i-m)^2}{2\sigma^2}}$$

Massimizzare la verosimiglianza corrisponde quindi a minimizzare  $\sum_{i=1,N} \frac{(x_i - m)^2}{2\sigma^2}$  al variare di  $m$ :

ponendo pari a 0 la derivata  $\frac{\partial}{\partial m} \sum_{i=1,N} (x_i - m)^2 = \sum_{i=1,N} 2(x_i - m)(-1)$ , si ottiene  $\sum_{i=1,N} (x_i - m) = 0$  e

quindi  $\sum_{i=1,N} x_i - N m = 0$  cioè  $m = \frac{\sum_{i=1,N} x_i}{N}$  !!!

<sup>3</sup> come spesso accade almeno in prima approssimazione. In realtà l'errore di inserzione è ineliminabile e quindi esiste sempre una pur debolissima correlazione fra le misure

## STIMA STATISTICA DEI PARAMETRI DI UNA DISTRIBUZIONE

Come detto, lo scopo della statistica è quello di determinare dall'analisi di un campione (una serie di misure) i parametri dell'intera popolazione (il valore vero della grandezza fisica in misura). Si tratta in ultima analisi di determinare la funzione degli  $N$  valori di un campione che meglio approssima il valore di tali parametri.

È possibile ottenere dallo stesso campione di grandezza  $N$  più stime  $\theta_N(q)$  di un particolare parametro  $q$ ; fra queste sono preferibili le stime:

- **consistenti**: quelle in cui la stima, al crescere della dimensione del campione, converge in probabilità verso il valore del parametro della popolazione:

$$\lim_{N \rightarrow \infty} P(|\theta_N(q) - q| < \varepsilon) = 1$$

Questa condizione assicura che al crescere della dimensione del campione, è praticamente certo che la stima del parametro disti dal parametro stesso per una quantità arbitrariamente piccola (del resto per  $N \rightarrow \infty$  il campione coincide con l'intera popolazione).

- **incondizionate** (o non distorte): quelle in cui il valore atteso della stima coincide col parametro indipendentemente dalla grandezza del campione:

$$E[\theta_N(q)] = q \quad \text{per ogni } N$$

Questa condizione assicura che se la stima del parametro venisse eseguita su più campioni di grandezza diversa, i risultati delle stime non si discosterebbero fra loro a causa della diversa grandezza dei campioni.

Vediamo ora come procedere per stimare da un campione alcune caratteristiche e riassunti della distribuzione della popolazione madre che segue una distribuzione con riassunti  $m$  e  $\sigma^2$ :

### STIMA DELLA PROBABILITÀ $p$ DI UN EVENTO

Dato un evento come è possibile stimare la probabilità  $p$  che esso ha di realizzarsi? Eseguiamo un numero  $N$  elevato di prove e contiamo quante volte  $k$  si realizza l'evento.

Calcoliamo la frazione di volte  $f = \frac{k}{N}$  (frequenza relativa) in cui l'esito è stato positivo:

la frequenza relativa è una stima della probabilità di realizzarsi dell'evento:

$$p \approx f = \frac{k}{N}$$

Questa stima è consistente in quanto  $\lim_{N \rightarrow \infty} P\left(\left|\frac{k}{N} - p\right| < \varepsilon\right) = 1$  (è il teorema di Bernoulli ...);

è anche non condizionata perché  $E\left(\frac{k}{N}\right) = \frac{E(k)}{N} = \frac{Np}{N} = p$

(il valore atteso di  $k$  si ottiene dalla distribuzione binomiale che valuta la probabilità che in  $N$  tentativi si ottengano  $k$  successi nell'ipotesi che ciascun evento abbia probabilità  $p$  di successo).

Come esempio di stima della probabilità di un evento si riveda quello del lancio di una moneta utilizzato nella definizione frequentistica della probabilità.

**STIMA DELLA MEDIA m DI UNA DISTRIBUZIONE**

Nel caso siano state ottenute N determinazioni  $x_i$  nelle **stesse condizioni** di misura, la

**miglior stima di m è la media aritmetica** delle N osservazioni:  $m \approx \bar{X} = \frac{\sum_{i=1,N} x_i}{N}$

- la media aritmetica è una stima consistente della media:

utilizziamo le note relazioni:  $E(\bar{X}) = m$ ,  $\sigma(\bar{X}) = \frac{\sigma(X)}{\sqrt{N}}$ ;

applicando la disuguaglianza di Chebychev alla variabile aleatoria  $\bar{X}$ :

$$P\left(|\bar{X} - m| \leq K \frac{\sigma(X)}{\sqrt{N}}\right) \geq 1 - 1/K^2 .$$

Passando al limite per  $N \rightarrow \infty$ , posto  $\epsilon = K \frac{\sigma(X)}{\sqrt{N}}$ , si ottiene  $\lim_{N \rightarrow \infty} P(|\bar{X} - m| \leq \epsilon) \geq 1 - \frac{\sigma^2(X)}{N \epsilon^2}$

e quindi, per K sufficientemente elevato, la verifica della consistenza della stima.

- la media aritmetica è una stima non distorta della media:

$$E(\bar{X}) = E\left(\frac{\sum_{i=1,N} X_i}{N}\right) = \frac{\sum_{i=1,N} E(X_i)}{N} = \frac{N E(X_i)}{N} = \frac{N m}{N} = m$$

$E(X_i) = m$  perché ricordiamo che le  $X_i$  del campione hanno tutte la stessa distribuzione della v.a. X della popolazione madre.

Si può notare l'analogia fra media e media aritmetica considerando che le frequenze relative  $f_j$  sono una stima delle probabilità  $P(X_j)$ :

$$\bar{X} = \frac{\sum_{j=1,v} n_j X_j}{N} = \sum_{j=1,v} \frac{n_j}{N} X_j = \sum_{j=1,v} f_j X_j \approx \sum_{j=1,v} P(X_j) X_j = E(X) = m$$

Nel caso in cui le N determinazioni  $x_i$  di  $X_i$  non siano state ottenute nelle stesse condizioni di misura, e quindi abbiano **varianze diverse**, la migliore stima di m è data dalla **media pesata**<sup>[4]</sup>

$$m \approx \frac{\sum \frac{x_i}{\sigma^2(X_i)}}{\sum \frac{1}{\sigma^2(X_i)}}$$

È sufficiente notare che se le varianze fossero tutte uguali fra di loro la formula della media pesata si trasformerebbe in quella della media aritmetica<sup>[5]</sup>.

<sup>4</sup> è possibile ricavarne l'espressione anche col metodo dei minimi quadrati

<sup>5</sup> sapete verificarlo? E nel caso di varianze uguali cosa vi aspettate succeda dell'incertezza da associare alla media pesata? Verificate anche questo: è molto istruttivo

**STIMA DELLA VARIANZA  $\sigma^2$  DI UNA DISTRIBUZIONE**

La **miglior stima di  $\sigma^2$**  è la **varianza sperimentale** (il quadrato della deviazione standard sperimentale)

$$\sigma^2 \approx \sigma_s^2(X) = \frac{\sum_{i=1,N} (x_i - \bar{X})^2}{N-1}$$

- la consistenza della stima della varianza si può intuire riscrivendo la varianza sperimentale:

$$\sigma_s^2(X) = \frac{\sum_{j=1,v} n_j (x_i - \bar{X})^2}{N-1} \text{ come } \sum_{j=1,v} \frac{n_j}{N-1} (x_i - \bar{X})^2$$

e considerando che per  $N \rightarrow \infty$  si ha che  $\bar{X} \rightarrow m$  e  $\frac{n_j}{N-1} \rightarrow \frac{n_j}{N} \rightarrow f \rightarrow p$  e quindi

$$\lim_{N \rightarrow \infty} \sigma_s^2(X) = \lim_{N \rightarrow \infty} \frac{\sum_{j=1,v} n_j (x_i - \bar{X})^2}{N-1} = \sum_{j=1,v} P_j (X_i - m)^2 = \sigma^2(X)$$

Più rigorosamente, considerando che la quantità  $\frac{\sum_{i=1,N} (x_i - \bar{X})^2}{\sigma^2(X)}$  segue la distribuzione del  $\chi^2$  con N-1

gradi di libertà e quindi ha media N-1 e varianza 2(N-1) si può ricorrere ancora alla disuguaglianza di Chebychev:

$$P\left(\left|\frac{\sum (x_i - \bar{X})^2}{\sigma^2} - (N-1)\right| < K\sqrt{2(N-1)}\right) \geq 1 - 1/K^2 \text{ che con qualche passaggio diventa:}$$

$$P\left(\left|\frac{\sum (x_i - \bar{X})^2}{N-1} - \sigma^2\right| < K \frac{\sigma^2 \sqrt{2(N-1)}}{N-1}\right) \geq 1 - 1/K^2$$

Passando al limite per  $N \rightarrow \infty$ , per K sufficientemente elevato si verifica la consistenza della stima.

- la stima della varianza non è condizionata:

se la varianza venisse stimata con  $\sigma_s'^2(X) = \frac{\sum_{i=1,N} (x_i - \bar{X})^2}{N}$  si otterrebbe per il valore atteso

$$\text{di } \sigma_s'^2(X): \quad E[\sigma_s'^2(X)] = E\left[\frac{\sum (x_i - \bar{X})^2}{N}\right] = E\left(\frac{\sum x_i^2}{N} - \bar{X}^2\right) = \frac{\sum E(x_i^2)}{N} - E(\bar{X}^2) \text{ dove}$$

il primo termine è pari a  $E(x_i^2) = \sigma^2(X) + m^2$  e il secondo è  $E(\bar{X}^2) = \frac{\sigma^2(X)}{N} + m^2$

$$\text{pertanto } E[\sigma_s'^2(X)] = \sigma^2(X) - \frac{\sigma^2(X)}{N} = \frac{N-1}{N} \sigma^2(X).$$

Definendo invece come stima di  $\sigma^2(X)$  la quantità  $\sigma_s^2(X) = \frac{N-1}{N} \sigma_s'^2(X) = \frac{\sum_{i=1,N} (x_i - \bar{X})^2}{N-1}$

si ottiene una stima non distorta:  $E[\sigma_s^2(X)] = E\left[\frac{N}{N-1} \sigma_s'^2(X)\right] = \sigma^2(X)$ .

In generale, per ottenere stime non distorte dovremo dividere non per N ma per il numero di v.a. indipendenti (gradi di libertà del problema) cioè N meno il numero di relazioni statistiche fra le diverse grandezze.

Per esempio nel caso di  $\sigma_s^2(X)$  l'uso di  $\bar{X}$  necessario per stimare m impone un vincolo fra le N variabili e quindi, fissate arbitrariamente le prime N-1, l'N-sima dipende numericamente (e non più solo probabilisticamente) dalle altre variabili.

Abbiamo visto che nel caso in cui le N determinazioni  $x_i$  di  $X_i$  non siano state ottenute nelle stesse condizioni di misura, e quindi abbiano **varianze diverse**<sup>[6]</sup>, la migliore stima di m è data dalla **media pesata**. In questo caso la varianza è stimata con l'espressione:

$$\sigma_s^2(\bar{X}_i) = \frac{1}{\sum \frac{1}{\sigma_s^2(X_i)}}$$

La varianza della media pesata risulta quindi matematicamente inferiore alla più piccola delle varianze  $\sigma_s^2(X_i)$ : ogni misura, anche se affetta da grosse incertezze, contiene informazioni importanti; il suo uso aumenta la conoscenza del fenomeno riducendo la varianza finale.

Come esempio valutiamo la media pesata delle misure:

127,4 ± 1,5  
 131,1 ± 1,3  
 129,2 ± 1,0

$$\bar{X} = \frac{\frac{127,4}{1,5^2} + \frac{131,1}{1,3^2} + \frac{129,2}{1,0^2}}{\frac{1}{1,5^2} + \frac{1}{1,3^2} + \frac{1}{1,0^2}} \pm \frac{1}{\sqrt{\frac{1}{1,5^2} + \frac{1}{1,3^2} + \frac{1}{1,0^2}}} = 129,36 \pm 0,70$$

Bisogna prestare attenzione al valore delle misure prima di utilizzare la media pesata: la teoria si basa sul fatto che le diverse misure si riferiscono alla stessa grandezza fisica e quindi siano trascurabili gli effetti sistematici. Diversamente l'uso della media pesata non è lecito.

<sup>6</sup> in cosa vi aspettate che si trasformi questa formula nel caso di varianze tutte uguali fra loro? ... provate

## STUDIO DELLA DIPENDENZA FUNZIONALE DI GRANDEZZE FISICHE

Supponiamo di voler studiare un sistema fisico ma di conoscerlo già a un livello tale da poter prevedere che se viene sottoposto alla sollecitazione X esso produrrà una risposta Y secondo a legge  $Y = f(X; q_1, q_2, \dots, q_m)$  dove  $q_i$  sono i parametri incogniti, scopo della nostra misura. Questo caso è assai frequente perché spesso lo studio di un sistema più o meno complesso si riduce alla determinazione della sua risposta a sollecitazioni note.

Si ricorre però a questo metodo anche quando una sola coppia di misure sarebbe in grado di fornire la risposta cercata perché lo studio della risposta a sollecitazioni diverse consente di evidenziare sia la presenza di errori sistematici e casuali sia eventuali anomalie nel comportamento del sistema.

Il **METODO DEI MINIMI QUADRATI** (che trova la sua giustificazione anche nel principio della massima verosimiglianza) consente di determinare statisticamente i valori dei parametri incogniti nella dipendenza funzionale.

In particolare, limitandoci al caso di una **dipendenza lineare**:

$$\boxed{Y = p X + q}$$
 con pendenza p e intercetta q incognite

è sufficiente, per ogni valore  $X_i$  della sollecitazione X (con  $i = 1, N$  valori diversi) eseguire una misura  $Y_i$  della grandezza Y ed eventualmente ottenere una stima della sua varianza  $\sigma_i^2$ .

●●● L'applicazione del metodo<sup>[7]</sup> fornisce le stime:

$$p_s = \frac{\sum \frac{X_i Y_i}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{Y_i}{\sigma_i^2}}{\sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2}}$$

$$\sigma^2(p_s) = \frac{\sum \frac{1}{\sigma_i^2}}{\left( \sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2} \right)}$$

$$q_s = \frac{\sum \frac{X_i^2}{\sigma_i^2} \sum \frac{Y_i}{\sigma_i^2} - \sum \frac{X_i Y_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2}}{\sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2}}$$

$$\sigma^2(q_s) = \frac{\sum \frac{X_i^2}{\sigma_i^2}}{\left( \sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2} \right)}$$

Per verificare se queste stime siano condizionate o meno calcoliamo il valore atteso di  $p_s$  ricordando che  $E(Y_i) = p X_i + q$ <sup>[8]</sup>:

$$E(p_s) = \frac{\sum \frac{X_i E(Y_i)}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{E(Y_i)}{\sigma_i^2}}{\sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2}} = \frac{\sum \frac{p X_i^2 + q X_i}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{p X_i + q}{\sigma_i^2}}{\sum \frac{X_i^2}{\sigma_i^2} \sum \frac{1}{\sigma_i^2} - \sum \frac{X_i}{\sigma_i^2} \sum \frac{X_i}{\sigma_i^2}} = p.$$

Analogamente si trova  $E(q_s) = q$ ; le stime dei minimi quadrati sono quindi non condizionate (non distorte).

<sup>7</sup> vedi PROBABILITÀ

<sup>8</sup> le formule sono state ricavate sotto l'ipotesi che l'incertezza relativa alle determinazioni  $X_i$  sia trascurabile

●●● Spesso per le misure di Y avviene che lo stesso operatore utilizzi lo stesso metodo e gli stessi strumenti. In questo caso è lecito ritenere che gli errori di misura e quindi le varianze delle  $Y_i$  siano tutte uguali fra loro<sup>9</sup>.

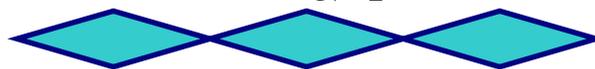
Le **formule si semplificano** (basta porre  $\sigma_i^2 = \sigma^2$ ) e diventano:

$$p_s = \frac{N \sum X_i Y_i - \sum X_i \sum Y_i}{N \sum X_i^2 - \sum X_i \sum X_i} \quad \sigma^2(p_s) = \frac{N}{N \sum X_i^2 - \sum X_i \sum X_i} \sigma^2$$

$$q_s = \frac{\sum X_i^2 \sum Y_i - \sum X_i Y_i \sum X_i}{N \sum X_i^2 - \sum X_i \sum X_i} \quad \sigma^2(q_s) = \frac{\sum X_i^2}{N \sum X_i^2 - \sum X_i \sum X_i} \sigma^2$$

●●● Nelle espressioni delle varianze compare la quantità  $\sigma^2$  che spesso **non è nota** (tipicamente si conosce solo il contributo determinato dalla strumentazione che è spesso un valore fortemente sottostimato). In questi casi è possibile determinarla dalle misure già ottenute:

$$\sigma^2 = \frac{\sum [Y_i - (p_s X_i + q_s)]^2}{N - 2}$$



**ESEMPIO**

Supponiamo di aver eseguito delle misure del periodo di oscillazione di un pendolo verticale in funzione della massa applicata all'estremità della molla che lo costituisce. Poiché esiste una relazione lineare fra il quadrato del periodo di oscillazione e tale massa si può valutare col metodo dei minimi quadrati la costante di proporzionalità fra  $T^2$  e  $M$ :  $T^2 = p M + q$ .

Sono state ottenute in laboratorio le seguenti 6 coppie di misure:

i	$M_i$ [kg]	$\bar{T}_i^2$ [s <sup>2</sup> ]	$\sigma(\bar{T}_i^2)$ [s <sup>2</sup> ]
1	0,400	0,388 00	0,000 75
2	0,480	0,459 7	0,001 8
3	0,560	0,530 7	0,002 0
4	0,640	0,599 70	0,000 93
5	0,720	0,664 2	0,002 2
6	0,800	0,735 3	0,002 3

Considerando che le incertezze delle misure dei quadrati dei periodi sono diverse fra loro decidiamo di utilizzare le formule più complete.

A partire dalle misure si calcolano le quantità:

$$p_s = 0,873\ 996\ s^2/kg \quad \text{con} \quad \sigma_{p_s} = 0,003\ 89\ s^2/kg;$$

$$q_s = 0,039\ 007\ s^2 \quad \text{con} \quad \sigma_{q_s} = 0,002\ 10\ s^2$$

e quindi

$$p_s = (0,874\ 0 \pm 0,003\ 9)\ s^2/kg$$

e

$$q_s = (0,039\ 0 \pm 0,002\ 1)\ s^2.$$

<sup>9</sup> non si tratta di un caso raro, anzi, nella grande maggioranza dei casi le incertezze non differiscono fra loro per più di un fattore 2-3 consentendo l'uso delle formule semplificate

- Se consideriamo più attentamente i valori delle  $\sigma(\bar{T}_i^2)$  notiamo che entro un fattore 2-3 sono simili e valgono circa  $\sigma(\bar{T}_i^2) \approx 0,0015 \text{ s}^2$ .

Eseguiamo nuovamente il calcolo considerando il caso di incertezze uguali fra loro ottenendo:

$$\begin{aligned}
 p_s &= 0,863\,929 \text{ s}^2/\text{kg} & \text{con } \sigma_{p_s} &= 0,004\,482 \text{ s}^2/\text{kg} \\
 q_s &= 0,044\,576 \text{ s}^2 & \text{con } \sigma_{q_s} &= 0,002\,758 \text{ s}^2 \\
 \text{e quindi} & & p_s &= (0,863\,9 \pm 0,004\,5) \text{ s}^2/\text{kg} \\
 & & \text{e } q_s &= (0,044\,6 \pm 0,002\,8) \text{ s}^2.
 \end{aligned}$$

Questi valori sono sostanzialmente equivalenti ai precedenti ma il loro calcolo è assai più rapido.

- Infine supponiamo, come spesso accade, di non aver potuto misurare le varianze relative ai periodi e di aver potuto considerare solo il contributo minimale dato dalla strumentazione. In questo caso, non avendo a disposizione le  $\sigma(\bar{T}_i^2)$  dalle misure:

i	$M_i[\text{kg}]$	$\bar{T}_i^2[\text{s}^2]$
1	0,400	0,388 0
2	0,480	0,459 7
3	0,560	0,530 7
4	0,640	0,599 7
5	0,720	0,664 2
6	0,800	0,735 3

si possono calcolare le quantità:

$$\begin{aligned}
 \sigma(\bar{T}_i^2) &= 0,002\,394 \text{ s}^2 \text{ anziché il precedente } \approx 0,0015 \text{ s}^2 \\
 p_s &= 0,863\,929 \text{ s}^2/\text{kg} & \text{con } \sigma_{p_s} &= 0,006\,854 \text{ s}^2/\text{kg} \\
 q_s &= 0,044\,576 \text{ s}^2 & \text{con } \sigma_{q_s} &= 0,004\,218 \text{ s}^2 \\
 \text{e quindi} & & p_s &= (0,863\,9 \pm 0,006\,8) \text{ s}^2/\text{kg} \\
 & & \text{e } q_s &= (0,044\,6 \pm 0,004\,2) \text{ s}^2.
 \end{aligned}$$

**Se non si possiede una calcolatrice già programmata per stimare la retta di regressione conviene calcolare preventivamente (accumulando<sup>10</sup> i dati in opportuni registri della calcolatrice o nelle celle di un foglio di calcolo elettronico) le quantità:**

$$\begin{aligned}
 N &= 6 & \Sigma X_i &= 3,6 \text{ kg} & \Sigma Y_i &= 3,3776 \text{ s}^2 \\
 \Sigma X_i Y_i &= 2,12332 \text{ kg s}^2 & \Sigma X_i^2 &= 2,272 \text{ kg}^2 & \Sigma Y_i^2 &= 1,984978 \text{ s}^4 \\
 \text{il denominatore di } p_s \text{ e } q_s: D &= 0,672 \text{ kg}^2 & \text{la deviazione standard delle } Y_i & \sigma &= 0,002\,394 \text{ s}^2
 \end{aligned}$$

••• **Considerando:**

- 1) il notevole risparmio di tempo introdotto dal secondo approccio,
- 2) il fatto che sostanzialmente i risultati coincidono

è naturale, almeno quando le precisioni richieste non sono particolarmente elevate, considerare la formula con le varianze uguali e ricavare la varianza della variabile dipendente direttamente dai dati.

<sup>10</sup> se, non avendo acquisito pratica col vostro strumento di calcolo, riportate su carta i risultati parziali fate attenzione a non introdurre approssimazioni pesanti limitando il numero di cifre significative

## TABELLE E GRAFICI

La potenza del metodo dei minimi quadrati ha lo svantaggio di richiedere il calcolo di formule che potrebbero distogliere l'attenzione degli studenti dagli scopi del corso: durante le esperienze ritengo che in laboratorio sia più importante concentrarsi sul metodo sperimentale che non sul rigore dell'elaborazione statistica dei dati. Tuttavia in laboratorio è anche di importanza fondamentale avere rapidamente una valutazione, anche se rozza, della qualità dei dati che si stanno raccogliendo.

Inoltre non è raro che si commetta uno sbaglio nell'immettere un dato nello strumento di calcolo ottenendo un risultato sbagliato. Sarebbe utile poter verificare rapidamente se il risultato è approssimativamente corretto; a questo punto sarà più affidabile considerare corretto il risultato calcolato.

Segue la discussione di un metodo approssimativo che consentirà di ottenere tali informazioni mediante pochissimi calcoli; durante la stesura della relazione ci sarà poi tempo a sufficienza per elaborare rigorosamente i risultati.

Per fissare le idee consideriamo un esempio: supponiamo di avere eseguito una serie di misure di allungamenti  $Y_i$  di una molla sottoposta a sollecitazioni crescenti mediante l'aggiunta di pesetti di massa  $X_i$  su un piattello.

Molto probabilmente gli allungamenti  $Y_i$  avranno le stesse incertezze perché per misurarli sarà stato utilizzato lo stesso strumento e lo stesso metodo di misura.

### INIZIAMO COL RIPORTARE LE MISURE



## TABELLE

Il modo più razionale di riportare molti dati è in forma di tabella perché permette confronti per colonne e/o righe, controlli di coerenza, di tendenze, etc.

Alcune raccomandazioni dettate in parte dalle convenzioni e in parte dalla praticità nell'uso:

**- ogni misura va riportata con l'unità di misura e col corretto numero di cifre significative:**

se si tratta di misure dirette:

fino al decimo di divisione per letture di strumenti analogici;

tutte le cifre per letture di strumenti digitali

riportando anche gli eventuali zeri terminali

se si tratta di misure derivate:

se è stata già valutata l'incertezza

lo stesso numero di cifre decimali dell'incertezza riportata con due cifre significative

se non è stata ancora valutata l'incertezza

un numero sufficiente di cifre significative che non richieda successivamente la ripetizione dei calcoli (ma senza esagerare<sup>[11]</sup>)

- indicazioni ausiliarie comuni a tutti i dati di una colonna (p.es.  $10^n$  o l'unità di misura) vanno riportate sull'intestazione: non devono ingombrare la tabella;
- evitare colonne di dati uguali; ricordarne l'esistenza in legenda;
- conviene impostare le tabelle in forma aperta (possibilità di aggiungere colonne/righe).
- evidenziare eventuali dati non graficati motivandone l'esclusione;

<sup>11</sup> difficilmente in laboratorio avrete a che fare con incertezze inferiori allo 0,1 % che implica al più 5 cifre significative

È **particolarmente importante** che almeno i dati graficati siano riportati in una tabella. In questo modo le informazioni per realizzare il grafico andranno prese dalla tabella e non ricercate all'interno della relazione. Inoltre, anche se non strettamente necessarie per la realizzazione del grafico, è importante che nella tabella compaiano anche le incertezze delle grandezze graficate: se un punto risultasse non allineato sarebbe assai rapido verificare se la discordanza è compatibile con l'incertezza della misura o è sinonimo di qualche altro effetto.

**GRAFICHIAMO LA SERIE DELLE N COPPIE DI MISURE  $Y_i$  IN FUNZIONE DI  $X_i$**



## GRAFICI

Anche se per la realizzazione professionale di un grafico è consigliabile l'utilizzo di opportuni programmi di calcolo e grafica, è fondamentale una fase di apprendimento che non può prescindere dalla loro realizzazione manuale. Per questo scopo sono a disposizione fogli di carta con divisioni ogni millimetro (carta millimetrata). Per il loro uso si consiglia:

- le scale sugli assi devono essere scelte in modo da **garantire la leggibilità delle coordinate di un qualsiasi punto** posto sul grafico (anche diverso dai punti risultati di misure);
- **non è indispensabile che il grafico si estenda a tutto il foglio** ma deve risultare leggibile
- per praticità d'uso le scale devono essere **multiple di 1, 2, 5 millimetri** (evitare, se possibile, 4; MAI 3, 7, 9 o valori che richiedano una calcolatrice)
- se non richiesto dal particolare problema in esame non è necessario che le scale inizino dall'origine
- ogni asse deve riportare il **simbolo della grandezza, l'unità di misura**, l'eventuale fattore  $10^n$
- su ogni asse vanno riportate a intervalli regolari poche (5-10) tacche con l'indicazione del valore: dato che il foglio è millimetrato non va riportata sul foglio la scala
- per non rendere difficoltosa ogni successiva elaborazione **non riportare sugli assi i valori delle misure e non collegare i punti** tra loro o con gli assi
- è utile<sup>[12]</sup> riportare sul grafico la relazione (titolo del grafico) che ci si attende occorrere fra le quantità riportate sugli assi

Graficata la serie si può tracciare con un righello la retta che secondo noi è mediamente equidistante da tutti i punti graficati. Poiché gli N punti su questo grafico rappresentano delle misure (quindi affette da errori) la retta (funzione analitica) non passa per tutti i punti sperimentali neanche se la nostra schematizzazione della legge fisica è corretta.

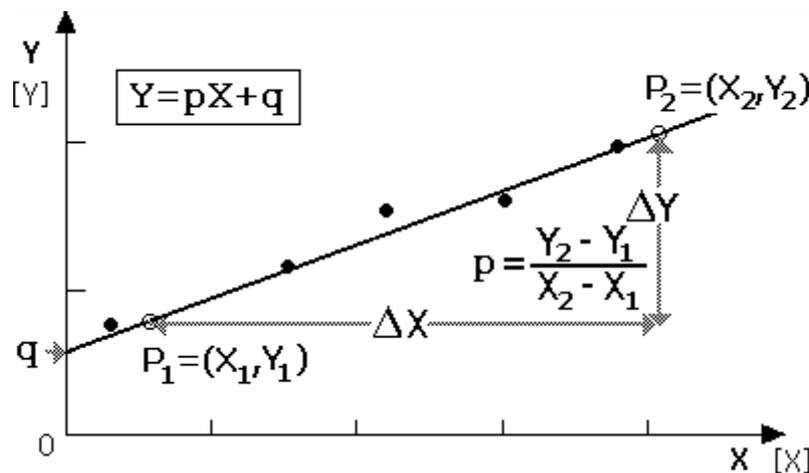
Dopo qualche tentativo si è però in grado di tracciare una retta che **non si discosta<sup>[13]</sup> per più di qualche percento da quella ottenibile mediante il metodo dei minimi quadrati** che si basa sullo stesso principio della minimizzazione<sup>[14]</sup> delle distanze.

<sup>12</sup> si tratta di un eufemismo: è obbligatorio (all'interno di questo corso, per fini didattici) perché consente di comprendere rapidamente il significato fisico del grafico. In questo corso ci limitiamo a studiare relazioni lineari. Non è una grossa limitazione dato che è spesso possibile trasformare relazioni non lineari in lineari (linearizzazione) graficando sugli assi opportune funzioni delle grandezze misurate

<sup>13</sup> sarebbe molto istruttivo se provaste a verificare quanto detto

<sup>14</sup> non è esattamente la stessa minimizzazione per due motivi, sapete individuarli ?

Un grafico delle N coppie di misure si presenterebbe così dopo aver tracciato la retta:  $Y = p X + q$  della quale per il momento ancora non sono noti i valori di  $p$  e  $q$ <sup>[15]</sup>



**STIMA DEI PARAMETRI DELLA RETTA**



A questo punto non resta che determinare il valore della pendenza<sup>[16]</sup>  $p$  e dell'intercetta  $q$ .

Per la pendenza  $p$  si prendono due punti sulla retta quanto più distanti possibile al fine di minimizzare l'effetto degli errori di lettura delle coordinate dal grafico.

Evidenziate, ad esempio cerchiandoli, i due punti e riportate i valori delle coordinate  $[x_1, y_1]$  e  $[x_2, y_2]$  (con le unità di misura!) al fine di calcolare la pendenza come rapporto<sup>[17]</sup> fra la differenza delle ordinate e la differenza delle ascisse:  $p = \frac{y_2 - y_1}{x_2 - x_1}$ .

L'intercetta  $q$  si legge direttamente dal grafico:  $q = Y(X=0)$ .

Per quel che riguarda il **calcolo delle incertezze** ...

Il metodo grafico per la stima dei parametri ottenuti consente di non utilizzare le formule dei minimi quadrati ma per stimare l'incertezza dei parametri ciò non è in generale evitabile. Tuttavia, se è sufficiente un'elaborazione approssimativa, si può snellire il calcolo. La prima approssimazione consiste nel considerare, se ragionevole, le incertezze tutte uguali fra loro (ma incognite dato che spesso delle grandezze non è nota l'incertezza). Le formule da utilizzare diventano quindi<sup>[18]</sup>:

$$\sigma_{ps} = \frac{\sigma_Y}{\sqrt{N} \sigma_X}$$

$$\sigma_{qs} = \frac{\sigma_Y}{\sqrt{N} \sigma_X} \sqrt{\sigma_X^2 + \bar{X}^2}$$

dove 
$$\sigma_X^2 = \frac{\sum X_i^2}{N} - \left( \frac{\sum X_i}{N} \right)^2$$

$$\sigma_Y^2 = \frac{\sum [Y_i - (p_s X_i + q_s)]^2}{N - 2}$$

<sup>15</sup> da questo momento in poi va dimenticata l'esistenza dei singoli punti: la migliore rappresentazione del fenomeno studiato è la retta che è stata tracciata

<sup>16</sup> analiticamente si parla di coseni direttori o di coefficienti angolari ma in questo caso c'è da notare che sugli assi sono riportate unità di misura in generale diverse e addirittura di grandezze fisiche diverse. L'angolo che si potrebbe misurare con un goniometro cambierebbe se lo stesso grafico venisse ripetuto con scale diverse: tale angolo non ha nessuna relazione col significato fisico del parametro  $p$ . Per evitare confusione si preferisce parlare di pendenza (con dimensioni fisiche pari a quelle di  $Y/X$ ).

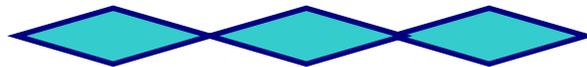
<sup>17</sup> la dimostrazione è banale ... trovatela

<sup>18</sup> vedi **PROBABILITÀ**

Ulteriori semplificazioni (quando applicabili):

1) spesso i punti sul grafico risultano pressoché equidistanti. Nell'esempio della sollecitazione di una molla è assai probabile che i pesetti siano uguali e che quindi i valori  $X_i$  siano uniformemente distribuiti fra un valore minimo  $X_m$  e uno massimo  $X_M$ . Allora è possibile considerare i valori  $X_i$  come determinazioni di una variabile aleatoria  $X$  distribuita uniformemente fra  $X_m$  e  $X_M$  e quindi:  $\sigma_X = \frac{X_M - X_m}{\sqrt{12}}$  e  $\bar{X} = \frac{X_M + X_m}{2}$

2) dato il significato di incertezza non è sempre fondamentale che questa sia quantificata esattamente; spesso è sufficiente conoscerne l'ordine di grandezza. In quest'ottica, dato che  $\sigma_Y$  quantifica lo scarto quadratico medio delle coordinate  $Y$  dalla retta tracciata, essa può essere stimata leggendo dal grafico considerando il massimo scarto  $d_M$  (positivo o negativo) e considerando che nel caso di una distribuzione uniforme  $2 d_M = \sqrt{12} \sigma_Y \rightarrow \sigma_Y = d_M / \sqrt{3}$ . Ovviamente quest'ulteriore approssimazione non è lecita qualora dal grafico non fosse possibile apprezzare  $d_M$  (punti perfettamente allineati, apparentemente).



## RELAZIONI NON LINEARI

Si è accennato alla linearizzazione delle funzioni da graficare. Come esempio si consideri il moto di un pendolo matematico in cui

$$T = 2\pi \sqrt{\frac{L}{g}}$$

non è una relazione lineare fra le misure dirette  $T$  e  $L$ . Ci sono due alternative equivalenti: graficare  $T$  vs  $\sqrt{L}$  o  $T^2$  vs  $L$ . Ovviamente  $\sqrt{L}$  o  $T^2$  non sono più misure dirette: nella tabella che riporta i dati da graficare vanno riportati i loro valori e le loro incertezze.

Altro esempio: oscillazione di una molla di costante elastica  $K$  con appesa una massa  $m + m_0$  (il contributo  $m$  può essere variato)

$$\frac{2\pi}{T} = \sqrt{\frac{K}{m + m_0}}$$

In questo caso conviene calcolare i reciproci e quadrare:

$$\frac{T^2}{(2\pi)^2} = \frac{m + m_0}{K} \rightarrow T^2 = \frac{(2\pi)^2}{K} m + \frac{(2\pi)^2 m_0}{K}$$

da questo titolo risulta evidente che la pendenza  $p = \frac{(2\pi)^2}{K}$  e l'intercetta  $q = \frac{(2\pi)^2 m_0}{K}$

E come trasformare relazioni del tipo di:  $T(t) = T_0 e^{-\frac{t}{\tau}}$ ? In questo caso, passando ai logaritmi, si ottiene:

$\ln [T(t)] = \ln(T_0) - t / \tau$  cioè una retta con  $p = -1 / \tau$  e intercetta  $\ln(T_0)$  ottenuta graficando le grandezze  $\ln [T(t)]$  vs  $t$

- Data la frequenza di studi di funzioni esponenziali è stata introdotta della carta in cui uno dei due assi riporta delle divisioni logaritmiche anziché lineari (carta semilogaritmica). In questo caso, **ANZICHÉ DOVER CALCOLARE I LOGARITMI DELLE QUANTITÀ IN ORDINATA, È SUFFICIENTE**

**RIPORTARE L'ARGOMENTO DELLA FUNZIONE;** sulla scala lineare delle ascisse andrebbero invece riportati i tempi. Il particolare tipo di carta produrrebbe una presentazione logaritmica e quindi nell'esempio citato si visualizzerebbe una retta a pendenza negativa (senza ricorrere all'uso di una calcolatrice).

Utilizzando la carta semilogaritmica la generica relazione  $Y = b e^{aX}$  verrebbe visualizzata come  $\ln(Y) = \ln(b) + a X$  e la retta tracciata sul grafico avrebbe intercetta  $q = \ln(b)$ . Il valore del coefficiente  $b$  verrebbe quindi letto direttamente sulla scala che riporta gli argomenti della funzione logaritmo:  $b = Y(X=0)$ .

Presi sulla retta due punti di coordinate  $P_1=[x_1; \ln(y_1)]$  e  $P_2 = [x_2; \ln(y_2)]$  la pendenza verrebbe

stimata da: 
$$p = \frac{\ln(y_2) - \ln(y_1)}{x_2 - x_1} = \frac{\ln(y_2/y_1)}{x_2 - x_1}$$

- Esiste, fra le altre, anche un tipo di carta (doppio-logaritmica) con entrambi gli assi logaritmici per la rappresentazione di funzioni del tipo  $Y = b X^a$  che verrebbero linearizzate passando ai logaritmi:

$$\ln(Y) = \ln(b) + a \ln(X)$$

In questo caso l'andamento lineare sarebbe fra le grandezze  $\ln(Y)$  e  $\ln(X)$  e la pendenza verrebbe

ottenuta da 
$$p = \frac{\ln(y_2) - \ln(y_1)}{\ln(x_2) - \ln(x_1)} = \frac{\ln(y_2/y_1)}{\ln(x_2/x_1)}$$
 mentre l'intercetta si ricaverebbe da  $b = \ln[Y(X=1)]$ .

- La linearizzazione delle funzioni è ovviamente valida sia per una rappresentazione grafica che per l'uso dei minimi quadrati.

In questo caso, pensando come esempio alla carta semilogaritmica, i dati da introdurre sarebbero le misure dirette  $X$  e quelle derivate  $\ln(Y)$ .

Sia per l'elaborazione grafica che per quella statistica mediante i minimi quadrati c'è da fare attenzione alle incertezze. Anche se la grandezza  $Y$  è misurata con lo stesso strumento e metodo e quindi ogni sua misura ha ragionevolmente la stessa incertezza consentendo l'uso delle formule semplificate, le misure, per esempio  $\sqrt{L}$ ,  $T^2$  o  $\ln(Y)$ , hanno incertezze diverse.

Tuttavia, sempre per il fatto che è spesso sufficiente una valutazione approssimata delle incertezze, è possibile utilizzare le formule con le incertezze uguali se queste distano fra loro per non più di un fattore due o tre.



## ISTOGRAMMA

A volte può essere necessario rappresentare graficamente una sola variabile, per esempio per analizzare la distribuzione di frequenze di particolari grandezze (controllo di produzione). Analizziamo questo problema mediante il seguente esempio:

Con un cronometro digitale in grado di apprezzare i centesimi di secondo vengono effettuate 40 misure di periodi  $T$  che risultano essere (espresse in secondi):

1-10:	1,80 1,71 1,74 1,90 1,78	1,72 1,69 1,70 1,73 1,73
11-20:	1,74 1,69 1,76 1,69 1,77	1,77 1,72 1,74 1,67 1,75
21-30:	1,79 1,79 1,71 1,77 1,83	1,81 1,73 1,75 1,71 1,76
31-40:	1,75 1,80 1,75 1,81 1,76	1,79 1,73 1,71 1,76 1,60

Elaborando questi dati si ottengono le quantità:

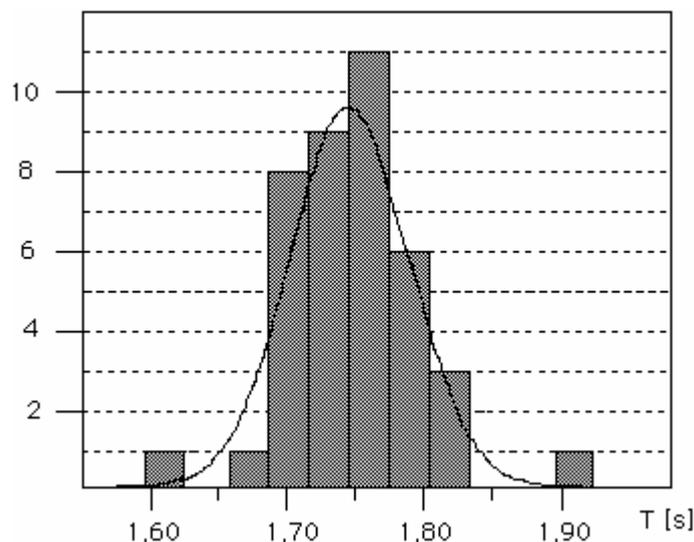
$$\bar{T} = \frac{\sum_{i=1,40} T_i}{40} = 1,748 \text{ s} \quad \text{e} \quad \sigma_s(T) = \sqrt{\frac{\sum_{i=1,40} (T_i - \bar{T})^2}{40 - 1}} = 0,0508 \text{ s}$$

Organizziamo ora i dati in modo da ottenere una loro rappresentazione che possa fornire maggiori informazioni: raggruppiamo i dati in una dozzina circa di classi (se i dati fossero distribuiti secondo la curva di Gauss - cosa che accade spesso - la quasi totalità di misure cadrebbe in un intervallo largo circa 6 deviazioni standard; per una buona rappresentazione dei dati la larghezza di una classe dovrebbe essere pari a mezza deviazione standard).

Per far ciò consideriamo la differenza 1,90 s - 1,60 s = 0,30 s fra il massimo e il minimo valore misurati; ogni classe dovrebbe corrispondere quindi 0,30 s / 12 = 0,25 s che arrotondiamo a 0,30 s (in ogni classe ci saranno quindi 3 possibili valori). Otteniamo pertanto la tabella:

	le misure sono espresse in secondi										
# classe:	1	2	3	4	5	6	7	8	9	10	11
dal valore:	1,60	1,63	1,66	1,69	1,72	1,75	1,78	1,81	1,84	1,87	1,90
al valore:	1,62	1,65	1,68	1,71	1,74	1,77	1,80	1,83	1,86	1,89	1,92
valore centrale:	1,61	1,64	1,67	1,70	1,73	1,76	1,79	1,82	1,85	1,88	1,91
misure nella classe:	1	0	1	8	9	11	6	3	0	0	1

e da questa il grafico (istogramma o diagramma a bastoni) seguente:



A partire dai dati raggruppati in classi si ottengono i valori approssimati:

$$\bar{T} = \frac{\sum_{j=1,11} n_j T_j}{40} = 1,748 \text{ s} \quad \text{e} \quad \sigma_s(T) = \sqrt{\frac{\sum_{j=1,11} n_j (T_j - \bar{T})^2}{40 - 1}} = 0,0507 \text{ s}$$

che, come si può notare coincidono con il calcolo esatto quando vengono applicate le convenzioni sul numero di cifre significative e decimali. Per questo motivo, una volta che sia stato effettuato un raggruppamento in classi, è più economico utilizzare i dati raggruppati per tutti i calcoli successivi.

L'istogramma di questo esempio consente di osservare l'andamento gaussiano delle misure; per confronto è riportata una gaussiana con  $m = 1,748 \text{ s}$  e  $\sigma = 0,0507 \text{ s}$

## TEST DI IPOTESI

Le espressioni ricavate precedentemente consentono di stimare su base statistica sia i riassunti di una distribuzione di probabilità come il valore atteso o la varianza, sia i parametri di un andamento funzionale (minimi quadrati).

Abbiamo introdotto anche gli elementi necessari per verificare la validità di alcune ipotesi. Ad esempio, considerando la distribuzione t-Student, è possibile quantificare la bontà dell'accordo fra la media aritmetica di una serie di misure e il valore atteso  $m$ . In questo caso lo scarto standardizzato:

$$t = \frac{(\bar{X} - m)}{\sigma_s(\bar{X})} \text{ segue una distribuzione di Student con } N-1 \text{ gradi di libertà.}$$

Se effettivamente il campione di misure ottenuto rappresenta la grandezza fisica in esame, sotto l'ipotesi di distribuzione gaussiana degli errori casuali, il valore della variabile  $t$  è prevedibile su base statistica.

Operativamente:

- si stabilisce un livello di confidenza
- si fissa il corrispondente intervallo di confidenza per lo scarto  $t$  considerando che il numero di misure effettuate
- se il valore  $t$  ottenuto calcolando la quantità  $\frac{(\bar{X} - m)}{\sigma_s(\bar{X})}$  cade all'interno dell'intervallo di confidenza prefissato l'esito del confronto è positivo (al livello di confidenza fissato)



## CONFRONTO FRA MISURE

In generale, però, non è nota la distribuzione di probabilità delle variabili aleatorie che caratterizzano i risultati di misure, anche se spesso è approssimativamente gaussiana; ancor più spesso si esegue una sola misura e quindi non è possibile calcolare una deviazione standard.

Ciononostante è ancora possibile verificare l'attendibilità della misura ricorrendo alla disuguaglianza di Chebychev: indipendentemente dalla distribuzione di probabilità è raro (<10%) che la differenza fra il valore ottenuto e il valore vero sia superiore a tre volte la deviazione standard<sup>[19]</sup>.

Questa considerazione ci permette di confrontare i risultati delle misure sia con valori di riferimento che fra di loro. Sono ovviamente possibili anche altri tipi di confronto quali lo scarto assoluto e relativo fra i valori. Esaminiamo i diversi casi:

<sup>19</sup> vedi PROBABILITÀ: il livello di confidenza è maggiore di  $1-1/t^2$ ; volendo aumentare la certezza del risultato andrebbe considerato un maggior numero di deviazioni standard. **Spesso, e in particolare all'interno di questo corso, si considera sufficiente un intervallo di confidenza pari a tre deviazioni standard**

**CONFRONTO MISURA – VALORE**

Si vuole confrontare la misura  $X \pm \sigma$  con il valore  $m$ . Se la misura  $X$  fosse rappresentativa del valore vero  $m$  e  $\sigma$  stimasse correttamente l'incertezza associata alla misura allora:

lo **scarto** (assoluto)  $\Delta = X - m$  dovrebbe essere "piccolo" (quanto?)  $\rightarrow$  qualitativo

lo **scarto relativo**  $s = (X - m)/m$  dovrebbe essere "piccolo" (pochi %)  $\rightarrow$  qualitativo

la **variabile t**<sup>[20]</sup>  $t = (X - m)/\sigma$  dovrebbe risultare  $|t| < 3$   $\rightarrow$  quantitativo

**CONFRONTO MISURA – MISURA**

Se non è dato un valore di riferimento, per confrontare fra loro le misure  $X_1 \pm \sigma_1$  e  $X_2 \pm \sigma_2$  occorrerà, a seconda del tipo di confronto, stimare il valore  $m$ :

lo **scarto** (assoluto)  $\Delta = X_1 - X_2$  dovrebbe essere "piccolo" (quanto?)  $\rightarrow$  qualitativo

lo **scarto relativo**  $s = \frac{X_1 - X_2}{\bar{X}}$  dove  $\bar{X} = \frac{X_1 + X_2}{2}$  è la migliore stima di  $m$   
dovrebbe essere "piccolo" (pochi %)  $\rightarrow$  qualitativo

la **variabile t** in questo caso viene costruita considerando che la differenza  $X_1 - X_2$  dovrebbe essere nulla e che la misura di  $X_1 - X_2$  ha un'incertezza  $\sigma(X_1 - X_2) = \sqrt{\sigma_1^2 + \sigma_2^2}$ :

$t = \frac{X_1 - X_2}{\sigma(X_1 - X_2)}$  dovrebbe risultare  $|t| < 3$   $\rightarrow$  quantitativo

Lo scarto assoluto, anche se a volte utile, richiede una certa "sensibilità" da parte di chi analizza i dati per riconoscere o meno la compatibilità fra le misure;

lo scarto relativo, riportando lo scarto assoluto al valore vero (o alla sua migliore stima), è di più semplice uso: all'interno di una particolare applicazione, un valore percentuale particolarmente basso o elevato è di facile interpretazione anche se non considera anche gli effetti degli errori di misura che possono falsare i risultati;

la variabile t quantifica maggiormente il significato dello scarto fra le misure:

se  $|t| < 3$  significa che lo scarto è compatibile con la presenza dei soli errori casuali (eventuali errori sistematici sono trascurabili rispetto a quelli casuali);

ripetendo più volte la/e misura/e è possibile ridurre, con le opportune medie aritmetiche, l'entità degli errori ed è possibile trovare una riduzione dello scarto.

Il confronto va considerato positivo e le misure compatibili (coincidenti)

se  $|t| > 3$  significa che lo scarto non è compatibile con la presenza dei soli errori casuali (sono presenti errori sistematici non trascurabili);

ripetendo più volte la/e misura/e non è possibile ridurre, con le opportune medie aritmetiche, l'entità degli errori: lo scarto è significativo di una differenza fra i valori ottenuti..

Il confronto va considerato negativo e le misure non compatibili (diverse)

<sup>20</sup> non esiste una convenzione sul nome di questa grandezza: spesso si indica con  $z$  lo scarto standardizzato per una variabile gaussiana e con  $t$  una variabile di Student. Ho preferito mantenere questo nome per scopi mnemonici anche se, a rigore, è corretto solo nel caso di una media aritmetica di misure distribuite gaussianamente.

## IL TEST DEL $\chi^2$

Vediamo ora come quantificare la bontà dell'ipotesi che una variabile aleatoria, della quale sono state eseguite delle misure, segua una particolare distribuzione.

Il metodo in esame si fonda sulla costruzione di una variabile  $U$  che quantifichi la differenza fra il campione di misure e la distribuzione ipotizzata e segua una distribuzione del  $\chi^2$  (da qui il nome del test). Per accettare o rifiutare l'ipotesi della particolare distribuzione è quindi sufficiente controllare se  $u$ , il particolare valore di  $U$  ottenuto, cade o meno all'interno di un opportuno intervallo di confidenza fissato a priori.

Perché  $U$  abbia una distribuzione del  $\chi^2$  occorre<sup>[21]</sup> che sia pari alla somma dei quadrati di variabili aleatorie indipendenti con distribuzione gaussiana, media nulla e varianza unitaria:

$$U = \sum u_i^2.$$

In questo caso si ipotizza di suddividere l'istogramma delle  $N$  misure effettuate in  $v$  intervalli. Per ogni intervallo si costruisce una

$$u_i = \frac{n_i - m_i}{\sigma(n_i)}$$

dove  $n_i$  e  $\sigma(n_i)$  rappresentano la frequenza e la deviazione standard della frequenza delle misure nell'intervallo  $i$ -mo e  $m_i$  la frequenza attesa per l' $i$ -mo intervallo.

Nel dettaglio, le  $N$  misure che costituiscono il campione in esame hanno ciascuna una probabilità  $p_i$  di cadere nell'intervallo  $i$ -mo e una probabilità  $1-p_i$  di esserne fuori; pertanto la frequenza di misure dell'intervallo  $i$ -mo segue una distribuzione binomiale di parametri  $N$  e  $p_i$  e quindi con valore medio  $m_i = N p_i$  e varianza  $\sigma^2 = N p_i (1-p_i)$ .

Se poi le  $n_i$  seguissero una distribuzione di Gauss allora la variabile

$$U = \sum_{i=1,v} \left( \frac{n_i - N p_i}{\sqrt{N p_i (1-p_i)}} \right)^2$$

seguirebbe una distribuzione del  $\chi^2$  con  $v-1$  gradi di libertà (c'è il vincolo  $\sum n_i = N$  che riduce di una unità le  $v$  variabili indipendenti).

Per ottenere, anche se in modo approssimativo, un andamento gaussiano occorre scegliere la larghezza, e quindi il numero, degli intervalli dell'istogramma in modo tale che in ognuno di essi cada un numero sufficientemente elevato di misure  $n_i$  tale da consentire l'approssimazione binomiale  $\rightarrow$  gaussiana. Ma  $n_i$  elevato si ottiene con  $N$  grande o realizzando un istogramma con pochi canali. In questo caso bisogna fare attenzione perché se i canali dell'istogramma fossero troppo pochi non sarebbe possibile riconoscere la forma del campione (per esempio nel limite di 1 o 2 canali non sarebbe più possibile distinguere un andamento esponenziale da uno gaussiano).

Non resta ora che determinare le  $p_i$ . Rappresentano la probabilità, come detto, che ha ognuna delle  $N$  misure, ritenute indipendenti, di cadere nell'intervallo  $i$ -mo. Come calcolarle dipende dalla distribuzione di probabilità della variabile  $X$  della quale sono state effettuate le  $N$  misure: a seconda di quale distribuzione ipotizziamo il calcolo delle  $p_i$  produrrà valori diversi e quindi un diverso valore  $u$  di  $U$  che quantificherà un diverso esito per il nostro test.

<sup>21</sup> vedi **PROBABILITÀ**

**• UN ESEMPIO DELL'USO DEL TEST DEL  $\chi^2$**

Applichiamo il test del  $\chi^2$  per valutare se il procedimento costruttivo di alcuni oggetti è tale per cui le loro masse siano distribuite secondo una gaussiana.

La seguente tabella riporta i risultati della misura della massa di 200 oggetti<sup>[22]</sup>; le misure sono state raggruppate in classi di 100g :

#classe	intervallo	valore centrale	frequenza
i	$M_i$ [kg]	$M_i$ [kg]	$n_i$
1	$0,15 \leq M < 0,25$	0,2	9
2	$0,25 \leq M < 0,35$	0,3	22
3	$0,35 \leq M < 0,45$	0,4	53
4	$0,45 \leq M < 0,55$	0,5	64
5	$0,55 \leq M < 0,65$	0,6	31
6	$0,65 \leq M < 0,75$	0,7	16
7	$0,75 \leq M < 0,85$	0,8	5

Calcoliamo la media aritmetica e la deviazione standard del campione a partire dalla tabella:

$$\bar{X} = 0,477 \text{ kg} \quad \sigma(X) = 0,133 \text{ kg.}$$

Vediamo se, applicando il test del  $\chi^2$ , risulta che il nostro campione è distribuito secondo una gaussiana con parametri pari, ad esempio, a  $M = 0,48 \text{ kg}$  e  $\sigma = 0,13 \text{ kg}$ ; vogliamo cioè verificare se risulta:

$$f(M) = \frac{1}{\sqrt{2\pi} 0,13} e^{-\frac{(M-0,48)^2}{2 \cdot 0,13^2}} \text{ kg}^{-1}$$

Iniziamo quindi col calcolare la probabilità  $p_1$  che la v.a.  $M$  sia compresa nell'intervallo  $[0,15 \text{ kg}; 0,25 \text{ kg}]$  che costituisce la prima classe. Integrando la distribuzione per tutti i valori di massa compresi nel primo intervallo si ha:

$$p_1 = \int_{0,15}^{0,25} \frac{1}{\sqrt{2\pi} 0,13} e^{-\frac{M-0,48}{2 \cdot 0,13^2}} dM = \int_{-2,54}^{-1,77} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \text{ERF}(2,54) - \text{ERF}(1,77) = 0,4945 - 0,4617 = 0,0328$$

Le 200 misure hanno quindi una probabilità del 3,28% di cadere nell'intervallo  $[0,15 \text{ kg}; 0,25 \text{ kg}]$  e una probabilità del 96,72% di caderne fuori: siamo di fronte ad un processo binomiale per il quale, in media, il contenuto di quel canale sarà  $N p_1 = 200 \times 0,0328 = 6,56$  misure.

Nel nostro campione, invece, abbiamo nel primo intervallo 9 misure; se la nostra ipotesi di distribuzione gaussiana con media 0,48 kg e deviazione standard 0,13 kg fosse valida, allora la differenza  $9 - 6,56$  sarebbe da imputare solo a fluttuazioni statistiche.

Lo scarto standardizzato vale quindi  $u_1 = \frac{9 - 6,56}{\sqrt{200 \times 0,0328 \times 0,9672}} = 0,969$ .

<sup>22</sup> i dati sono stati prodotti da una simulazione al calcolatore: si tratta della media aritmetica di quattro variabili che seguono una distribuzione uniforme fra 0 e 1

Procedendo analogamente per gli altri intervalli si ottiene:

i	n <sub>i</sub>	z <sub>i</sub>	z <sub>i+1</sub>	ERF(z <sub>i</sub> )	ERF(z <sub>i+1</sub> )	p <sub>i</sub>	Np <sub>i</sub>	N p <sub>i</sub> (1-p <sub>i</sub> )	u <sub>i</sub> = $\frac{n_i - Np_i}{\sqrt{Np_i(1-p_i)}}$	u <sub>i</sub> <sup>2</sup>
1	9	1,77	2,54	0,4617	0,4945	0,0328	6,56	6,344832	0,969	0,938
2	22	1,00	1,77	0,3413	0,4617	0,1204	24,08	21,180768	- 0,452	0,204
3	53	0,23	1,00	0,0910	0,3413	0,2503	50,06	37,529982	0,480	0,230
4	64	0,54	0,23	0,2054	0,0910	0,2964	59,28	41,709408	0,731	0,534
5	31	1,31	0,54	0,4049	0,2054	0,1995	39,90	31,939950	- 1,575	2,480
6	16	2,08	1,31	0,4813	0,4049	0,0764	15,28	14,112608	0,192	0,037
7	5	2,85	2,08	0,4978	0,4813	0,0165	3,30	3,245550	0,943	0,890
$\sum_i$	200					0,9923				5,313

Come si può notare  $\sum p_i < 100\%$ ; infatti poiché la curva di Gauss si estende fino all'infinito, è ovvio che la probabilità che una misura cada nella zona coperta dal nostro campione sia inferiore all'unità.

La somma dei quadrati degli scarti standardizzati  $u = \sum_{i=1,v} \left( \frac{n_i - Np_i}{\sqrt{Np_i(1-p_i)}} \right)^2$  segue, con le approssimazioni viste, una distribuzione del  $\chi^2$  con  $v - 1 = 6$  gradi di libertà.

Se scegliamo un intervallo di confidenza  $E(\chi^2) \pm \sigma(\chi^2) = 6 \pm \sqrt{12} = [2,54; 9,46]$ , dato che il valore u trovato (5,313) cade nell'intervallo, possiamo concludere che il nostro campione segue una distribuzione gaussiana con media 0,48 kg e deviazione standard 0,13 kg.

Ci saremmo però potuti chiedere se il campione dei dati ottenuto seguisse una distribuzione gaussiana, senza esplicitare i valori della media e della varianza. In questo caso la variabile U avrebbe dovuto quantificare l'accordo fra il campione di misure e una distribuzione gaussiana con media pari alla media aritmetica delle N misure e deviazione standard pari alla deviazione standard sperimentale e quindi il calcolo delle p<sub>i</sub> si sarebbe basato sulla:

$$f(M) = \frac{1}{\sqrt{2\pi} \sigma_s(X)} e^{-\frac{(M-\bar{X})^2}{2 \sigma_s^2(X)}}$$

L'uso, nell'espressione della densità di probabilità, delle formule della media aritmetica e della deviazione standard sperimentale avrebbe ridotto a 4 il numero dei gradi di libertà portando ad un intervallo di confidenza pari a  $4 \pm \sqrt{8} = [1,17; 6,83]$  nel quale, tuttavia, in questo esempio ancora cade il valore di U trovato.



## IL METODO DEL $\chi^2$ MINIMO

Il test del  $\chi^2$  può essere utilizzato anche per determinare quale sia la distribuzione madre di un campione di misure.

In questo caso viene ipotizzata una distribuzione della v.a. X nota a meno di parametri q<sub>i</sub>:

$$f(x; q_1, q_2, \dots, q_K)$$

In questo caso i parametri possono essere sia dei riassunti di una distribuzione, sia più in generale dei coefficienti che indicano per esempio la frazione di distribuzione che è uniforme, quella che è esponenziale e quella che è gaussiana. È compito dello sperimentatore scegliere, sulla base della conoscenza del problema in esame, gli ingredienti della "miscela".

Quindi si variano i  $K$  parametri fino a trovare quel loro insieme di valori per il quale è minima la

$$\text{quantità } U(q_1, q_2, \dots, q_K) = \sum_{i=1, v} \left( \frac{n_i - Np_i}{\sqrt{Np_i(1-p_i)}} \right)^2 \text{ dove } p_i = p_i(q_1, q_2, \dots, q_K)$$

La minimizzazione corrisponde alla migliore approssimazione della distribuzione ipotizzata al campione (minor distanza fra il contenuto dei canali dell'istogramma e il valore atteso in base alla distribuzione ipotizzata).

Trovato l'insieme dei  $K$  parametri corrispondente al minimo (le quantità dei diversi ingredienti nella miscela) si può applicare ancora una volta il test del  $\chi^2$  per quantificare la bontà dell'approssimazione.

In questo caso, però, ogni parametro che è stato utilizzato durante la minimizzazione diminuisce di una unità il numero di v.a. indipendenti e quindi il confronto andrà effettuato con una distribuzione del  $\chi^2$  con  $v - 1 - K$  gradi di libertà.

Ovviamente data la complessità dei calcoli questo metodo richiede l'uso di un calcolatore, ma è un metodo molto potente e largamente utilizzato perché produce risultati notevoli a partire da scarse conoscenze del sistema (e, ovviamente, da un nutrito campione di misure).

## STIMA DELLE INCERTEZZE

Nel 1993 l'Organizzazione Internazionale per la Standardizzazione (ISO) ha pubblicato<sup>(\*)</sup> una "Guida all'espressione dell'incertezza di misura" basata sulle raccomandazioni dell'Ufficio Internazionale di Pesi e Misure (BIPM). La Guida riassume i fondamenti teorici, le definizioni e le procedure elaborate dalle più autorevoli organizzazioni mondiali di metrologia:

BIPM	Bureau International de Poids et Mesures
IEC	International Electrotechnical Commission
IFCC	International Federation of Clinical Chemistry
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
IUPAP	International Union of Pure and Applied Physics
OIML	International Organization of Legal Metrology

Aderiscono all'iniziativa anche altri istituti nazionali affiliati come, ad esempio:

DIN (D)	Deutsches Institut für Normung
NIST (USA)	National Institute of Standards and Technology
UNI (I)	Ente Italiano per l'Unificazione

1) *The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated:*

*A: those which are evaluated by statistical methods*

*B: those which are evaluated by other means.*

*There is not always a simple correspondence between the classification into categories A or B and the previously used classification into "random" and "systematic" uncertainties. The term "systematic uncertainty" can be misleading and should be avoided.*

2) *The components in category A are characterized by the estimated variances or the estimated "standard deviations" and the number of degrees of freedom.*

3) *The components in category B should be characterized by quantities which may be considered as approximations to the corresponding variances, the existence of which is assumed.*

4) *The combined uncertainty should be characterized by the numerical value obtained by applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of "standard deviations".*

5) *If, for a particular application, it is necessary to multiply the combined uncertainty by a factor to obtain an overall uncertainty, the multiplying factor used must always be stated.*

*For estimate of an input quantity  $X_i$  that has not been obtained from repeated observations, the standard uncertainty is evaluated by scientific judgement based on all the available information on the possible variability of  $X_i$ . The pool of information may include:*

- previous measurement data;*
- experience with or general knowledge of the behaviour and properties of relevant materials and instruments;*
- manufacturer's specifications;*
- data provided in calibration and other certificates;*
- uncertainties assigned to reference data taken from handbooks.*

---

<sup>\*</sup> International Organization for Standardization (ISO), "Guide to the expression of uncertainty in measurement", Geneva, Switzerland, 1993.

... inoltre ...

*Lo scopo di ogni misurazione è la determinazione del valore vero di un misurando  $M$ . Le imperfezioni degli strumenti, le variazioni delle condizioni ambientali e l'influenza dell'osservatore (nonché i suoi sbagli) provocano errori di misura. Essi sono il motivo per cui non è possibile trovare il valore vero  $M$ .*

*Si assume che i valori  $x_i$  ottenuti da diverse misurazioni individuali di una serie di misurazioni altro non sono che le determinazioni di una variabile casuale  $X$ . Questa variabile aleatoria  $X$  obbedisce ad una distribuzione di probabilità caratterizzata in particolare da due parametri che sono il valore atteso  $m$  e la deviazione standard  $\sigma$ .*

*In assenza di errori sistematici il valore atteso  $m$  coincide col valore vero  $M$  del misurando. La deviazione standard  $\sigma$  è una misura della variabilità, per via dell'errore casuale, di un singolo valore misurato dal valore atteso del misurando.*

*I parametri  $m$  e  $\sigma$  della distribuzione di probabilità non sono generalmente noti. Il problema consiste nel determinare delle loro stime a partire da una serie di misurazioni. Usualmente per la stima di  $m$  si utilizza la media aritmetica  $\bar{X}$  e per quella di  $\sigma$  si utilizza la deviazione standard sperimentale  $\sigma_s$ . Poiché i valori misurati sono realizzazioni di una variabile aleatoria,  $\bar{X}$  fluttuerà statisticamente intorno a  $m$  e  $\sigma_s$  intorno a  $\sigma$ .*

*Sulla base delle assunzioni riguardanti il tipo di distribuzione (spesso si assumerà quella gaussiana) con l'aiuto dei valori  $\bar{X}$  e  $\sigma_s(\bar{X})$  è possibile determinare un intervallo di confidenza all'interno del quale è contenuto, con un livello di confidenza fissato, il valore  $m$ .*

*Gli effetti noti degli errori sistematici vengono eliminati applicando correzioni. Per gli altri, sconosciuti, un approccio per la loro riduzione consiste nell'allargamento dell'intervallo di confidenza in base alle assunzioni applicabili al tipo di misurazione effettuata.*

*Adattato dalle norme DIN 1319 parte 3*

... pertanto<sup>[23]</sup>:

## **l'incertezza da associare ad ogni misura può essere:**

---

<sup>23</sup> *anziché riportare accuratamente tutta la normativa relativa al trattamento dei dati sperimentali ho preferito rielaborarla prendendone solo alcuni elementi e adattandoli ai contenuti del corso. L'arbitrarietà (e in alcuni casi, purtroppo, anche la superficialità) di questa operazione non rende sempre utilizzabile il materiale per scopi professionali ma a volte ho preferito sacrificare il rigore formale all'efficacia didattica. Me ne scuso con i professionisti e gli operatori del settore. Gli studenti interessati possono contattarmi per conoscere la documentazione originale*

## INCERTEZZE DI TIPO A

Sono quelle ottenute a partire dalle deviazioni standard sperimentale della media aritmetica: che quantifica quanto bene  $\bar{X}$  stimi il valore atteso  $m$  della grandezza fisica  $X$ :

$$\sigma_s(\bar{X}) = \frac{\sigma_s(X)}{\sqrt{N}} = \sqrt{\frac{\sum_{i=1, N} (x_i - \bar{X})^2}{N(N-1)}}$$

In base al teorema del limite centrale è anche possibile conoscere la distribuzione di probabilità della media aritmetica: la curva di Gauss con media pari al valore vero  $m$  e deviazione standard pari a  $\frac{\sigma}{\sqrt{N}}$ . Pertanto se il valore vero di una misura è  $m$  e l'entità degli errori casuali è  $\sigma$ , ci si aspetta che la media aritmetica di  $N$  misure sia compresa nel 68% dei casi nell'intervallo

di confidenza<sup>[24]</sup>:  $\left[ m - \frac{\sigma}{\sqrt{N}}; m + \frac{\sigma}{\sqrt{N}} \right]$ .

••• Inversione: con un livello di confidenza del 68% si ha che:  $m - \frac{\sigma}{\sqrt{N}} \leq \bar{X} \leq m + \frac{\sigma}{\sqrt{N}}$

che corrisponde alle disuguaglianze:  $\bar{X} - \frac{\sigma}{\sqrt{N}} \leq m \leq \bar{X} + \frac{\sigma}{\sqrt{N}}$ .

Dopo questa inversione si può stabilire che il valore vero  $m$  è all'interno dell'intervallo:

$$[\bar{X} - \sigma_s(\bar{X}); \bar{X} + \sigma_s(\bar{X})]$$

pertanto: **usualmente<sup>[25]</sup> il risultato di una serie di misure viene espresso da:**

$$x = \bar{X} \pm \sigma_s(\bar{X})$$

**indicando in questo modo che con alta probabilità il valore vero cercato è compreso nell'intervallo:**

$$[\bar{X} - \sigma_s(\bar{X}); \bar{X} + \sigma_s(\bar{X})]$$

È chiaro quindi perché, per ridurre l'effetto degli errori casuali, sia sufficiente aumentare  $N$  (cioè ripetere più volte la misura della stessa grandezza): in questo modo si riduce la larghezza dell'intervallo di confidenza  $\sigma_s(\bar{X}) \propto 1/\sqrt{N}$ .

Ovviamente  $\sigma_s(X)$  non dipende da  $N$ : ripetendo la misura si determina meglio la forma della distribuzione di probabilità della  $X$  ma non la si altera; è il meccanismo di compensazione insito nella somma degli scarti che riduce l'entità residua degli errori casuali nella media aritmetica riducendo  $\sigma_s(\bar{X})$ .

<sup>24</sup> perché si è confidenti che con un livello di confidenza del 68% il valore cercato è all'interno dell'intervallo

<sup>25</sup> per motivi di tempo talora si trasmette l'informazione relativa all'incertezza delle misure utilizzando un opportuno numero di cifre significative p.es. quello determinato dalla sensibilità della strumentazione utilizzata.

••• Il numero  $N$  di osservazioni deve essere sufficientemente elevato non tanto affinché  $\bar{X}$  sia una stima affidabile di  $m$  (bastano 4-5 misure) ma soprattutto perché  $\sigma_s(\bar{X})$  sia una stima affidabile della deviazione standard.

Infatti se la v.a. è gaussiana si può approssimativamente ritenere che

$$\frac{\sigma(\sigma)}{\sigma} = \frac{1}{\sqrt{2(N-1)}}$$

(cioè se  $N = 5, 10, 20$  l'incertezza relativa della stima di  $\sigma$  vale 35%, 24%, 16%).

Bisogna quindi tener conto della differenza fra  $\sigma_s(\bar{X})$  e  $\sigma(\bar{X})$  quando si calcolano livelli di confidenza: se la distribuzione di probabilità di  $X$  è normale (gaussiana) della differenza fra la deviazione standard e la sua stima tiene conto la distribuzione t-Student.

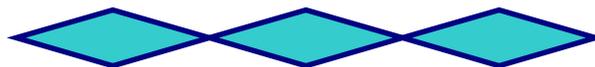
Tuttavia se  $X$  è distribuita gaussianamente (e spesso lo è) già con solo 5 misure c'è una probabilità di circa il 90% che  $x$  sia compreso nell'intervallo  $[\bar{X} - 2\sigma_s(\bar{X}); \bar{X} + 2\sigma_s(\bar{X})]$ .

## INCERTEZZE DI TIPO B

**non** sono ottenute come deviazione standard sperimentali della media aritmetica.

Se data una grandezza  $X$  viene eseguita una sola misurazione ottenendo il valore  $x_1$  l'incertezza da associare a  $x_1$  come stima di  $X$  non può essere ricavata dall'unica misura a disposizione ma dalla conoscenza delle caratteristiche degli strumenti utilizzati o (in loro assenza) di risultati precedenti eseguiti nelle stesse condizioni e con gli stessi strumenti.

Se non c'è una conoscenza specifica della distribuzione di  $X$  all'interno di un certo intervallo, si può solo assumere che essa sia costante nell'intervallo e nulla all'esterno (distribuzione uniforme)<sup>[26]</sup>. In questo caso la deviazione standard è pari a  $\frac{1}{\sqrt{12}}$  volte la larghezza dell'intervallo (p. es. 0,29 volte la divisione sulla scala dello strumento o l'intervallo definito da una tolleranza).



## IL RISULTATO DI UNA MISURA PUÒ ESSERE ESPRESSO NEI SEGUENTI MODI:

- $m = (100,021\ 47 \pm 0,000\ 35) \text{ g}$
- $m = 100,021\ 47 (35) \text{ g}$
- $m = 100,021\ 47 (0,000\ 35) \text{ g}$
- $m = 100,021\ 47 \text{ g}$  con  $\sigma = 0,35 \text{ mg}$
- $m = 100,02 \text{ g}$

All'interno di questo corso verrà preferita la prima modalità; la terza e la quarta sono scarsamente diffuse; l'ultima viene vietata (per motivi didattici) all'interno di questo corso ma è diffusissima nei settori in cui l'incertezza della misura non riveste un ruolo critico.

<sup>26</sup> Notare che, nel caso non sia nota la distribuzione di  $X_i$  all'interno dell'intervallo che ne contiene "tutti" i valori, l'assunzione di una distribuzione uniforme implica che la deviazione standard è pari a  $\frac{1}{\sqrt{12}}$  volte l'intervallo mentre una distribuzione normale, approssimando il livello del 99,73% al 100%, implica una deviazione standard pari a  $\frac{1}{6}$  dell'intervallo. La differenza è trascurabile se si pensa a quanta informazione è necessaria per giustificare l'una anziché l'altra.

## ESEMPIO

Vengono eseguite 20 misure indipendenti della temperatura  $t$  (in "gergo": viene estratto casualmente dalla popolazione madre un campione di 20 individui).

Tali valori sono riportati nella seguente tabella raggruppati in intervalli di  $1^\circ\text{C}$  per ottenere un istogramma (la preparazione di un istogramma non è necessario per l'analisi statistica dei dati ma, come detto, è utile per visualizzare l'andamento dei dati).

$t_1$ ( $^\circ\text{C}$ )	$t_2$ ( $^\circ\text{C}$ )	$t$ ( $^\circ\text{C}$ ) con $t_1 \leq t < t_2$
94,5	95,5	-
95,5	96,5	-
96,5	97,5	96,90
97,5	98,5	98,18; 98,25
98,5	99,5	98,61; 99,03; 99,49
99,5	100,5	99,56; 99,74; 99,89; 100,07; 100,33; 100,42
100,5	101,5	100,68; 100,95; 101,11; 101,20
101,5	102,5	101,57; 101,84; 102,36
102,5	103,5	102,72
103,5	104,5	-
104,5	105,5	-

La media aritmetica delle 20 osservazioni è  $\bar{t} = 100,145^\circ\text{C}$  (miglior stima del valor medio di  $t$ ); la deviazione standard sperimentale è pari a  $1,489^\circ\text{C}$  (incertezza associata ad una misura di  $t$ ) e la deviazione standard sperimentale della media aritmetica è pari a  $0,333^\circ\text{C}$  (incertezza associata alla misura della media aritmetica); pertanto  $t = (100,1 \pm 1,5)^\circ\text{C}$  e  $\bar{t} = (100,15 \pm 0,33)^\circ\text{C}$

**(incertezza di tipo A)**

Se l'unica informazione a disposizione fosse stata che una misurazione aveva fornito il valore  $100,5^\circ\text{C}$  e che in generale la temperatura era compresa fra i due valori  $96^\circ\text{C}$  e  $104^\circ\text{C}$ , la stima della deviazione standard sarebbe stata  $8^\circ\text{C} / \sqrt{12} \approx 2,3^\circ\text{C}$ ; pertanto  $t = \bar{t} = (100,5 \pm 2,3)^\circ\text{C}$

**(incertezza di tipo B).**

Se invece fossero state ottenute le due misure  $96^\circ\text{C}$  e  $104^\circ\text{C}$  si sarebbe ottenuta la media aritmetica  $\bar{t} = 100^\circ\text{C}$ , la deviazione standard sperimentale  $5,7^\circ\text{C}$  e quella della media aritmetica<sup>[27]</sup>  $4,0^\circ\text{C}$ . Pertanto  $t = (100,0 \pm 5,7)^\circ\text{C}$  e  $\bar{t} = (100,0 \pm 4,0)^\circ\text{C}$

**(incertezza di tipo A)**

<sup>27</sup> è istruttivo dimostrare che date le misure  $x_1$  e  $x_2$  la deviazione standard sperimentale della loro media aritmetica vale  $|x_1 - x_2|/2$ , cioè la metà della loro distanza (semi-dispersione massima) ... provate

## CIFRE SIGNIFICATIVE

I metodi di misura, gli strumenti e l'osservatore possono essere classificati in base a:

**sensibilità:** capacità di apprezzare piccole variazioni delle grandezze in esame

**precisione:** capacità di produrre lo stesso risultato ripetendo più volte la stessa osservazione  
(errori casuali trascurabili)

**accuratezza:** capacità di produrre un risultato esente da errori sistematici

Quanto più una misura è spinta (è necessario determinare i particolari) (sensibilità)  
tanto meglio (piccoli errori casuali) si dovrà osservare il fenomeno (precisione)  
e tanto più si dovrà porre attenzione agli errori sistematici (accuratezza)

Di conseguenza il valore numerico della misura andrà riportato con un maggior numero di cifre  
(cifre significative)

Più sono le cifre significative, migliore deve essere la qualità della misura che implica quindi un maggior *costo*: acquisto(€), manutenzione, strumentazione, personale, tempo, facilità d'uso, ...

**IL NUMERO DI CIFRE SIGNIFICATIVE DEL RISULTATO DI UNA MISURA È QUINDI STRETTAMENTE CORRELATO ALLA BONTÀ DELLA MISURA E NON PUÒ ESSERE SCELTO ARBITRARIAMENTE.**

Per ottenere il numero di cifre significative occorre togliere l'eventuale virgola al numero e contare tutte le cifre a partire dalla prima a sinistra non nulla:

numero	cifre	cifre decimali	cifre significative
12 300	5	0	5
1 230,0 x 10	5	1	5
123,0 x 10 <sup>2</sup>	4	1	4
1,23 x 10 <sup>4</sup>	3	2	3
0,123 0 x 10 <sup>5</sup>	5	4	4

Usualmente **L'INCERTEZZA VIENE RIPORTATA CON DUE CIFRE SIGNIFICATIVE.**

I valori dei risultati devono essere arrotondati opportunamente per essere consistenti con le incertezze: ad esempio se  $R = 10,057\ 62\ \Omega$  con  $\sigma = 27\ m\ \Omega$   
allora sarà  $R = (10,058 \pm 0,027)\ \Omega$ .

Se un numero ha più cifre significative di quante ne occorrono, il numero va arrotondato (approssimando all'unità superiore se l'ultima cifra è 5 o superiore)

0,445 945    0,445 95    0,445 9    0,446    0,45    0,4

0,445 95

0,446 0

0,446

0,45

0,5

<--- arrotondamenti successivi: differiscono al più  
per la meno significativa delle cifre dai valori  
ottenuti direttamente

Per alcune applicazioni (è vivamente sconsigliato all'interno di questo corso) è sufficiente rappresentare l'incertezza mediante il numero di cifre significative del risultato; in tal caso il risultato deve essere riportato con un numero di cifre significative tale per cui l'incertezza corrisponda all'ultima di esse:

$m = (100,021\ 47 \pm 0,000\ 35)\ g$     ->     $m = 100,021\ 5\ g$

$R = (10,058 \pm 0,027)\ \Omega$     ->     $R = 10,06\ \Omega$

## ESEMPIO

Calcolare la media, la deviazione standard e la deviazione standard della media delle misure:

**I(mA) 124 136 142 117 140 138 125**

$$\bar{I} = 131,71429 \text{ mA}$$

$$\sigma_s(I) = 9,6040 \text{ mA}$$

$$\sigma_s(\bar{I}) = 3,62997 \text{ mA}$$

A) Riportare la deviazione standard della media aritmetica con due cifre significative

$$\sigma_s(\bar{I}) = 3,6 \text{ mA}$$

B) Arrotondare la media in modo da avere lo stesso numero di cifre decimali della deviazione standard

$$\bar{I} = 131,7 \text{ mA}$$

Per riportare correttamente il risultato di una serie di misure utilizzare le regole A) e B) dell'esempio ed utilizzare la notazione:

$$\mathbf{I = (131,7 \pm 3,6) \text{ mA}}$$



Eseguendo delle operazione fra risultati di misure, la considerazione da ricordare è che aggiungere ad una cifra incerta un'altra cifra non aumenta la quantità di informazione. Pertanto:

1) il valore di una misura diretta deve avere tante cifre significative quante sono quelle ottenute all'aver eseguito la lettura di uno strumento analogico al decimo di divisione o al digit nel caso di uno strumento digitale;

2) nelle quantità ottenute dalle operazioni di somma o differenza di misure il risultato deve avere tante cifre decimali quante ne ha la misura col minor numero:

$$\begin{array}{r} 5,94 + \\ 789,1 + \\ 15,426 = \\ 810,466 \Rightarrow \\ \Rightarrow 810,5 \end{array}$$

3) nelle quantità ottenute come prodotto o rapporto di misure il risultato deve avere tante cifre significative quante ne ha la misura col minor numero:

$$\begin{array}{r} 2,43 \times \\ 6,9 = \\ 16,767 \text{ che va arrotondato a } 17 \text{ in} \\ \text{modo tale che il risultato abbia due cifre} \\ \text{significative come } 6,9. \end{array}$$

## INCERTEZZE ASSOLUTE E RELATIVE

La quantificazione assoluta dell'incertezza non consente di valutare la bontà di una misura. Per esempio l'incertezza 0,12 cm è indice di una ottima misura se il valore misurato è di 730,23 cm ma rappresenta una misura di scarsa qualità se il valore misurato è - 0,35 cm.

Per questo motivo spesso si rapporta l'incertezza al valore della misura:  $\frac{\sigma(X)}{|X|}$

(**incertezza relativa di X**)

per cui le due misure precedenti presentano incertezze relative rispettivamente di

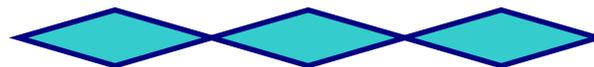
$$\frac{0,12}{730,23} = 0,0001643 = 0,01643 \% \approx 0,016 \% \quad \text{e} \quad \frac{0,12}{0,35} = 0,3429 = 34,29 \% \approx 34 \%$$

dove nell'ultimo passaggio si utilizza la convenzione di utilizzare due cifre significative per esprimere le incertezze. Il confronto sulla qualità delle due misure è immediato.

Nella formula dell'incertezza relativa  $\sigma(X)$  rappresenta la deviazione standard sperimentale di tipo A o B (cioè la migliore stima dell'incertezza della misura) e X la media aritmetica delle misure o il valore dell'unica misura effettuata (cioè la migliore stima del valore vero).

La strumentazione di uso comune produce incertezze relative<sup>[28]</sup> dell'ordine del %. Valori superiori al 10% non sono sempre accettabili<sup>[29]</sup>. Ottenere incertezze inferiori allo 0,1 % senza aver utilizzato strumentazione di qualità o metodi particolari è spesso sinonimo di sbagli di calcolo.

Va prestata attenzione al denominatore della formula dell'incertezza relativa: le incertezze relative perdono significato se le misure sono pressoché nulle. In questo caso eventuali confronti possono basarsi solo sulle incertezze assolute.



## MISURE INDIRETTE

spesso un misurando Y non è misurato direttamente ma è determinato a partire da altre M grandezze (misurate direttamente)  $X_1, X_2, \dots, X_i, \dots, X_M$  attraverso una relazione

$$Y = f(X_1, X_2, \dots, X_i, \dots, X_M) \quad (1)$$

che può essere una definizione o una legge geometrica, fisica o di qualunque altro tipo che lega fra loro diverse grandezze fisiche.

Una stima (misura derivata) y della grandezza Y è data da:

<sup>28</sup> dato che l'incertezza relativa è una quantità adimensionale consente il confronto fra le prestazioni di strumenti che misurano grandezze fisiche diverse: un metro con divisioni al mm è migliore di una bilancia con portata 1 kg e sensibilità 10 g/div

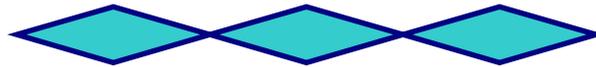
<sup>29</sup> in applicazioni particolari sono tollerate anche incertezze del 100% perché sinonimo di quantità diverse da zero e indicatori dell'ordine di grandezza del valore della grandezza in esame: cosa preferireste se vi dicessero che vincerete ad una lotteria (250±250)€ oppure (10000±10000)€ ?

$$Y \approx y = f(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_i, \dots, \bar{X}_M) \quad (1)$$

dove col simbolo  $\bar{X}_i$  si intende la media aritmetica delle  $N_i$  misure  $x_{ik}$  ( $k=1, N_i$ ):  $\bar{X}_i = \frac{\sum_{k=1, N_i} x_{ik}}{N_i}$

o l'unica determinazione  $x_i$  nel caso di una sola misura.

Questa stima  $y$  di  $Y$  risulta corretta solo se la funzione (1) è lineare<sup>[30]</sup>; altrimenti costituisce un'approssimazione valida al primo ordine di uno sviluppo della funzione in serie di Taylor (sviluppo nell'intorno del punto determinato dai risultati delle misure:  $[\bar{X}_1, \bar{X}_2, \dots, \bar{X}_i, \dots, \bar{X}_M]$ ).



### PROPAGAZIONE DELLE INCERTEZZE ASSOLUTE

La deviazione standard sperimentale del risultato di una misurazione ottenuto dai valori di altre grandezze fra loro indipendenti è detta **incertezza standard combinata**:

$$\sigma(y) \approx \sigma_s(y) = \sqrt{\sum_{i=1, M} \left( \frac{\partial f}{\partial X_i} \right)_{X_i = \bar{X}_i}^2 \sigma_s^2(\bar{X}_i)} \quad (2)$$

Essa è basata sullo sviluppo della funzione (1) in serie di Taylor al primo ordine effettuato nell'intorno del punto determinato dai valori veri delle grandezze misurate direttamente (valori veri approssimati dalle rispettive medie aritmetiche delle misure:  $[\bar{X}_1, \bar{X}_2, \dots, \bar{X}_i, \dots, \bar{X}_M]$ ).

Anche le derivate parziali  $\partial f / \partial X_i$  andrebbero calcolate, in linea di principio, nei valori attesi delle  $X_i$ . Queste derivate sono dette **coefficienti di sensibilità** perché descrivono di quanto varia la grandezza  $Y$  al variare delle  $X_i$ . La varianza combinata  $\sigma^2(y)$  viene quindi vista come somma di  $M$  termini ognuno dei quali rappresenta la stima della varianza di  $Y$  dovuta alla variazione del particolare  $X_i$ .

Quando la funzione (1) non è nota i coefficienti di sensibilità sono ricavabili sperimentalmente misurando la variazione di  $Y$  al variare di una particolare  $X_i$  mentre le altre grandezze  $X_{j \neq i}$  restano costanti.

Anche in questo caso con i simboli  $\bar{X}_i$  e  $\sigma(\bar{X}_i)$  si intendono le migliori determinazioni del valore vero della grandezza  $X_i$  e la stima dell'incertezza della sua misura.

Pertanto:

- tipo A) media aritmetica e deviazione standard sperimentale della media aritmetica
- tipo B) unica determinazione e incertezza di tipo B

Qualora la non linearità della  $f$  nell'intorno dei valori medi  $\bar{X}_i$  sia significativa, all'equazione (2) vanno sommati termini di ordine più elevato.

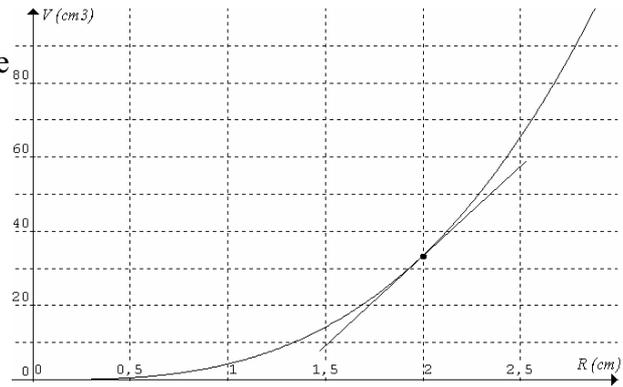
Se invece le grandezze non sono indipendenti ma correlate l'espressione più appropriata per la varianza combinata richiede l'uso delle covarianze. Va notato che se viene utilizzato uno stesso strumento o dato di riferimento o calibrazione con una incertezza significativa, possono introdursi correlazioni non trascurabili fra grandezze altrimenti indipendenti.

**Nelle esperienze svolte durante il corso la non linearità e la correlazione saranno in generale trascurabili e la (2) verrà ritenuta sufficientemente accurata per i nostri scopi.**

<sup>30</sup> la stima sarebbe distorta perché  $E(f(X_1, X_2, \dots, X_M)) = f(E(X_1), E(X_2), \dots, E(X_M))$  solo se la funzione  $f$  è lineare

Questa stima è quindi esatta solo se la funzione (1) è lineare; altrimenti costituisce una approssimazione al primo ordine dello sviluppo in serie di Taylor

**ESEMPIO**<sup>[31]</sup>:  $V = 4/3 \pi R^3$   
 $\sigma_s(V) = 4\pi R^2 \sigma_s(R)$



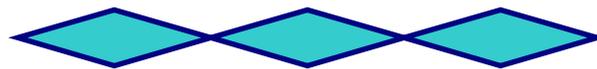
Nel grafico si vede quanto sia linearizzabile la funzione  $V = 4/3 \pi R^3$  nell'intorno della misura  $R = 2$  cm: è sufficiente che  $\sigma_s(R)$  non faccia variare i risultati di una misura di  $R$  fuori dell'intervallo [1,8 cm; 2,2 cm] in cui la retta tangente approssima "bene" la funzione.

**SBAGLIO FREQUENTE**

$Y = a X + b X^2$   ~~$\sigma(Y) = \sqrt{a^2 \sigma^2(X) + 4b^2 x^2 \sigma^2(X)} = \sqrt{a^2 + 4b^2 x^2} \sigma(X)$~~

anziché  $\sigma(Y) = |(a + 2 b x)| \sigma(X)$

la formula di propagazione è ottenuta considerando variabili indipendenti;  $X$  e  $X^2$  non lo sono !



**PROPAGAZIONE DELLE INCERTEZZE RELATIVE:**

se la funzione (1) è un monomio:  $Y = c X_1^{p_1} X_2^{p_2} \dots X_M^{p_M}$  l'incertezza standard combinata può essere calcolata tenuta più rapidamente mediante le incertezza relative.

Considerando che i coefficienti di sensibilità valgono  $\partial f / \partial X_i = Y p_i X_i$  sostituendo l'espressione del monomio nella (2) si ottiene, col solito significato dei simboli:

$$\frac{\sigma_s(y)}{y} = \sqrt{\sum_{i=1,M} p_i^2 \left( \frac{\sigma_s(\bar{X}_i)}{\bar{X}_i} \right)^2} \quad (3)$$

**VALIDA SOLO PER MONOMI** <sup>[32]</sup>

È consigliato l'uso, quando possibile, della (3) al posto della (2) per due motivi:

- il calcolo delle derivate, appena la funzione diventa un poco complessa, diventa pesante
- non è facile distinguere i contributi dati dalle misure dirette e quindi verificare la bontà del calcolo.

**SBAGLIO FREQUENTE**

$Y = a X + b X_0$   ~~$\sigma(Y)/|y| = |a| \sigma(X)/|x|$~~

anziché  $\sigma(Y)/|y| = |a| \sigma(X) / |a x + b X_0|$

$Y$  non è sotto forma di monomio: l'incertezza relativa è data dal rapporto  $\sigma(Y) = |a| \sigma(X)$  diviso per  $|y| = |a x + b X_0|$  !

<sup>31</sup> se la (1) è funzione della sola  $X$  allora la (2) diventa banalmente  $\sigma_s(Y) = |dY/dX|_x \sigma_s(X)$

<sup>32</sup> se la (1) è funzione della sola  $X$  allora la (3) diventa banalmente  $\sigma_s(Y)/|y| = |p| \sigma_s(X)/|x|$

Il seguente esempio mostra come la (3) possa essere utilizzata per scegliere, in fase di progetto di una misura, la strumentazione più idonea:

Una barra di metallo ( $\rho \approx 8 \text{ g/cm}^3$ ) ha le dimensioni nominali 5 mm x 5 cm x 5 dm.  
 Si vuole misurare  $\rho$  avendo a disposizione una bilancia con  $1/s = 10 \text{ g/div}$ , un Palmer, un calibro e un metro a nastro (con divisioni di 1 mm).  
 Ottimizzare l'uso degli strumenti (riferirsi a quelli utilizzati nel corso) e stimare la minima incertezza ottenibile nella misura di  $\rho$ .

Considerando rispettivamente a, b e c i tre lati della barra si ha:

$$V = a b c \approx 0,5 \times 5 \times 50 = 125 \text{ cm}^3; \quad m = V \rho \approx 125 \times 8 = 1000 \text{ g}$$

Quindi la misura di massa avrà un'incertezza relativa  $\frac{\sigma(m)}{m} = 0,29 \%$  dove  $\sigma(m) = 10\text{g}/\sqrt{12}=2,9 \text{ g}$

L'incertezza relativa complessiva sarà invece data dalla (3):

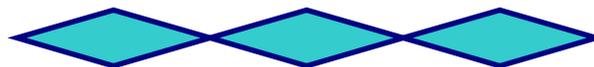
$$\frac{\sigma(\rho)}{\rho} = \sqrt{\left(\frac{\sigma(m)}{m}\right)^2 + \left(\frac{\sigma(a)}{a}\right)^2 + \left(\frac{\sigma(b)}{b}\right)^2 + \left(\frac{\sigma(c)}{c}\right)^2}$$

la scelta ottimale degli strumenti è quella per cui i vari termini sono dello stesso ordine di grandezza:

	Palmer ( $\sigma = 0,0029 \text{ mm}$ )	calibro ( $\sigma = 0,014 \text{ mm}$ )	metro ( $\sigma = 0,29 \text{ mm}$ )
a 5 mm	0,058 %	0,28 %	5,8 %
b 5 cm	fuori portata	0,028 %	0,58 %
c 5 dm	fuori portata	fuori portata	0,058 %

utilizzando rispettivamente il Palmer, il calibro e il metro  $\frac{\sigma(\rho)}{\rho} = 0,30\%$  mentre utilizzando il

calibro (meno "costoso") anche per il lato a  $\frac{\sigma(\rho)}{\rho} = 0,41\%$  (sostanzialmente equivalente).



●●● **Nel riportare il risultato di una misura e la sua incertezza occorre:**

- dare la descrizione completa di come è definita la grandezza Y;
- fornire il valore della stima della grandezza e la deviazione standard (combinata se la misura è indiretta) con le opportune unità di misura e, qualora sia appropriato, anche l'incertezza relativa;
- descrivere il metodo utilizzato per ricavare dai dati il valore del risultato e la sua incertezza;
- nel caso di misure derivate riportare il valore e l'incertezza (anche relativa) di ogni grandezza che contribuisce alla grandezza Y e la descrizione di come sono stati ottenuti;
- presentare l'analisi dei dati in modo tale che sia possibile comprendere rapidamente i passi più importanti e si possa ripetere il calcolo del risultato se necessario;
- fornire tutti i termini correttivi e le costanti utilizzate.

## ESEMPI

- 1) Date le misure  $L1 = (20,42 \pm 0,26)$  cm  
 $L2 = (10,11 \pm 0,43)$  cm

### 1) misurare la lunghezza $L_s = L1 + L2$

$$L_s = 30,53 \text{ cm} \pm \sigma_s(L_s)$$

$$\sigma_s(L_s) = \sqrt{\left(\frac{\partial L_s}{\partial L1}\right)^2 \sigma_s^2(L1) + \left(\frac{\partial L_s}{\partial L2}\right)^2 \sigma_s^2(L2)} = \sqrt{(+1)^2 \sigma_s^2(L1) + (+1)^2 \sigma_s^2(L2)} = \sqrt{0,26^2 + 0,43^2} = 0,5025 \text{ cm}$$

$$L_s = (30,53 \pm 0,50) \text{ cm} \quad (1,6\%)$$

### 2) misurare la lunghezza $L_d = L1 - L2$

$$L_d = 10,31 \text{ cm} \pm \sigma_s(L_d)$$

$$\sigma_s(L_d) = \sqrt{\left(\frac{\partial L_d}{\partial L1}\right)^2 \sigma_s^2(L1) + \left(\frac{\partial L_d}{\partial L2}\right)^2 \sigma_s^2(L2)} = \sqrt{(+1)^2 \sigma_s^2(L1) + (-1)^2 \sigma_s^2(L2)} = \sqrt{0,26^2 + 0,43^2} = 0,5025 \text{ cm}$$

$$L_d = (10,51 \pm 0,50) \text{ cm} \quad (4,8\%)$$

### 3) misurare l'area $A = L1 \times L2$

$$A = 206,4462 \text{ cm}^2 \pm \sigma_s(A)$$

$$\sigma_s(A) = \sqrt{\left(\frac{\partial A}{\partial L1}\right)^2 \sigma_s^2(L1) + \left(\frac{\partial A}{\partial L2}\right)^2 \sigma_s^2(L2)} = \sqrt{(L2)^2 \sigma_s^2(L1) + (L1)^2 \sigma_s^2(L2)} = \sqrt{10,11^2 0,26^2 + 20,42^2 0,43^2} = 9,166 \text{ cm}^2$$

$$A = (206,4 \pm 9,2) \text{ cm}^2 \quad (4,6\%)$$

- 2) Misurare la grandezza derivata "densità di una sfera":  $\rho = \frac{m}{\frac{4\pi}{3} r^3}$  (a)

Supponiamo di aver misurato la massa  $m$  ( $m = \bar{m} \pm \sigma_s(\bar{m})$ ) e il raggio  $r$  ( $r = \bar{r} \pm \sigma_s(\bar{r})$ ).

Propagando le incertezze assolute si ha:

$$\sigma(\rho) = \sqrt{\left(\frac{1}{\frac{4\pi}{3} \bar{r}^3}\right)^2 \sigma_s^2(\bar{m}) + \left(-3 \frac{\bar{m}}{\frac{4\pi}{3} \bar{r}^4}\right)^2 \sigma_s^2(\bar{r})} \quad (b)$$

In questo caso però è possibile, e quindi preferibile, utilizzare la formula di propagazione per le incertezze relative che è di uso più immediato (non richiede calcoli di derivate):

$$\frac{\sigma(\rho)}{\rho} = \sqrt{\left(\frac{\sigma_s(\bar{m})}{\bar{m}}\right)^2 + (-3)^2 \left(\frac{\sigma_s(\bar{r})}{\bar{r}}\right)^2} \quad (c)$$

da cui si ottiene:

$$\sigma(\rho) = \rho \sqrt{\left(\frac{\sigma_s(\bar{m})}{\bar{m}}\right)^2 + 9 \left(\frac{\sigma_s(\bar{r})}{\bar{r}}\right)^2} \quad (d)$$

Va notato che matematicamente le due formule (b) e (d) sono coincidenti: entrambe derivano dalla linearizzazione della (a) mediante sviluppo in serie di Taylor nell'intorno del punto  $(\bar{m}, \bar{r})$ <sup>[33]</sup>

<sup>33</sup> anche se  $\rho$  varia col cubo di  $r$ , nell'intorno di  $\bar{r}$  la funzione è linearizzabile: è sufficiente considerare valori di  $r$  che si discostino di poco da  $\bar{r}$ :  $\frac{r-\bar{r}}{\bar{r}} \ll 1$ . Nel nostro caso questo implica  $\frac{\sigma(\bar{r})}{\bar{r}} \ll 1$ , relazione generalmente verificata in tutte le misure.

- 3) Misurare la densità di un cilindro di massa  $M$ , altezza  $h$  e diametro  $d$  a partire dai dati:

$$M = 13,2 \ 12,4 \ 14,0 \ 12,8 \ 14,6 \ 13,5 \text{ g}$$

$$h = 10,02 \text{ cm (una sola misura con strumento analogico)}$$

$$d: \text{ in misure precedenti si era ottenuto } 8,500 \text{ mm mediante calibro (50}\mu\text{m/div)}$$

$$\mathbf{M:}$$
 dai dati si ricava  $\sum_{i=1,6} m_i = 80,5 \text{ g}$  e quindi  $\bar{M} = \frac{\sum_{i=1,6} m_i}{6} = 13,427 \text{ g}$ ; inoltre

$$\sum_{i=1,6} m_i^2 = 1083,25 \text{ g}^2 \text{ e quindi } \sigma_M = \sqrt{\frac{\sum_{i=1,6} m_i^2 - 6 \cdot \bar{M}^2}{5}} = 0,801 \text{ g} \text{ da cui } \sigma_{\bar{M}} = \frac{\sigma_M}{\sqrt{6}} = 0,326 \text{ g}$$

$$\text{Pertanto } M = (13,43 \pm 0,33) \times 10^{-3} \text{ kg}$$

**h:** l'unica misura è riportata al decimo di mm; quindi la divisione dello strumento è il mm;

$$\sigma_B = 1 \text{ mm} / \sqrt{12} = 0,29 \text{ mm}$$

$$\text{Pertanto } h = (100,20 \pm 0,29) \times 10^{-3} \text{ m}$$

**d:**  $\sigma_B = 50 \mu\text{m} / \sqrt{12} = 14,43 \mu\text{m}$

$$\text{Pertanto } d = (8,500 \pm 0,014) \times 10^{-3} \text{ m}$$

$$\text{e quindi } \rho = \frac{\bar{M}}{\frac{\pi}{4} \bar{d}^2 \bar{h}} \pm \sqrt{\left( \frac{1}{\frac{\pi}{4} \bar{d}^2 \bar{h}} \right)^2 \sigma_M^2 + \left( -2 \frac{M}{\frac{\pi}{4} \bar{d}^3 \bar{h}} \right)^2 \sigma_d^2 + \left( -\frac{M}{\frac{\pi}{4} \bar{d}^2 \bar{h}^2} \right)^2 \sigma_h^2}$$

$$\rho = 2362 \pm \sqrt{(175875)^2 (0,33 \times 10^{-3})^2 + (-555766)^2 (0,014 \times 10^{-3})^2 + (-23573)^2 (0,29 \times 10^{-3})^2} =$$

$$= 2362 \pm \sqrt{(58,039)^2 + (7,781)^2 + (6,836)^2} = (2362 \pm 59) \text{ kg/m}^3$$

Alternativamente, poiché la relazione  $\rho = \frac{M}{\frac{\pi}{4} d^2 h}$  è un monomio, usando le incertezze relative:

$$M = (13,43 \pm 0,33) \times 10^{-3} \text{ kg} \rightarrow \frac{\sigma_M}{M} = \frac{0,33 \times 10^{-3} \text{ kg}}{13,43 \times 10^{-3} \text{ kg}} = 0,02457 = 2,5\%$$

$$d = (8,500 \pm 0,014) \times 10^{-3} \text{ m} \rightarrow \frac{\sigma_d}{d} = \frac{0,014 \times 10^{-3} \text{ m}}{8,5 \times 10^{-3} \text{ m}} = 0,0016471 = 0,16\%$$

$$h = (100,20 \pm 0,29) \times 10^{-3} \text{ m} \rightarrow \frac{\sigma_h}{h} = \frac{0,29 \times 10^{-3} \text{ m}}{100,2 \times 10^{-3} \text{ m}} = 0,0028942 = 0,29\%$$

$$\text{quindi } \rho = \frac{\bar{M}}{\frac{\pi}{4} \bar{d}^2 \bar{h}} = 2362 \text{ kg/m}^3 \text{ e } \sigma_\rho = \left( \frac{\sigma_\rho}{\rho} \right) \rho = \sqrt{(+1)^2 \left( \frac{\sigma_M}{M} \right)^2 + (-2)^2 \left( \frac{\sigma_d}{d} \right)^2 + (-1)^2 \left( \frac{\sigma_h}{h} \right)^2} \rho$$

pertanto:

$$\rho = 2362 \left( 1 \pm \sqrt{(0,025)^2 + 4(0,0029)^2 + (0,0016)^2} \right) = 2362(1 \pm 0,02571) = (2362 \pm 61) \text{ kg/m}^3$$

La piccola differenza nei due casi sulla seconda cifra delle incertezze è dovuta al fatto che le incertezze vengono riportate con due cifre significative e quindi la meno significative risente dell'approssimazione

## ESERCIZI

1) Arrotondare a 3,2,1 cifre significative ciascuno dei valori:

54,46    23,00    0,04445    995,4

2) Riportare con due cifre significative i valori:

$\pi$  e  $\sqrt{2}$   $\bar{g}$  0,432    324,43     $3 \times 10^7$     2,3% di 342,4    26% di 0,421    0,13% di  $723 \times 10^4$

3) A Calcolare T avendo ottenuto le misure:  $(100 \pm 10) \mu\text{s}$ ;  $(110,0 \pm 5,0) \mu\text{s}$ ;  $(90 \pm 10) \mu\text{s}$

B Quale sarebbe il risultato se le tre incertezze fossero pari a  $5,0 \mu\text{s}$ ?

C E se fossero pari a  $3,0 \mu\text{s}$ ?

4)  $\theta = \frac{C T' + C_s T''}{C + C_s}$  misurare indirettamente  $\theta$  a partire dalle misure di  $T'$  e  $T''$

supporre note  $C$  e  $C_s$  e che siano  $\sigma_{T'} = \sigma_{T''} = \sigma_T$

5) Sapendo che il prodotto della frequenza per la lunghezza d'onda delle radiazioni elettromagnetiche è pari alla velocità della luce, se misura all'1% la frequenza  $540 \times 10^{12}$  Hz, quanti  $\mu\text{m}$  misura la lunghezza d'onda? Quante cifre significative servono per  $c$ ?

6) In una serie di 100 misure di  $L$  si ottiene  $\Sigma L_i = 2700$  mH e  $\Sigma L_i^2 = 82\,800$  mH<sup>2</sup>.

Qual è il valore dell'induttanza?

7) Due resistenze  $R_1 = 200 \Omega$  al 2,0% e  $R_2 = 800 \Omega$  al 2,0% vengono collegate prima in serie e poi in parallelo. Quanto vale la resistenza complessiva nei due casi?

8) Due capacità  $C' = 0,2 \mu\text{F}$  ( $\frac{\sigma(C')}{C'} = 1,5\%$ ) e  $C'' = (800,0 \pm 4,0)$  nF vengono collegate prima in serie e poi in parallelo. Quanto vale la capacità complessiva nei due casi?

9) Determinare la misura di una sbarra di lunghezza nominale 150 mm con uno strumento che nell'intervallo misurato richiede un termine correttivo di  $+60 \mu\text{m}$ . Vengono effettuate  $n = 20$  misure sotto condizioni di ripetibilità (le misure sono espresse in mm):

150,14 150,04 149,97 150,08 149,93    149,99 150,13 150,09 149,89 150,01

149,99 150,04 150,02 149,94 150,19    149,93 150,09 149,83 150,03 150,07

10) Determinare il valore di  $P = 200 \frac{T}{\mu d^3}$  kPa a partire dai seguenti dati:

$T = 10,00$  K (con incertezza del 2,0%)

$\mu_1 = (23,0 \pm 2,0)$  e  $\mu_2 = (27,0 \pm 4,0)$

$d$  ([m]  $\pm 0,0020$  m) = 0,970; 0,990; 1,000; 1,010; 1,030

Quale misura influisce maggiormente sull'incertezza del risultato finale?

Quali sono le dimensioni della costante 200 ?

11) Quanto vale l'incertezza assoluta delle seguenti g.f. misurate indirettamente:

$$Y = \ln X \quad a = \frac{F}{m} \quad z = a X + b Y^2 + c X Y$$

12) Il valore della costante elastica di una molla  $K$  dipende dal modulo di scorrimento  $G$  del materiale, dal diametro  $d$  del filo che la costituisce, dal diametro  $D$  delle spire e dal numero  $N$  di spire secondo la

$$\text{relazione } K = \frac{G d^4}{8 N D^3}.$$

Quali sono le dimensioni di  $G$  nel S.I.? Qual è l'espressione dell'incertezza relativa di  $G$ ?

13) La misura, effettuata con un calibro ventesimale, dello spessore di un blocco di 400 fogli di carta è pari a 4 cm (con tutte le cifre significative che occorrono). Qual è lo spessore  $s$  dei fogli supponendo che siano uguali? E senza questa ipotesi?

14) Le due misure  $I_1 = (7,40 \pm 0,50)$  mA e  $I_2 = (8,50 \pm 0,50)$  mA possono essere ritenute uguali? Se l'incertezza fosse di 0,10 mA quali conclusioni si potrebbero trarre riguardo gli errori sistematici?

15) Con tre diversi strumenti (di cui uno è uno solo è starato) sono state effettuate le seguenti misure della stessa grandezza:

A)	210	211	212	213	214
B)	213	214	214	214	215
C)	215,8	215,9	216,0	216,1	216,2

Evidenziare lo strumento starato.

16)  $P = V I \cos \phi$        $V = (5,00 \pm 0,10)$  V     $I = (2,00 \pm 0,10)$  mA     $\phi = (0,20 \pm 0,10)$  rad  
Misurare P.

17) Se  $V = (6,0 \pm 1,0)$  V;  $R = 4,9$  k $\Omega$  al 10% e  $I = (1,00 \pm 0,10)$  mA le tre grandezze soddisfano la legge di Ohm:  $V = R I$ ?  
Se i valori indicati non fossero incertezze ma tolleranze cosa se ne potrebbe dedurre?

18) Determinare la densità di una sfera di massa  $m$  e diametro  $d$  a partire dalle misure effettuate con una bilancia con  $1/s = 5$  g/div e un calibro ventesimale:

$m$  [kg] 0,71 0,72 0,7 0,72       $d$  [cm] 8,11 8,115 8,13 8,125

[lascio a voi il compito di riscrivere i risultati in modo corretto!]

È sufficiente l'approssimazione  $\pi = 3,14$ ?

19) Con un calibro decimale si misurano il diametro esterno  $D$  di un tubo ( $D \approx 2$  cm) e il suo spessore  $\delta \approx 1$  mm allo scopo di misurare l'area della sua sezione.

La bontà della misura aumenterebbe se venisse misurato direttamente il diametro interno  $d$ ?

20) Si vuol misurare  $g$  utilizzando un pendolo di lunghezza  $L = (100,00 \pm 0,20)$  cm il cui periodo di

oscillazione è ben descritto da  $T = 2\pi \sqrt{\frac{L}{g}}$ . Si ottengono le misure:

$T$  (s  $\pm 0,010$  s) 2,00 1,80 1,90 2,20 2,00 2,10.

Quanto vale  $g$ ? Se viene fornita un'ulteriore misura  $T = (1,98 \pm 0,10)$  s cosa fornisce la media pesata fra questo risultato e il precedente?

21) Di un disco si misura la massa  $M = (100,00 \pm 0,20)$  g e si eseguono 100 misure del diametro  $D$  (con incertezza di tipo B trascurabile) ottenendo dai dati le quantità:

$$\sum_{i=1,100} D_i = 100,00 \text{ cm e } \sum_{i=1,100} D_i^2 = 101,00 \text{ cm}^2. \text{ Calcolare il momento di inerzia } I = M \frac{D^2}{8}.$$

22) Un tondino di acciaio viene suddiviso in 200 provini che vengono inviati a tre ditte diverse per misurarne il modulo di Young:

ditta	# provini	$\bar{E}$ [ $10^{10}$ Nm $^{-2}$ ]	$\sigma(E)$ [ $10^{10}$ Nm $^{-2}$ ]
A)	36	20,0	1,5
B)	64	20,2	2,0
C)	100	20,1	2,0

Quanto vale  $E$ ? Ricavare i valori  $E_{MAX}$  e  $E_{min}$  che delimitano un intervallo di confidenza di  $2 \sigma$ .

23) Misurare indirettamente la quantità di calore Q ceduta da un corpo di massa m e calore specifico c quando la sua temperatura scende da T a T<sub>0</sub>: Q = m c (T – T<sub>0</sub>). Dati:

T<sub>0</sub> = 250 K (misura ottenuta con incertezza trascurabile)

m = 24,75 g (misura effettuata con bilancia con sensibilità 1 g/div; a piatto scarico la bilancia indica 0,25 g)

c = (500±20) J/kg/K (con un livello di confidenza del 95% ricavato da incertezze di tipo A)

T: sono stati ottenuti i seguenti risultati (raggruppati in classi)

T [K]	270	280	290	300	310	320	330
#	3	2	5	8	6	5	1

$$[\sum T_i = 9\ 010\text{ K}; \sum T_i^2 = 2\ 713\ 500\text{ K}^2]$$

Calcolare la probabilità (con una sola cifra significativa) che Q sia maggiore di 0,6 kJ (assumere che il risultato sia distribuito gaussianamente).

24) 100 componenti scartati vengono analizzati e classificati in base al numero di porte logiche guaste. 20 componenti presentano da 1 a 5 porte guaste

20 componenti presentano da 6 a 10 porte guaste

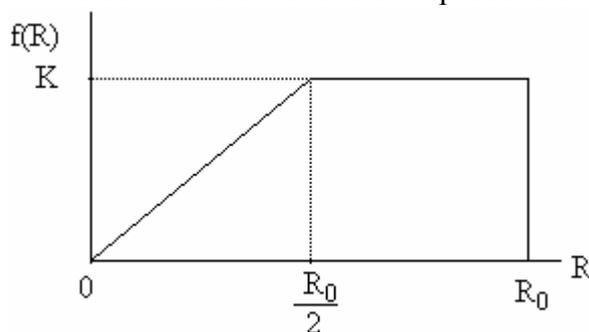
40 componenti presentano da 11 a 15 porte guaste

20 componenti presentano da 16 a 20 porte guaste.

Si vuol verificare se la distribuzione è uniforme applicando il test del  $\chi^2$  per accettare l'ipotesi se il risultato è compreso nell'intervallo di confidenza  $E(\chi^2) \pm \sigma(\chi^2)$ . Qual è il risultato del test?

25) La v.a. X è distribuita uniformemente. 750 misure di X vengono raggruppate in 10 classi contigue di uguale larghezza contenenti n<sub>i</sub> misure. Avendo calcolato  $\sum n_i^2 = 57150$ ; il campione risulta in accordo con la distribuzione madre?

26) La densità di probabilità del valore delle resistenze prodotte da una fabbrica è la seguente:



- A) Determinare (con le corrette unità di misura) K, E(R) e la varianza di R (R<sub>0</sub> = 1kΩ)
- B) Scrivere esplicitamente la densità di probabilità della media aritmetica di 100 misure che seguono la precedente distribuzione.
- C) suddividere l'intervallo (0-R<sub>0</sub>) in 10 classi di larghezza R<sub>0</sub>/10; negli stessi intervalli le frequenze di 300 misure sono: 6, 18, 24, 21, 36, 48, 35, 32, 43, 37. Utilizzando il test del  $\chi^2$  dire se è lecito assumere che queste 300 misure seguano la distribuzione analizzata. Usare un intervallo di confidenza largo 2 deviazioni standard e centrato intorno alla media attesa.

# SOLUZIONI

- 1) 

54,5	23,0	0,044 5	995
54	23	0,044	$1,0 \times 10^3$
$0,05 \times 10^3$	$0,02 \times 10^3$	0,04	$10^3$
- 2) 3,1 2,7 1,4 9,8 0,43  $0,32 \times 10^3$   $30 \times 10^6$  7,9 0,11  $9,4 \times 10^3$
- 3) A le tre misure sono compatibili ( $|t|=1,79$ ) media pesata:  $T = (105,0 \pm 4,1) \mu s$   
 B le tre misure sono compatibili ( $|t|=2,83$ ) media pesata:  $T = (100,0 \pm 2,9) \mu s$   
 C le ultime due misure non sono compatibili ( $|t|=4,71$ ) media aritmetica:  $T = (100,0 \pm 5,8) \mu s$
- 4) si tratta di calcolare  $\sigma_\theta = \sigma_T \sqrt{\frac{C^2 + C_s^2}{(C + C_s)^2}}$
- 5) Essendo  $c \approx 2,998 \times 10^8$  m/s noto senza incertezza  $\rightarrow \lambda = (0,5556 \pm 0,0056) \mu m$   
 3 cifre introducono un errore trascurabile rispetto all'incertezza dell'1%
- 6)  $L = \frac{\sum L_i}{N} \pm \sqrt{\left( \frac{\sum L_i^2}{N} - \left( \frac{\sum L_i}{N} \right)^2 \right) \frac{1}{N-1}} = (27,0 \pm 1,0) \text{ mH}$
- 7) da  $R_1 = (200,0 \pm 4,0) \Omega$  e  $R_2 = (800 \pm 16) \Omega \rightarrow R_s = (1000 \pm 16) \Omega$ ;  $R_p = (160,0 \pm 2,6) \Omega$
- 8) da  $C' = (200,0 \pm 3,0) \text{ nF}$  e  $C'' = (800,0 \pm 4,0) \text{ nF} \rightarrow C_p = (1000,0 \pm 5,0) \text{ nF}$   $C_s = (160,0 \pm 2,5) \text{ nF}$
- 9)  $\bar{X} = 150,02 \text{ mm}$ ;  $\bar{X}$  corretta = 150,08 mm;  $\sigma_s(X) = 0,0895 \text{ mm}$ ;  $\sigma_s(\bar{X}) = 0,020 \text{ mm}$   
 $X = (150,020 \pm 0,020) \text{ mm}$
- 10)  $T = (10,00 \pm 0,20) \text{ K}$  0,20 %  
 $\mu = (23,8 \pm 1,8)$  (media pesata) 7,6 %  $\leftarrow$   
 $d = (1,000 \pm 0,010) \text{ m}$  (media aritmetica)  $3 \times 1,0$  %  
 $P = (84,0 \pm 6,9) \text{ kPa}$  8,2 %  
 $[M][L]^2[t]^{-2}[\theta]^{-1} = [\text{energia}]/[\text{temperatura}]$
- 11)  $\sigma(Y) = \frac{\sigma(X)}{X}$   $\sigma(a) = \sqrt{\left(\frac{1}{m}\right)^2 \sigma_F^2 + \left(\frac{F}{m^2}\right)^2 \sigma_m^2}$   $\sigma(Z) = \sqrt{(a + cY)^2 \sigma^2(X) + (2bY + cX)^2 \sigma^2(Y)}$
- 12)  $[L]^{-1}[t]^{-2}[M] = [\text{forza}]/[\text{superficie}]$   $\frac{\sigma_G}{G} = \sqrt{\left(\frac{\sigma_K}{K}\right)^2 + 16\left(\frac{\sigma_d}{d}\right)^2 + \left(\frac{\sigma_N}{N}\right)^2 + 9\left(\frac{\sigma_D}{D}\right)^2}$
- 13)  $S = (40\,000 \pm 14) \mu m$ ; considerando  $S = 400 s_i \rightarrow s = (100,000 \pm 0,035) \mu m$   
 considerando  $S = \sum s_i \rightarrow s = (100,00 \pm 0,72) \mu m$
- 14) Nel primo caso i valori distano  $1,6 \sigma$  e quindi sono uguali;  
 nel secondo  $7,8 \sigma$  (la differenza non è dovuta a soli effetti casuali)
- 15) A)  $212,00 \pm 0,71$  B)  $214,00 \pm 0,32$  C)  $216,000 \pm 0,071$   
 A) e B) sono compatibili ( $t=2,6 < 3$ );  
 C) non è compatibile né con A ( $t=5,4$ ) né con B ( $t=6,2$ ): è lo strumento starato
- 16)  $P = (9,80 \pm 0,56) \text{ mW}$
- 17) A)  $V = (6,0 \pm 1,0) \text{ V}$   $R I = (4,90 \pm 0,69) \text{ V}$   $t = 0,91 < 3$  soddisfano la legge  
 B)  $\sigma_V = \frac{2V}{\sqrt{12}} = 0,58 \text{ V}$ ;  $\sigma_R = \frac{0,98 \text{ k}\Omega}{\sqrt{12}} = 0,28 \text{ k}\Omega$ ;  $\sigma_I = \frac{0,2 \text{ mA}}{\sqrt{12}} = 0,058 \text{ mA}$  e quindi

$V = (6,00 \pm 0,58) \text{ V}$ ;  $R I = (4,90 \pm 0,40) \text{ V}$  i due valori disterebbero  $1,6 \sigma$  la legge sarebbe verificata anche in questo caso.

18)  $m [\text{kg}] \pm 0,0014 \text{ kg}$     0,7100 0,7200 0,7000 0,7200     $m = (0,7125 \pm 0,0048) \text{ kg}$   
 $d [\text{cm}] \pm 0,0014 \text{ cm}$     8,1100 8,1150 8,1300 8,1250     $d = (8,1200 \pm 0,0046) \text{ cm}$   
 $\rho = (2,542 \pm 0,018) 10^3 \text{ kg/m}^3$

si: è sufficiente perché  $\frac{3,14 - \pi}{\pi} = 0,051\%$  è trascurabile rispetto a  $\frac{\sigma(\rho)}{\rho} = 0,69\%$

19) si:  $\sigma[S(D,d)] < \sigma[S(D,\delta)] < \sigma[S(d,\delta)]$  con  $\sigma = 0,029 \text{ mm}$  e  
 $S(D,\delta) = \pi/4 D^2 - \pi/4 (D-\delta)^2$ ;     $S(d,\delta) = \pi/4 (d+\delta)^2 - \pi/4 d^2$ ;     $S(D,d) = \pi/4 D^2 - \pi/4 d^2$

20)  $T = (2,000 \pm 0,058) \text{ s} \rightarrow g = (987 \pm 57) \text{ cm/s}^2$ ;  $T = (1,995 \pm 0,050) \text{ s} \rightarrow g = (992 \pm 50) \text{ cm/s}^2$

21)  $I = (1,250 \pm 0,025) \times 10^{-6} \text{ kg m}^2$

22) dai valori  $(20,00 \pm 0,25) 10^{10} \text{ Nm}^{-2}$ ,  $(20,20 \pm 0,25) 10^{10} \text{ Nm}^{-2}$  e  $(20,10 \pm 0,20) 10^{10} \text{ Nm}^{-2}$  si ottengono la misura  $E = (20,10 \pm 0,13) 10^{10} \text{ Nm}^{-2}$  e l'intervallo  $(19,84 ; 20,36) 10^{10} \text{ Nm}^{-2}$

23)  $T_0 = 250 \text{ K}$      $m = (24,50 \pm 0,29) \text{ g}$      $95\% \Leftrightarrow z = 1,96 \rightarrow c = (500 \pm 10) \text{ J/kg/K}$

$$\sigma_s(\bar{T}) = \sqrt{\left(\frac{\sum T_i^2}{N} - \bar{T}^2\right) \frac{1}{N-1}} = 2,935 \text{ K} \rightarrow T = (300,3 \pm 2,9) \text{ K} \rightarrow Q = (616 \pm 38) \text{ J}$$

$P(Q > 600 \text{ J}) = 0,5 + \text{ERF}(13/38) = 0,63 \rightarrow P = 0,6$

24) Dato che gli intervalli hanno la stessa larghezza la probabilità è la stessa per tutti  $p_i = 25\%$  e quindi  $Np_i = 25$ . Pertanto:  $u_1 = u_2 = u_4 = -1,1557$   $u_3 = 3,464$  da cui  $u = 16$  che è al di fuori dell'intervallo  $3,00 \pm 2,45$ . Pertanto la distribuzione non è uniforme

25) si:  $u = \sum_{i=1,10} \frac{(n_i - 750 \times 0,1)^2}{750 \times 0,1 \times 0,9} = \frac{\sum n_i^2 - 2 \times 750 \times 0,1 \sum n_i + 10(750 \times 0,1)^2}{750 \times 0,1 \times 0,9} =$   
 $= \frac{57150 - 150 \times 750 + 10 \times 75^2}{750 \times 0,1 \times 0,9} = 13,3$  che è compreso in  $(9 \pm 8,5)$

26) A) Area unitaria  $\rightarrow K (R_0/2 * 1/2 + R_0/2 * 1) = 1 \rightarrow K = 4/(3R_0)$

$f(R) = 8R/(3R_0^2)$  se  $R < R_0/2$ ;  $f(R) = 4/(3R_0)$  se  $R > R_0/2$

$E(R) = 8(R_0^3/24)/(3R_0^2) + 4(R_0^2/2 - R_0^2/8)/(3R_0) = R_0/9 + R_0/2 = 11/18 R_0$

$\sigma^2(R) = E(R^2) - E^2(R) = 8(R_0^4/64)/(3R_0^2) + 4(R_0^3/3 - R_0^3/24)/(3R_0) - (11/18)^2 R_0^2 = 7/9 R_0^2$

B)  $f(R) = \frac{30}{\sqrt{2\pi} \sqrt{7R_0}} e^{-\frac{(\bar{R} - 11/18 R_0)^2}{2 \times 7/900 R_0^2}}$

C)

	$p_i$	$n_i$	$u_i$
$p_1 = P(0 \leq R < 0,1 R_0) = 4/300$	4/300	6	-1,00
$p_2 = P(0,1 R_0 \leq R < 0,2 R_0) = 12/300$	12/300	18	-1,68
$p_3 = P(0,2 R_0 \leq R < 0,3 R_0) = 20/300$	20/300	24	-0,93
$p_4 = P(0,3 R_0 \leq R < 0,4 R_0) = 28/300$	28/300	21	1,39
$p_5 = P(0,4 R_0 \leq R < 0,5 R_0) = 36/300$	36/300	36	0
$p_6 = P(0,5 R_0 \leq R < 0,6 R_0) = 40/300$	40/300	48	-1,36
$p_7 = P(0,6 R_0 \leq R < 0,7 R_0) = 40/300$	40/300	35	0,85
$p_8 = P(0,7 R_0 \leq R < 0,8 R_0) = 40/300$	40/300	32	1,36
$p_9 = P(0,8 R_0 \leq R < 0,9 R_0) = 40/300$	40/300	43	-0,51
$p_{10} = P(0,9 R_0 \leq R < 1,0 R_0) = 40/300$	40/300	37	0,51

$= 300/300$     300    si:  $u = 11,56$  compreso in  $(9 \pm 8,5)$