Capitolo 2

Gli errori sperimentali

2.1 L'errore di sensibilità (misura singola)

Definiamo **sensibilità** *di uno strumento la più piccola quantità che è possibile misurare.*

Con una riga da disegno possiamo valutare la misura di una lunghezza fino al millimetro. Con un "calibro" da laboratorio possiamo arrivare al decimo di millimetro e con un "laser" spingerci addirittura al centesimo di millimetro. Una sensibilità infinitamente piccola é di fatto impossibile da raggiungere.

Per questo motivo lo strumento usato, qualunque esso sia, ci consentirà di misurare una grandezza fisica solo fino ad un certo limite.

Supponiamo, per esempio, di usare una riga da disegno per calcolare la lunghezza L di un banco scolastico. Tale valore sia:

L = 78,2 cm

Poiché lo strumento non ci permette di apprezzare i decimi di millimetro (e qualunque valutazione "ad occhio" é da evitare perché poco scientifica e poco seria), dobbiamo esprimere il risultato ottenuto nel modo seguente:

L = 78,2 cm ± 0,1 cm

Questa notazione vuol dire che, nell'impossibilità di valutare il decimo di millimetro, dobbiamo supporre che:

- la lunghezza L considerata cade sicuramente tra 78,1 cm e 78,3 cm
- il valore centrale dell'intervallo L = 78,2 cm risulta essere il valore più ragionevole da attribuire alla lunghezza del banco

- l'errore dovuto alla sensibilità dello strumento ammonta a ± 0,1 cm (che rappresenta quindi il limite con cui noi possiamo conoscere la grandezza studiata).

Se invece eseguissimo la stessa misura con uno strumento di sensibilità diversa, potremmo scrivere:

L = 78 cm ± 1 cm	con uno strumento di sensibilità pari a 1 cm
L = 78, 24 cm ± 0,01 cm	con uno strumento di sensibilità pari a 0,01 cm
L = 78, 249 cm ± 0,001 cm	con uno strumento di sensibilità pari a 0,001 cm.

2.2 Gli errori casuali (misure ripetute)

Misuriamo, quindi, la lunghezza L del banco scolastico utilizzando la riga da disegno (sensibilità di 0,1 cm). Ripetendo la procedura una decina di volte, ci aspettiamo di ottenere sempre lo stesso risultato che, per quanto appena detto, va espresso in questa forma:

L = 78,2 cm ± 0,1 cm

Tutto ciò sembra seguire le regole dettate dal buon senso. Ma se nel processo di misura potessimo usare uno strumento di *sensibilità molto più elevata della riga*, in grado di valutare grandezze inferiori al millesimo di centimetro, accadrebbe qualcosa di inatteso. I risultati, con nostra grande sorpresa, inizierebbero a differenziarsi leggermente gli uni dagli altri e, ad esempio, potrebbero essere così distribuiti:

78, 243 cm   78, 242 cm   78, 249 cm   78, 245 cm   78, 247 cm   78, 241 cm   ...

Come spiegare un fatto del genere ?

Notiamo subito come la differenze tra un valore e l'altro siano molto piccole, ed è per questo motivo che uno strumento dalla sensibilità limitata come la riga da disegno non è stata in grado di evidenziarle.

Tuttavia quello che è successo risulta decisamente contrario al nostro senso comune e ci verrebbe da dire che qualcosa NON sta funzionando come dovrebbe.

Ma non é così: l'aumento di sensibilità del nuovo strumento ha fatto emergere di colpo i cosiddetti "errori casuali" o "statistici". Il fenomeno é ben osservabile utilizzando in laboratorio una "rotaia a cuscino d'aria" e collegandola a un cronometro elettronico in grado di valutare tempi inferiori al millesimo di secondo. Se ora studiamo il tempo impiegato da un carrellino a percorrere sulla rotaia uno spazio fissato, e ripetiamo la misura più volte confrontando i risultati così ottenuti, potremo accorgerci facilmente che essi sono tutti leggermente diversi tra loro (anche se, vale la pena ripeterlo, le differenze sono piccolissime ...).

Diamo ora la seguente definizione: si dicono **casuali** *quegli errori la cui origine é attribuibile a una somma di fattori non sempre ben identificabili, e che alterano la misura sia in eccesso che in difetto in modo perfettamente "casuale".*

I motivi di queste variazioni sono molteplici. Evidenziamo i più importanti:

1. sensibilità non infinita, ma limitata, degli strumenti utilizzati che impedisce di eseguire misure al di sotto di un certo valore, come spiegato in precedenza

2. imperfezione intrinseca dello strumento, che non è in grado di ripetere due volte la stessa misura in modo assolutamente identico

3. modo di operare del tecnico di laboratorio, che in una ripetizione di misure identiche non può agire perfettamente sempre allo stesso modo e che, quindi, altera con la sua sola presenza (anche se di pochissimo. . .) il valore di una misura

4. variazioni imponderabili delle condizioni ambientali: cambiamenti di umidità, di temperatura e di pressione atomosferica, impercettibili spostamenti d'aria, presenza di polvere, attriti meccanici dentro e fuori gli strumenti, presenza di campi elettrici e magnetici variabili, vibrazioni...

Gli errori che intervengono hanno le seguenti caratteristiche:

- sono perfettamente casuali, cioè si distribuiscono in ugual modo sia in eccesso che in difetto rispetto al valore "vero" della misura
- non possono essere eliminati, perché connaturali ad ogni esperimento di laboratorio
- possono e devono, però, essere correttamente analizzati con uno studio di tipo statistico in cui si prenda in considerazione un numero molto elevato di valori ottenuti con la stessa strumentazione e, per quanto possibile, nelle stesse condizioni ambientali. Vedremo più avanti come fare.

2.3 Il valore "vero" di una misura non è conoscibile

Da quanto emerso possiamo così concludere che ogni processo di misura di una grandezza fisica, anche se condotto nel modo più attento possibile e con lo strumento più raffinato a disposizione, non consente di conoscere con assoluta precisione il valore cercato.

Il **"valore vero"** *di una grandezza fisica è quindi, di fatto, non conoscibile e ogni misura é inevitabilmente accompagnata anche da un "errore sperimentale".*

La parola **errore sperimentale**, tuttavia, assume in questo caso un significato completamente diverso da quello di "sbaglio" o di "procedimento non corretto", volendo qui intendere *l'inevitabile limite alla precisione con cui ogni risultato di laboratorio può essere ottenuto.*

Tutto ciò evidenzia una caratteristica degli strumenti detta **precisione**, definita come *la capacità di ripetere la stessa misura fornendo risultati che differiscono poco (elevata precisione) o tanto (bassa precisione) l'uno dall'altro.*

Per quanto detto in precedenza a proposito della sensibilità e della precisione degli strumenti, entrambe limitate in misura maggiore o minore, dovrebbe risultare ancora più chiaro allo studente perché lo sperimentatore non sarà mai in grado di ottenere il **"valore vero"** di una grandezza fisica. Per ovviare a questo limite, si é soliti considerare "*valore vero*" la misura ottenuta con lo strumento migliore a disposizione e con la migliore attrezzatura sperimentale possibile, oppure il valore determinato attraverso un raffinato procedimento teorico.

Se ora rappresentiamo tale valore ideale come il centro di un bersaglio e le singole misure (ripetute più volte nelle stesse condizioni) come i colpi andati a segno, possiamo definire **accuratezza** di uno strumento *la sua proprietà di ripetere le misure in modo tale che la loro media sia il più vicino possibile al cosiddetto "valore vero" della grandezza studiata* (cioè il centro del bersaglio in Fig.2.1). Si potranno così verificare le seguenti situazioni:

1. misure accurate, ma poco precise (Fig.2.1 - caso A)
2. misure accurate e molto precise (caso B)
3. misure poco accurate e poco precise (caso C)
4. misure precise, ma poco accurate (caso D).

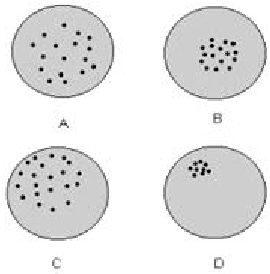

Fig.2.1 - *Precisione ed accuratezza di una serie di misure "ripetute".*

2.4 Gli errori sistematici (misure singole o ripetute)

Di tutt'altro genere sono i cosiddetti "errori sistematici". Si definiscono **sistematici** *gli errori dovuti a una errata taratura di uno strumento oppure al suo uso improprio.*

Vediamo qualche esempio:

- Un caso tipico di uso improprio é l'imprecisione che accompagna il cronometraggio manuale di intervalli di tempo estremamente brevi. L'inevitabile ritardo dovuto ai riflessi dello sperimentatore introduce un errore, che nel migliore dei casi, é compreso tra 2 e 5 decimi di secondo. Per questo motivo nelle gare sportive si fa largo uso di cronometraggi elettronici.
- Un altro caso di uso improprio si ha quando lo strumento é difettoso o tarato male. Una bilancia non perfettamente azzerata altera sempre nello stesso senso (cioè sempre in eccesso o sempre in difetto) ogni misura effettuata. Prima di usarla, quindi, bisogna procedere ad una corretta taratura. Uno strumento difettoso, invece, deve essere assolutamente cambiato.
- L' errore di parallasse, *dovuto a un cattivo allineamento tra l'occhio dell'osservatore e la scala graduata dello strumento utilizzato*, é una delle più frequenti cause di errori sistematici. Prendiamo, ad esempio, un termometro che sia posto sul muro a un'altezza di circa un metro sopra della testa dell'osservatore. Poiché la colonnina di mercurio scorre in leggero rilievo rispetto alla scala graduata che ne dà posizione, se la lettura non é effettuata lungo una direzione ad essa perpendicolare ma é ottenuta, ad esempio, guardando dal basso verso l'alto, la rilevazione della temperatura risulta inevitabilmente sovrastimata perché il mercurio sembrerà arrivare in un punto diverso da quello effettivamente raggiunto. Lo stesso avviene se consideriamo una bilancia, dove la misura è ricavata dalla posizione di un ago mobile (lancetta) sullo sfondo di una scala graduata: se l'osservatore non si pone in asse, la lettura non sarà mai corretta (**Fig.2.2**).

In conclusione possiamo affermare che gli errori sistematici, a differenza dagli errori casuali e di sensibilità, sono errori veri, nel senso che introducono nel processo di misura uno "sbaglio". Il vantaggio, però, è che possono (e devono) essere eliminati migliorando lo strumento utilizzato, correggendone i difetti e le modalità d'impiego, controllandone con attenzione il funzionamento e la taratura.

Per questo motivo, d'ora in avanti, supporremo che nelle nostre misure sia assente ogni tipo di errore strumentale.

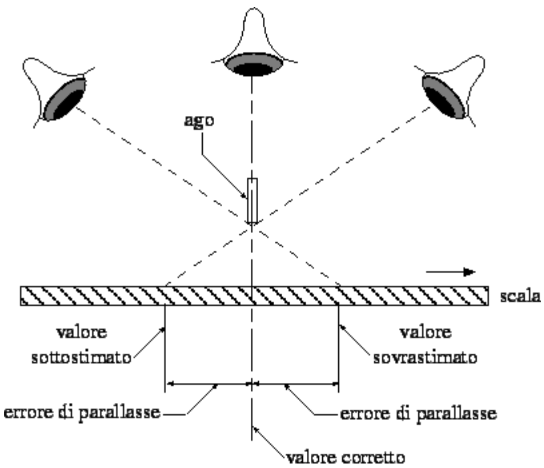

Fig.2.2 - *L'errore di parallasse.*

2.5 Come scrivere un risultato sperimentale

Per tutto quanto detto in precedenza, dobbiamo sempre comunicare il risultato di una misura sperimentale con l'errore a esso associato. La scrittura assumerà la seguente forma:

$$risultato = misura \pm errore \tag{2.1}$$

Vediamo come determinare l'errore nelle varie situazioni che si possono presentare e come comportarci per approssimare in modo corretto il risultato finale.

2.6 Errori assoluti e percentuali

Quando l'errore è accompagnato dall'unità di misura, come nel caso della sensibilità dello strumento, l'errore si chiama **errore assoluto** ϵ_{ass}.

Questa scrittura ha il pregio di essere particolarmente semplice e di facile comprensione. Per contro, ha il difetto di rendere difficilmente confrontabili misure eseguite su oggetti diversi. Dobbiamo, infatti, considerare migliore la misura della statura di un uomo alto 1,80 m con un errore assoluto di un millimetro o la rilevazione dell'altezza di una montagna stimata in 8 800 m, con l'errore assoluto di un metro?

Per rispondere a tale domanda introduciamo il concetto di **errore percentuale**, *definito come il rapporto tra l'errore assoluto e la misura, moltiplicato per cento.*

In formule:

$$\varepsilon_{\%} = \frac{\epsilon_{ass}}{misura} \cdot 100 \tag{2.2}$$

Abbiamo così un errore dello 0, 056% sulla statura dell'uomo e dello 0, 01% sull'altezza della montagna. Infatti, eseguendo i conti:

$$\varepsilon_{\%}(uomo) = \frac{0,056\,m}{1,80\,m} \cdot 100 = 0,056\%$$

$$\varepsilon_{\%}(montagna) = \frac{1\,m}{8\,800\,m} \cdot 100 = 0,011\%.$$

Ora i due errori possono essere facilmente confrontati e appare evidente come la seconda misura sia migliore della prima (di circa cinque volte).

Nel caso in cui si conosca l'errore percentuale e si voglia calcolare l'errore assoluto è necessario usare la seguente formula inversa:

$$\varepsilon_{ass} = \frac{\epsilon\%}{100} \cdot misura \tag{2.3}$$

2.7 Gli errori assoluti (in misure singole e ripetute)

Vediamo ora come calcolare l'errore assoluto in funzione del tipo di misura che eseguiamo.

Misura singola: l'errore di sensibilità

Se di una grandezza fisica abbiamo a disposizione una sola misurazione non abbiamo molte possibilità di scelta: l'errore assoluto da associare a tale valore deve necessariamente essere *l'errore dato dalla sensibilità dello strumento.*

Nel caso in cui sia possibile ripetere più volte l'esperimento abbiamo visto che può accadere di ottenere misure tutte leggermente diverse tra loro. Si può allora dimostrare che la "miglior stima" della grandezza studiata è semplicemente fornita dal *valore medio delle misure ottenute*, e che tale valore è tanto più attendibile quanto più è elevato il numero di misure effettuate. Vediamo ora come deve essere calcolato in questo caso l'errore assoluto che accompagna il valore medio.

Misure ripetute: la semidifferenza (o semidispersione)

Supponiamo di ripetere una misura più volte mettendoci nelle stesse condizioni ambientali ed usando lo stesso strumento dall'elevata sensibilità. Definiamo **semidifferenza** (o semidispersione) **SD** di una serie di misure ripetute il valore:

$$SD = \frac{Misura\,maggiore - misura\,minore}{2} \tag{2.4}$$

I valori che si ottengono siano, ad esempio, i seguenti:

0,111 m   0,112 m   0,108 m   0,108 m   0,116 m   0,109 m   0,113 m

Abbiamo quindi:

SD = (0,116 m – 0,108 m) : 2 = 0,004 m

Se la misura che abbiamo eseguito è stata ripetuta un numero non elevato di volte, *la semidispersione rappresenta l'errore assoluto, di origine casuale, da associare al valore medio.*

Nel nostro caso, essendo il valore medio $\mathrm{x}_{medio} = 0,111\,m$, dobbiamo scrivere:

$x_{medio} = 0,111\,m \pm 0,004\,m.$

Misure ripetute molte volte: la deviazione standard σ[1]

Se i dati a nostra disposizione sono molti e tutti ottenuti dalla ripetizione della stessa misura nelle identiche condizioni sperimentali, il concetto di semidispersione è statisticamente troppo debole per assolvere il compito di rappresentare correttamente l'errore assoluto da associare al valor medio x_{medio} di una serie di valori x_1, x_2, x_3, . . . x_3.

A tal fine bisogna introdurre la **deviazione standard** σ , *definita come la radice quadrata della media dei quadrati degli scarti tra le singole misure e il valore medio.* [2]

La definizione è complessa solo apparentemente. Analizziamola passo passo per comprenderla al meglio e, soprattutto, per imparare a usarla.[3]

Siano date le seguenti cinque misure:

$x_1 = 22 \qquad x_2 = 24 \qquad x_3 = 18 \qquad x_4 = 19 \qquad x_5 = 27$

Con la scrittura x_i che compare nella definizione di deviazione standard si è voluto identificare la generica misura *i-esima.* Cioè, quando $i = 3$ vuol dire che sto considerando la terza misura, $x_3 = 18$.

Calcoliamo ora il valore medio: $x_{medio} = 22$.

Creiamo una tabella in questo modo:

- nella prima colonna mettiamo la serie delle misure
- nella seconda colonna gli *scarti rispetto al valore medio*, cioè la differenza $(x_i - x_{medio})$ che compare nella formula
- nella terza colonna il *quadrato dei singoli scarti*, quindi il valore $(x_i - x_{medio})^2$
- nella quarta colonna scriviamo *la media di tali scarti*, $\sum \frac{(x_i - x_{medio})^2}{N}$
- nella quinta colonna, la radice quadrata di questo valore, che è la *deviazione standard* σ cercata.

La procedura di calcolo può essere automatizza e velocizzata utilizzando un foglio elettronico, come quello fornito da "Microsoft Excel" o da "Open Office". In questo caso si devono ricopiare le cinque misure nelle celle A1, A2. . . A5, e poi si deve scrivere, in un'altra cella a scelta, l'espressione:[4]

= DEV.ST.POP(A1: A5)

Confermata la procedura con il tasto RETURN, nella stessa cella comparirà il valore cercato.

Quando le misure da analizzare sono molte, ci sono delle profonde giustificazioni teoriche per preferire la *deviazione standard* alla *semidispersione*. Tali motivazioni saranno trattate in un prossimo capitolo.

[1] Il simbolo σ rappresenta la lettera "s minuscola" dell'alfabeto greco, e si legge *sigma*.

[2] In formule:

$$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(x_i - x_{medio})^2}{N}} \tag{2.5}$$

[3] Esistono argomenti teorici, che a questo livello del corso trascuriamo, in base ai quali al denominatore della frazione sarebbe meglio considerare (N-1) al posto di N.

[4] Si veda, al proposito, il capitolo sull'analisi dati svolta con *Microsoft Excel* o con *Open Office*.

misure	*valore medio*	*scarti*	*quadrato degli scarti*	*media del quadrato degli scarti*	*radice quadrata della media del quadrato degli scarti*
22		0	0		
24		2	2		
18	22	-4	16	10,8	3,286
19		-3	9		
27		5	25		

Fig. 2.3 - *Calcolo della deviazione standard.*

2.8 Errore assoluto e cifre significative

Consideriamo come al solito una serie di misure ripetute nelle stesse condizioni sperimentali, il loro valore medio e il corrispondente errore assoluto. Sia, per esempio:

$misura$ = 5,728273 cm ± 0,0479219 cm

Come possiamo esprimere il risultato nel modo corretto, tenendo conto solo delle cifre significative e non di quelle fornite dalla calcolatrice nei vari calcoli ??

Notiamo come la prima cifra diversa da zero nell'errore assoluto sia il 4: questo vuol dire che l'errore sulla misura inizia con la seconda cifra decimale, cioè il centesimo. Come diretta conseguenza di ciò, anche le cifre della misura inizieranno a diventare incerte a partire dal secondo decimale. E' allora ragionevole eseguire l'approssimazione in questo punto: le cifre che seguono la seconda posizione decimale non hanno più significato, perché non sono state realmente misurate.

Il risultato può essere così espresso:

$misur$a = 5,73 cm ± 0,05 cm

dove, in questo caso, si è resa necessaria una approssimazione per eccesso sia sull'errore che sul valore della misura. Se invece, con un altro esempio, avessimo una scrittura del tipo:

$misura$ = 0,6954 cm ± 0,00425 cm

dovremmo approssimare in questo modo

$misura$ = 0,695 cm ± 0,004 cm.

Concludendo, possiamo affermare che a meno di eccezioni vale la seguente **regola**:

la scrittura delle cifre significative di una misura finisce dove inizia l'errore, cioè nella posizione decimale della prima cifra diversa da zero dell'errore assoluto.

2.9 La legge di propagazione degli errori (misure indirette)

Consideriamo ora misure indirette, ottenute cioè tramite un calcolo algebrico, e proponiamoci di calcolare l'errore assoluto sul risultato.

Calcoliamo il perimetro e l'area di un rettangolo i cui lati b e h siano, rispettivamente:

$b = 12{,}6$ cm $\pm$ 0,4 cm $\qquad h = 25{,}84$ cm $\pm$ 0,05 cm

Un rapido conto porta ai risultati:

$perimetro = 76,\ 88$ cm $\pm\ \epsilon_{ass}$

$Area = 325,\ 584$ cm^2 $\pm\ \epsilon_{ass}$

Come calcolare il valore assoluto associato all'area e al perimetro e con quante cifre significative si devono esprimere i due risultati?? La risposta si basa sulla **legge di propagazione degli errori** che si riassume in tre semplici regole (qui non dimostrate):

1. in una somma o una differenza di misure, l'errore totale associato al risultato è dato dalla somma dei singoli errori assoluti

2. in un prodotto o in un rapporto di misure, l'errore totale associato al risultato è la somma dei singoli errori percentuali

3. in una potenza o in un'estrazione di radice di una misura, l'errore totale associato al risultato è dato dall'errore percentuale sulla base, moltiplicato per l'esponente.

Applichiamo quanto detto all'esempio in questione.

Poiché il perimetro è la somma dei quattro lati del triangolo, l'errore assoluto sul totale sarà dato dalla somma degli errori assoluti sui quattro lati:

$perimetro = 76,\ 88$ cm $\pm$ (0,4 + 0,4 + 0,05 + 0,05) cm

$perimetro = 76{,}88$ cm $\pm$ 0,9 cm

$perimetro = 76{,}9$ cm $\pm$ 0,9 cm $\qquad$ (*scritto con la corretta approssimazione*).

Per quanto riguarda l'area, il procedimento è un po' più complicato perché al risultato si perviene attraverso una moltiplicazione. Dovremo quindi calcolare gli errori percentuali su ogni misura di partenza e poi sommarli per ottenere l'errore percentuale totale. Quindi:

$$\epsilon_{\%\,base} = \frac{0,4\,cm}{12,6\,cm} \cdot 100 = 3,17\,\%$$

$$\epsilon_{\%\,altezza} = \frac{0,05\,cm}{25,84\,cm} \cdot 100 = 0,19\,\%$$

$$\epsilon_{\%\,totale} = 3,17\,\% + 0,19\,\% = 3,36\,\%$$

$$Area\ = 325,584\,cm^2 \pm 3,36\,\%$$

A questo punto trasformiamo l'errore percentuale in errore assoluto. Solo così potremo eseguire correttamente le approssimazioni sul risultato. Otteniamo:

$$\epsilon_{ass\,Area} = \frac{\epsilon\%}{100} \cdot misura = 10,6140384\ cm^2$$

$Area\ = 326\ cm^2 \pm 11\ cm^2$ $\qquad$ (*scritto con la corretta approssimazione*).

2.10 Misure uguali: l'intervallo di confidenza

Se ogni misura di laboratorio è, per sua natura, sempre accompagnata da un errore sperimentale, come possiamo regolarci per sapere se due grandezze sono uguali?

Facciamo l'esempio di oggetti apparentemente identici che siano stati pesati con lo stesso strumento e nelle stesse condizioni ambientali. Le misure delle loro masse sono le seguenti:

A = 250 $g \pm 6$ g

B = 247 $g \pm 4$ g

Come possiamo capire se le masse sono identiche?

Per dare una risposta, dobbiamo considerare il cosiddetto **intervallo di confidenza**, definito come *l'intervallo centrato sul valore medio e di raggio pari a 2σ (deviazione standard) in cui "si confida", in termini statistici, che sia compreso il "valore vero" della grandezza fisica studiata.*

Nel caso di A tale intervallo, calcolato a *2σ*, parte da 238 g ed arriva fino a 262 g. Nel caso di B, invece, l'intervallo inizia a 239 g e finisce a 255 g.

Introduciamo ora la seguente regola, rimandando la spiegazione al prossimo paragrafo: *due misure sperimentali possono essere ritenute statisticamente uguali se i loro intervalli di confidenza a 2σ si sovrappongono almeno in parte.*

Nel nostro caso la sovrapposizione esiste: le due masse A e B possono quindi essere considerate uguali (in senso statistico).

Vale anche la seguente regola: *due misure sperimentali possono essere ritenute statisticamente diverse se i loro intervalli di confidenza a 3σ non si sovrappongono nemmeno in parte.*

Per comprendere il significato degli intervalli di confidenza occorre affrontare la spiegazione della *curva di Gauss.*

2.11 La curva di Gauss

Prendiamo dieci monete e per ognuna di esse individuiamo la faccia che presenta una "testa" e quella che riporta una "croce" (oppure due simboli analoghi). Mettiamo le monete in un bicchiere, agitiamole e riversiamole su un tavolo. Prendiamo nota di quante sono le monete che presentano il simbolo che abbiamo identificato con "testa". Ripetiamo la procedura un centinaio di volte, sempre prendendo nota del numero di monete che dopo il lancio presentano il simbolo "testa" e raccogliamo i risultati in una tabella come la seguente:

numero di "testa"	0	1	2	3	4	5	6	7	8	9	10
frequenza dell'evento "testa"	0	2	2	9	17	27	21	12	5	4	1

Se ora, a partire da questi dati, costruiamo *l'istogramma delle frequenze* (cioè il diagramma a barre che esprime con che "frequenza" una misura si è ripetuta), otteniamo il grafico riportato in Fig.2.4.

Con questa procedura abbiamo simulato il processo di misura di una variabile che può assumere valori da 0 a 10 (il numero di teste) con valore medio pari a 5 (5 è il valore più probabile, quello per il quale le probabilità di ottenere testa o croce sono equivalenti).

La situazione descritta è perfettamente simile al caso in cui si considerino i valori assunti da una grandezza fisica misurata sperimentalmente. L'unica differenza è che nella realtà una generica variabile fisica potrebbe assumere un qualsiasi valore compreso tra 0 e 10 e non solo valori interi: quindi anche risultati pari a 1,77 oppure 8,7654 sono possibili. In tale situazione, se ripetiamo la misura un numero infinitamente grande di volte, le barre dell'istogramma tendono ad assottigliarsi perché gli intervalli in cui raccogliere le misure diminuiscono di ampiezza: infatti, posso raccogliere le misure negli insiemi (0,01-0,02); (0,02-0,03); (0,03-0,04)...

Il grafico viene ad assumere l'aspetto di una curva continua dalla caratteristica forma a campana: **la curva di Gauss**. L'approssimazione è tanto migliore quanto più sono i punti utilizzati (vedi Fig.2.5).

La curva così ottenuta, detta anche **gaussiana** oppure **curva degli errori**, ha le seguenti importanti proprietà:

1. ogni punto della curva individua una singola misura (il valore riportato sull'asse delle ascisse) e la frequenza con cui si è ripetuta (asse delle ordinate)
2. il vertice della gaussiana rappresenta la misura che si è ottenuta più frequentemente: si può dimostrare matematicamente che tale valore tende a coincidere con il *valore medio* della serie delle misure
3. la gaussiana rappresenta la dispersione delle misure ad opera degli *errori casuali*
4. la larghezza della gaussiana è un indice di precisione delle misure effettuate. Se abbiamo due serie di misure della stessa grandezza fisica, *la serie a cui corrisponde una gaussiana più stretta è quella più precis*a
5. *la larghezza della curva di Gauss a metà della sua altezza* vale circa $\pm\sigma$
6. *la larghezza della curva di Gauss alla base,* vale circa $\pm 3\sigma$

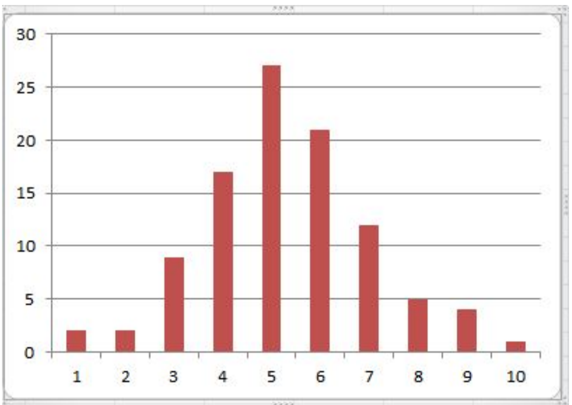

Fig.2.4 - *Istogramma delle frequenze dopo 100 lanci di una moneta.*

Possiamo capire, adesso, perché la deviazione standard sia il miglior indicatore degli errori casuali. La spiegazione risiede nella curva di Gauss: la *deviazione standard* ne rappresenta la larghezza circa a metà altezza, e quindi il miglior indice della dispersione delle misure ad opera degli errori casuali.[5]

Inoltre, un intervallo di confidenza ampio $\pm 2\sigma$ abbraccia di fatto quasi tutta la curva e si può dimostrare che ad esso corrisponde una probabilità del 95,5% di comprendere il valore "vero" della grandezza fisica studiata.

[5] La deviazione standard è la semilarghezza L della curva di Gauss a "circa" metà della sua altezza massima. Molti testi in lingua inglese fanno riferimento a questo parametro con la sigla FWHM (full width half maximum). La relazione corretta tra semilarghezza e deviazione standard è la seguente: L =1,17 σ.

Se invece dovesse capitare che anche costruendo intervalli di confidenza di ampiezza pari a "quasi" tutta la gaussiana, cioè di larghezza $\pm 3\sigma$, questi NON dovessero sovrapporsi, allora vorrebbe dire che le due misure analizzate sono da considerarsi ragionevolmente diverse: le rispettive gaussiane, infatti, non si toccano.

Si può realizzare anche un caso ambiguo: gli intervalli di confidenza di larghezza pari a $\pm 2\sigma$ NON si sovrappongono, ma quelli calcolati a $\pm 3\sigma$ invece sì: in una situazione critica come questa, purtroppo, non possiamo essere sicuri né dell'uguaglianza né della diversità delle misure. L'esperimento dovrebbe essere ripetuto cercando di ottenere risultati più precisi, migliorando, ad esempio, le procedure seguite o la qualità degli strumenti usati.

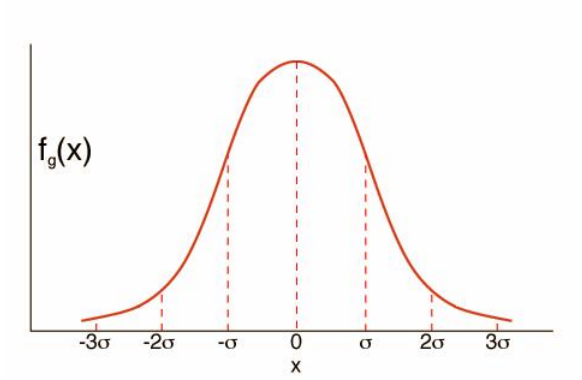

Fig.2.5 - *Relazione tra la larghezza della curva di Gauss e la deviazione standard* σ.

2.12 ... in conclusione

Come detto nel paragrafo 2.5, il modo corretto per esprimere una misura sperimentale è il seguente:

$$risultato = misura \pm errore.$$

Abbiamo poi visto nei paragrafi successivi che, per determinare l'errore assoluto, si possono presentare vari casi:

Misura singola	$\rightarrow$	l'errore assoluto è la sensibilità dello strumento
Misure ripetute (poche)	$\rightarrow$	l'errore assoluto è la semidispersione
Misure ripetute (molte)	$\rightarrow$	l'errore assoluto è la deviazione standard σ
Misure indirette - laboratorio	$\rightarrow$	usare la propagazione degli errori
Misure indirette - esercizi	$\rightarrow$	usare il metodo rapido delle cifre significative

Ricordiamo ancora una volta come la conoscenza dell'errore assoluto sia fondamentale per la scrittura del risultato finale *con la corretta approssimazione,* mentre, per tale fine, la conoscenza dell'errore percentuale non sia affatto sufficiente.

2.13 Tipi diversi di media

Uno studente ha preso i seguenti voti nella stessa disciplina:

7	6	6	9	8	4	5	5	7	7	6	7	3

Se si vuole calcolare la media, il procedimento è noto: bassa sommare tutti i valori tra loro e dividere per il numero delle misure: 13, nel nostro caso. Si ha, quindi:

$$media\, aritmetica \;=\; \frac{x_1 + x_2 + x_3 + \ldots\ldots\ldots\ldots x_N}{N} \tag{2.6}$$

$$media\, aritmetica = \frac{7+6+6+9+8+4+5+5+7+7+6+7+3}{13} = \frac{80}{13} = 6,15$$

Procedendo in questo modo, lo studente ha ottenuto la cosiddetta **media aritmetica** dei suoi voti.

Ci sono altri modi di eseguire una media. Supponiamo che dei 13 voti considerati i primi 5 siano relativi a prove scritte (che il professore considera più importanti) e gli ultimi 8 a prove orali. Il professore di questa materia attribuisce un'importanza 1,5 volte superiore alle prove scritte: in statistica si dice che questi numeri hanno un "**peso**" maggiore, e nel nostro caso il loro "peso" è 1,5 mentre il "peso" delle prove orali è 1. Nel calcolare la media dello studente bisogna allora tenere conto che i voti da analizzare NON hanno tutti la stessa importanza. Per far ciò si deve eseguire quella che si chiama **media pesata**, così costruita:

$$media\, pesata \;=\; \frac{x_1{\cdot}peso_1 + x_2 \cdot peso_2 + x_3{\cdot}peso_3 + \ldots\ldots\ldots\ldots x_N{\cdot}peso_N}{somma\, dei\, pesi} \tag{2.7}$$

Sostituendo i valori, otteniamo:

$$\frac{7{\cdot}1,5+6{\cdot}1,5+6{\cdot}1,5+9{\cdot}1,5+8{\cdot}1,5+4{\cdot}1+5{\cdot}1+5{\cdot}1+7{\cdot}1+7{\cdot}1+6{\cdot}1+7{\cdot}1+3{\cdot}1}{15,5} = 6,32$$

Si può notare facilmente come i risultati ottenuti NON siano identici, ancorché simili.

Un tipo diverso di media è la cosiddetta **media geometrica**, definita come *la radice ennesima del prodotto dei valori considerati.* Calcoliamo, ad esempio, la media geometrica dei numeri 7 e 8:

$$media\, geometrica \;=\; \sqrt{7 \cdot 8} = \sqrt{56} = 7,48 \tag{2.8}$$

valore che risulta essere ancora molto vicino alla media aritmetica (7,5) vista in precedenza.

Un altro importante indicatore statistico è la **moda**, definito come *il valore che, in una serie di misure, compare più frequentemente.* Per eseguire questo nuovo calcolo dobbiamo riprendere in esame i voti dello studente e disporli, per comodità, in ordine crescente:

3	4	5	5	6	6	6	7	7	7	7	8	9

La *moda* di questo insieme numerico è il numero 7, perché è quello, tra tutti, che compare con maggiore frequenza.

L'ultimo indicatore statistico che trattiamo è la **mediana**, *definita come il valore centrale di un insieme ordinato di elementi.* Riprendiamo in esame la succesione di voti, anche in questo caso ordinata dal minore al maggiore. Il valore "centrale" è rappresentato dal terzo

6 della serie, perché questo elemento si trova in posizione tale da avere sei numeri sia alla sua destra che alla sua sinistra. Esso è la *mediana* cercata.

Nel caso in cui i valori considerati siano in numero pari (14 invece che 13, ad esempio), esistono "due" elementi centrali, non uno solo: la *mediana* sarà correttamente calcolata eseguendo la media algebrica di questi due valori.