Errori : di solito vengono effettuate molte misure *indipendenti* della stessa grandezza, fino a collezionare un numeroso campione di misure

ma a causa degli errori di misura il risultato varia sensibilmente da misura a misura

si definisce **errore = | valore misurato – valore vero |**

cause di errore: → Limiti strumentali → Metodi di misura errati → Cause accidentali

categorizzazione degli errori:

→ Sistematici misura di un intervallo di tempo usando un orologio che va troppo lento, o troppo veloce.

misura della lunghezza di un oggetto non in modo perpendicolare all'oggetto (errore di parallasse)

→ Casuali o "statistici" incertezze dovute a cause accidentali e alla limitatezza del campione di misure

gli errori statistici sono riducibili aumentando il numero di misure indipendenti della stessa grandezza → aumentando la dimensione del campione

ma , ammesso che esista, quale e' il **" valore vero "** se alla base del fenomeno in esame vi fosse una variabile aleatoria sarebbe meglio parlare di errore

come di **errore = | valore misurato – valore medio |**

ma quanto vale il valor medio ? per saperlo con certezza si dovrebbero fare una infinita' di misure ripetute

se si hanno a disposizione solo un numero finito di misure si puo' solo tentare di " stimarlo " con il minimo margine di errore possibile si assume come *stima* del valor medio (vero) μ della grandezza in esame la media aritmetica dei risultati ottenuti nelle varie misure

es. : si siano effettuate n misurazioni della stessa grandezza fisica $x_1, x_2 \ldots x_n$ la media aritmetica delle n misure e' :

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

la media aritmetica **stima** il valor "vero" μ , ma con un certo errore ε $\mu = (\overline{x} \pm \varepsilon)$ *unita' di misura*

problema : come stimare l'errore statistico ε ?

come indicatore di dispersione di una distribuzione intorno alla sua media si usa la deviazione standard campionaria o " errore quadratico medio",

in inglese "Root Mean Square" o *r.m.s.* che e' definito come :

$$dev.\ st. = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

commento sull'uso di n o di $n-1$

una tra le proprieta' piu' importanti della media aritmetica e' che l'errore statistico della media aritmetica stessa e' dato da:

$$\varepsilon(\overline{x}) = \frac{dev.\ st.}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Intervallo di confidenza

si siano effettuate n misurazioni indipendenti della stessa grandezza fisica x, ottenendo : $x_1, x_2 \ldots x_n$

la miglior stima campionaria del valor medio incognito e' la media aritmetica $\overline{x}$ delle n misure dove $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

ma ripetendo una seconda volta n misure della stessa grandezza fisica, si otterrebbe un diverso insieme di risultati : $x'_1, x'_2 \ldots x'_n$

la miglior stima campionaria del valor medio incognito continuerebbe ad essere la media aritmetica $\overline{x}'$ dove $\overline{x}' = \frac{1}{n}\sum_{i=1}^{n} x'_i$

ma essendo gli x_i' diversi dagli x_i la nuova media aritmetica sarebbe diversa $\overline{x}' \neq \overline{x}$ dunque anche la media aritmetica varia, *imprevedibilmente*,

da campione a campione ossia e' essa stessa una *variabile aleatoria*

- come stima della dispersione della media si assume la deviazione standard campionaria e
- come errore sulla media, si assume la deviazione standard campionaria divisa per la radice di n

rilevanza della distribuzione gaussiana

- se la distribuzione di una serie di n variabili aleatorie x_i con $i = 1, n$ segue la forma funzionale gaussiana si puo' dimostrare che la media aritmetica delle n v. a. sara' distribuita in modo gaussiano
- anche se la distribuzione delle n v. a. x_i non fosse gaussiana se $n >> 1$ grazie al teorema del limite centrale, si puo' assumere che la distibuzione della media campionaria sia gaussiana

infine

- istogrammando i risultati di molte misure ripetute ed **indipendenti** tra loro risulta, molto spesso ma non sempre, che le misure si distribuiscono in modo gaussiano dunque nella maggior parte dei casi, ma non sempre,

- se si stima il valor medio (vero) come $\mu = \overline{x} \pm \frac{dev.\ st.}{\sqrt{n}}$ si ha il 68% di probabilita' di fare una stima esatta

- se si stima il valor medio (vero) come $\mu = \overline{x} \pm\ 2\frac{dev.\ st.(\overline{x})}{\sqrt{n}}$ si ha il 95% di probabilita' di fare una stima esatta

- se si stima il valor medio (vero) come $\mu = \overline{x} \pm\ 3\frac{dev.\ st.(\overline{x})}{\sqrt{n}}$ si ha il 99.7% di probabilita' di fare una stima esatta

il valore della percentuale di probabilita' che si desidera ottenere, ossia l' attendibilita' della stima del valor "vero" che si desidera raggiungere, e' detto " livello di confidenza "

supponiamo siano state molte fatte misure di precisione ed indipendenti tra loro, di una stessa grandezza fisica, e' ragionevole attendersi che i risultati non si riproducano perfettamente postulando che il fenomeno stesso in esame abbia un carattere aleatorio ossia che l'esito della misura sia descrivibile in termini di una variabile aleatoria si potranno applicare i metodi statistici per il trattamento dei dati sperimentali vale a dire per stimare i parametri della variabile aleatoria incognita

in conclusione:

una misura sperimentale e' assimilabile al verificarsi di uno tra i tanti possibili risultati che una v.a. puo' assumere la statistica predittiva, utilizzando i risultati rigorosi della teoria della probabilita', e' in grado di suggerire:

- quale sia il miglior stimatore possibile del valor medio , o valor “vero”, → di solito, ma non sempre, la media aritmetica,

- quale sia il margine di errore con cui si puo' fare la stima in funzione della numerosita' del campione, ossia di determinare quale sia l'errore sulla media aritmetica

- di valutare quale sia l'attendibilita' di questa misura in termini di probabilita' , ossia quale sia il livello di confidenza della stima

e' chiaro che si tratta di <u>previsioni</u> e che per la aleatorieta' stessa del fenomeno in esame e' <u>impossibile</u> sapere con certezza quale risultato si presentera' all' atto di ogni singola prova (misura)

Ricapitolando: postulando che il processo di misurazione sia assimilabile ad effettuare un esperimento aleatorio possiamo ''recuperare'' la non riproducibilita' degli esperimenti

Esempio

sono state fatte 25 misure ripetute ed indipendenti tra loro di una grandezza fisica, ad es. il peso di un oggetto misurato con una bilancia precisa al per mille

1.72, 1.65, 1.81, 1.72, 1.72, 1.67, 1.71, 1.72, 1.74, 1.70, 1.73, 1.70, 1.76, 1.72, 1.75, 1.71, 1.71, 1.72, 1.69, 1.79, 1.74, 1.73, 1.76, 1.73, 1.71.

i risultati , in gm, sono : e' evidente che la misura non si riproduce perfettamente

non avendo altre informazioni a disposizione si dovra' stimare il valor medio, impropriamente detto valor "vero" della grandezza incognita, usando i dati del

campione di misure calcoliamo la media campionaria e l' errore sulla media

se x_i e' la i-esima misura $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ ⇨ $\overline{x} = \frac{1}{25}\sum_{i=1}^{25} x_i = 1.7244$

$$dev.\ st. = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{1}{25-1}\sum_{i=1}^{25}(x_i - 1.7244)^2} = 0.03368$$

l'errore sulla media vale $\varepsilon(\overline{x}) = \frac{dev.\ st.}{\sqrt{n}} = \frac{0.03368}{\sqrt{25}} = 0.0067$ arrotondando l'errore ad una sola cifra $\varepsilon = 0.007$

se il livello di confidenza prescelto e' il 68 % il risultato della misura e' : $\mu = (1.724 \pm 0.007)\ gm$

oppure $\mu = (1.72 \pm 0.01)\ gm$ al 95% di livello di confidenza o $\mu = (1.72 \pm 0.02)\ gm$ al 99% di livello di confidenza

da notare la relazione tra la precisione e il grado di fiducia, o livello di confidenza a parita' di numerosita' del campione, ossia a parita' di n, se una cresce l'altra cala

se si utilizzasse la convenzione delle cifre significative il risultato ottenuto con il 68% andrebbe presentato come m = 1.724 gm mentre se avessimo operato al 95 e 99 % di livello di confidenza andrebbe presentato come m = 1.72 gm per costruire un istogramma disponiamo le misure in ordine crescente

⇨

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1.65	1.67	1.69	1.70	1.70	1.71	1.71	1.71	1.71	1.72	1.72	1.72	1.72	1.72	1.72	1.73	1.73	1.73	1.74	1.74	1.75	1.76	1.76	1.79	1.81

⇨ calcoliamo quale sia la frequenza con la quale si presenta un particolare risultato

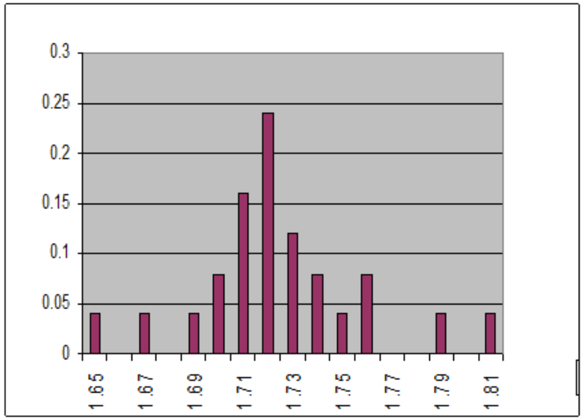
⇨ grafichiamo la frequenza relativa , ossia la frequenza diviso il numero totale di misure

⇨ la frequenza relativa e' **normalizzata** all'unita di modo che l'istogramma rappresenti una distribuzione di probabilita'

Misure (x_i)	Frequenza (F_i)	Frequenza relativa = Frequenza / N_{tot} (Fr_i)
1.65	1	0.04
1.66	0	0
1.67	1	0.04
1.68	0	0
1.69	1	0.04
1.7	2	0.08
1.71	4	0.16
1.72	6	0.24
1.73	3	0.12
1.74	2	0.08
1.75	1	0.04
1.76	2	0.08
1.77	0	0
1.78	0	0
1.79	1	0.04
1.8	0	0
1.81	1	0.04

$N_{tot} = \Sigma F_i = 25$

istogramma delle frequenze relative



Backup Slides