11. GLI ERRORI DI MISURA

Prendi un metro da muratore (di quelli pieghevoli, formati da più aste collegate da cerniere, totale 2 m) e prova a misurare, al centimetro, la lunghezza del corridoio della tua scuola.
Ripeti l'operazione più volte, segnando sempre su di un taccuino il valore ottenuto.

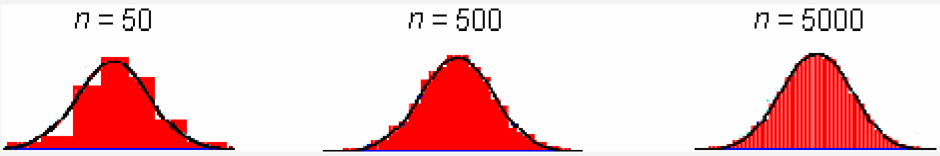
Certamente non otterrai la stessa misura ad ogni prova:
infatti, nel disporre il metro sul pavimento, ti capiterà di non iniziare esattamente dallo stesso punto, di riportare il metro non sempre con precisione quando devi spostarlo per ricollocarne un'estremità nella posizione alla quale eri giunto al passo precedente, di piegarlo leggermente, e così via.

Adesso coraggio, perché ho bisogno che tu faccia TANTE misurazioni, diciamo 100
(sono certo che i tuoi compagni di classe si presteranno a collaborare … ognuno potrebbe fare 4-5 misurazioni).

Ora hai a disposizione 100 numeri.
Può darsi che alcuni di questi numeri coincidano, ma in generale saranno invece un poco diversi fra loro.
Considera il minimo e il massimo valore rilevato, e suddividi l'intervallo $[x_{\text{min}}, x_{\text{MAX}}]$
in un certo numero di sottointervalli, diciamo otto-dieci
(in generale, se le misure sono *n*, si consiglia di far sì che il numero di intervalli non superi $\sqrt{n}$) :
ad esempio, se la minima e la massima delle misure registrate sono state di m 23,92 e di m 24,11,
avremo $24{,}11-23{,}92=0{,}19$ e questo intervallo di metri 0,19 (19 cm) potrà portarci a definire
10 sottointervalli di 2 cm ciascuno: $[23{,}92; 23{,}94), [23{,}94; 23{,}96), [23{,}96; 23{,}98), \ldots , [24{,}10; 24{,}12)$.
Ora, per ciascun sottointervallo, conta la rispettiva "frequenza",
ossia conta il numero di misure, fra le 100 registrate, che cadono in quel sottointervallo;
traccia, con un foglio elettronico, un istogramma con le *classi di misura* in *orizzontale* e le *frequenze* in *verticale*.

Potrai osservare che le misure "centrali" della distribuzione saranno in linea di massima più frequenti, e quelle estreme meno. In effetti, nell'atto pratico della misurazione, si commettono sempre **errori "casuali"** talvolta in difetto talvolta in eccesso, e **se il numero di misurazioni effettuate diventa alto, l'istogramma tenderà ad assomigliare a una curva "a campana"** detta "**gaussiana**" (F. Gauss, 1777-1855).

Ecco qui di seguito un "fumetto" di possibili configurazioni dell'istogramma delle frequenze al crescere del numero n di misure effettuate.



La **Gaussiana** è una curva la cui equazione è nientemeno che $y=\dfrac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, dove:

$\pi=3{,}14159\ldots$ (ben noto); $e=2{,}71828\ldots$ (numero di Nepéro);
$\boldsymbol{\mu}$**,** $\boldsymbol{\sigma}$ sono due numeri fissi che, nel caso in cui la curva abbia a che fare con il problema da noi esaminato, ossia quello delle misure ripetute di una quantità (affette da errori "casuali" o "statistici"), sono interpretabili come rispettivamente la **media aritmetica** e lo **scarto quadratico medio che si otterrebbero facendo un numero colossale (= *tendente all'infinito*) di misure**.

Trovi la cosa complicata? In effetti, lo è … ☹
Questi studi richiedono nozioni matematiche più avanzate (la teoria delle "*distribuzioni di probabilità*")
e non è facile, in una trattazione di carattere non specialistico,
mantenere il discorso su di un livello che sia nel contempo accessibile e rigoroso …

… ma noi ci proviamo ☺.

Se le misurazioni effettuate, affette da errore casuale, sono tante
(di solito, detto *n* il numero di misure, "tante" significa perlomeno $n>30$,
ma alcuni Autori scrivono $n>50$ o $n>60$, altri $n>100$ … e insomma, più sono, meglio è),
allora l'istogramma delle frequenze tende ad assomigliare ad una gaussiana;
e *quanto più* tale somiglianza sussiste, *tanto più,*

detta $\bar{x}$ la media di queste misure $\boxed{\bar{x} = \textbf{MEDIA} = \dfrac{x_1+x_2+\ldots+x_n}{n}}$

e detto s il loro scarto quadratico medio $\boxed{s = \textbf{S.Q.M.} = \sqrt{\dfrac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n}}}$,

sono corrette le affermazioni seguenti:

a) **$\overline{x}$ è un valore prossimo al vero valore della grandezza in questione, dove per "vero" valore si intende quello che si otterrebbe come media su di un numero enorme di misure**

b) **circa il 68% delle misure effettuate rientra nell'intervallo $(\overline{x}-s,\ \overline{x}+s)$**
circa il 95% delle misure effettuate rientra nell'intervallo $(\overline{x}-2s,\ \overline{x}+2s)$
circa il 99,7% delle misure effettuate rientra nell'intervallo $(\overline{x}-3s,\ \overline{x}+3s)$

c) **se facessi un'ulteriore misura, questa avrebbe**
circa il 68% di probabilità di cadere nell'intervallo $(\overline{x}-s,\ \overline{x}+s)$
circa il 95% di probabilità di cadere nell'intervallo $(\overline{x}-2s,\ \overline{x}+2s)$
circa il 99,7% di probabilità di cadere nell'intervallo $(\overline{x}-3s,\ \overline{x}+3s)$

Di solito, per misurare una grandezza fisica, si effettua un certo numero *n* di operazioni di misura, si calcolano la media $\overline{x}$ e lo scarto quadratico medio s degli *n* valori così trovati, poi si scrive che la grandezza in gioco vale

$$\boxed{\overline{x}\pm s}\ ,$$

dove per la piena comprensione di questa scrittura occorre tenere presenti le 3 considerazioni a), b), c).

Torniamo soltanto a ribadire alcuni concetti davvero fondamentali.
Affinché le affermazioni precedenti siano corrette, il numero *n* delle misure deve essere "GRANDE"…

Inoltre LE AFFERMAZIONI CONTENGONO DEGLI AVVERBI "CIRCA", NON SOLO PER IL FATTO CHE I VALORI 68%, 95%, 99,7% SONO TUTTI APPROSSIMATI, MA SOPRATTUTTO PER IL FATTO CHE SI STA PENSANDO AD UNA CONFIGURAZIONE PROBABILISTICA IDEALE ALLA QUALE SI TENDE AD AVVICINARSI (SENZA PERO' RAGGIUNGERLA) AL CRESCERE DI *n*.
La veridicità di a), b), c) è tanto maggiore quanto più $\overline{x}$ (media delle n misure realmente effettuate) è prossimo a μ (vero valore della grandezza, media su un numero di misure che tende all'infinito) e quanto più *s* (s. q. m. delle n misure realmente effettuate) è prossimo a σ (s. q. m. su "infinite" misure); ed è all'aumentare del numero delle misure che effettivamente $\overline{x}, s$ tendono a identificarsi con μ, σ !!!

ESEMPIO

Qui sotto riportiamo 96 misure in mm della larghezza della lavagna di un'aula, rilevate dai 24 studenti, che hanno effettuato 4 misurazioni ciascuno:

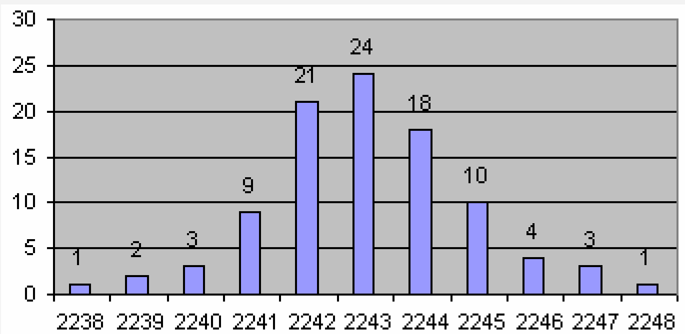
2242	2240	2244	2243	2244	2244	2242	2247	2244	2242	2246	2244	2241	2244	2242	2241	2244	2243	2242	2241
2242	2242	2243	2242	2246	2243	2245	2238	2246	2244	2244	2244	2243	2242	2245	2241	2243	2239	2244	2243
2245	2243	2247	2243	2244	2245	2242	2243	2245	2239	2246	2242	2243	2241	2244	2245	2244	2241	2242	2241
2243	2242	2243	2244	2243	2248	2243	2242	2241	2245	2243	2242	2240	2245	2244	2242	2243	2242	2241	2243
2243	2244	2243	2242	2245	2244	2243	2242	2243	2245	2242	2240	2243	2242	2243	2247				

Il calcolo ci dà

$$media=\overline{x}=2243{,}06...;\quad scarto\ quadratico\ medio=s=1{,}79...$$

Se ora andiamo a contare il numero di misure che sono comprese nell'intervallo $(\overline{x}-s,\ \overline{x}+s)$, vediamo che tali misure sono $21+24+18=63$.
Bene, 63 è assai prossimo al 68% di 96 (che vale circa 65). [vedi NOTA]

Ecco l'istogramma della distribuzione di frequenza, che in effetti presenta, pur con irregolarità, il tipico andamento "a campana".



NOTA - Per la precisione, quello che abbiamo inizialmente indicato come il 68% avrebbe potuto essere meglio approssimato come 68,3%, e il 95% come 95,4%. Oppure, si sarebbe potuto scrivere 95% ma sostituendo il fattore 2 con un più preciso 1,96.

Lo diciamo per scrupolo, e tuttavia insistiamo: **non dobbiamo confondere la configurazione probabilistica ideale, teorica, alla quale ci si avvicinerebbe se *n* tendesse a infinito, con la situazione reale, che è approssimata bene, ma non certo alla perfezione, quando *n* comincia ad esser >30, o meglio ancora >100.**

Un'ultima puntualizzazione.
La Statistica Inferenziale insegna che, per meglio stimare lo scarto quadratico medio σ relativo alle "infinite" misure, è più giusto calcolare lo scarto quadratico medio s del "campione di n misure" attraverso la formula "corretta" che si ottiene prendendo come denominatore $n-1$ anziché n:

$$\textbf{scarto quadratico medio "corretto"} = \hat{s} = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n-1}} = \sqrt{\frac{n}{n-1}}\cdot s$$

E' pur vero che **quando n è grande, scarto quadratico medio "corretto" e "non corretto" differiscono di pochissimo; e quando n non è grande, la teoria esposta non vale più!**
Già per valori di n dell'ordine di qualche decina, la differenza è assai piccola.
Ad esempio, con $n=30$, il fattore $\sqrt{n/(n-1)}$ vale $\sqrt{30/29}\approx 1{,}017$ che è molto vicino a 1!

In **EXCEL** e in **OPENOFFICE CALC**
lo **scarto quadratico medio "non corretto"** è **dev.st.pop()** mentre quello "**corretto**" è **dev.st()**

APPROFONDIMENTO (NON SEMPLICE): INTERVALLI DI CONFIDENZA, ERRORE STANDARD

In realtà, quando andiamo a calcolare la media $\bar{x}$ e lo scarto quadratico medio s sulle n misure che abbiamo effettuato, il nostro interesse è puntato, più che a *quelle* particolari n misure, al valore "*vero*" – che ci è sconosciuto – della grandezza in esame.
Ora, abbiamo già detto che quest'ultimo può essere pensato come "quel valore μ che si otterrebbe come media su di un numero sterminato di misure".
Ma fino a che punto possiamo ritenere che la media $\bar{x}$ da noi calcolata sia prossima al "vero" valore μ?
La "statistica inferenziale" ci insegna che se noi effettuiamo una serie di n misure, ed n è grande (certi Autori scrivono $n>30$, altri $n>50$ o 60, altri ancora $n>100$; … in realtà … quanto stiamo dicendo tende ad essere tanto più veritiero quanto più n è alto), allora, determinando per queste n misure la media $\bar{x}$ e lo scarto quadratico medio s, il *vero* valore μ avrà una probabilità

- **del 68% circa di rientrare nell'intervallo** $\left(\bar{x}-\frac{s}{\sqrt{n}},\ \bar{x}+\frac{s}{\sqrt{n}}\right)$
- **del 95% circa di rientrare nell'intervallo** $\left(\bar{x}-2\frac{s}{\sqrt{n}},\ \bar{x}+2\frac{s}{\sqrt{n}}\right)$
- **del 99,7% circa di rientrare nell'intervallo** $\left(\bar{x}-3\frac{s}{\sqrt{n}},\ \bar{x}+3\frac{s}{\sqrt{n}}\right)$

anche se per *maggiore precisione concettuale*, poiché il "vero valore" è … quello che è, è *costante*, mentre *a variare è invece l'insieme delle n misure e con esso l'intervallo* che ne deriva (è come se noi "estraessimo a sorte un intervallo, per poi domandarci se comprende o no il valore 'vero' "), bisognerebbe piuttosto partire "dal punto di vista dell'intervallo", dicendo che

- **il 68% circa degli intervalli** $\left(\bar{x}-\frac{s}{\sqrt{n}},\ \bar{x}+\frac{s}{\sqrt{n}}\right)$
  costruiti ciascuno facendo n misure e calcolandone i relativi $\bar{x}$ e s contiene al suo interno il "vero" valore (e il 32% circa lo lascia invece al suo esterno)
- **il 95% circa degli intervalli ecc. ecc.**
- **il 99,7% circa degli intervalli ecc. ecc.**

Questi intervalli di cui abbiamo parlato vengono chiamati "INTERVALLI DI CONFIDENZA".
Ad es., $\left(\bar{x}-3\frac{s}{\sqrt{n}},\ \bar{x}+3\frac{s}{\sqrt{n}}\right)$ **è un "intervallo di confidenza al 99,7%" per il vero valore della grandezza.**

Osserviamo l'uso del termine "confidenza" (= fiducia) al posto di "probabilità".

La quantità $\frac{\sigma}{\sqrt{n}}$ viene detta "**ERRORE STANDARD DELLA MEDIA**" (brevemente: "errore standard"),
e, se n è grande, così come $\bar{x}$ è una buona approssimazione per μ, allo stesso modo s è una buona approssimazione per σ e quindi $\frac{s}{\sqrt{n}}$ è una buona approssimazione per $\frac{\sigma}{\sqrt{n}}$.

Il discorso è intrigante, ma complicato. RICAPITOLIAMO LE PREMESSE E LA SIMBOLOGIA.

Stiamo supponendo di ricercare il valore "vero" di una determinata grandezza, tramite una misura, anzi: tramite una serie di *n* misure, di cui faremo poi la media.
μ è il vero valore della grandezza. μ è incognito e viene approssimato con la media $\overline{x}$ delle *n* misure realizzate.

Se noi avessimo la possibilità di effettuare un numero grandissimissimissimo di misure, al tendere all'infinito di questo numero, la media delle misure tenderebbe a μ.

Ma noi per forza di cose ci dobbiamo accontentare delle nostre *n* misure.
n è grande, ma non colossale: prenderemo *n* almeno maggiore di 30, preferibilmente maggiore di 100 …
tuttavia le nostre misure, pur essendo tante, saranno *n* e basta.

Calcoleremo dunque la media $\overline{x}$ e lo scarto quadratico medio s delle nostre *n* misure.
Bene, avendo preso *n* piuttosto grande abbiamo fiducia che $\overline{x}$ sia già una approssimazione piuttosto precisa per μ, e che s sia già prossimo a quello che sarebbe lo s. q. m. σ se noi potessimo effettuare "infinite" misure.

Possiamo anzi "quantificare" questa nostra "fiducia".

Se consideriamo, ad esempio, l'intervallo $\left(\overline{x}-2\frac{s}{\sqrt{n}},\ \overline{x}+2\frac{s}{\sqrt{n}}\right)$,

la nostra fiducia che questo intervallo contenga μ è all'incirca del 95%, perché la Statistica Inferenziale insegna che, qualora andassimo a barbosissimamente effettuare 100 serie, o 1000 serie, … , di *n* misure ciascuna, calcolando per ognuna di queste il relativo $\overline{x}$ e il relativo s,

all'incirca il 95% degli intervalli $\left(\overline{x}-2\frac{s}{\sqrt{n}},\ \overline{x}+2\frac{s}{\sqrt{n}}\right)$ così costruiti conterrebbero μ.

Questo è tanto più vicino al vero quanto più *n* è grande, ma a partire da $n>30$ cominciamo già ad andar benino!

ESEMPIO

Prendiamo in prestito un esempio dal testo "**Essential medical statistics**" di B. R. Kirkwood e J. A. C. Sterne, dove ogni cosa è spiegata con calma, precisione, e ottimi riferimenti concreti (*hats off*, tanto di cappello!)

In realtà qui si ragiona in un ambito più generale del nostro.
Viene infatti esaminata non una singola grandezza misurata più volte, bensì una "*popolazione*" limitata (l'insieme delle 10000 case), nonché un suo "*campione*" (le 100 case che vengono visitate).
Ma lo stesso discorso fatto per le misure vale, nei suoi tratti essenziali, anche in questo contesto, perché si può osservare che la quantità di cui ci si sta occupando (la superficie da disinfestare nelle case) è uno degli svariati fenomeni della realtà che presentano una *Gauss-like distribution*, vale a dire: una distribuzione simile alla gaussiana.

Nell'ambito di un piano per l'eradicazione della malaria si progetta di trattare con insetticida tutte le 10000 case di una certa area rurale. Problema: quanto insetticida acquistare? Per deciderlo, si estrae da quelle 10000 case un campione casuale di 100 case, e le si ispeziona per misurare in ciascuna casa la superficie che richiede di essere bonificata.

In quelle 100 case la superficie media su cui spruzzare l'insetticida risulta essere di $\overline{x}=24,2\ m^2$ *con uno scarto quadratico medio* $s=5,9\ m^2$.

Non è realistico a questo punto supporre che la superficie media $\overline{x}$ *rilevata nel campione di 100 case coincida con la media* μ *della superficie da disinfestare nell'intera "popolazione" delle 10000 case; tuttavia, è possibile valutare quanto sia da ritenere affidabile la media campionaria* $\overline{x}=24,2\ m^2$

se si va a calcolare l'errore standard, approssimabile con $\frac{s}{\sqrt{100}}=\frac{s}{10}=\frac{5,9}{10}=0,59\approx 0,6$.

A questo punto, infatti, si può dire che l'intervallo $(24,2\pm 0,6)\ m^2$ *ha una probabilità del 68% circa di contenere il valore incognito* μ *della media di tutta la "popolazione" delle 10000 case; e che l'intervallo* $(24,2\pm 2\cdot 0,6)\ m^2=(24,2\pm 1,2)\ m^2$ *ha una probabilità del 95% circa di contenere* μ.

Allora l'intervallo $(24,2\pm 1,2)\ m^2$ *è un intervallo di confidenza al 95% per* μ*; se quindi ipotizziamo che questo intervallo contenga* μ*, abbiamo una probabilità del 95% circa di ipotizzare il vero.*

μ *dovrebbe perciò, al 95% di "confidenza", di "fiducia", non essere superiore a* $(24,2+1,2)\ m^2=25,4\ m^2$ *per cui se acquistiamo una quantità di insetticida tale da poter coprire* $25,4\cdot 10000\ m^2=254000\ m^2$ *abbiamo il 95% di probabilità che questo sia sufficiente al bisogno.*

Tutto il discorso fatto regge bene perché la numerosità del nostro campione (n = 100) è decisamente alta. Coraggio, allora: abbiamo stimato quanto insetticida plausibilmente ci serve, andiamo a procurarcelo.

E se volessimo comprare l'insetticida sulla base di una confidenza del 99,7% circa? Per quanti metri quadrati dovremmo attrezzarci? Fai tu il semplice calcolo: troverai circa $260000\ m^2$.

***MEAN ± SD* oppure *MEAN ± SEM* ???**

Il ruolo dello *scarto quadratico medio* (**SD, Standard Deviation**) e quello dell'*errore standard della media* o semplicemente *errore standard* (**SEM, Standard Error of the Mean**) non devono essere confusi.

Sovente alcuni risultati, ad esempio in Medicina, vengono scritti con un'incertezza uguale al SEM, che è sempre per definizione minore della SD, proprio per dare l'idea di una minore variabilità … ma ciò può essere fonte di fraintendimenti gravi, se il lettore poi confonde questo SEM con la SD.

Cerco di spiegarmi. Supponiamo che una certa caratteristica quantitativa x relativa al sangue umano venga testata su di un campione di 400 individui presi a caso dalla popolazione generale, e si trovi che in questi individui la caratteristica in gioco vale 235 ± 42, essendo 235 la media calcolata sui 400 individui osservati, e 42 la SD delle 400 osservazioni.

Supponiamo inoltre che si sappia che la caratteristica studiata si distribuisce nella popolazione secondo la "campana" di Gauss o comunque una sua buona approssimazione (NOTA ♥)

Bene, se si scrive che la caratteristica in esame è stata osservata, in quel campione di 400 soggetti, con un valore dato da 235 ± 42 $(mean,\ SD)$, allora un medico che legge l'articolo scientifico potrà dire:

> in quel campione di 400 persone, pressappoco il 95%
> aveva quel valore compreso fra $235 - 2 \cdot 42 = 151$ e $235 + 2 \cdot 42 = 319$, e siccome quel campione (essendo abbastanza numeroso) è un'immagine piuttosto fedele dell'intera popolazione,
> se si presenta da me un paziente che ha quel valore minore di 151 o maggiore di 319,
> sono portato a classificare quel caso come anomalo e tale da richiedere ulteriori indagini cliniche;
> se invece in un paziente il valore è esterno all'intervallo $235 \pm 2 \cdot 2{,}1$
> (2,1 è il valore approssimato dell'Errore Standard della Media o SEM, che si desume da una SD di 42 con $n = 400$: $42/\sqrt{400} = 42/20 = 2{,}1$), questo non mi preoccuperà affatto!
>
> Piuttosto, l'intervallo $235 \pm 2 \cdot 2{,}1$ è un *intervallo di confidenza* al 95% per x, nel senso che ha il 95% di probabilità di contenere il "vero valore" di x, ossia la media dei valori di x nell'intera popolazione.

Quindi

♪ il **SEM** mi interessa per valutare con quale probabilità un dato intervallo intorno alla media campionaria contenga la media dell'intera popolazione, ossia per la **STIMA DELLA MEDIA INCOGNITA** μ,

♫ mentre la **SD** mi interessa per quantificare la **DISPERSIONE delle rilevazioni NEL MIO CAMPIONE**, considerazioni che poi posso estendere tali e quali all'intera popolazione, perché, dato il numero elevato di elementi del campione e dato che erano stati estratti casualmente dalla popolazione, **il campione rappresenterà abbastanza fedelmente la popolazione intera**.

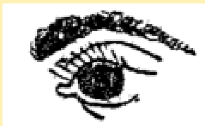
NOTA IMPORTANTE ♥

Questa richiesta è essenziale, perché

PARECCHI FENOMENI DELLA REALTA' PRESENTANO UNA "GAUSS-LIKE DISTRIBUTION", MA CIÒ NON VALE PER ALTRI!

Ad esempio, hanno una distribuzione più o meno sovrapponibile alla gaussiana

- gli errori di misura, come abbiamo visto (ma, a dire il vero, *non proprio sempre*)
- le distanze dal centro di un bersaglio per una serie di tiri
- i quozienti di intelligenza
- le altezze degli adulti di una stessa etnìa e sesso



… mentre per la distribuzione dei *pesi* delle persone la differenza rispetto alla gaussiana è già più marcata.

SINONIMO di "**DISTRIBUZIONE GAUSSIANA**" è "**DISTRIBUZIONE NORMALE**".

Le considerazioni sopra riportate possono rendere una prima idea di alcune fra le questioni di cui si occupa la **STATISTICA INFERENZIALE**.

Essa **interviene quando si cerca di studiare una caratteristica dell'intera popolazione tramite osservazioni condotte su di un suo sottoinsieme ("campione"), e occorre quantificare il grado di attendibilità di questo procedimento**.

Come nei **sondaggi elettorali**.

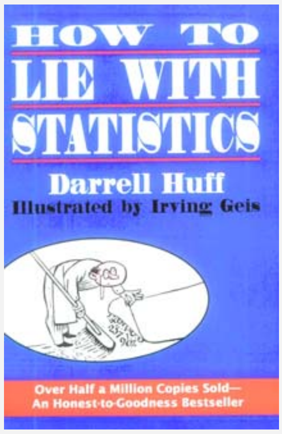
Come nelle **ricerche farmacologiche**, dove si va a confrontare l'evoluzione clinica di due gruppi di malati, a uno dei quali viene somministrata la sostanza attiva e all'altro, invece, un preparato inerte (il "*placebo*").

Come nei **test finalizzati a verificare** (in un determinato contesto) **la bontà di una ipotesi**.

La statistica inferenziale considera anche il caso in cui siano disponibili solo **piccoli campioni**.

Noi però, nei limiti del nostro corso, ci dobbiamo fermare ai pochi cenni dati, senza approfondire oltre.

How to Lie with Statistics



E' un libretto di divulgazione, scritto da Darrell Huff nel lontano 1954, che ha avuto uno straordinario successo di vendite, e conserva ancor oggi piena attualità.
Aiutandosi con garbate illustrazioni, passa in rassegna i modi attraverso i quali la pubblicità e la politica manipolano e presentano in modo parziale e distorto le statistiche, per spingere il consumatore o l'elettore a conclusioni sbagliate.

Il campione con l'errore incorporato
Alle interviste sulle letture abituali probabilmente la gente risponderà mentendo, almeno parzialmente, perché "confessare" letture frivole o imbarazzanti non fa fare bella figura. Analogo il discorso per l'igiene personale.
Poi le persone contattate hanno tendenza a dare quelle risposte che pensano possan far piacere a chi conduce l'intervista: l'autore riferisce, come esempio significativo, di un'analisi statistica a soggetto politico che aveva avuto esiti radicalmente diversi con intervistatori di pelle bianca o rispettivamente nera.
Le persone in una stazione ferroviaria sono rappresentative della popolazione generale?
Probabilmente no: le madri di bambini piccoli, ad esempio, potrebbero scarseggiare in quel campione.
E gli incaricati a svolgere sondaggi per strada potranno tendere a scegliere persone più pulite o più gradevoli, o chi intuiscono sia più disponibile a rispondere, specie se devono terminare il loro compito in tempi ristretti.

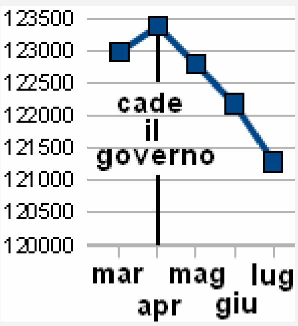
La media ben scelta
Quando si parla di "media", in realtà ci si sta riferendo a una *media aritmetica*, a una *mediana* o a una *moda*?
Dire che lo stipendio medio annuo dei dipendenti di un'azienda è, poniamo, di 38.500 dollari, può essere comodo per i dirigenti. Ma questa media, che è l'ordinaria *media aritmetica*, è comprensiva anche dei compensi stratosferici dei pochissimi manager strapagati, e i sindacalisti potrebbero invece considerare come stipendio "medio" la *mediana* degli stipendi, pari a 20.000 dollari (in pratica: solo metà dei dipendenti percepisce uno stipendio superiore a 20.000 dollari, l'altra metà inferiore). E' possibile che dirigenza e sindacati usino dunque il medesimo termine "media", in relazione a indicatori ben diversi.

Quei piccoli numeri che non ci sono
E' purtroppo frequente che si utilizzino (senza segnalarlo), per un'indagine statistica, campioni troppo piccoli per poter dare risultati attendibili; che si ometta la specificazione del grado di "dispersione" dei dati …

Molto rumore per praticamente nulla
Del tutto inutile confrontare due dati non molto differenti fra loro, senza specificare quale sia l'intervallo di "incertezza" di questi dati! Oppure: se abbiamo l'elenco completo delle marche di sigarette in commercio, elencate per grado di pericolosità decrescente ma tutte pressappoco allo stesso livello di tossicità, è insensato e ingannevole pubblicizzare la marca che sta in fondo all'elenco dicendo che è la "più raccomandabile"!

Il grafico fantasmagorico
"Tagliare" in modo scaltro i grafici, e/o scegliere furbamente le lunghezze dei segmenti che rappresentano date quantità in orizzontale e in verticale, può favorire forti distorsioni nella percezione di chi osserva.
Il diagramma in basso a sinistra nella pagina mira maliziosamente a suggerire che il numero di copie vendute da una rivista di politica sia crollato dopo la caduta del governo, mentre si è mantenuto pressoché stabile.

L'immagine monodimensionale
Rappresentare una quantità con ideogrammi ha un'insidia: se due raffigurazioni "in scala" di altezza una doppia dell'altra vengono utilizzate per illustrare il fatto che un certo valore è raddoppiato, l'osservatore ha comunque un'impressione diversa: l'area della seconda figura è *quadrupla* rispetto alla prima, e se anzi le figure vengono pensate come tridimensionali un raddoppio dell'altezza comporta un volume che è addirittura 8 volte tanto.
Quindi disegni di questo tipo possono essere impiegati per indurre la sensazione di una crescita (o diminuzione) più forte di quella reale. L'immagine in basso a destra dà un esempio di questo effetto psicologico.



Tracollo … o sostanziale stabilità delle vendite?

Il numero pseudoconnesso
"Se qualcuno non può dimostrare ciò che vorrebbe dimostrare, può dimostrare qualcos'altro e far finta che sia la stessa cosa" …

Il vecchio *post hoc* ritorna in sella
Se B segue, in ordine di tempo, A, ciò non implica che A sia causa di B.
Nelle Nuove Ebridi si era convinti che i pidocchi facessero bene alla salute ☺; in realtà, è ben facile che una persona ammalata sviluppi la febbre, e l'aumento di temperatura … scaccia i pidocchi! Causa ed effetto completamente ribaltati.

Il signore il basso guadagna… il doppio, il quadruplo, o 8 volte tanto?

ALTRI MODI DI QUANTIFICARE L'INCERTEZZA DELLA MISURA (per la distinzione fra la parola "errore" - spesso adoperata impropriamente - e la parola "incertezza", vedi l'importante NOTA a pag. 421)

a) **SCARTO ASSOLUTO MEDIO (SCARTO MEDIO, DEVIAZIONE MEDIA, ERRORE MEDIO)**
Al posto dello scarto quadratico medio s, si può prendere lo "scarto assoluto medio" δ

ossia la **media dei valori assoluti degli scarti dalla media**: $\boxed{\delta = \dfrac{|x_1 - \overline{x}| + |x_2 - \overline{x}| + \ldots + |x_n - \overline{x}|}{n}}$.

Si scriverà allora che la grandezza in esame vale $\boxed{\overline{x} \pm \delta}$

b) **SEMIDISPERSIONE (INCERTEZZA ASSOLUTA, ERRORE ASSOLUTO, ERR. MASSIMO)**
Effettuate le n misure $x_1, x_2, \ldots, x_n$ e calcolata la media $\overline{x}$ di queste, si va a determinare
la "semidispersione" (da alcuni detta "incertezza assoluta" o "errore assoluto" o "errore massimo")

cioè la **semidifferenza fra la più grande e la più piccola delle misure rilevate**: $\boxed{d = \dfrac{x_{MAX} - x_{min}}{2}}$

poi si scrive semplicemente che il valore della grandezza in questione è $\boxed{\overline{x} \pm d}$.

Questo metodo molto elementare della semidispersione viene impiegato più che altro QUANDO IL NUMERO DELLE MISURE A DISPOSIZIONE È BASSO O MOLTO BASSO.
La semidispersione è sovente indicata col simbolo Δx (naturalmente, se la grandezza è t, si userà Δt !)
Si legge "delta *x*"; quel Δ è un simbolo utilizzato, in questo e in altri casi, come "operatore di differenza".

c) **IL CASO DELLA MISURA UNICA, AD ES. PERCHE' LO STRUMENTO E' POCO SENSIBILE**
Quando, infine, lo strumento di misura è poco sensibile, cosicché
gli errori "casuali" o "statistici" non emergono e si rileva dunque sempre la stessa, grossolana, misura;
oppure anche quando l'operazione di misura viene effettuata una sola volta,
si scrive, detta $\overline{x}$ la misura trovata, che il valore della grandezza è $\boxed{\overline{x} \pm a}$,
essendo a l'**ampiezza** dell'**intervallo che corrisponde a due "tacche" consecutive del misuratore**
(o la **semiampiezza** nel caso le tacche siano abbastanza distanziate).

In qualsiasi caso, l'incertezza dichiarata riguardo a una misura non dovrebbe mai essere inferiore a quella dovuta alla sensibilità dello strumento.

"*In generale, la presenza di errori casuali nella misura fa sì che l'errore statistico risulti maggiore dell'errore strumentale (la sensibilità dello strumento), ma talvolta può accadere il contrario! Si stabilisce allora che l'incertezza nelle misure è data dal maggiore tra questi due errori*"
(prof. Aurelio Agliolo Gallitto, Dipartimento di Fisica, Università di Palermo, http://portale.unipa.it)

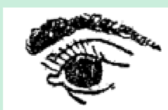
Quando poi si è scelto quale tipo di incertezza (**si trova spesso scritto, impropriamente: di *errore*** ☹)
si vuol scrivere accanto alla media delle misure, sarebbe bene **indicare questa scelta ESPLICITAMENTE!**

Vediamo un **ESEMPIO**.
40 misurazioni del periodo T di oscillazione di un pendolo hanno fatto registrare questi valori (in secondi):

4,80	4,82	4,84	4,83	4,79	4,83	4,86	4,86	4,82	4,83	4,87	4,88	4,87	4,89
4,83	4,75	4,86	4,82	4,84	4,87	4,81	4,78	4,85	4,86	4,84	4,79	4,84	4,88
4,85	4,80	4,84	4,85	4,89	4,85	4,83	4,79	4,84	4,81	4,85	4,84		

- La **media** delle misure è stata quindi $4{,}83625$, arrotondata a $4{,}84$
- la **semidispersione** è stata $0{,}07$ per cui potremo scrivere, tenendo conto di essa, $T = 4{,}84 \pm 0{,}07$
- lo **scarto assoluto medio ("errore medio")** è stato $0{,}0243125$ arrotondabile a $0{,}02$ o a $0{,}024$
per cui, tenendo conto di esso, $T = 4{,}84 \pm 0{,}02$ o in alternativa $T = 4{,}836 \pm 0{,}024$
- lo **scarto quadratico medio** è stato $s = 0{,}0309586...$ arrotondato a $0{,}03$ da cui $T = 4{,}84 \pm 0{,}03$
(verifica che la percentuale dei valori compresi fra $4{,}84 - 0{,}03$ e $4{,}84 + 0{,}03$ non si discosta molto dal 68%!)

Come si vede,
l'intervallo, intorno alla media, che si utilizza per esprimere il valore di una grandezza, dipende dal modo col quale viene espressa l'incertezza; e la corretta interpretazione della scrittura "... ± ..." sarà legata alla conoscenza del significato delle varie quantità δ, d, a, s, $s/\sqrt{n}$ … tenendo in debito conto il numero di misure effettuate.



ERRORI RELATIVI / INCERTEZZE RELATIVE
(♥ noiosamente ribadiamolo: è brutta consuetudine ☹ della letteratura scientifica scrivere, tendenzialmente, la parola "errore" anche nei casi in cui il termine corretto sarebbe "incertezza")

L' "errore/incertezza relativo/a" è il quoziente, il rapporto, fra un errore/incertezza (di qualsiasi tipo!) e il valore della grandezza da misurare (valutato tramite la media delle misure rilevate; se tale valore fosse negativo, si intende di ignorarne il segno, cioè di prenderlo in valore assoluto).

Consideriamo nuovamente le 40 misure del periodo di un pendolo elencate a pag. 414.

- La media delle misure è stata $4{,}83625$, arrotondata a $4{,}84$
  La semidispersione è stata $0{,}07$, da cui la possibilità di scrivere $T = 4{,}84 \pm 0{,}07$.
  Dunque l'incertezza assoluta viene qui valutata in $0{,}07$:
  bene, l'incertezza relativa sarà allora all'incirca di $\frac{0{,}07}{4{,}84} \approx 0{,}014$.
  In forma percentuale, l'incertezza relativa è (circa) dell' $1{,}4\,\%$
- L'errore medio è stato $0{,}0243125$ arrotondato a $0{,}02$ per cui, tenendo conto di esso, $T = 4{,}84 \pm 0{,}02$:
  l'errore medio relativo è (circa) $\frac{0{,}02}{4{,}84} \approx 0{,}004$, e l'errore medio relativo percentuale circa dello $0{,}4\%$
- Lo scarto quadratico medio è stato $s = 0{,}0309586...$ arrotondato a $0{,}03$ da cui $T = 4{,}84 \pm 0{,}03$
  quindi lo **scarto quadratico medio relativo** - detto, come sappiamo, "**coefficiente di variazione**" -
  è (circa) $\frac{0{,}03}{4{,}84} \approx 0{,}006$, e lo scarto quadratico medio relativo percentuale è all'incirca dello $0{,}6\%$

L'incertezza relativa può essere impiegata per confrontare la precisione di misure di quantità diverse.

Ad esempio, se nella misura dell'altezza di una parete A c'è l'incertezza di 10 cm
mentre nella misura dell'altezza di un'altra parete B l'incertezza è di 20 cm,
non possiamo affermare che la misura di A sia più precisa di quella di B
se non conosciamo quanto valgono, all'incirca, le altezze di A e di B …
Poniamo che A sia una casa a due piani alta pressappoco 6 metri e B un grattacielo di circa 130 metri:
l'incertezza relativa su A sarà di $\frac{0{,}10}{6} \approx 0{,}017$ mentre l'incertezza relativa su B di $\frac{0{,}20}{130} \approx 0{,}0015$ (meno della decima parte della precedente!), quindi in questo caso va considerata di gran lunga più precisa la misura di B.

GLI ERRORI "SISTEMATICI"

Nel valutare la misura di una grandezza fisica,
oltre agli **errori "CASUALI" (detti anche "ACCIDENTALI" o "STATISTICI")**
(ossia: oltre agli errori **legati a circostanze *imprevedibili* e mai completamente controllabili,**
le quali possono influire sul risultato della misura *ora per difetto, ora per eccesso*),
si possono commettere anche errori cosiddetti **SISTEMATICI**.

Questi **influiscono sempre per difetto o sempre per eccesso sul valore rilevato**, e derivano:

- ❑ dall'**inadeguatezza dello strumento di misura** (esempi: un orologio che "ritardi", un termometro che con la propria temperatura vada a modificare in modo sensibile la temperatura dell'oggetto in esame …)
- ❑ dall'**uso non appropriato di tale strumento** (es.: dimenticarsi di "azzerarlo", quando ciò sia necessario)
- ❑ da **applicazione di leggi sbagliate o metodi sbagliati di indagine**
  (ad esempio cercare di determinare la profondità di un pozzo lasciandovi cadere una pietra e annotando dopo quanti secondi si sente "splash", per poi utilizzare la formula nota che regola spazi e tempi nella caduta dei gravi … ma senza tener conto che il suono dell'impatto con l'acqua ci mette a sua volta un certo tempo per salire dal fondo del pozzo alle nostre orecchie).

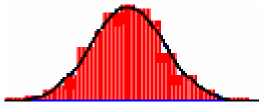
Gli errori sistematici possono essere individuati ed eliminati o perlomeno minimizzati, mentre sugli errori accidentali non possiamo far nulla
(a parte, è ovvio, cercare di effettuare l'operazione di misura con tutta l'attenzione di cui siamo capaci);
l'incertezza legata agli errori accidentali è ineliminabile: può solo essere quantificata coi metodi visti sopra, e ridotta facendo, se possibile, un numero elevato di misure.

Alcuni testi introducono come categoria a sé stante gli "**ERRORI DI SENSIBILITA'** ",
ossia quelli **legati alla sensibilità dello strumento**.
Se misuro la larghezza di un foglio di carta con un righello le cui tacche più ravvicinate siano quelle dei mm,
a ogni misura sarà comunque associata un'incertezza di $0{,}5$ mm (secondo alcuni, di 1 mm)

Gli errori casuali si presentano solo quando sono maggiori della sensibilità dello strumento!!!

ESERCIZI (risposte a pag. 426)

1) Calcola, per il seguente insieme di dati: 0 1 2 2 5
   a) la media ………… b) la semidispersione ………….. c) lo scarto assoluto medio ……………
   d) la varianza ………… e) lo scarto quadratico medio (arrotondato a 1 cifra dopo la virgola) …………

2) VERO O FALSO?
   a) "*Scarto quadratico medio*" e "*deviazione standard*" sono sinonimi
   b) Se effettuo tantissime misure di una grandezza *G*,
   e calcolo la loro media $\overline{x}$ e la loro deviazione standard s,
   nello scrivere $G = \overline{x} \pm s$ io intendo che l'intervallo da $\overline{x} - s$ a $\overline{x} + s$
   ha una probabilità del 68% circa di contenere il vero valore della grandezza
   c) La media $\overline{x}$ fra un numero elevato *n* di misure è una buona approssimazione del valore vero μ
   della grandezza, e se a questo punto faccio *k* misure in più e vado a calcolare la media fra tutte
   le $n + k$ misure, certamente tale nuova media sarà ancora più vicina al vero valore della grandezza
   d) Nella figura  che si riferisce alle ripetute misurazioni di una quantità fisica,
   le altezze dei rettangoli rappresentano le frequenze
   e) Nella stessa figura di prima, le basi dei rettangoli rappresentano le classi di misura
   f) Lo scarto quadratico medio "corretto" è minore di quello "non corretto"
   g) La funzione "scarto quadratico medio" (non corretto) si indica, nel foglio elettronico, con dev.st.()
   h) Per dimezzare l' "errore standard della media" occorre raddoppiare il numero delle misure

3) Sono state rilevate 625 misure. La media di queste è stata $\overline{x} = 152{,}4$ e lo scarto quadratico medio $s = 2{,}5$.
   a) Un intervallo nel quale rientrerà, pressappoco, il 95% delle misure effettuate è quello che va
   da …………… a ……………
   b) Un intervallo di confidenza al 95% per il vero valore della grandezza in esame
   (ossia, un intervallo che ha una probabilità intorno al 95% di contenere il vero valore della grandezza)
   è invece quello compreso fra …………… e ……………

4) La media fra 64 misurazioni di una grandezza risulta essere 173,5 e il loro scarto quadratico medio 2,3.
   a) Determina un intervallo nel quale dovrebbe rientrare pressappoco il 68% di questi 64 dati
   b) Determina un intervallo di confidenza al 68% per il valore della grandezza in esame

5) Misurando 80 volte il tempo di caduta di un grave da una data altezza, in secondi, sono stati trovati i valori:

43,3	41,2	42,5	42,4	43,4	43	44	42	41,6	40,8	42,4	44	43,8	42	43,4	42,2
42,8	41,5	44,5	43,3	42,5	44,2	41,8	42,4	42,2	43,2	42,2	41,8	42,1	43,1	41,7	42,1
42,8	44	43,3	42,7	44,3	44,1	41,4	42,5	42,8	42,8	43,1	42,1	43	42	42,2	42
42,5	41	43,6	43,3	42,9	43,2	42,3	42,9	42,3	41,8	42,2	42,7	41,3	44,4	42,8	42,8
42,1	40,9	43,7	43,6	43,4	43	42,4	44,3	41,7	43,7	42,7	43,4	42,3	42,1	43,3	42

Foglio elet-tro-nico!

   a) Esprimi quel tempo come … ± … utilizzando la deviazione standard e arrotondando media e dev. standard
   a 1 cifra dopo la virgola. Conta il numero di dati tra $\overline{x} - s$ e $\overline{x} + s$ e il numero di quelli tra $\overline{x} - 2s$ e $\overline{x} + 2s$.
   b) Determina un intervallo di confidenza al 95% (cosa significa?) per il valore della grandezza in esame.

6) Clicca sulla freccia per un altro esercizio di questo tipo, con 400 dati già pronti ⇨

7) Si vuole stimare l'età media in cui si presenta una data patologia.
I 400 pazienti seguiti da un famoso centro specializzato hanno contratto la malattia all'età media di 44 anni.
La distribuzione di queste 400 età è *Gauss-like*, con scarto quadratico medio uguale a 10 anni.
Se ne deduce allora che l'intervallo di età che va da … anni a … anni ha una probabilità valutabile intorno al 95% di contenere l'età media di insorgenza della malattia, qualora venisse calcolata sui malati di tutta Italia.
Si può anche dire che, in quel campione di 400 malati, pressappoco il 95% avrà contratto la malattia nell'intervallo di età che va da … anni a … anni;
e siccome il campione, piuttosto numeroso, rispecchia la popolazione generale, tale intervallo di età sarà anche quello entro il quale sviluppa la malattia il 95% circa degli italiani che si ammalano.

8) Dal sito www.regentsprep.org:
Battery lifetime is normally distributed (= segue la distribuzione *normale*, cioè *gaussiana*)
for large samples (*sample* = *campione*).
The mean lifetime is 500 days and the standard deviation is 61 days.
What percent of batteries have lifetimes longer than 561 days?

9) Supponi che in una grande città sudamericana sia stata rilevata l'altezza di 420 ragazzi quattordicenni, ottenendo una media di m 1,67 e uno scarto quadratico medio di cm 10.
Allora un intervallo di altezze che ha il 95% di probabilità di contenere l'altezza media di tutti i ragazzi di quell'età, residenti in quella città, è quello che va da ……. a …….

10) I moderni test per attribuire il cosiddetto "quoziente di intelligenza" (Q.I.) sono progettati in modo che nella curva, approssimativamente gaussiana, ottenuta disponendo in orizzontale i vari punteggi realizzabili e in verticale il numero di persone, tutte di una stessa età e, che hanno realizzato quella determinata fascia di punteggio, la media risulti uguale a 100 e la deviazione standard a 15.
In questo modo, pressoché il 95% delle persone della stessa età avrà un Q.I. che si collocherà fra … e …

11) Per una certa popolazione di rane allo stato naturale, si è visto che la lunghezza della vita è distribuita normalmente (cioè, segue una distribuzione gaussiana) con media 10 anni e deviazione standard di 3 anni (www.cli.di.unipi.it).
Quale percentuale di queste rane sopravvive oltre i 16 anni?

12) Sono state effettuate solo 5 misure, che hanno fornito gli esiti seguenti: 85 86,5 85,5 88 86
Se vogliamo esprimere il valore della grandezza con una scrittura del tipo ... ± ... , come faremo?

13) Misurando ripetutamente una grandezza sono stati trovati i valori
2,60 2,59 2,58 2,59 2,59 2,54 2,58.
a) Esprimi quella grandezza come ... ± ... utilizzando la semidispersione.
b) Se scriviamo $G = \overline{x} \pm d$, dove d è la semidispersione,
in generale siamo sicuri che tutte le osservazioni effettuate rientrino fra $\overline{x} - d$ e $\overline{x} + d$?

14) In un sito Internet troviamo che la misura di un dato tempo è $(3,27 \pm 0,02)\,s$ e che la misura di una data velocità è $(24,4 \pm 0,3)$ $(mean \pm SD)$. Ma nessuna delle due scritture è scientificamente corretta: perché?

15) Stabilisci quale delle due scritture seguenti esprime una misura di velocità più precisa:
$(3,24 \pm 0,04)\,m/s$ $(mean;\ SD)$; $(40,5 \pm 0,5)\,m/s$ $(mean;\ SD)$

16) Esprimendo un tempo come $(8,0 \pm 0,2)\,s$, qual è l'incertezza relativa percentuale?

17) Se si prendono i due insiemi di dati seguenti: a) 0 3 3 3 6 b) 0 1 2 3 4 per confrontarli, onde stabilire se i dati sono più "sparpagliati" nel primo caso o nel secondo, cosa occorrerebbe calcolare?

18) Lo scarto quadratico medio di n misure è risultato uguale a $2,0$ e calcolando l'errore standard della media si è ottenuto $0,1$. Quante misure sono state effettuate? …………

19) La tabella qui a destra riporta, in tre casi, il valore di una misura con a fianco l'incertezza da cui questo dato è affetto.
Stabilisci quale delle tre misure può essere considerata la più precisa.

a)	1,25	0,05
b)	10,0	0,3
c)	0,0040	0,0001

20) Una ditta produce camomilla in bustine da 5 grammi.
Si vuole controllare che non troppe bustine abbiano un peso sensibilmente diverso dal valore ottimale.
Pesando 40 bustine prodotte consecutivamente da un macchinario, si trovano i seguenti valori in grammi:

4,89	5,21	5,20	4,76	4,78	5,16	4,84	4,78	4,86	4,88
5,04	5,26	4,74	5,14	4,88	4,82	4,80	5,08	5,20	5,18
5,03	5,18	4,81	4,77	5,20	5,19	5,25	4,75	4,77	4,78
4,90	4,80	4,86	5,18	4,85	4,87	5,05	5,21	5,11	4,82

Quali sono la media e la deviazione standard di questo campione?
Se si è osservato che i dati in esame presentano una *Gauss-like distribution*, delle circa 48000 bustine prodotte in una giornata lavorativa, quante si può presumere che andranno a pesare non più di 4,60 grammi?

21) SOLO ALCUNE DISTRIBUZIONI SONO GAUSSIANE O *GAUSS-LIKE* , ALTRE NO!!!

Se consideriamo ad esempio la distribuzione dei *pesi delle persone*, o la distribuzione dei *tempi di attesa* dei clienti in una filiale bancaria, capiremo che *le differenze rispetto alla distribuzione normale sono notevoli*.
a) Spiega in che senso queste due distribuzioni, se confrontate con la normale, presentano una "coda verso destra" nella campana.

❑ Nelle **distribuzioni "non normali"**
non si ha necessariamente la coincidenza fra media, moda, mediana.
b) Per fare un esempio semplice, quanto valgono media, mediana e moda nella distribuzione della figura a destra, relativa ai voti dati in una classe di 21 studenti da un insegnante generosissimo? ……. ….... …....