



Corso di Chimica Generale
CCS Ingegneria CHI-MAT
A.A. 2018/2019 (I° Semestre)
LABORATORIO 1

 POLITECNICO DI MILANO



Esperimenti “Multitecnica”

Prof. Attilio Citterio
Dipartimento CMIC “Giulio Natta”



Misure di Peso e Volume

Parte A : Uso della bilancia analitica

Parte B : Trasferimento quantitativo

Parte C : Effettuazione di un prelievo

Parte D : Calibrazione di una pipetta

Parte E : Uso e lettura di una buretta



Obiettivo

- **L'esperienza introduce procedure e tecniche impiegate nell'analisi qualitativa e quantitativa**
- **L'apprendimento deve essere individuale**
- **Abituarsi ad associare ad ogni misura di grandezza fisica (massa, volume, concentrazione, ecc.) l'unità di misura e l'errore relativo (sistematico o meno), analizzando in termini statistici l'insieme dei dati ottenuti.**



Note Importante per la Partecipazione alle Esperienze a alla Stesura delle Relazioni

- 1) La frequenza ai laboratori è obbligatoria. L'esperienza deve essere partecipata attivamente e preceduta da una attenta lettura della traccia fornita e una programmazione delle attività da effettuarsi in laboratorio.
- 2) Copiare i risultati riportati da altri, inventare dati, modificare i valori sperimentali misurati nell'ambito di questo o altro laboratorio è disonestà accademica.

Gli studenti identificati dagli addetti di laboratorio perché non osservanti dei punti 1 e 2 verranno penalizzati nella votazione finale!!



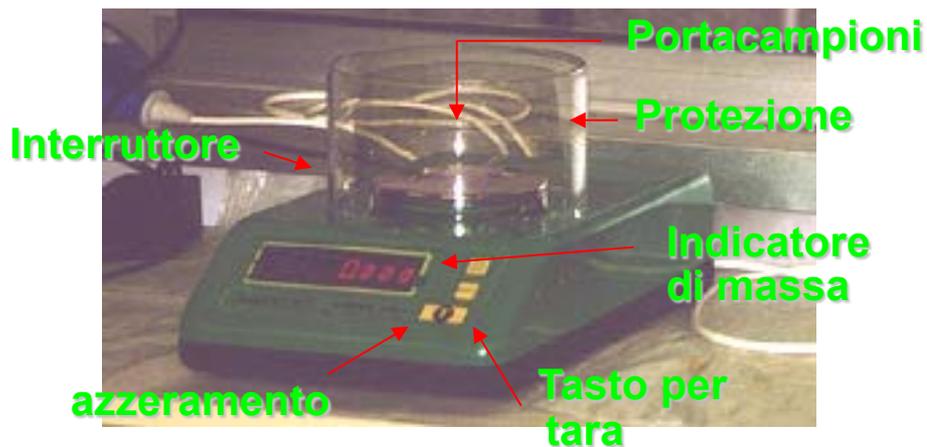
Note di Sicurezza

- Tutti i reagenti usati in questa esperienza devono essere considerati nocivi (vedere i relativi MSDS se si hanno dubbi) .
- Le operazioni richiedono l'uso di occhiali e dei guanti.
- Evitare l'ingestione di soluzioni o solidi.
- Fare attenzione nell'uso della vetreria che è fragile e la cui rottura può provocare ferite.
- Non inquinare i vari reagenti e le soluzioni ri-introducendo nelle confezioni i campioni prelevati.
- Il KMnO_4 è un forte ossidante ed è tossico: maneggiarlo con cautela senza versamenti (come solido o sue soluzioni non versarlo nel lavandino ma raccoglierlo nell'apposito contenitore per lo smaltimento dei rifiuti tossico-nocivi). L' H_2O_2 è anch'essa un forte ossidante.





tipologie



**Bilancia
tecnica elettronica**

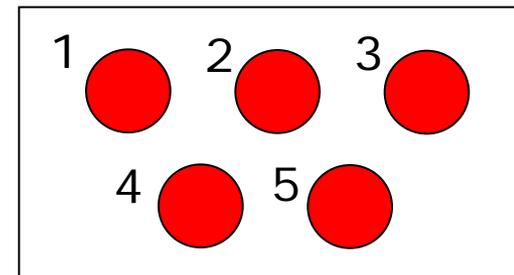


**Bilancia
analitica**



A. Peso di Monete da 0.50 €

- Istruzione preliminare sull'uso di bilance (taratura e azzeramento)
- Pesare 5 monete da 0.50 euro una alla volta (identificandole) - Riportare il peso su un foglio opportunamente predisposto
- Pesare tutte le 5 monete e, prelevandone in successione una alla volta con una pinza, pesarle per differenza.
- Determinare la media, la mediana e la deviazione standard di ogni moneta

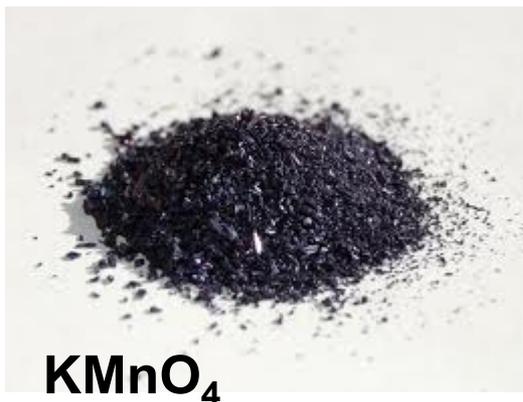




Parte B - Trasferimento Quantitativo

Via fase liquida

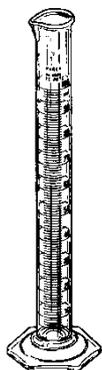
- Pesare un beaker vuoto da 50 mL su una bilancia a 3 cifre.
- Mediante una spatola trasferire nel beaker del KMnO_4 solido (circa 0.4 g) e pesarlo esattamente sulla stessa bilancia.
- Dalla differenza di peso del beaker prima e dopo l'aggiunta del KMnO_4 recuperare il peso esatto alla terza cifra decimale del KMnO_4 caricato.





Misure di Volumi di Liquidi

vetreria



**Cilindro
graduato**



Buretta



**Pipetta
graduata**



**Matraccio
graduato**



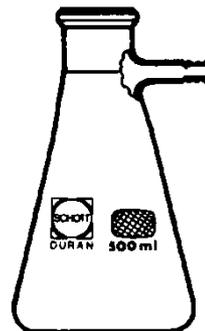
Contenitori per Liquidi



Beaker



Beuta



Beuta da vuoto



Pallone



Boccetta per lavaggio



Provetta



Provetta da centrifuga

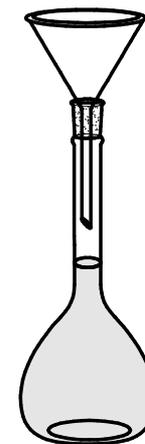


Pipetta



Preparazione della Soluzione

- Sciogliere il KMnO_4 nel beaker in circa 20 *mL* di acqua distillata. Per evitare perdite mescolare cautamente. La soluzione è quasi satura, e la dissoluzione dei cristalli è lenta.
- Trasferire quantitativamente la soluzione in un matraccio tarato da 100 *ml* mediante un imbuto.





Portata a Volume

- Far scorrere la soluzione su una bacchetta di vetro appoggiata al becco del beaker.
- Sciacquare il beaker e ripetere la procedura (**valutare** la quantità di acqua di lavaggio necessaria per trasferire quantitativamente).
- Riprendere le gocce sulla bacchetta con l'acqua di una spruzzetta, trasferendo accuratamente al matraccio. Risciacquare analogamente l'imbuto.
- **Diluire la soluzione nel matraccio fino alla tacca (livello inferiore del menisco). Mettere il tappo, capovolgere e scuotere il matraccio. Rimetterlo verticale e lasciare tornare in superficie le bolle d'aria. Ripetere (8-10 volte) fino ad omogeneizzazione. Conservare la soluzione per la parte C.**





Obiettivo:

*il liquido che si vuole erogare
deve possedere la stessa composizione di quello
prima della misura*

Campionamento Accurato



- Riempire una pipetta con la soluzione di KMnO_4 e lasciarla scolare.
- Introdurre poca acqua distillata nella pipetta (da un beaker da 50 mL), risciacquare e smaltire nell'apposito bidone di recupero (non nel lavandino!) la soluzione di risciacquo.
- (vedi filmato)
- Disporre orizzontalmente la pipetta e farla ruotare per sciacquare le pareti. Determinare:
 - Il minimo di risciacqui per rimuovere il colore dalla pipetta (< 3)
 - Il minimo volume di acqua di risciacquo ($< 5 \text{ mL}$ in cilindro grad.).
 - Durante le operazioni di risciacquo controllare il colore dell'acqua nel beaker di prelievo. Se è rosa la si è contaminata con KMnO_4 . Ripetere con più attenzione fino a che non si verifica più l'inconveniente.



Omogeneizzazione

- Sciacquare bene una pipetta da 10 *mL* con la soluzione di KMnO_4 .
- Pipettare 5 *mL* in un matraccio da 100.
- Portare a volume senza rimescolarne il contenuto.
- Mescolare ripetutamente capovolgendo e scuotendo il matraccio (*notare la difficoltà di disperdere il colore uniformemente*).
- Sciacquata la pipetta, pipettare 10 *mL* della soluzione in una provetta.



Parte D - Calibrazione

- Pulire una pipetta da 5 o 10 mL
- Pesare un beuta da 50 mL con tappo
- Pipettare 10.0 mL di acqua distillata nella beuta da un beaker di acqua distillata, equilibrato a T ambiente
- Ripesare la beuta con il tappo
- Ripetere per 3 volte
- Valutare Media (\bar{M}) e deviazione standard (σ) delle misure effettuate



Parte E - Uso della Buretta

- Riempire la buretta con H_2O distillata
- Portare alla tacca zero agendo sul rubin.
- Spillare 3 ml di H_2O raccogliendola in un beaker pesato. Ripesare.
- Ripetere 2 volte
- (*valutare M e σ*)
- Spillare 30 o 40 gocce di H_2O in un beaker pesato. Ripesare. Ripetere 2 volte.
- Rifare con mezze gocce
- (*valutare M e σ*)





Parte F. Riproducibilità di una Buretta e Titolazione

- Pesare una beuta da 50 mL con tappo
- Gocciolarvi con la migliore precisione 5 mL di acqua. Tappare la beuta e pesare.
- Valutare per differenza il peso di H₂O.
- Ripetere 2 volte (valutare M e σ)
- Calcolare il valore della densità dell'acqua
- Verificare le misure con quelle di altri 2 gruppi.



Titolazione Redox

- Pipettare 50 ml della soluzione di KMnO_4 in un beaker
- Aggiungere con una pipetta 20 gocce di H_2SO_4 al 30% e caricare la soluzione ottenuta nella buretta.
- Titolare una soluzione di H_2O_2 (concentrazione = 0.80%), procedendo piano per evitare di superare il punto di viraggio (da *incolore* – a *bruno/viola*).



Titolazione - Procedura

- In una titolazione si misura la quantità di un reagente (il volume di liquido erogato dalla buretta $\Delta V = V_2 - V_1$) necessario a consumare tutto un altro reagente.
- Si valuta il livello iniziale del liquido viola (V_1). Si posiziona il contenitore del campione da titolare sotto la buretta. Come il titolante viola viene a contatto con il campione il suo colore sparisce (lo ione Mn^{2+} è rosa tenue, mentre lo ione MnO_4^- è viola intenso).
- Si aggiunge la soluzione del titolante della buretta alla soluzione del campione nella beuta. Finché rimane incolore, il titolante è in difetto sul reagente.
- Al punto di fine titolazione, quando si è aggiunto tanto titolante da far sparire tutto il reagente, la soluzione vira al viola debole per un lieve eccesso di titolante.
- A questo punto si fa la lettura finale (V_2).





Titolazione di Soluzioni di Acqua Ossigenata

Utilizzo del metodo della titolazione volumetrica per determinare la concentrazione di una soluzione di acqua ossigenata fornita.

CONCETTI / TECNICHE

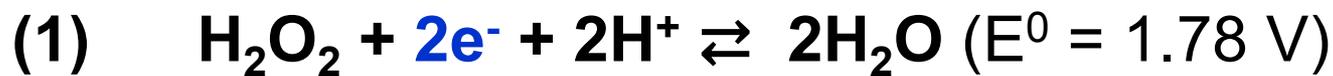
Titolazione, Punto di Fine = Punto di Equivalenza
Standardizzazione, Standard Primario

Si vuol determinare **quanta acqua ossigenata** (*Percento in Peso*) è contenuta in un campione titolando:

La Soluzione di **Acqua Ossigenata (H_2O_2)** con **Permanganato di Potassio ($KMnO_4$)** di concentrazione nota precisamente (**$KMnO_4$** standardizzato)



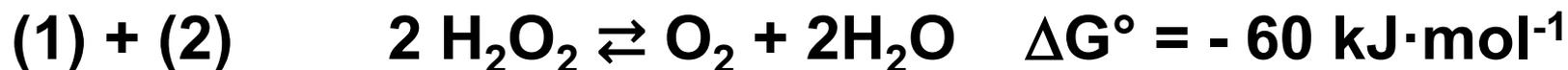
- **Acqua Ossigenata**: è normalmente considerata un **AGENTE OSSIDANTE** (cioè, una sostanza capace di *accettare* elettroni)



- Però, H_2O_2 può comportarsi anche da **agente riducente** ed essere **ossidata** ad **ossigeno molecolare (O_2)** da forti agenti ossidanti, quale lo **ione Permanganato, MnO_4^-** secondo la reazione :

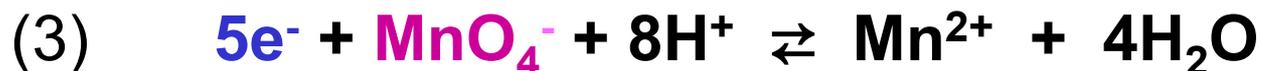


- *In effetti, se non stabilizzata (per aggiunta di inibitori), l' H_2O_2 decompone reagendo con se stessa lentamente secondo la reazione:*

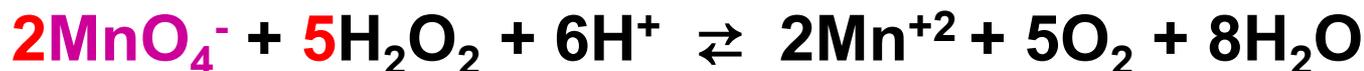




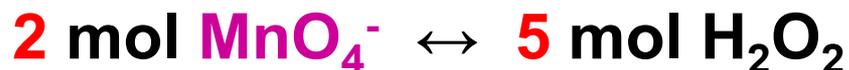
- **Lo ione permanganato, MnO_4^-** normalmente reagisce formando ioni **MANGANOSI Mn^{2+}** in SOLUZIONE ACIDA:



- Perciò, l'equazione bilanciata per la reazione di KMnO_4 con H_2O_2 , in un ambiente ACIDO è $[2 \times (5) + 5 \times (1)]$:



- Un reagente, MnO_4^- è colorato (**viola scuro**), gli altri reagenti e prodotti sono quasi incolori. Perciò, il colore viola **SPARISCE** finché non c'è più H_2O_2 con cui reagire.
- Il punto di fine si ha quando il colore di MnO_4^- **PERSISTE**.
- La stechiometria complessiva della reazione è:





STANDARDIZZAZIONE

- La concentrazione $[\text{KMnO}_4]$ si determina per standardizzazione.
- Lo Standard è **l'OSSALATO di SODIO ($\text{Na}_2\text{C}_2\text{O}_4$)**.
- **Lo ione ossalato** è ossidato tramite la semireazione:



- La reazione complessiva tra gli ioni MnO_4^- e $\text{C}_2\text{O}_4^{2-}$ è:



- La stechiometria complessiva della standardizzazione



- Consultare un testo di matematica o statistica per una visione d'insieme dell'uso di funzioni statistiche nelle misure sperimentali. Di seguito vengono dati elementi minimi per la valutazione dei parametri statistici più rilevanti e utili per la stesura delle relazioni.



Statistica Descrittiva

Usata per descrivere la natura dei dati.

Statistica Induttiva

Usata per dedurre affermazioni, predizioni o prendere decisioni.

Le statistiche descrittive sono comunemente riportate ma entrambe sono necessarie per interpretare i risultati.



Misure Sperimentali

- Cifre Significative
- Errori
- Incertezza
- Distribuzione Gaussiana
- Intervalli di Confidenza



Cifre Significative

- Metodo usato per esprimere l'accuratezza e la precisione.
- Non si possono riportare numeri migliori del metodo usato per misurarli.

67.2 unità = tre cifre significative

**Cifre
certe**

**Cifra
incerta**



Cifre Significative: Regole per gli Zero

Gli zero iniziali **non sono** significativi.

Zero iniziali

$.0421$ - tre cifre significative

Gli zero interni **sono** significativi.

Zero interno

4.021 - quattro cifre significative

Gli zero terminali **sono** significativi.

Zero terminali

4.210 - quattro cifre significative



Cifre Significative nei Calcoli

Regole per l'Addizione e la Sottrazione

- Si guarda il *numero di cifre decimali*.
- Allineare le cifre decimali.
- Il risultato ha il numero di cifre decimali del numero con le cifre inferiori.
 - Esempio: $25.5 + 0.01 = 25.5$; $6.579 - 6.575 = 0.004$

Regole per la Moltiplicazione e Divisione

- Si guarda il *numero di cifre significative*.
- Eseguire la moltiplicazione o divisione
- Il risultato ha il numero di cifre significative del numero con cifre significative più basse.



Errore - differenza tra la risposta ottenuta e quella “vera”.

- **Sistematico** - problema connesso al metodo, tutti gli errori sono della stessa entità, grandezza e direzione - Individuare gli errori.
- **Casuale** - basato su limiti e precisione della misura - Errori indeterminati si possono trattare statisticamente.
- **Sbagli** - operazioni scorrette. Meglio ripetere il lavoro.

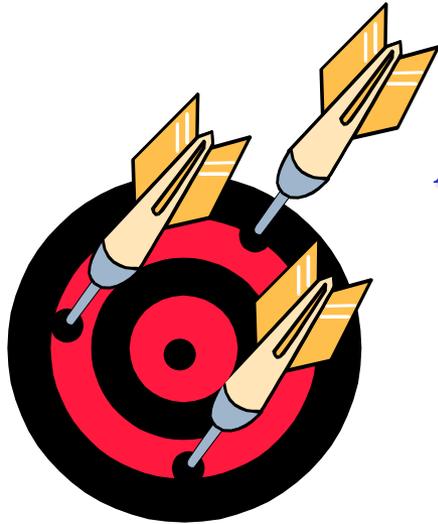


Accuratezza e Precisione

Accuratezza - quanto vicino è il “valore medio” a quello **corretto** (“vero”)

Precisione - quanto riproducibile (quanto sono in **accordo tra loro i valori misurati**)

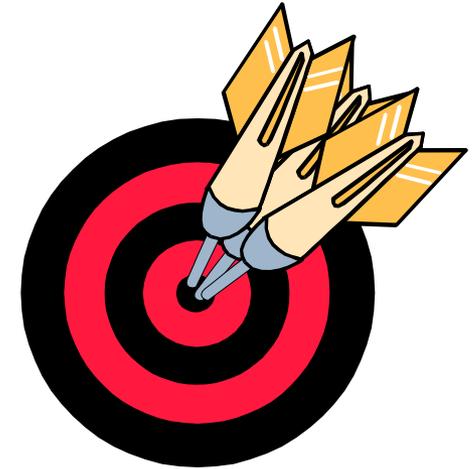
Accuratezza e precisione Il vero obiettivo! Valori accurati che sono ben in accordo tra loro. I due parametri non sono correlati, cioè una misura può essere precisa e non accurata o accurata e non precisa.



Accuratezza



Precisione



Accuratezza
e Precisione
Il vero obiettivo!



Potenziali errori strumentali

- **Variazione della temperatura**
- **Contaminazione della strumentazione**
- **Fluttuazioni di tensione**
- **Rottura di componenti**

Tutti questi errori si possono correggere tramite calibrazioni o una adeguata manutenzione dello strumento.

Errori del Metodo

- **Reazioni lente o incomplete**
- **Specie instabili**
- **Reagenti non specifici**
- **Reazioni collaterali**

Questi errori si correggono sviluppando un metodo adatto.



Identificazione degli Errori

Errori Personali

- Sbagli di lettura di uno strumento o scala
- Calibrazione impropria
- Pessima tecnica/preparazione del campione
- Pregiudizi personali
- Calcoli non corretti dei risultati

Errori Casuali

Questi sbagli si possono minimizzare o del tutto eliminare con un appropriato allenamento ed esperienza.

Ogni valore ottenuto è il risultato di molti fattori e variabili. Molte di queste variabili sono fuori controllo e sono di natura casuale.

Esempio - Lettura di una bilancia elettrica.

Un campione riportato come 1.0023 g presenta in realtà un peso nell'intervallo 1.0022 - 1.0024 g. Non se ne conosce il valore esatto alla quarta cifra decimale.



Riassunto Cifre Significative

- numero minimo di cifre per scrivere un numero senza perdita di accuratezza
- zero significative solo quando in mezzo o ultimo numero a destra del decimale
- +/- = numero di **cifre decimali** pari al numero w/ il minimo di **cifre decimali**
- \times/\div = numero di **cifre** pari al numero con le più basse **cifre**
- \log = numero di cifre nella mantissa



Calibrazione di una pipetta da 100 ml.

Fasi nella procedura:

- 1 - Pesare un contenitore pulito e asciutto.**
- 2 - Riempire la pipetta con acqua fino alla tacca.**
- 3 - Scaricare l'acqua nel contenitore.**
- 4 - Ri-pesare il contenitore.**
- 5 - Determinare il volume sulla base del peso e della densità dell'acqua.**



Fonti di errore

1 - *Peso del contenitore.*

Ogni bilancia ha un limite e, per esempio, una tipica bilancia analitica presenta un errore minimo di ± 0.1 mg.

Tale valore dovrebbe essere piccolo in confronto agli altri pesi che si devono misurare.



Fonti di errore

2 - *Riempire la pipetta con acqua alla tacca.*

Quanto si può valutare accuratamente il livello dell'acqua nella pipetta?

3 - *Scaricare l'acqua nel contenitore.*

Quanto è eseguibile bene l'operazione. L'angolo della pipetta e la viscosità del liquido provocheranno piccole variazioni.



Fonti di errore

4 - *Ri-pesare il contenitore.*

Qui si introduce un altro errore a causa del limite della bilancia.

L'errore ora è associato sia all'acqua che al contenitore.



Fonti di errore

5 - *Determinare il volume in base al peso e densità dell'acqua.*

Ora, la temperatura influisce sulla densità dell'acqua.

Si deve anche tenere conto che esistono limiti a quanto accuratamente è nota la densità ad una certa temperatura e che può variare in base alla purezza dell'acqua.



Per questa semplice procedura, si ha un minimo di cinque fonti di errore.

Ognuno può comportare un errore + o -.

L'errore totale sarà:

$$E_T = \pm E_{\text{peso1}} \pm E_{\text{vol}} \pm E_{\text{scar}} \pm E_{\text{peso2}} \pm E_{\text{densità}}$$

E questo si verifica in un semplice esempio!



Un Semplice Esempio (7)

- **Identificare le fonti di errore può aiutare a ridurre alcune.**
- **Non sarà mai possibile eliminare tutte le fonti di errore.**
- **Queste fonti sono di natura casuale.**
- **Ci si deve basare sul trattamento statistico dei dati per valutare questi errori.**



assoluta (ex)

- » incertezza stimata associata alla misura (p.es. ± 0.2 mm)
- » si noti che l'incertezza ha le stesse unità della misura



relativa (e_x/x)

relativa percentuale $(\%e_x)$

- » $\%e_x = (e_x/x) \times 100$
- » per una incertezza assoluta costante, l'incertezza relativa diminuisce con il valore della misura



Addizione e sottrazione

- » $x + y + z = w \pm ?$
- » $e_w = (e_x^2 + e_y^2 + e_z^2)^{1/2}$ (stesso numero di cifre decimali)

Moltiplicazione e divisione

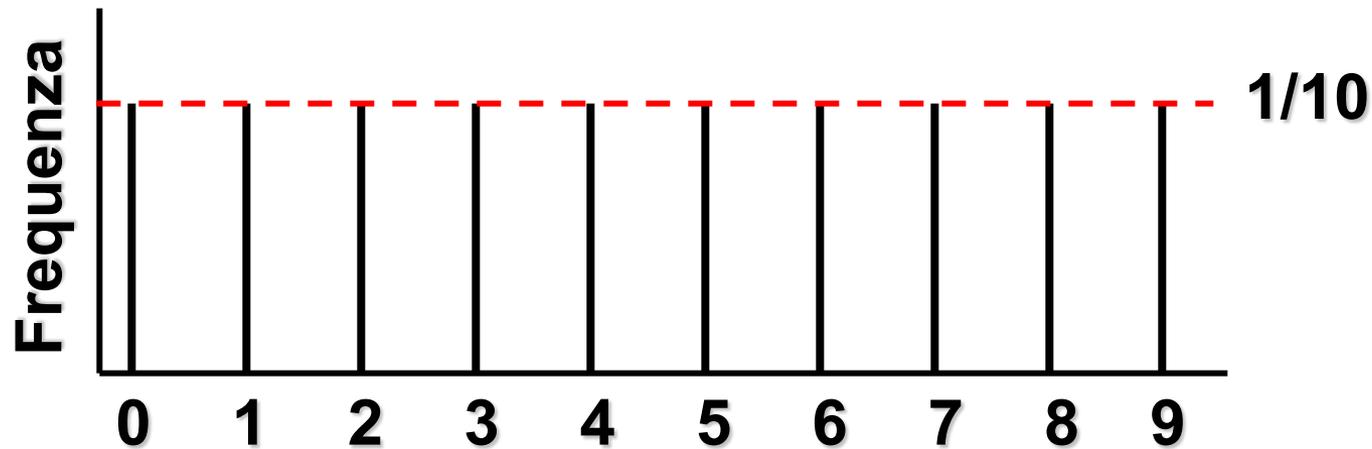
- » $(x)(y)(z) = w \pm ?$
- » $\%e_w = (\%e_x^2 + \%e_y^2 + \%e_z^2)^{1/2}$
(stesso numero di cifre significative del numero con il minor numero di c.f.)



- **Un carattere associato ad un evento**
 - **la sua tendenza ad avere luogo**
- **Per vedere di che cosa si sta parlando, usiamo un esempio basato su un dado a 10 facce.**
 - **Se molte persone hanno ognuna un dado identico, e ognuno lo lancia una volta, quale sarà il risultato atteso?**



Lancio di un Dado

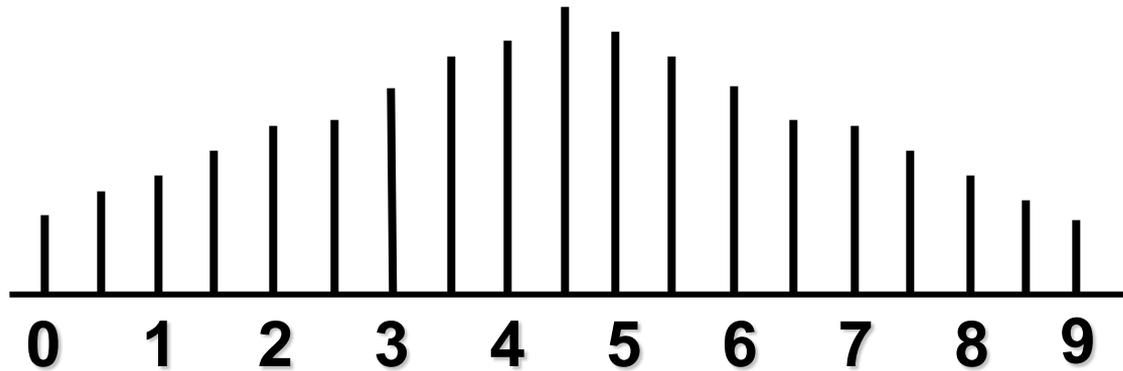


Per un lancio solo, ogni valore ha la stessa probabilità di comparire.

Cosa succede se ogni persona ha due dadi?



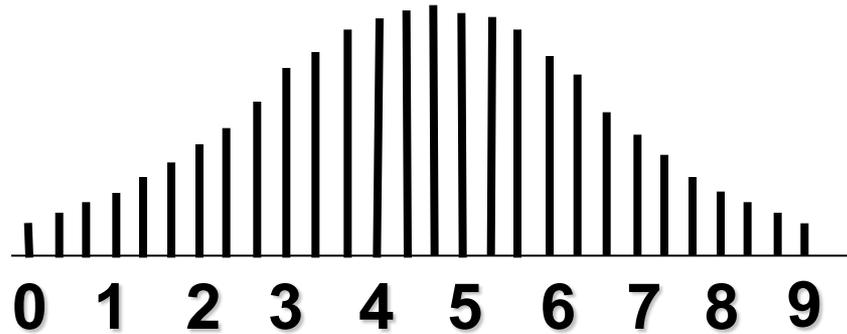
Media di due dadi - un lancio



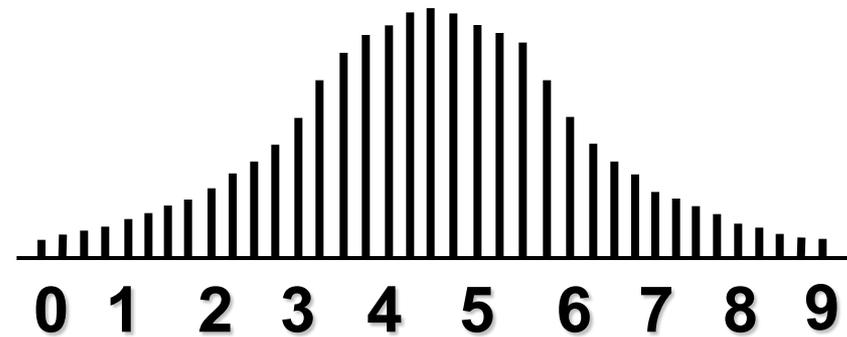
Si può continuare così, usando più dadi e un singolo lancio.



Più Dadi



Tre dadi

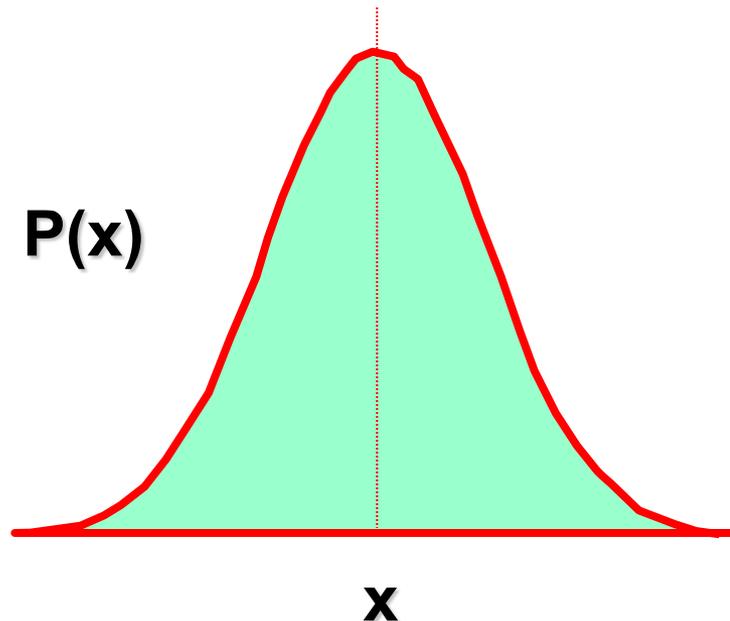


Quattro dadi



Come il numero di dadi tende ad infinito, i valori diventano continui, organizzandosi in una curva:

La curva diventa una distribuzione normale o Gaussiana





Distribuzione Normale (o Gaussiana)

Questa distribuzione è descritta dalla relazione :

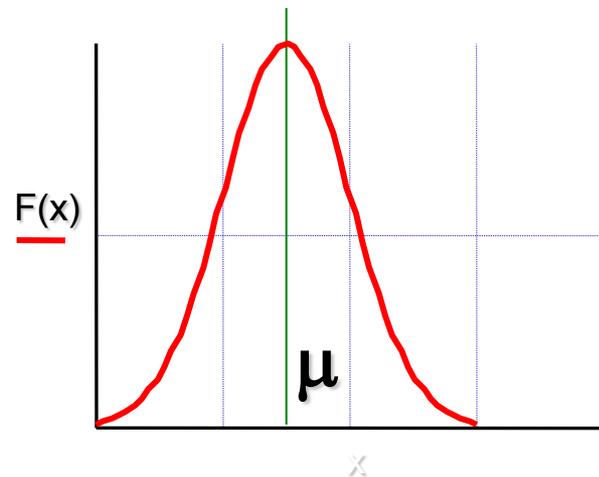
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x - \mu)^2}{2\sigma^2}$$

$$y = \left(\frac{1}{(2\pi)^{1/2} \sigma} \right) e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$-\infty < x < +\infty$$

σ - deviazione standard

μ - media universale



Naturalmente si richiede un insieme infinito di dati.



Distribuzione Gaussiana

1. errore normale, o curva ben distribuita
2. Una curva che predice la distribuzione dei dati
- solo errori casuali
3. DEVE definire i limiti della base di dati



- Curva caratterizzata da due parametri
 - media aritmetica $\mu = \bar{x}$
 - deviazione standard σ



◆ Media Aritmetica

$$\bar{x} = \frac{\sum x_i}{n}$$

x_i = *valore della misura i*

n = *numero di misure*



- Deviazione standard

$$\sigma = \left(\frac{\sum (x_i - \bar{x})^2}{n - 1} \right)^{1/2}$$

$n - 1 =$ gradi di libertà

$\sigma^2 =$ varianza



Distribuzione Gaussiana

- La deviazione standard è una misura dell'ampiezza della distribuzione - quanto sono raggruppati i valori attorno alla media
 - ◆ deviazione standard della media

$$\bar{\sigma} = \frac{\sigma}{\sqrt{n}}$$

- ◆ Si riduce la deviazione standard della media eseguendo più misure



Distribuzione Gaussiana

- ◆ per $x = -\infty$ a $+\infty$, la probabilità deve essere 1
- ◆ la probabilità che un valore stia in un certo intervallo è data dall'area sottesa dalla curva in quell'intervallo
 - ◆ 95.5% di tutti i valori cadono entro $\pm 2\sigma$ della media
 - ◆ 99.7% di tutti i valori cadono entro $\pm 3\sigma$ della media



Distribuzione Normale (2)

- I termini si possono calcolare come :

$$\mu = \sum_{i=1}^N x_i / N$$

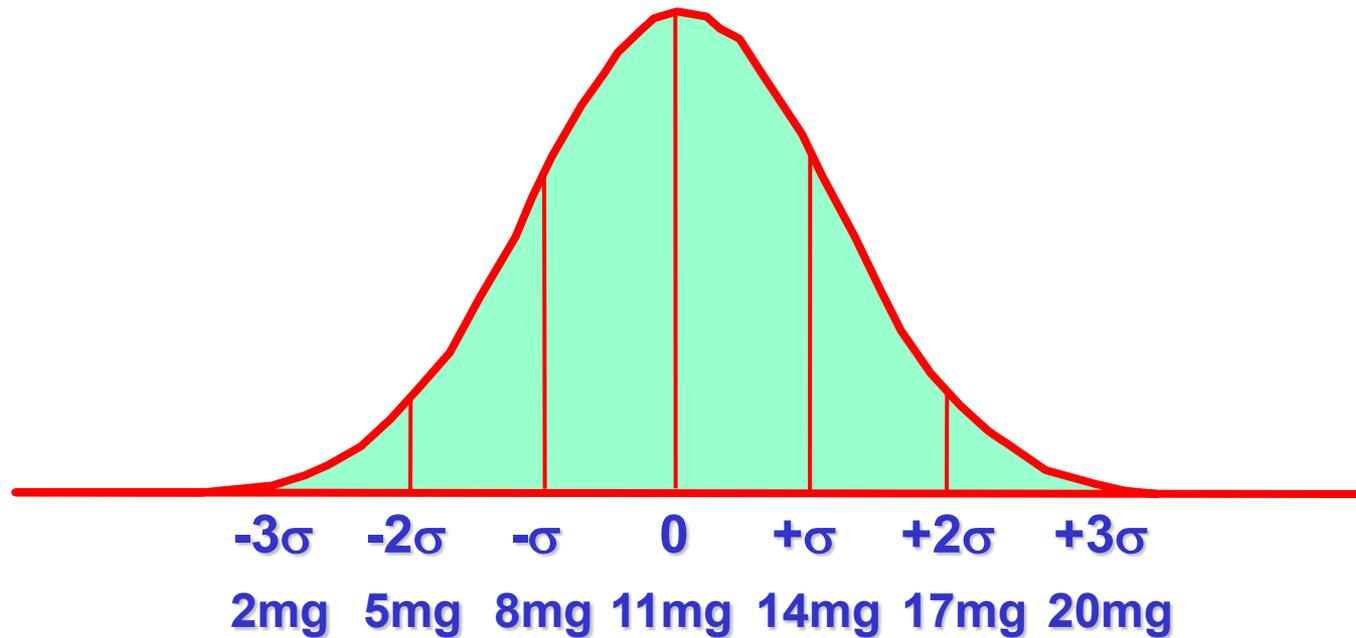
$$\sigma^2 = \text{varianza} = \sum_{i=1}^N (x_i - \mu)^2$$

$$\sigma = \sqrt{\sigma^2}$$

Questi si applicano ad un insieme infinito di dati ma si possono usare per insiemi di dati in cui $N > 100$ e dove ogni variazione è di natura completamente casuale.



Distribuzione Normale (3)

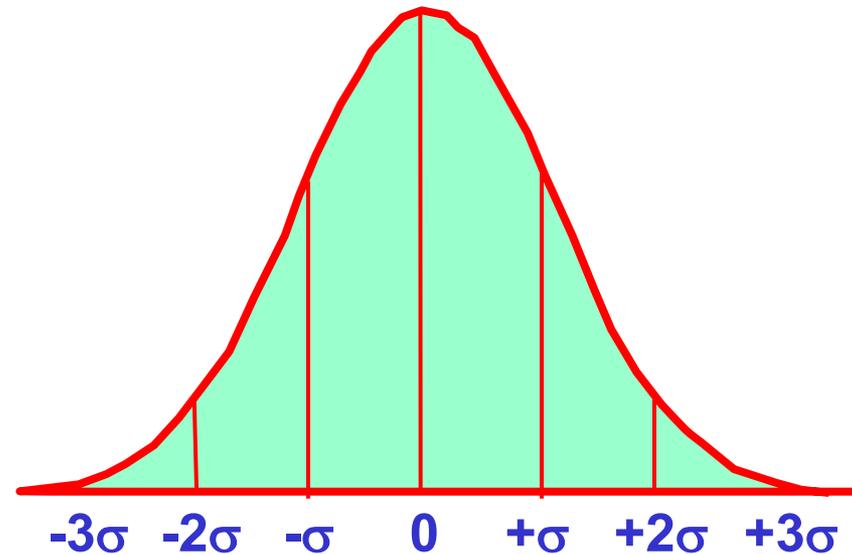


Quando si determina σ , si assume una distribuzione normale e si correlano i dati in unità σ .



Aree della Distribuzione Normale

L'area sottesa da ogni porzione della curva indica la probabilità di accadimento di un evento.



$\pm 1\sigma = 68.3\%$ dei dati

$\pm 2\sigma = 95.5\%$

$\pm 3\sigma = 99.7\%$



Insiemi con Numeri Elevati di Dati

- **Si può usare la curva di distribuzione normale per predire la probabilità di accadimento di un evento.**
- **Questo approccio è valido solo per insiemi molto grandi di dati ed è utile per settori quali il controllo di qualità di prodotti di massa.**
- **Negli esempi seguenti, si assumerà che si abbia a che fare con insiemi molto grandi di dati e che μ e σ siano stati valutati.**



La Variabile Ridotta

- Assumendo che si conosca μ e σ per un insieme di dati, si può calcolare u (la variabile ridotta) come:

$$u = \frac{(x - \mu)}{\sigma}$$

- Ciò si realizza semplicemente convertendo i valori di test dalle normali unità (mg, ore, ...) alla deviazione standard.



Uso della Variabile Ridotta

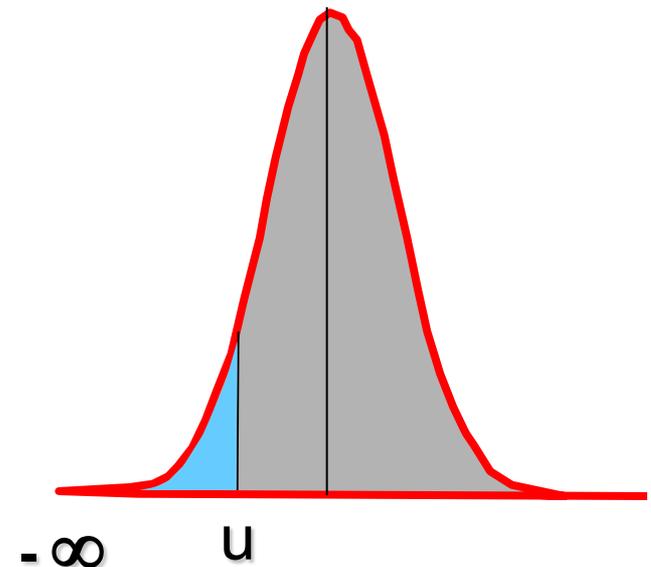
- **Assumendo che i dati seguano la distribuzione normale, si può usare u per predire la probabilità di accadimento di un evento.**
- **La probabilità si può dedurre dai valori di u riportati in adatte tabelle - modello A e B.**
- **Il modello da usare dipende dal problema a cui si deve dare una risposta.**



Modello A

- Questo modello fornisce l'area sottesa dalla curva da u ad ∞ .

$ u $	area	$ u $	area
0.0	0.5000	2.0	0.0227
0.2	0.4207	2.2	0.0139
0.4	0.3446	2.4	0.0082
0.6	0.2743	2.6	0.0047
0.8	0.2129	2.8	0.0026
1.0	0.1587	3.0	1.3×10^{-3}
1.2	0.1151	4.0	3.2×10^{-5}
1.4	0.0808	6.0	9.9×10^{-10}
1.6	0.0548	8.0	6.2×10^{-16}
1.8	0.0359	10.0	7.6×10^{-24}

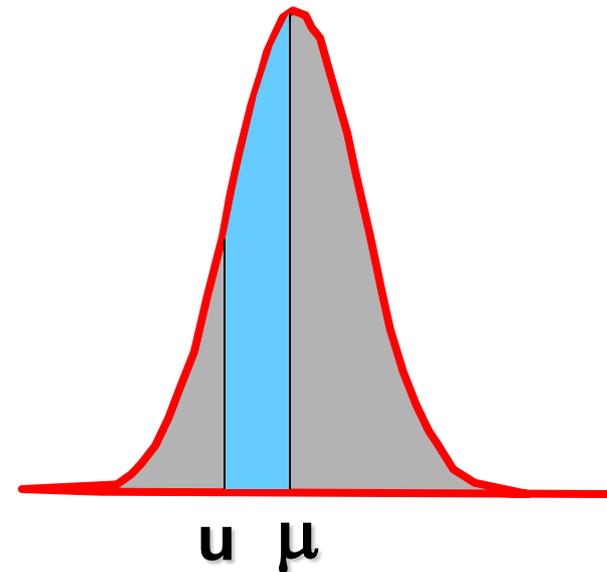




Modello B

- Questo modello fornisce l'area sottesa dalla curva da u ad μ .

$ u $	area	$ u $	area
0.0	0.0000	1.6	0.4452
0.2	0.0793	1.8	0.4641
0.4	0.1554	2.0	0.4473
0.6	0.2258	2.2	0.4861
0.8	0.2881	2.4	0.4918
1.0	0.3413	2.6	0.4953
1.2	0.3849	2.8	0.4974
1.4	0.3192	3.0	0.4987





Esempio

Un pneumatico viene prodotto con la seguente statistica di durata su strada.

$$\mu = 93300 \text{ km}$$

$$\sigma = 16100 \text{ km}$$

Quanto percorso in miglia si potrebbe garantire perché meno del 5% dei pneumatici siano da sostituire?



Esempio (2)

- **Con un occhio al modello A, si trova che un area di 0.05 è molto vicina a 1.6σ .**
- **Si usa ora l'equazione variabile ridotta per un σ di -1.6 (si vuole il valore $<$ della media).**

$$- 1.6 = (x - 93300) / 16100$$

$$x = 67540 \text{ km}$$



Un Altro Esempio

- **Si installa un elettrodo per pH per controllare una corrente di processo. Il fornitore lo fornisce con le seguenti specifiche di vita media dell'elettrodo:**

$$\mu = 8000 \text{ ore}$$

$$\sigma = 200 \text{ ore}$$

Se è risultato necessario sostituire l'elettrodo dopo 7200 ore di uso, l'elettrodo è “scadente”?



Un Altro Esempio (2)

Si calcola u :

$$u = \frac{(7200 - 8000)}{200} = -4.0 \sigma$$

Dal modello A, si trova che la probabilità ad un valore di 4 è 3.2×10^{-5} .

Ciò indica che solo lo 0.0032 % di tutti gli elettrodi dovrebbero rompersi a 7200 ore o prima. L'elettrodo è evidentemente scadente.



Quando si usano insiemi di dati più piccoli, si deve tenere conto di :

- 1. I campioni sono rappresentativi della popolazione? I valori devono essere completamente casuali.**
- 2. Se si prende un insieme non-casuale come tutti gli uomini o le donne - le differenze sono significative?**



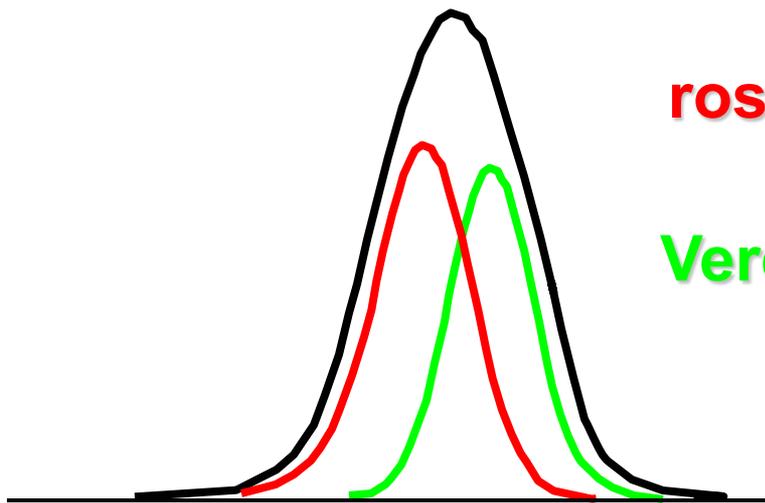
- Per insiemi di dati più piccoli, si usano le seguenti espressioni:

$$\text{Media} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}$$



- Questo esempio mostra cosa può succedere se si introduce un “bias” nel corso del campionamento.



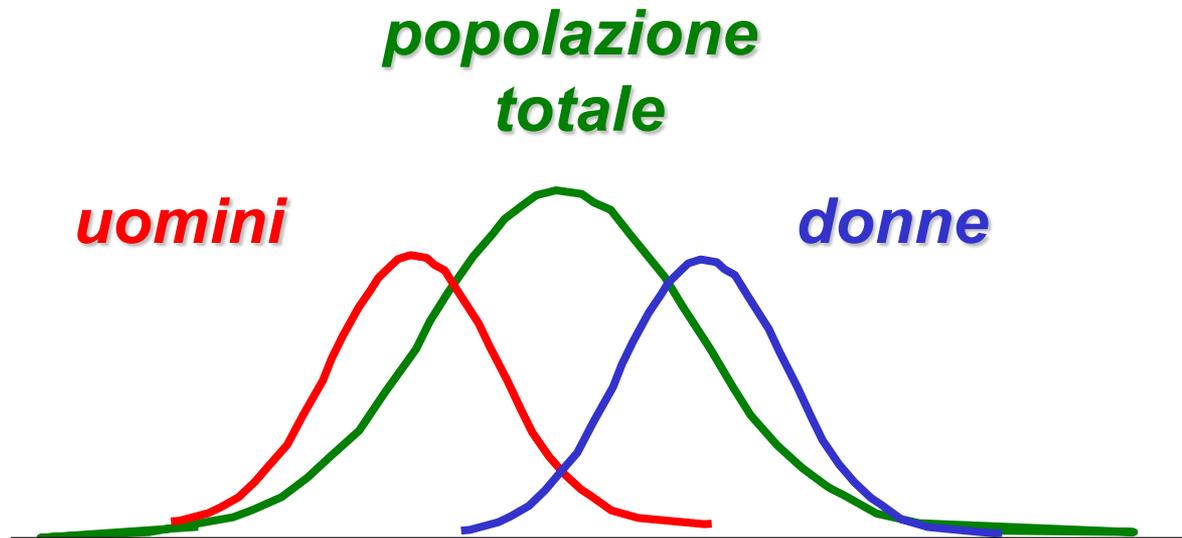
rosso - campionamento casuale

Verde - “bias” nel campionamento



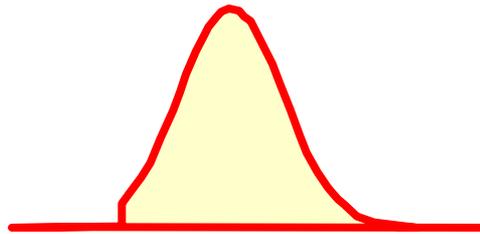
Insiemi di Dati più Piccoli

- In questo esempio, si introduce un “bias” semplicemente valutando separatamente gli uomini dalle donne.

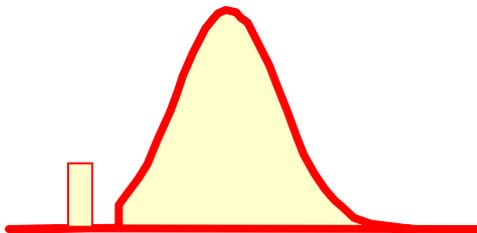




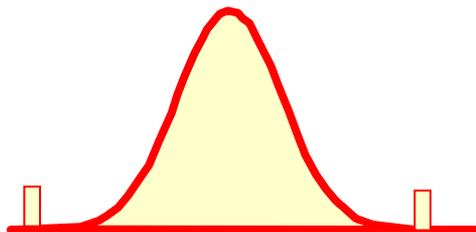
Altri Tipi di “bias”



“bias” a limite di rivelazione



“bias” a valore nominale



spalle



- **L'introduzione di un bias non è sempre un fatto negativo.**
- **Il bias può esser dovuto ad una vera differenza fra le popolazioni.**
- **Può anche essere dovuto ad un cattivo campionamento di una popolazione.**
- **Occorrono strumenti per valutare la differenza.**



Per una distribuzione normale :

Media - media numerica dei valori

Mediana - tendenza centrale
- valore del centro di dati ordinati
- valore reale se N è dispari o
media di due valori centrati se
è pari.

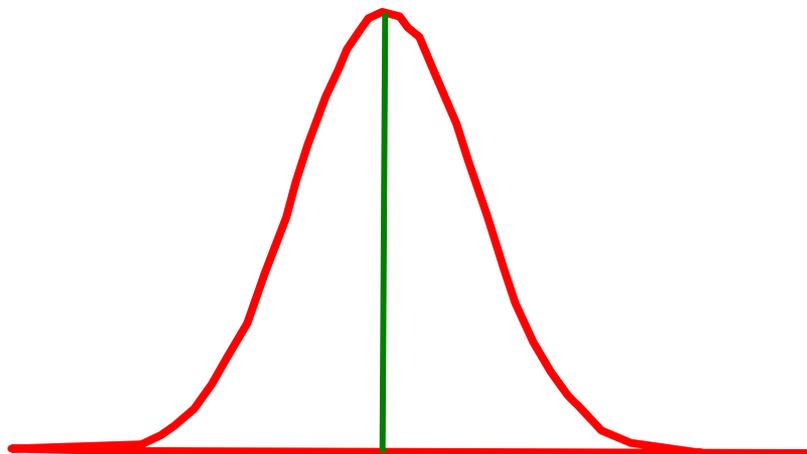
Moda - il valore più frequente.



Idealmente, i tre valori dovrebbero coincidere.

Se non sono uguali, dovrebbero essere molto vicini.

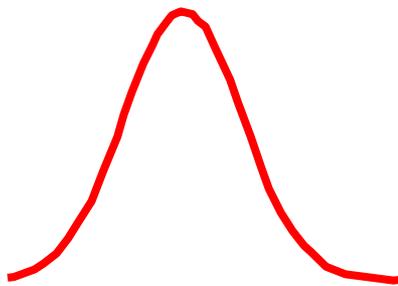
media, mediana e moda



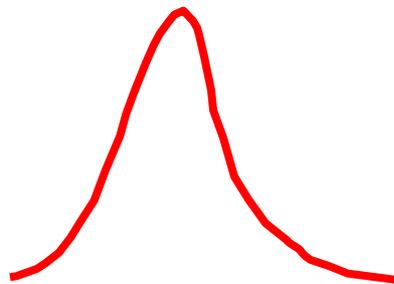


E' un test per vedere se una popolazione è Gaussiana.

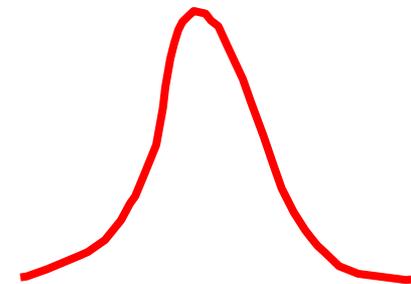
$$g = \frac{\sum_{i=1}^N (x - \bar{x})^3}{N \cdot \sigma_x^3}$$



$g = 0$



$g < 0$



$g > 0$



Per piccoli insiemi di dati, si usano le espressioni:

Varianza =
$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}$$

Deviazione Standard =
$$\sigma = \sqrt{\sigma^2}$$

Gradi di libertà = $df = n - 1$ (tipicamente)



Altri Termini (2)

Deviazione Standard della media

$$\sigma_{\bar{x}} = \frac{\sigma_x}{N}$$

Coefficiente di variazione

$$cv = \frac{\sigma_x}{\bar{x}}$$

Deviazione standard relativa

$$RSD = \frac{100\sigma_x}{\bar{x}}$$



$df = n - \#$ di parametri richiesti.

Esempio.

Se si hanno 10 misure, si possono usare a coppie per ottenere 9 diverse misure della media.

Se si tenta di farne una decima, una delle coppie sarà già stata usata - la deviazione standard è cioè la stessa.



Gradi di Libertà

- **Quando si esegue una regressione lineare di una linea usando una coppia di dati X, Y , il modello usato ($Y = mX + b$) impiega due parametri (m e b).**
- **I gradi di libertà saranno in questo caso (o modello) $N-2$.**
- **Pertanto il modello usato determina i gradi di libertà.**



In molti casi è necessario combinare i risultati:

- **Valori da laboratori diversi**
- **Dati raccolti in giorni diversi**
- **Uso di uno strumento differente**
- **Uso di un metodo di analisi differente**

L'operazione di combinazione dei dati, viene riferita come “accorpamento” dei dati.



Statistiche di Accorpamento (2)

- **Non si può combinare semplicemente tutti i valori e calcolarne la media e gli altri valori statistici.**
- **Ne possono risultare differenze con i risultati ottenuti.**
- **Il numero dei campioni raccolti potrebbe essere differente per ciascun insieme.**
- **E' anche auspicabile individuare un modo per valutare se i risultati sono significativamente diversi.**



$$\sigma_p = \sqrt{\frac{\sigma_1^2 \cdot x \cdot df_1 + \sigma_2^2 \cdot x \cdot df_2 + \dots + \sigma_k^2 \cdot x \cdot df_k}{df_1 + df_2 + \dots + df_k}}$$

Perciò la deviazione standard accorpata è pesata per i gradi di libertà. Ciò tiene conto del numero dei dati e di qualsiasi parametro usato nell'ottenere i risultati.



Statistiche di Accorpamento - Esempio

Insieme	σ	n	df	σ^2	df· σ^2
1	1.35	10	9	1.82	16.4
2	2.00	7	6	4.00	24.0
3	2.45	6	5	6.00	30.0
4	1.55	12	11	2.40	26.4

$$\sigma_p = \frac{16.4 + 24.0 + 30.0 + 26.4}{9 + 6 + 5 + 11} = 1.77$$



Semplice Analisi della Varianza

- **Finora, si è assunto che tutta la varianza osservata deriva da una sorgente singola, casuale.**
 - **non probabile**
 - **possono esistere molte fonti di varianza**
- **Si introdurrà una analisi della varianza in più insieme di campioni.**



Semplice Analisi della Varianza

- **In generale, quando le sorgenti della varianza sono linearmente indipendenti (non correlate), le varianze sono additive.**

$$\sigma_{totale}^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_k^2$$

Occorre spesso impostare esperimenti per valutare l'entità e le fonti di varianza.



Determinazione delle Fonti di Varianza

Partiamo da un semplice esempio in cui sono previste due sole sorgenti di varianza.

Modello semplice a due livelli:

σ_m^2 = varianza del componente del campione

σ_a^2 = varianza del metodo analitico

Con adatte sperimentazioni e analisi si possono individuare ed attribuire.



Semplice Modello a 2 Livelli

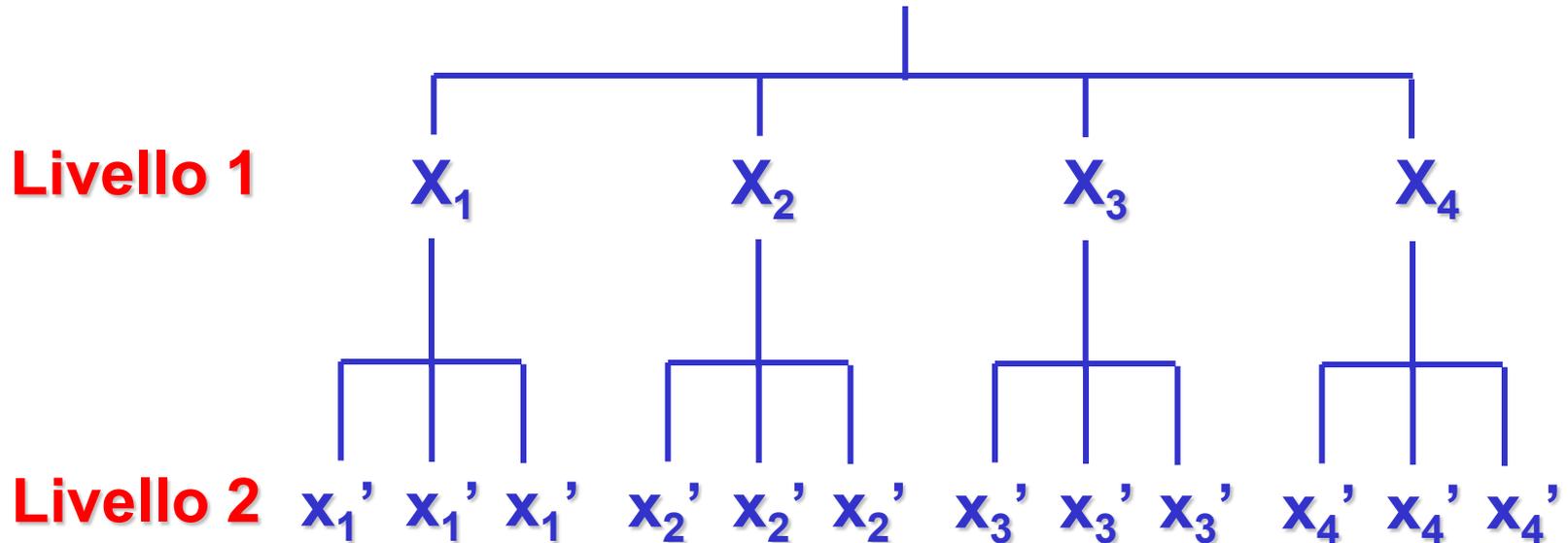
Campione	Repliche	Deviazione	Media
1	16.0, 16.1, 16.2	0.2, 0.0, -0.2	16.1
2	14.9, 15.1, 15.3	0.4, 0.0, -0.4	15.1
3	15.8, 15.8, 15.8	0.0, 0.0, 0.0	15.8
4	16.2, 16.0, 15.9	0.2, 0.0, -0.2	16.0

Questo insieme di dati è del tipo a due livelli.

- campione**
- replica**
- primo livello**
- secondo livello**



Semplice Modello a Due Livelli (2)



Il livello 1 dà una idea della variabilità del metodo

Il livello 2 ci dice qualcosa sulla variabilità del campione



Determinazione della Varianza

Prima - calcolare σ_T

$$\sigma_T = \frac{\sum (x_i - x_T)^2}{N_{\text{campioni}} - 1} = 0.381$$

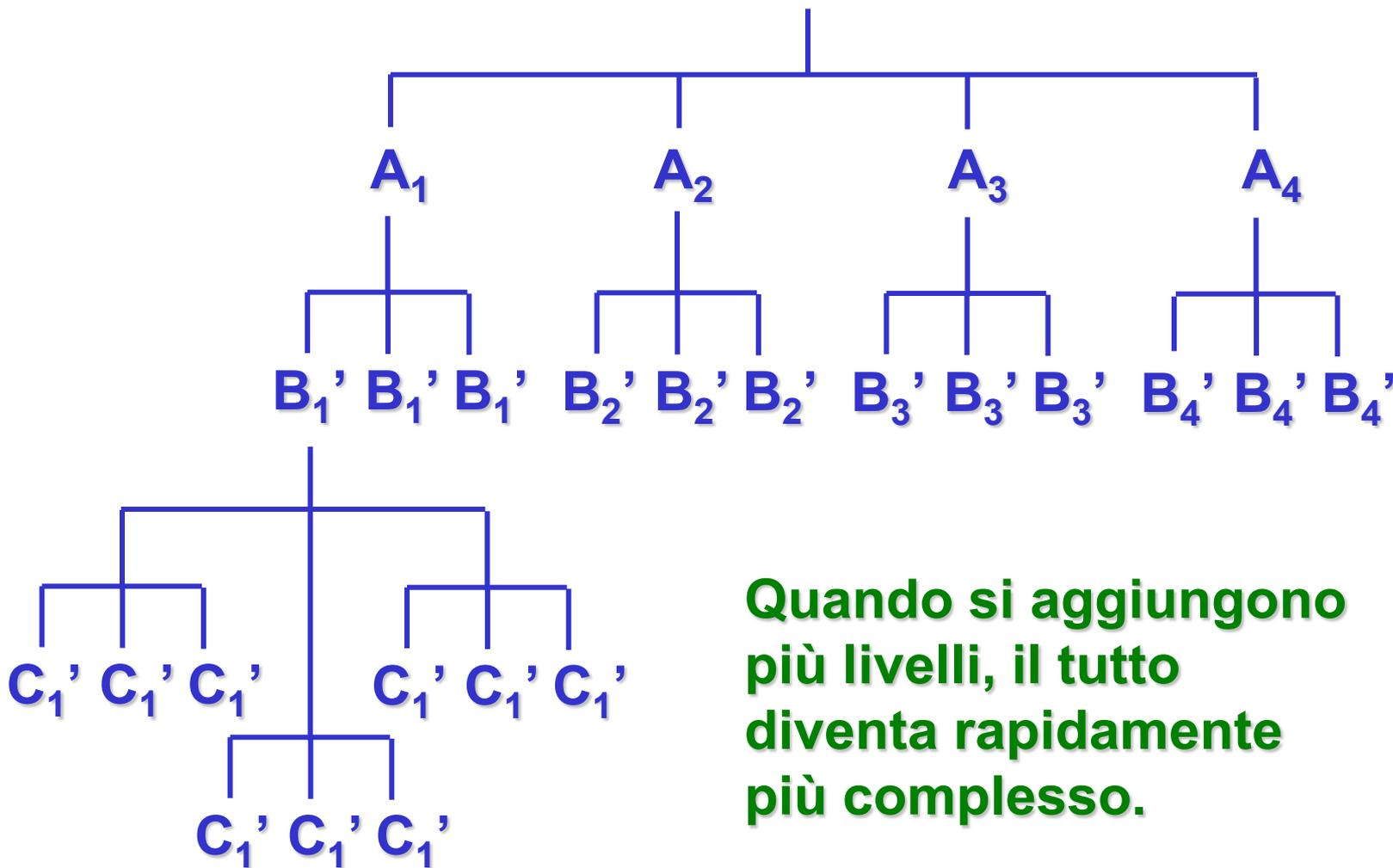
Quindi - calcolare σ_M^2

$$\sigma_M^2 = \frac{\sum d^2}{N_{\text{misura}} - N_{\text{campioni}}} = 0.218$$

Infine $\sigma_a^2 = \sigma_T^2 - \sigma_M^2 = 0.163$



Aggiunta di Più Livelli



Quando si aggiungono più livelli, il tutto diventa rapidamente più complesso.



Aggiunta di Più Livelli (2)

- **Quando si trattano molte variabili, la progettazione diventa molto importante.**
- **Progettazione Fattoriale**

Un metodo per predisporre esperimenti

Un'opportuna progettazione aiuta a minimizzare il numero di esperimenti da condursi per ottenere risultati accettabili.



- Qual è la possibilità che il valore “vero” (μ) cada in un certo intervallo attorno alla media (\bar{x})?

$$\mu = \bar{x} \pm \frac{t \cdot \sigma}{\sqrt{n}}$$



- ◆ “ t ” è detta “**t di Student**” ed è un valore derivato per via statistica
- ◆ il valore di “ t ” è dipendente dal numero di misure e dal livello di confidenza desiderato
 - ◆ Aumentando il livello di confidenza cresce l’intervallo dei valori
 - ◆ Aumentando il numero di misure diminuisce l’intervallo dei valori



- ◆ Tabella “ t student”
- ◆ Base di dati (12.6, 11.9, 13.0, 12.7, 12.5)
- ◆ media = 12.5, dev. std. = 0.4
- ◆ C'è un 50% di possibilità che il valore vero stia in quell'intervallo?



Intervalli di Confidenza

◆ $t_{50\%} = 0.741$

$$\mu = 12.5 \pm \frac{(0.741)(0.4)}{\sqrt{5}}$$

◆ $12.4 < \mu < 12.6$

◆ al 90% di confidenza ($t_{90\%} = 2.132$)

$12.1 < \mu < 12.9$



Intervallo di Confidenza della Media

- **Ci dice dove dovrebbe cadere la maggior parte dei dati.**
- **Un calcolo molto comune per riportare la variabilità dei dati.**
- **Un modo rapido per identificare valori singolari.**



Intervallo di Confidenza della Media

Per ampi campioni di dati

$$CI = \mu \pm \frac{Z \cdot \sigma}{\sqrt{N}}, \quad Z = \text{fattore di probabilità}$$

<i>CI</i>	2 volte Z
90	1.645
95	1.960
99	2.575
99.99995	5.000

Questo approccio assume che N è un valore molto grande.

Z deriva dalla riga infinita della tabella t



Intervallo di Confidenza della Media

- Talvolta si raccoglie un insieme quasi infinito di dati. E' più comune però lavorare con insiemi più piccoli.
- In questi casi si può far riferimento al test t .

$$CI = \bar{x} + t (\sigma_x / N^{1/2})$$

Deviazione Standard
della media



I valori t tengono conto degli errori introdotti a seguito della dimensione del campione, i gradi di libertà e la potenziale scodatura del campione.

Si usa in pratica la distribuzione χ^2 - chi quadro.

$$\chi_{n-1}^2 = \frac{(n-1)\sigma^2}{s^2}$$

Essa permette di stimare la varianza della popolazione dalla varianza del campione.

Tutto ciò è fuso assieme nei valori di t .



Valori di t

Gradi di libertà	livello di confidenza		
	90%	95%	99%
1	6.31	12.7	63.7
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	2.78	4.60
5	2.02	2.57	4.03
6	1.94	2.45	3.71
7	1.90	2.36	3.50
8	1.86	2.31	3.36
9	1.83	2.26	3.25
10	1.81	2.23	3.17



Esempio

Dati; 1.0, 1.02, 1.10, 0.95, 1.00

$$\begin{aligned} \text{media} &= 1.016 \\ \sigma_x &= 0.0541 \\ \sigma_{\bar{x}} &= 0.0242 \end{aligned}$$

Valori di t per 4 gradi di libertà

95 % di confidenza = 2.13

99 % di confidenza = 2.78



Esempio

$$CI_{95\%} = 1.016 \pm \frac{2.13 \times 0.0541}{5^{1/2}}$$
$$= 1.02 \pm 0.05 \quad (\pm 5\%)$$

$$CI_{99\%} = 1.016 \pm \frac{2.78 \times 0.0541}{5^{1/2}}$$
$$= 1.02 \pm 0.07 \quad (\pm 7\%)$$



Quando due Medie Differiscono Significativamente

- **Se si hanno due insiemi di numeri**
 - **da campioni differenti**
 - **da prelievi differenti dello stesso campione**
- **Sono in realtà diversi.**
- **Se le medie sono identiche, è più probabile che si tratti di un caso che di altro.**
- **Si possono eseguire dei test per vedere se i risultati (identici o no) sono gli stessi o meno.**



Quando due Medie Differiscono Significativamente

- **Per verificare se le due medie in realtà differiscono, si deve prima calcolare la media e la deviazione standard per ogni campione.**
- **Quindi si deve valutare in base a due possibili casi:**
 - **Caso 1 - A e B non differiscono significativamente**
 - **Caso 2 - A e B differiscono significativamente.**
- **L'assunto fatto influenza come si affronta il problema.**



Fasi da seguire

1. Accorpare i valori per ottenere σ_p
2. Calcolare la varianza della media per entrambi gli insiemi.

$$V_A = \sigma_p^2 / n_A, \quad V_B = \sigma_p^2 / n_B \quad \bar{x}_A > \bar{x}_B$$

3. Scegliere un livello di probabilità per la decisione (95%, 99%, ...).



Caso 1

4. Calcolare l'incertezza della differenza fra le due medie.

$$V_{\Delta} = t \sqrt{(V_A + V_B)}$$

Valutare t in base al livello di confidenza e ai gradi di libertà desiderati.

$$df = n_A + n_B - 2$$

Se $|x_A - x_B| > V_{\Delta}$, allora le medie differiscono.



Fasi da seguire

1. Calcolare la varianza della media per entrambi gli insiemi.

$$V_A = \sigma_A^2/n_A; \quad V_B = \sigma_B^2/n_B \quad x_A > x_B$$

2. Calcolare il numero effettivo di gradi di libertà. Non si può sommare semplicemente df_A e df_B se gli insiemi sono nettamente differenti.

$$df = \frac{(V_A + V_B)^2}{\left(\frac{V_A^2}{n_A - 1}\right) + \left(\frac{V_B^2}{n_B - 1}\right)}$$



4. Calcolare l'incertezza della differenza fra le due medie.

$$V_{\Delta} = t (V_A + V_B)$$

Valutare t in base al livello di confidenza e ai gradi di libertà desiderati.

$$df = n_A + n_B - 2$$

Se $|x_A - x_B| > V_{\Delta}$, allora le medie differiscono.



Caso 1 vs. Caso 2

Entrambi i metodi forniscono essenzialmente gli stessi risultati.

Verifichiamolo con un esempio:

A - media = 50 mg/l, $\sigma = 2.0$ mg/l, n = 5

B - media = 45 mg/l, $\sigma = 1.5$ mg/l, n = 6

$$|x_A - x_B| = 5 \text{ mg/l} = \Delta x$$



Esempio, Caso 1

$$1. \quad \sigma_p = \sqrt{\frac{2^2 \times 4 + 1.5^2 \times 5}{4 + 5}} = 1.74$$

$$2. \quad V_A = 1.74^2 / 5 = 0.6058$$
$$V_B = 1.74^2 / 6 = 0.5047$$

3. Si usi un limite di confidenza del 95% ($t = 2.262$)

$$4. \quad V_D = 2.262 \sqrt{0.6056 + 0.5047} = 2.38$$

$\Delta x > V_\Delta$ ($5 > 2.38$) pertanto le medie sono differenti ad un livello di confidenza del 95%.



Esempio, Caso 2

1. $V_A = 2^2 / 5 = 0.800$, $V_A = 1.5^2 / 6 = 0.375$

2. Usando ancora un limite di confidenza del 95%

3.
$$df = \frac{(0.8 + 0.375)^2}{\frac{0.8^2}{4} + \frac{0.375^2}{5}} = 7.34 \cong 7$$

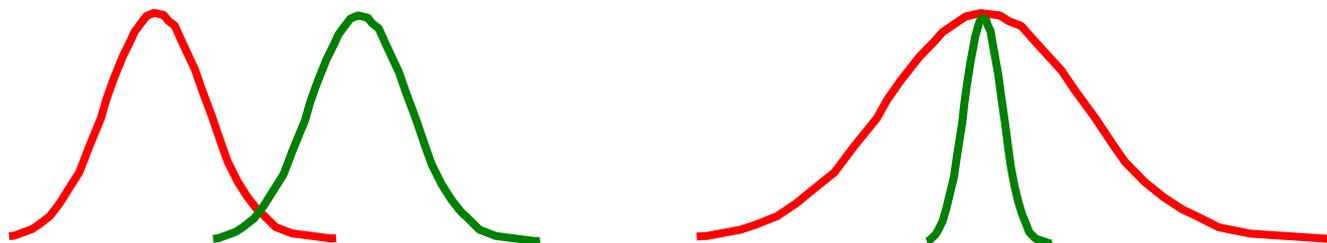
4. $V_{\Delta} = 2.365 \sqrt{0.800 + 3.250} = 2.6$

$$\Delta x > V_{\Delta}$$

Si trova di nuovo che le medie sono differenti.



- **Si possono avere campioni che si considerano significativamente differenti pur avendo la stessa media.**



In entrambi gli esempi, le popolazioni si devono considerare differenti - benché le medie, mediane e moda siano identiche nell'esempio di destra.



Questo test si può usare per verificare se due popolazioni sono differenti in base alle variazioni della varianza.

Esempi

- ◆ **La precisione della misura è variata?**
- ◆ **Il metodo è stato alterato?**
- ◆ **Esistono variazioni significative connesse al laboratorio o all'analista?**



Il Test F (2)

Calcolo di F:

$$F = \frac{\sigma^2_{\text{maggiore}}}{\sigma^2_{\text{minore}}}$$

F è sempre 1 o superiore e dipende dal livello di confidenza e dai gradi di libertà.

Si può dedurre il valore F_c per i livelli desiderati.



Ritornando all'esempio precedente

A - media = 50 mg/l, $\sigma = 2.0$ mg/l, n = 5

B - media = 45 mg/l, $\sigma = 1.5$ mg/l, n = 6

$$F = 2^2 / 1.5^2 = 1.78$$

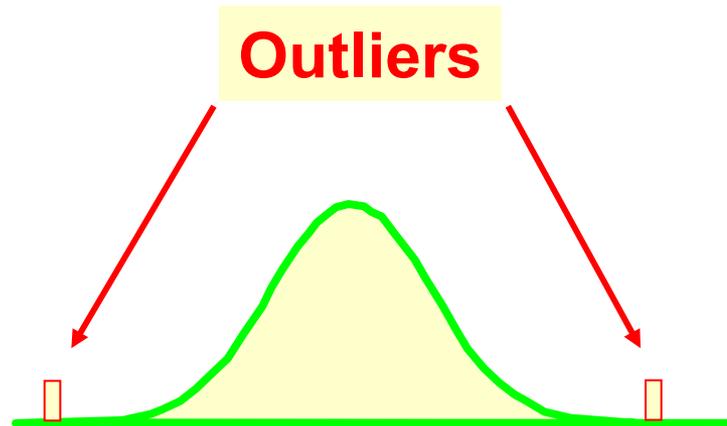
F_c è 7.39 al 95% di confidenza e 9 df.

Perciò, i valori di varianza sono essenzialmente gli stessi e le medie devono realmente differire.



Non Accettazione di Dati

- Talvolta si sa che un valore dei dati sembra brutto (outliers). Non si può scartarlo subito - per rifiutare un dato deve esistere una base.





Regola dell'errore enorme

test di Dixon - test Q

test di Grubbs

Ognuno ha i suoi vantaggi / svantaggi.



Regola dell'Errore Enorme

- **Assume che si abbia una certa idea di quanto dovrebbe essere il valore della deviazione standard o i possa calcolarla.**

$$M = \frac{|\text{sospetto} - \text{media}|}{S}$$

Se $M > 4$ il punto può essere scartato.

Ciò corrisponde ad un brutale test t. Si usa solo per eliminare dati ovviamente “brutti”.



Assume

- \bar{x} e σ_x sono sconosciute.
- I dati sono distribuiti normalmente.

Fasi

1. Ordinare i dati: $x_1 < x_2 < \dots < x_n$
2. Scegliere un livello di confidenza
3. Calcolare il rapporto (in base a n)
4. Valutare un valore adatto
5. Se il rapporto è $>$ del valore in tabella, scartarlo



Test di Dixon - Rapporti

- Usare il rapporto τ_{10} per 3 - 7 dati.

$$\tau_{10} = \frac{X_n - X_{n-1}}{X_n - X_1} \quad \text{o} \quad \frac{X_2 - X_1}{X_n - X_1}$$

*per verificare
valori alti*

*per verificare
valori bassi*

Due versioni basate sui suoi valori più alti e più bassi sottoposti al test.



Test di Dixon - Rapporti

- Per 8 - 10 punti, usare τ_{11}

$$\frac{X_n - X_{n-1}}{X_n - X_2}$$

$$\frac{X_2 - X_1}{X_{n-1} - X_1}$$

- Per 11 - 13 punti, usare τ_{21}

$$\frac{X_n - X_{n-2}}{X_n - X_2}$$

$$\frac{X_3 - X_1}{X_{n-1} - X_1}$$

- Per 14 - 25 punti, usare τ_{22}

$$\frac{X_n - X_{n-2}}{X_n - X_3}$$

$$\frac{X_3 - X_1}{X_{n-2} - X_1}$$



Test di Dixon - Tabella Parziale

Statistica	Rischio di errata eliminazione				
	n	0.5%	1%	5%	10%
t_{10}	3	.994	.988	.941	.886
	4	.926	.889	.765	.679
	5	.821	.780	.642	.557
	6	.740	.698	.560	.482
	7	.680	.637	.507	.434
t_{11}	8	.725	.683		
	9	.677	.635		
	10	.639	.579		
t_{21}	11	.713	.642		
	12	.675	.615		
	13	.649			
t_{22}	14	.674			
	15	.674			



Test di Grubbs

- Questo approccio richiede il calcolo della media e della deviazione standard.
 1. Ordinare i punti
 2. Selezionare i punti sospetti
 3. Calcolare la media e la deviazione standard usando tutti i punti.
 4. Calcolare T. $T = | \text{media} - \text{sospetti} | / \sigma_x$
 5. Valutare T sulla tabella.

Se T è > del valore in tabella scartarlo.



Test di Grubbs - Tabella Parziale

n	Rischio di errata eliminazione		
	0.1%	1%	5%
3	1.155	1.155	1.153
4	1.496	1.492	1.463
5	1.780	1.749	1.672
6	2.011	1.944	1.822
7	2.201	1.097	1.938
8	2.358	2.221	2.032
9	2.492	2.323	2.110
10	2.606	2.410	2.176