

# Trattamento Statistico dei dati (Richiami e Approfondimenti)

**Fernando Palombo**

<http://www.mi.infn.it/~palombo/didattica/Lab-OFM/Lezione2.pdf>

# Stima Puntuale dei Parametri

- ❑ Supponiamo di aver fatto  $n$  misure di una variabile casuale  $X$ :  $x_1, x_2, \dots, x_n$ .  
Noi assumiamo che le  $n$  misure siano **indipendenti** tra di loro e **identicamente distribuite** (i.i.d.).
- ❑ Noi vogliamo determinare la funzione densità di probabilità (p.d.f.) della variabile casuale  $X$  a partire dalle  $n$  misure fatte (il nostro campione di dati). Questa p.d.f. in generale non è conosciuta.
- ❑ Spesso è noto a priori che la p.d.f. della variabile studiata appartiene ad una famiglia parametrica (per esempio una gaussiana) con uno o più parametri liberi  $\theta$

$$f(x) = f(x | \theta) \theta \in \Theta$$

- ❑ A partire dalle  $n$  misure fatte noi vogliamo determinare i parametri  $\theta$  (**inferenza statistica**). Per stimare questi parametri devo introdurre gli stimatori. Si dice fit di un parametro il processo che porta alla stima puntuale di un parametro.

# Stimatori Puntuali

- ❑ Si chiama statistica **qualunque** funzione  $t$  delle  $n$  misure effettuate:

$$t = t(x_1, x_2, \dots, x_n)$$

- ❑ Una statistica usata per stimare il valore puntuale di un parametro si dice stimatore puntuale del parametro.
- ❑ Uno stimatore è “buono” se è: **consistente**, **non distorto** ed **efficiente**
- ❑ **Consistente** vuol dire che all’aumentare della statistica, il valore stimato tende al valore vero
- ❑ Si dice **distorsione** (detta anche bias) la differenza tra il valore stimato ed il valore vero del parametro. Se il bias è nullo la misura si dice non distorta.
- ❑ Tra i vari stimatori si dice **efficiente** quello che stima il parametro con la minore varianza

# Stimatori della Media

- ❑ Esistono diversi estimatori di media (aritmetica, geometrica, armonica, ...).
- ❑ Lo stimatore **media aritmetica** è “buono” ed quello **più usato**.
- ❑ Se la distribuzione delle misure è simmetrica, allora un altro stimatore della media consistente e non distorto è la mediana.
- ❑ La varianza sulla media aritmetica  $\mu$  è minore della varianza sulla mediana  $x_{\text{med}}$  :  $\sigma_{\mu}^2 = \sigma^2/n$  mentre  $\sigma_{\text{med}}^2 = \sigma^2/n \cdot \pi/2$  (quindi tra i due lo stimatore media aritmetica è migliore).
- ❑ Però la mediana è più **robusta** della media aritmetica (cioè risente meno delle misure nelle code della distribuzione).

# Stimatori della Varianza

- ❑ Per la stima della varianza  $\sigma^2$  di un campione di  $n$  misure  $x_1, \dots, x_n$  (**i.i.d.**) di una variabile casuale  $X$ , si possono avere due situazioni:
  - è noto il valore vero  $\mu$  della variabile  $X$  (fatto non usuale)
  - NON è noto il valore vero.

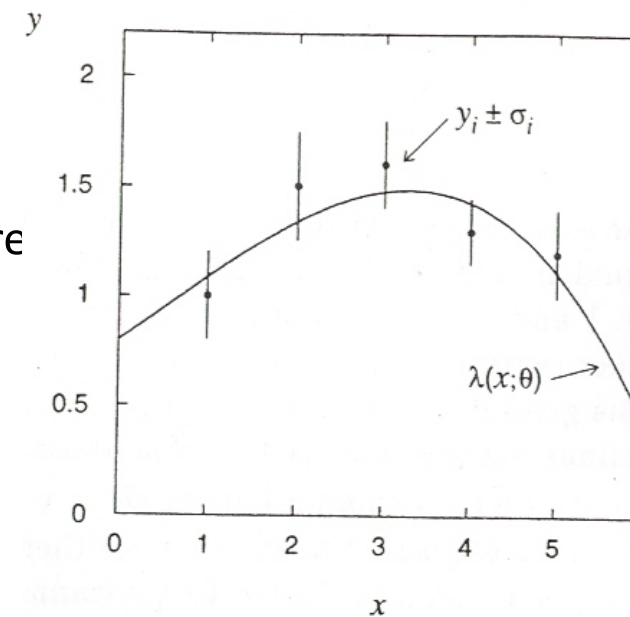
- ❑ Nel primo caso stimo il valore della varianza così:  $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$   
Questo stimatore è consistente e non distorto.

- ❑ Nel secondo caso stimo come valore vero la media aritmetica  $\bar{x}$  dei dati  $x_i$ .  
In questo caso lo stimatore è distorto. Infatti non ci sono più  **$n$**  dof ma  **$n-1$**   
(un grado di libertà è perso perché la media è estratta dai dati).

- ❑ Questa distorsione può essere corretta così:  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$   
 $s^2$  è la **varianza del campione (varianza campionaria)**.

# Stimatori di Minimi Quadrati

- ❑ Supponiamo di misurare due variabili casuali  $X$  e  $Y$  : ad ogni valore di  $X$  misuro il valore di  $Y$ . Per esempio negli istanti  $x_1, x_2, \dots, x_n$  misuro la posizione  $y_1, y_2, \dots, y_n$ . Ognuna di queste misure avrà una propria deviazione standard  $\sigma_i$
- ❑ Supponiamo di conoscere la relazione funzionale  $\lambda(x; \theta)$  che per ogni  $x$  mi permette di determinare il corrispondente valore di  $y$ .
- ❑ La funzione  $\lambda$  contiene un parametro (o più parametri) che devo stimare a partire dalle misure sperimentali.
- ❑ Ad ogni misura  $x_i$  associamo un valore misurato  $y_i$  ed un valore stimato  $\lambda(x_i; \theta)$ .
- ❑ La differenza  $y_i - \lambda(x_i; \theta)$  è detta residuo. Sommiamo i quadrati dei residui di tutte le misure pesati con l'inverso della deviazione standard  $\sigma_i$



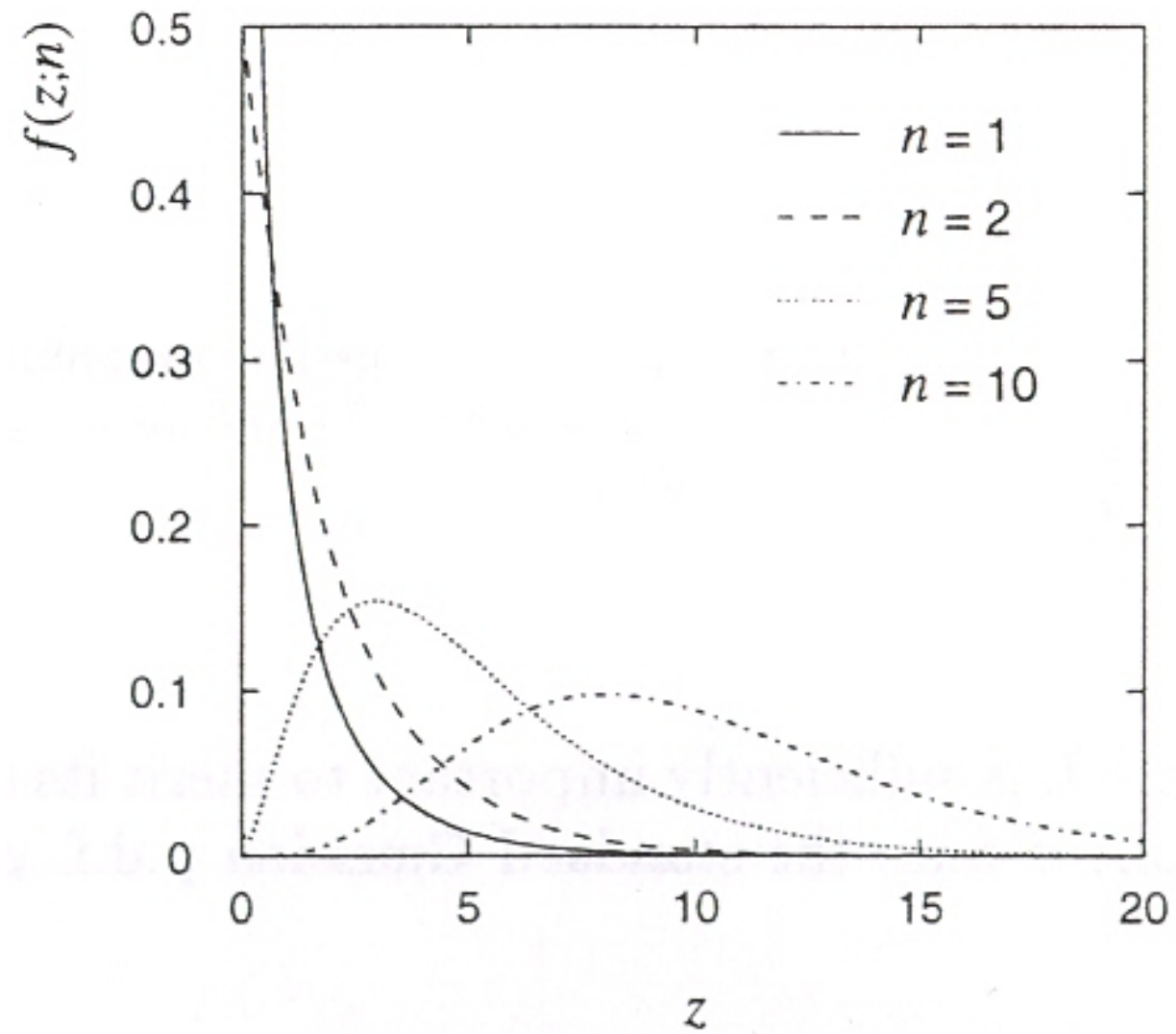
# Stimatori di Minimi Quadrati

- ❑ Questa somma è chiamata  $\chi^2$

$$\chi^2 = \sum_{i=1}^n \left[ \frac{y_i - \lambda(x_i; \theta)}{\sigma_i} \right]^2$$

- ❑ Questa funzione segue infatti la distribuzione del  $\chi^2$  con un numero di gradi di libertà uguale al numero  $n$  di misure meno il numero  $m$  di parametri liberi nel fit (dof =  $n - m$ ).
- ❑ Il parametro incognito  $\theta$  da stimare è il valore che minimizza questa funzione. Questo stimatore è detto dei **minimi quadrati**.
- ❑ Questo stimatore ha caratteristiche ottimali solo in presenza di modelli lineari -> regressione lineare.
- ❑ Questo stimatore è molto usato anche in presenza di funzioni non lineari.

## Distribuzione $\chi^2$



La media di una variabile  $Z$  distribuita secondo il  $\chi^2$  è uguale al numero di gradi di libertà

## Combinazione di Più Misure

- ❑ Si abbiano n misure indipendenti della stessa variabile casuale Y:  $y_1 \pm \sigma_1, y_2 \pm \sigma_2, \dots, y_n \pm \sigma_n$  e sia  $\lambda$  il valore vero aspettato.

- ❑ Mi calcolo la quantità : 
$$\chi^2(\lambda) = \sum_{i=1}^n \frac{(y_i - \lambda)^2}{\sigma_i^2}$$

- ❑ Se azzero la derivata prima rispetto a  $\lambda$  e risolvo per  $\lambda$ , ottengo: 
$$\lambda = \frac{\sum_{i=1}^n y_i / \sigma_i^2}{\sum_{i=1}^n 1 / \sigma_i^2}$$

- ❑ La **media pesata** si ottiene pesando le misure con le loro varianze.

- ❑ Passando alle derivate seconde si ottiene la varianza della media pesata:

$$V[\lambda] = \frac{1}{\sum_{i=1}^n 1 / \sigma_i^2}$$

- ❑ Se le misure che si stanno combinando hanno lo stesso errore, allora la media pesata coincide con la media aritmetica.

## Esempio: Fit lineare nella misura di e/m

- Misura del rapporto e/m. Si misura il campo magnetico B ed il raggio R della traiettoria degli elettroni per una determinata differenza di potenziale  $\Delta V$ .
- La funzione lineare da fittare è  $(BR)^2 = 2 m/e \Delta V$  che possiamo scrivere così:  
 **$Y = a X$**  con  $X = 2 \Delta V$ ,  $Y = B^2 R^2$  e  $a$  (parametro da stimare) =  $m/e$
- Tabella di misure (di un gruppo di studenti in laboratorio):

N	$x = 2\Delta V$ (V)	$\sigma_{2\Delta V}$ (V)	$y = B^2 R^2 (T^2 m^2)$	$\sigma_y (T^2 m^2)$
1	410.6	0.2	$1.97296 \cdot 10^{-09}$	$1.9735 \cdot 10^{-10}$
2	647.8	0.2	$3.19070 \cdot 10^{-09}$	$3.1915 \cdot 10^{-10}$
3	576.0	0.2	$2.88407 \cdot 10^{-09}$	$2.8848 \cdot 10^{-10}$
4	535.2	0.2	$2.57291 \cdot 10^{-09}$	$2.5736 \cdot 10^{-10}$
5	748.6	0.2	$3.69151 \cdot 10^{-09}$	$3.6924 \cdot 10^{-10}$
6	795.6	0.2	$3.94540 \cdot 10^{-09}$	$3.9463 \cdot 10^{-10}$
7	548.8	0.2	$2.82670 \cdot 10^{-09}$	$2.8274 \cdot 10^{-10}$
8	468.4	0.2	$2.29469 \cdot 10^{-09}$	$2.2953 \cdot 10^{-10}$
9	707.0	0.2	$3.41219 \cdot 10^{-09}$	$3.4130 \cdot 10^{-10}$
10	501.0	0.2	$2.73823 \cdot 10^{-09}$	$2.7389 \cdot 10^{-10}$
11	676.8	0.2	$3.41176 \cdot 10^{-09}$	$3.4126 \cdot 10^{-10}$
12	618.6	0.2	$3.28271 \cdot 10^{-09}$	$3.2835 \cdot 10^{-10}$
13	718.6	0.2	$4.20530 \cdot 10^{-09}$	$4.2063 \cdot 10^{-10}$

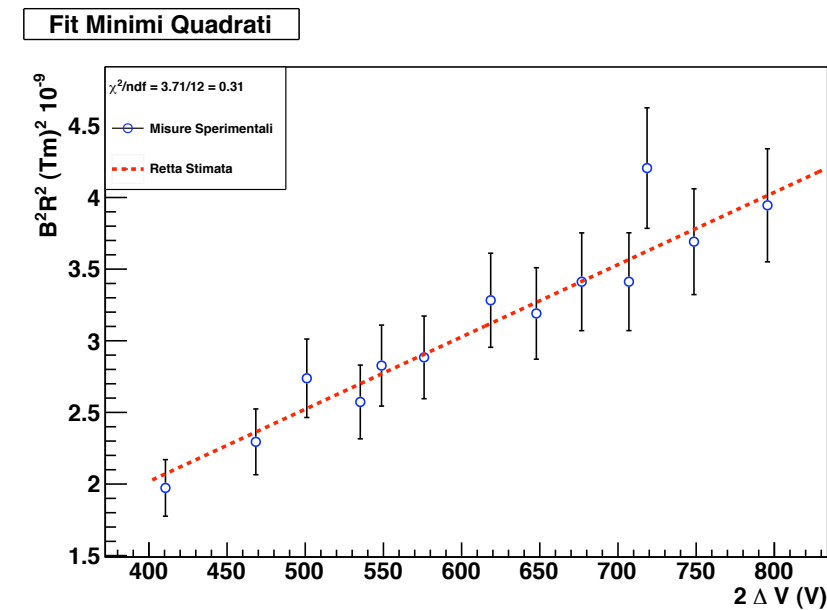
# Fit lineare nella misura di e/m

- ❑  $m/e = (5.1 \pm 1.4) 10^{-12} \text{ Kg/C}$   
invertendo e propagando gli errori si ha:

$$e/m = (2.0 \pm 0.5) 10^{11} \text{ C/Kg}$$

- ❑ Sono soddisfatto di questo fit ? È buono?

- ❑ Il primo controllo che faccio è di tipo visivo. Controllo se la curva stimata si adatta bene ai dati sperimentali



- ❑ La quantità che ho definito  $\chi^2$  ha come valore di aspettazione il numero di gradi di libertà (dof). Quindi mi aspetto come  $\chi^2$  ridotto ( $\chi^2/\text{dof}$ ) un numero circa uguale ad uno.
- ❑ Se il  $\chi^2$  ridotto è parecchio più piccolo di 1 (come nel nostro caso 0.31) il risultato del fit è addirittura migliore di quanto mi sarei aspettato (dati gli errori di misura). Comunque devo controllare che gli errori non siano sovrastimati o che non siano correlati
- ❑ Se il  $\chi^2$  ridotto fosse stato molto maggiore di 1, allora ci sarebbero stati buoni motivi per mettere in dubbio la scelta del modello di fit lineare. **Attenzione:** un buon modello può essere messo in dubbio da errori sottostimati.

## Esempio per Incertezze Sistematiche

- ❑ Come esempio di stima dell'incertezza sistematica consideriamo la misura della velocità della luce.

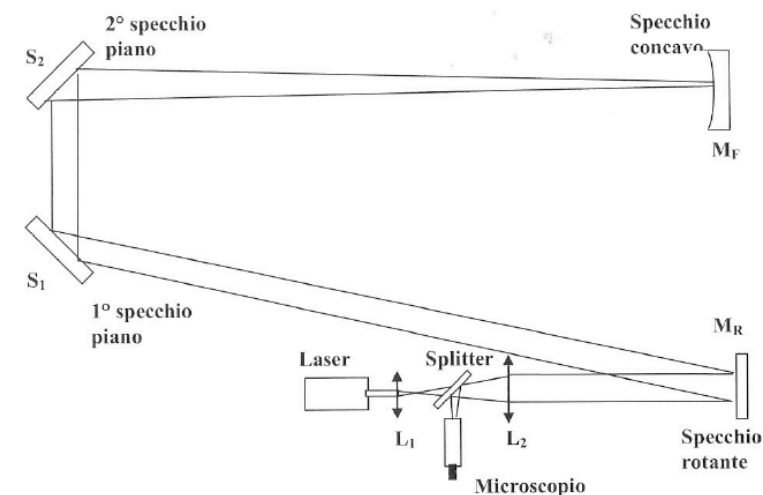
- ❑ La velocità della luce viene misurata a partire dalla formula

$$c = \frac{4f_2 D^2 \Delta\omega}{(D + a - f_2) \Delta\delta}$$

- ❑ D è pari a 12.47 m ed è somma di tre segmenti ognuno misurato con incertezza  $\sigma_x = 0.010$  m  $\rightarrow \sigma_D = 0.017$  m  $\approx 0.02$  m

- ❑ La misura di a è  $(0.452 \pm 0.002)$  m mentre  $f_2$  è quantità specifica del rivelatore e il suo valore è fornito dal costruttore:  $f_2 = (0.252 \pm 0.001)$  m

- ❑ Nella misura di c si usano i valori nominali di D ( 12.47 m), di a (0.452 m) e di  $f_2$  (0.252 m). Un gruppo di studenti ha ottenuto come media delle misure della velocità fatte in laboratorio:  **$c_0 = (2.996 \pm 0.009) 10^8$  m/s**



## Esempio per Incertezze Sistematiche

$$\square \quad c_0 = \frac{4f_2 D^2}{D + a - f_2} \left( \frac{\Delta\omega}{\Delta\delta} \right)_{medio} = (2.996 \pm 0.009) \cdot 10^8 \text{ m/s}$$

dove  $0.009 \cdot 10^8$  è l'incertezza statistica sul valore stimato  $c_0$ .

- Le incertezze su  $D$ ,  $a$  e  $f_2$  determinano una incertezza (sistematica)  $\sigma_{sist}$  su  $c_0$ . Questa incertezza  $\sigma_{sist}$  può essere stimata propagando su  $c_0$  le incertezze dei valori nominali di  $D$ ,  $a$  e  $f_2$  usati nel calcolo della velocità :

$$\sigma_{sist} = \sqrt{\left( \frac{\partial c_0}{\partial D} \right)^2 \sigma_D^2 + \left( \frac{\partial c_0}{\partial a} \right)^2 \sigma_a^2 + \left( \frac{\partial c_0}{\partial f_2} \right)^2 \sigma_{f_2}^2}$$

dove le derivate parziali vanno calcolate con i valori nominali di  $D$ ,  $a$  ed  $f_2$  mentre  $\sigma_D$ ,  $\sigma_a$ , e  $\sigma_{f_2}$  sono le incertezze con cui sono note  $D$ ,  $a$  e  $f_2$ .

- In questo esempio l'incertezza sistematica è pari a  $0.01 \cdot 10^8$  m/s (con due cifre decimali) e quindi la misura  $c_0$  è data così:  $c_0 = (3.00 \pm 0.01 \pm 0.01) \cdot 10^8$  m/s o anche, sommando in quadratura gli errori:  $c_0 = (3.00 \pm 0.01) \cdot 10^8$  m/s

## Esempio per Incertezze Sistematiche

- ❑ L'incertezza  $\sigma_{\text{sist}}$  nella misura di  $c$  originata dalle incertezze dei valori nominali di  $D$ ,  $a$  e  $f_2$  può essere anche stimata in questo modo:
- ❑ Una volta stimato il valore  $c_0$  dalle misure sperimentali con i valori nominali di  $D(12.45)$ ,  $a(0.452)$  e  $f_2(0.252)$ , ricalcolo i valori di  $c_0$  sostituendo al valore nominale di  $D$  il suo valore aumentato e diminuito di  $1\sigma$  (sempre con gli stessi valori nominali di  $a$  e  $f_2$ ). Ottengo così due valori di  $c_0$ :  $c_{01}$  e  $c_{02}$ . il modulo della differenza  $|c_{01} - c_{02}|$  lo prendo come valore dell'incertezza sistematica su  $c_0$  dovuta all'incertezza su  $D$ . Talvolta come incertezza sistematica si prende la massima incertezza tra  $c_{01}$  e  $c_{02}$
- ❑ In modo analogo stimo il contributo sistematico a  $c_0$  dovuto ad  $a$  e ad  $f_2$ . Sommo i tre contributi in quadratura e ottengo l'incertezza sistematica totale su  $c_0$ .
- ❑ Quando si ha qualche riserva sul grado di fiducia nelle incertezze di  $D$ ,  $a$  ed  $f_2$ , si può preferire allargare l'incertezza sistematica.

Va sempre descritto come è stata calcolata l'incertezza sistematica

# Verifica di Ipotesi

- ❑ La Verifica di Ipotesi (detta anche Teoria delle Decisioni) è l'altra parte fondamentale della Statistica Inferenziale.
- ❑ È molto comune avere problemi di verifica di ipotesi (non solo in Fisica). Per poter decidere su una ipotesi è da rigettare o scegliere quale tra due ipotesi è la più favorita dalle misure fatte devo fare un **test statistico**.
- ❑ Noi facciamo una certa assunzione che chiamiamo ipotesi. Tradizionalmente questa ipotesi è detta **ipotesi nulla  $H_0$** . In genere si fa anche una **ipotesi alternativa  $H_1$** . Il test statistico mi permette di valutare viste le misure fatte il livello di accordo dei miei dati con quanto aspettato sotto l'ipotesi nulla o di scegliere tra due ipotesi.
- ❑ Io mi riferirò qui a test statistici che considerano solo l'ipotesi nulla.

# Test di Bontà del Fit

- ❑ Si definisce **valore-P (P-value)** la probabilità, sotto l'ipotesi nulla  $H_0$  assunta, di avere un valore della statistica di test che corrisponde ad un risultato compatibile o meno compatibile di quanto effettivamente osservato nel fit sui dati. Il P-value talvolta è detto **livello di confidenza** (o **livello di significatività**)
- ❑ Nel nostro caso di test di bontà del fit come statistica di test usiamo il  $\chi^2$ .
- ❑ Risultati compatibili o meno compatibili di quanto ho trovato nel fit corrispondono a fit con  $\chi^2$  uguali o maggiori (peggiori) del  $\chi_0^2$  che ho trovato nei dati. Quindi il valore P si ottiene integrando la distribuzione del  $\chi^2$  con dof =  $n - m$  dal valore  $\chi_0^2$  a  $+\infty$  ( $n$  è il numero di misure nel fit e  $m$  è il numero di parametri liberi nel fit):

$$P = \int_{\chi_0^2}^{+\infty} f(z; n_{dof}) dz$$

- ❑ Questo significa che: se rifaccio molte volte lo stesso esperimento (nuovi campioni di dati con la stessa numerosità di dati e stesse condizioni), il P-value rappresenta la frazione di esperimenti in cui il fit ha un  $\chi^2$  uguale o maggiore (peggiore) di quello che ho effettivamente trovato nel mio campione di dati.<sup>16</sup>

## Esempio di Test di Bontà del Fit

- ❑ Nell'esempio di fit lineare visto nelle precedenti slide il fit ha un  $\chi^2$  di 3.71 con 12 dof ed il P-value vale 0.99.
- ❑ In questo esempio l'ipotesi nulla  $H_0$  è il modello di fit lineare .
- ❑ Se rifacessi molte volte il mio esperimento nel 99% dei casi otterrei un  $\chi^2$  uguale o maggiore del  $\chi_0^2$  che ho trovato nel mio fit. Quindi il mio fit è (particolarmente) buono !
- ❑ Si tenga presente che in statistica frequentista (classica) un test statistico può **solo rigettare l'ipotesi  $H_0$  nulla** con un certo P-value. Ma si ferma qui.
- ❑ **NON** si può dare una probabilità che  $H_0$  sia vera. Talvolta erroneamente il P-value è preso come probabilità che  $H_0$  sia vera.

# Compatibilità tra le Misure

- Con un campione di misure io posso stimare il valore di certe grandezze fisiche (velocità della luce, rapporto  $e/m$ , ecc ).
- La misura può essere fatta da un solo gruppo (esperimento) o da gruppi diversi (che usano campioni casuali diversi, indipendenti e anche con diversa numerosità).
- Mi aspetto che fluttuazioni statistiche portino a risultati diversi. Noi siamo interessati a vedere se le differenze dei risultati sono compatibili con **fluttuazioni statistiche** o se sono dovute ad altri effetti (come per esempio misure sbagliate, campioni non casuali, sorgenti di sistematici non considerate, uso improprio della statistica, altri effetti).
- In definitiva si vuole controllare per esempio se una misura è compatibile con il suo valore vero noto a priori (o considerato tale), se due misure della stessa grandezza sono compatibili tra loro o se una singola misura è compatibile con un campione di misure.

# Compatibilità tra Misure

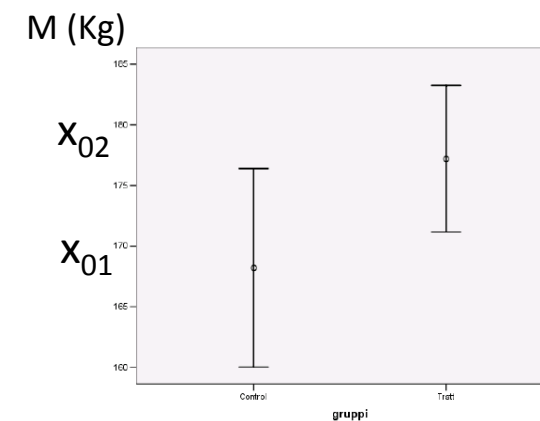
- Supponiamo di voler controllare se due misure della stessa grandezza fatte da gruppi diversi sono tra loro compatibili. Siano le due misure :

$$x_1 = x_{01} \pm \sigma_{\text{stat1}} \pm \sigma_{\text{sist1}} \quad \text{e} \quad x_2 = x_{02} \pm \sigma_{\text{stat2}} \pm \sigma_{\text{sist2}}$$

- Sommo in quadratura gli errori:  $x_1 = x_{01} \pm \sigma_{\text{tot1}}$  con  $\sigma_{\text{tot1}} = \sqrt{\sigma_{\text{stat1}}^2 + \sigma_{\text{sist1}}^2}$

$$x_2 = x_{02} \pm \sigma_{\text{tot2}} \quad \text{con} \quad \sigma_{\text{tot2}} = \sqrt{\sigma_{\text{stat2}}^2 + \sigma_{\text{sist2}}^2}$$

- In figura due misure con le rispettive barre di errore ( $1 \sigma_{\text{tot}}$ ). Se le barre di errore “si toccano” (come succede in figura), diciamo che le due misure sono compatibili entro  $1\sigma$ . Se non si toccassero diremmo che sono incompatibili (entro  $1\sigma$ )



# Compatibilità tra Misure

- Vogliamo ora vedere la compatibilità di due misure in modo più quantitativo e preciso.
- Per fare questo semplifichiamo **per ora** il problema supponendo di aver **corretto TUTTE** le sorgenti (scovate) di incertezze sistematiche. Le correzioni siano con incertezza nulla. Quindi non abbiamo errori sistematici.
- Confrontiamo una misura  $x_0$  di una grandezza fisica col valore vero e noto (o supposto tale)  $\mu$  di questa grandezza.
- La grandezza nota la supponiamo con incertezza nulla.

# t-Test

- ❑ W. S. Gosset (1876-1937) , chimico inglese assunto dalla birreria Guinness Brewery a Dublino per una attività che oggi chiameremmo controllo di qualità. È stato un pioniere della Statistica.
- ❑ Con test a campionamento di assaggi di boccali di birra doveva stimare quale malto o processo di produzione portava ad una birra di miglior qualità.
- ❑ Gosset aveva notato che applicare le conoscenze statistiche di quel tempo (distribuzioni gaussiane) quando la numerosità dei campioni era piccola portava a risultati non soddisfacenti.
- ❑ Bisognava utilizzare un diverso tipo di distribuzione, ora nota come **t di Student** . Non potendo pubblicare i risultati col proprio nome Gosset assunse lo pseudonimo Student. Test che usano questa distribuzione si chiamano t-Test
- ❑ Abbiamo misurato una grandezza  $X$  con un campione di  $n$  misure indipendenti e fatte nelle stesse condizioni. Supponiamo che questo campione sia estratto da una popolazione distribuita in modo gaussiano con media  $\mu$  e deviazione standard  $\sigma$  .

# Distribuzione t di Student

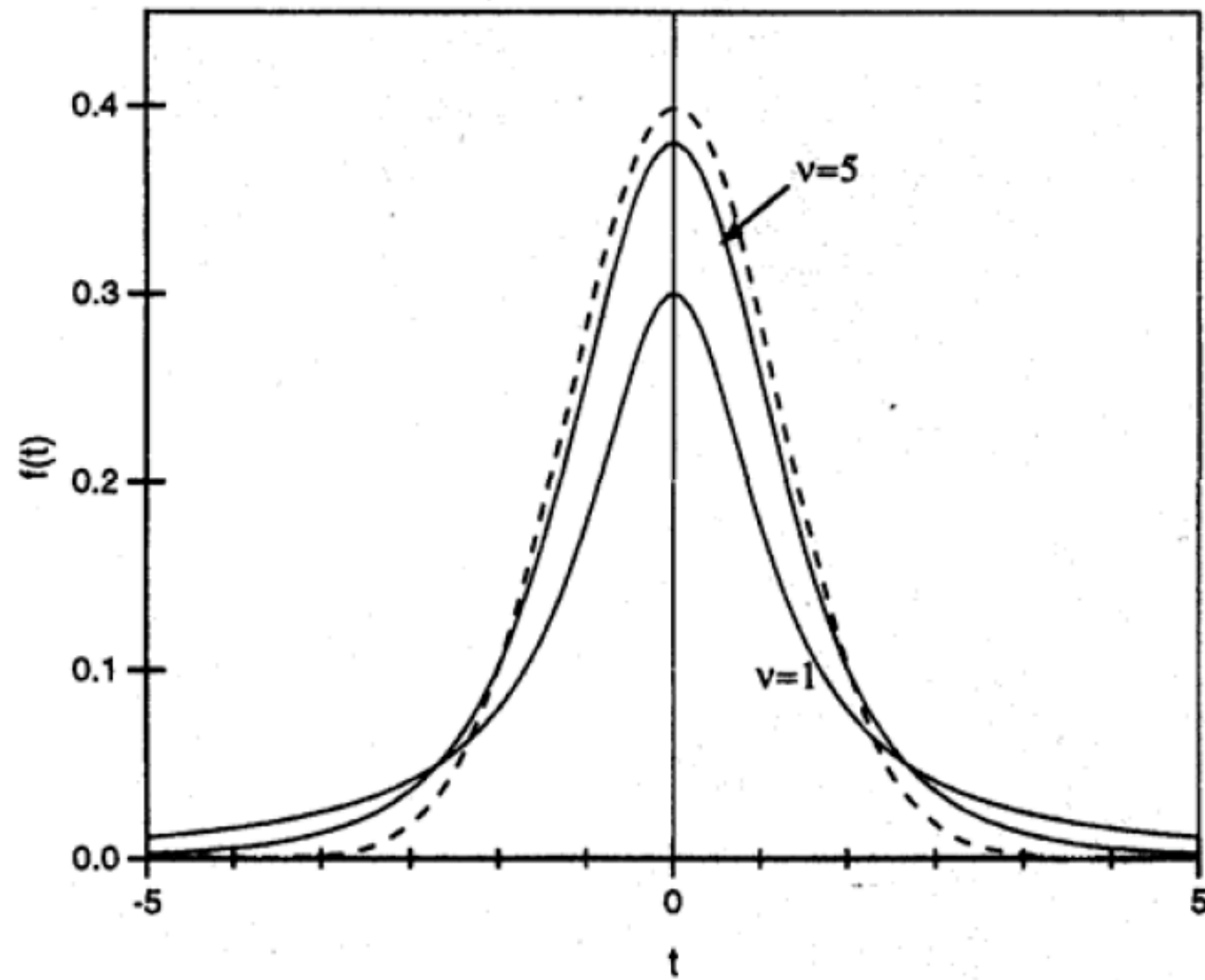
- ❑ Se  $X \sim N(\mu, \sigma^2)$  allora detta  $z=(x-\mu)/\sigma$  segue che  $Z \sim N(0,1)$
- ❑ Se da questa popolazione estraggo diversi campioni di  $n$  misure indipendenti, ognuno con una certa media  $\bar{x}$ , la distribuzione degli scarti (tra la media ottenuta e la media attesa  $\mu$ ) normalizzati

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \text{ segue anch'essa una distribuzione gaussiana standard}$$

- ❑ Sorge qui un problema perché in generale la  $\sigma$  della popolazione **NON** è nota. Trattiamo questo caso, approssimando la deviazione standard  $\sigma$  della popolazione col valore della deviazione standard campionaria  $s$ :

$$t_{(n-1)} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- ❑ Questa variabile segue una distribuzione t di Student con  $n-1$  gradi di libertà



La distribuzione t di Student è una curva simmetrica attorno all'origine.  
La curva tratteggiata è la gaussiana standard. Al crescere del numero di gradi di libertà  $v$ ,  $t_{(v)}$  tende alla gaussiana standard  $N(0,1)$

# Compatibilità di Due Misure

- Supponiamo di aver misurato una quantità di cui si conosce a priori il valore esatto (o supposto tale)  $\mu$  con una serie di  $n$  misure ripetute indipendenti. Sia  $x_0$  il valore misurato. Mi chiedo:

**$x_0$  e  $\mu$  sono tra loro compatibili?**

- Come ipotesi nulla  $H_0$  assumiamo  $x_0 = \mu$
- Lo scarto  $D = x_0 - \mu$  normalizzato (studentizzato) con l'errore standard della media segue la distribuzione t di Student con  $n-1$  dof. Mi calcolo il valore di questa variabile corrispondente alla mia media  $x_0$

$$t_{(n-1)}^0 = \frac{x_0 - \mu}{\frac{s}{\sqrt{n}}}$$

- Mi calcolo quindi il P-value cioè la probabilità che sotto l'ipotesi  $H_0$  ripetendo più volte la misura con diversi campioni ottengo un valore di  $t_{(n-1)} \geq t_{(n-1)}^0$  oppure valori  $-t \leq -t_{(n-1)}^0$
- Notare che nella definizione di scarto io includo come possibili discrepanze sia quelle positive che quelle negative.

# Compatibilità

- ❑ Il P-value rappresenta la probabilità con cui io rigetto l'ipotesi nulla  $H_0$ .
- ❑ Noi possiamo anche considerare la gaussiana standard e determinare il punto  $z_0$  in corrispondenza del quale la probabilità da quel punto a  $+\infty$  è uguale al valore P trovato.
- ❑ La distanza di  $z_0$  dall'origine esprime la probabilità in termini di deviazioni standard (gaussiane) con cui si rigetta l'ipotesi nulla. Per esempio se  $z_0$  dista 4.1 dall'origine si dice che l'ipotesi nulla  $H_0$  è rigettata con  $4.1 \sigma$  (gaussiane).
- ❑ **Fin qui la statistica.** Nella pratica si fissa una soglia in numero di sigma, (per esempio 2) e se l'ipotesi nulla è rigettata con meno di 2 sigma, allora si dice che  $x_0$  e  $\mu$  sono compatibili entro  $2 \sigma$  e la discrepanza osservata è frutto di una fluttuazione statistica. Se l'ipotesi nulla è rigettata con più di  $2\sigma$ , allora  $x_0$  e  $\mu$  si ritengono incompatibili al livello di  $2 \sigma$ .
- ❑ La soglia in numero di sigma è arbitraria e dipende dal settore di utilizzo.

# Esempio

- ❑ Ho misurato una grandezza 8 volte. So che in precedenza questa grandezza è stata misurata con grande precisione. Sia 1.250, il suo valore (nella sua unità di misura) già noto ed esatto. Ho trovato  $\bar{x}_0 = 1.234$  come media (aritmetica) delle 8 misure con una deviazione standard campionaria  $s$  di 0.049 da cui ottengo un valore  $t_{(7)}^0 = 0.938$ .
- ❑ Integro la  $t_{(7)}$  da 0.938 a  $+\infty$ , ottenendo 0.190 così come integrando da -0.938 a  $-\infty$ . Per calcolare questo valore faccio comodamente uso di un calcolatore statistico. Io rigetto l'ipotesi nulla con un p-value di 0.380.
- ❑ Io posso riportare la probabilità 0.190 nella coda a destra della gaussiana standard, ottenendo il valore  $z_0 = 0.878$  e analogamente nell'altra coda, ottenendo il valore simmetrico  $-z_0$ . Poiché  $|z_0| < 2$  io assumo che la mia misura e il valore atteso sono compatibili entro  $2\sigma$  (gaussiane) e associo la discrepanza osservata tra questi due valori a fluttuazioni statistiche.
- ❑ Test statistici di questo tipo si dicono a **due code** (bilaterale). In taluni casi si è interessati a test statistici ad **una coda** (o unilaterale) a destra o a sinistra.

# Uso di un Calcolatore Statistico

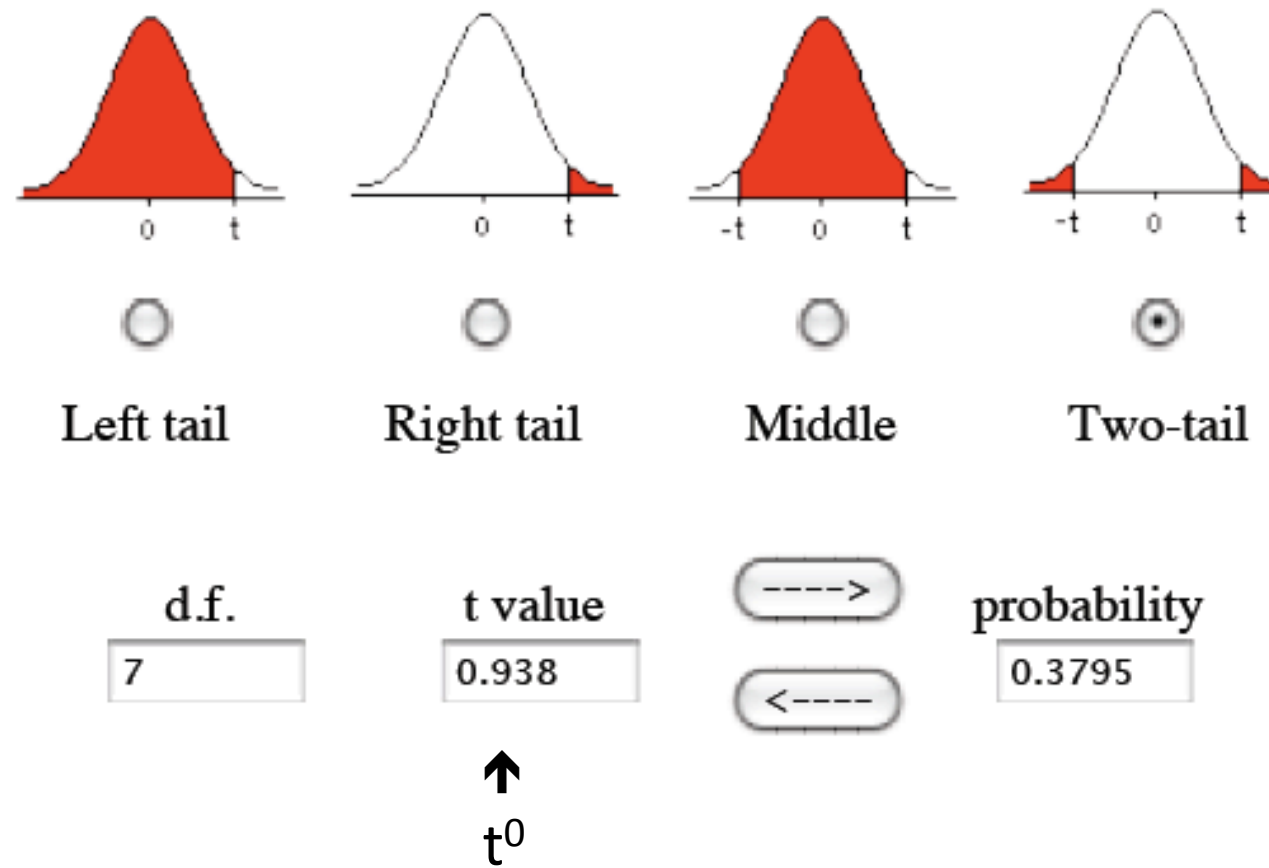
- In rete esistono molti calcolatori statistici che rendono i calcoli di tipo statistico molto semplici.
- È possibile anche scaricare sul proprio computer dalla rete una applicazione per fare questi calcoli.
- Comunque in caso di impossibilità (di rete o di computer) uno può sempre utilizzare le Tavole Statistiche del libro.
- Le Tavole Statistiche con un Calcolatore Statistico le trovate ad esempio a questo link:

[http://www.tutor-homework.com/statistics\\_tables/statistics\\_tables.html](http://www.tutor-homework.com/statistics_tables/statistics_tables.html)

- Vediamo come funziona in pratica.

# Calcolatore Statistico

**Table 2: Student's t-Distribution**



# Compatibilità di due Medie Campionarie

- ❑ Supponiamo di voler confrontare le misure di una stessa grandezza fatte da due diversi gruppi. I due campioni indipendenti sono estratti da una stessa popolazione di media  $\mu$ .
- ❑ Un gruppo ottiene un valore  $x_1$  avendo fatto  $n_1$  misure mentre il secondo gruppo ottiene un valore  $x_2$  avendo fatto  $n_2$  misure. Le deviazioni standard campionarie ottenute sono  $s_1$  e  $s_2$ . I due risultati  $x_1$  e  $x_2$  sono tra loro compatibili?

- ❑ In questo caso la variabile da utilizzare nel test è la seguente distribuzione t di Student:

$$t_{(n_1+n_2-2)} = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

dove  $n_1 + n_2 - 2$  è il numero di gradi di libertà della distribuzione.

- ❑ La compatibilità delle due misure viene calcolata tramite questa variabile seguendo il procedimento già visto.

# z-Test

- Supponiamo ora che la varianza  $\sigma^2$  sia nota oppure che la numerosità  $n$  delle misure nel campione sia tale che la varianza campionaria  $s^2$  sia una buona approssimazione di  $\sigma^2$ :  $\sigma^2 \approx s^2$

- In questi casi la distribuzione degli scarti normalizzati del valore delle medie  $\bar{x}$  ottenute con più campioni indipendenti provenienti dalla stessa popolazione e del valore atteso  $\mu$  segue una distribuzione gaussiana standard  $Z \sim N(0,1)$  
$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Misurato  $x_0$  e l'errore standard della media  $\sigma/\sqrt{n}$  mi calcolo il valore  $z_0$  
$$z_0 = \frac{x_0 - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Se  $|z_0| < 2$  allora assumo la mia misura compatibile col valore atteso (entro le  $2\sigma$ ). Se  $|z_0| > 2$  assumo la mia misura incompatibile col valore atteso.

# Compatibilità Includendo Incertezze Sistematiche

- ❑ Una misura sperimentale ha in generale una incertezza di tipo statistico e una incertezza di tipo sistematico:  $x_{\text{mis}} \pm \sigma_{\text{stat}} \pm \sigma_{\text{sist}}$
- ❑ Nelle slide precedenti abbiamo considerato la compatibilità di due misure supponendo nulli o trascurabili le incertezze sistematiche.
- ❑ Nella pratica e in generale questo non è vero. Anzi talvolta a dominare sono proprio le incertezze sistematiche.
- ❑ Un metodo semplice (e **approssimato**) di introdurre le incertezze sistematiche nel confronto delle due misure è di sostituire nelle formule viste l'errore totale all'errore statistico.
- ❑ Poiché l'errore totale è maggiore dell'errore statistico, due misure non compatibili sulla base del solo errore statistico possono essere compatibili sulla base dell'errore totale.

## Stime Intervallari di Parametri

- ❑ In talune misure il valore misurato è parecchio più piccolo dell'errore di misura. Ad esempio  $x = 3 \pm 21$  (nella sua unità di misura). In questi casi il valore misurato ha scarso significato e si preferisce presentare la misura in altro modo.
- ❑ Supponiamo di avere un campione di  $n$  misure ripetute i.i.d.  $x_1, x_2, \dots, x_n$ , tratte da una popolazione la cui distribuzione gaussiana ha valore medio  $\mu$  **NON NOTO** e deviazione standard  $\sigma$  **NOTA**.
- ❑ Detta  $x_0$  la media (aritmetica) delle mie  $n$  misure, la quantità  $z = \frac{x_0 - \mu}{\frac{\sigma}{\sqrt{n}}}$  segue una gaussiana standard  $N(0,1)$ .
- ❑ Qual è la probabilità che  $z$  cada tra  $-1.960$  e  $1.960$ ? Questo avviene quando il **valore di  $\mu$**  si trova nell'intervallo  $\left[ x_0 - 1.960 \frac{\sigma}{\sqrt{n}}, x_0 + 1.960 \frac{\sigma}{\sqrt{n}} \right]$
- ❑ Sulla gaussiana standard l'area compresa nell'intervallo  $-1.960$  e  $1.960$  costituisce lo 0.95 (95%) dell'area totale.

## Stime Intervallari di Parametri

- ❑ In questo caso noi diciamo che  $[x_0 - 1.960 \sigma/\sqrt{n}, x_0 + 1.960 \sigma/\sqrt{n}]$  costituisce un Intervallo (centrale) di Confidenza (CI) al Livello di Confidenza (CL) di 0.95 (del 95%).

- ❑ Se estraiamo campioni di  $n$  misure dalla popolazione considerata ( di cui non conosciamo il valore vero ma conosciamo la varianza), allora nel 95% dei casi il valore **vero e incognito** è all'interno dell'intervallo:

$$\left[ x_0 - 1.960 \frac{\sigma}{\sqrt{n}}, x_0 + 1.960 \frac{\sigma}{\sqrt{n}} \right]$$

- ❑ Se invece di 1.960 consideriamo **1.645 e 2.576**, allora i CI ottenuti sono rispettivamente al CL del **90%** e **99%**, rispettivamente.
- ❑ Se invece di 1.960 prendiamo 1 si ha un CI al 68.3% CL ( ad  $1\sigma$  ), se usiamo il valore 2 si ha un CI al 95.4% CL (a  $2\sigma$ ) mentre con 3 si ha un CL di 99.7%.
- ❑ La stima puntuale  $3 \pm 21$  la scriverei meglio come  $[-18, +24]$  al 68.3% CL.  
Gli intervalli di confidenza più usati (in Fisica) sono quelli al 90%, 95% e 99% CL.

# Intervalli di Confidenza per Campioni Piccoli

- In un campione di 15 misure si è misurato un valore medio di 128.5 con errore standard della media 5.4. Determinare l'intervallo di confidenza della media al 90% CL.

Valor medio  $x_0 = 128.5$ , dof  $n = 14$ , errore standard della media  $\frac{s}{\sqrt{n}} = 5.4$   
Il numero delle misure è piccolo quindi l'intervallo di confidenza al 90% CL va determinato sulla distribuzione t di student a 14 dof.

L'intervallo di confidenza al 90% CL si ottiene determinando il valore b tale che l'area racchiusa nell'intervallo

$$\left[ x_0 - b \frac{s}{\sqrt{n}}, x_0 + b \frac{s}{\sqrt{n}} \right)$$

sia uguale a 0.9. b risulta uguale a 1.761. L'intervallo di confidenza al 90% CL è  $[128.5 - 1.761 * 5.4, 128.5 + 1.761 * 5.4] = [119.0, 138.0]$

- Se utilizzassi una distribuzione gaussiana, b sarebbe uguale a 1.645 e l'intervallo di confidenza al 90% CL sarebbe  $[119.6, 137.4]$ .