

Medie

Numero progressivo	valore
1	10.01588
2	10.01496
3	9.998057
4	9.995584
5	10.00627
6	10.01139
7	10.00563
8	10.0208
9	10.00385
10	9.991494
11	9.999579
12	10.00182
13	10.0074
14	9.99659
15	10.00657
16	9.997447
17	10.00552
18	10.00084
19	10.00145
20	9.99211
21	10.01315
22	10.00807
23	10.01458
24	10.00425
25	10.01572
26	9.995653
27	10.01475
28	9.981818
29	10.01426
30	9.998887

$$m_{1,5} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}$$

$$m_{2,5} = \frac{x_6 + x_7 + x_8 + x_9 + x_{10}}{5}$$

$$m_{3,5} = \frac{x_{11} + x_{12} + x_{13} + x_{14} + x_{15}}{5}$$

$$m_{i,n} = \frac{x_{(i-1)n+1} + x_{(i-1)n+2} + x_{(i-1)n+3} + \dots + x_{in}}{n}$$

Distribuzione delle
medie di numerosità n

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Distribuzione delle medie

- supponendo nota la media μ e la varianza σ della popolazione delle singole misure x_i
- A) qual'è il valore atteso della distribuzione delle medie ?
- B) qual'è la varianza della distribuzione delle medie ?
- per qualsiasi distribuzione valgono le relazioni:

$$\mu(X + Y + Z) = \mu(X) + \mu(Y) + \mu(Z)$$

$$\mu(kX) = k \cdot \mu(X)$$

$\mu(X)$ =valore atteso per X

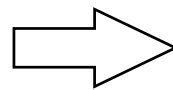
A)

$$\mu(M_n) = \mu\left(\frac{X_1 + X_2 + X_3 + \dots + X_n}{n}\right)$$

□ poichè X_1, X_2, \dots, X_N costituiscono la stessa distribuzione X ,
si ottiene:

$$\mu(M_n) = \frac{1}{n} [\mu(X_1) + \mu(X_2) + \mu(X_3) + \dots + \mu(X_n)]$$

$$\mu(M_n) = \frac{1}{n} \cdot n \cdot \mu(X)$$



$$\underline{\mu(M_n) = \mu(X)}$$

B)

□ ricordando che per una
combinazione lineare di
variabili casuali indipendenti

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots$$

vale la relazione: $\sigma^2(Y) = b_1^2 \cdot \sigma^2(X_1) + b_2^2 \cdot \sigma^2(X_2) + \dots$

$$\sigma^2(M_n) = \sigma^2\left(\frac{X_1 + X_2 + X_3 + \dots + X_n}{n}\right) =$$

$$= \frac{1}{n^2} \sigma^2(X_1) + \dots + \frac{1}{n^2} \sigma^2(X_n)$$

$$\Rightarrow \sigma^2(M_n) = \frac{\sigma^2(X)}{n}$$

Errore standard della media

- ❑ la variabilità delle medie è minore della variabilità delle singole misure

$$\sigma^2(M_n) = \frac{\sigma^2(X)}{n} \quad \Rightarrow \quad \sigma(M_n) = \sigma_m = \frac{\sigma_x}{\sqrt{n}}$$

- ❑ la riduzione di variabilità è “lenta”:
 - ♦ per ridurre l'errore standard di un fattore n devo mediare n^2 misure consecutive

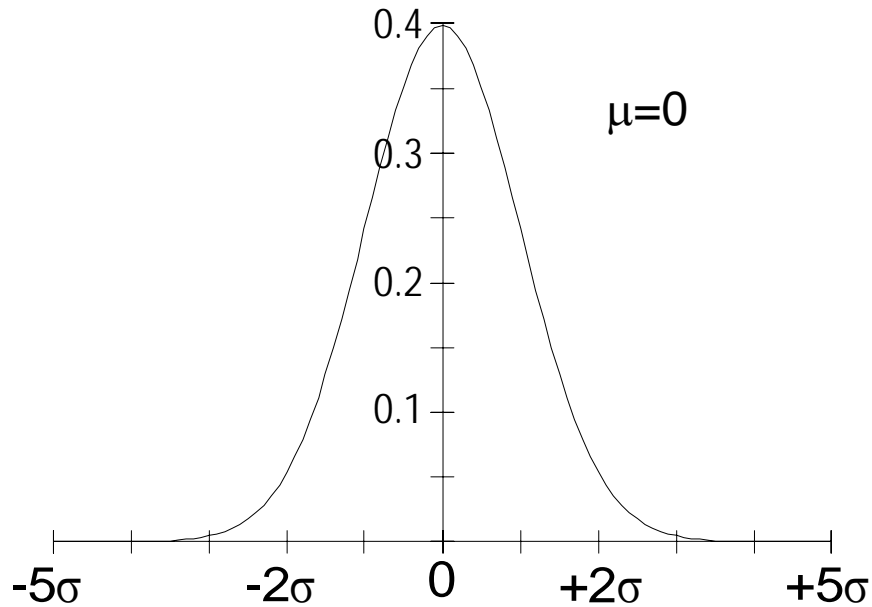
Distribuzione delle medie

- ☐ esiste una relazione fra il tipo di distribuzione delle singole misure x_i ed il tipo di distribuzione delle medie ?
- ☐ NO (sotto opportune, ampie condizioni)
- ☐ la distribuzione delle medie è sempre di tipo normale o gaussiano

Teorema centrale

- ☐ ***per qualsiasi distribuzione, quando le medie vengono calcolate su un numero sufficiente di singole misure, la distribuzione delle medie approssima una distribuzione normale***
- ☐ ***l'operazione di media deve essere di natura casuale: per eseguire una media, non posso scegliere i campioni, essi devono essere estratti casualmente***
- ☐ ***già per $n=4$, la distribuzione delle medie è ben approssimata da una distribuzione normale***

Distribuzione normale



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

$$E(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = \mu$$

$$Var(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx = \sigma^2$$

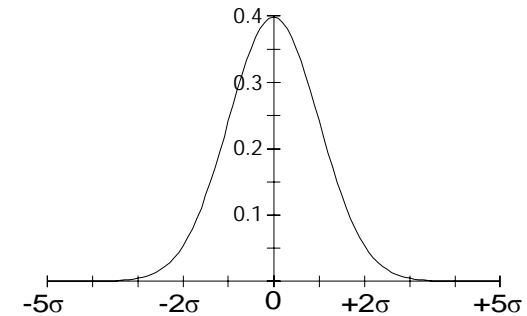
Distribuzione normale

la frequenza di ripetizione dei valori
compresi nell'intervallo: $A \leq x \leq B$

coincide con la frequenza di ripetizione dei valori
nell'intervallo:

$$\frac{A - \mu}{\sigma} \leq z \leq \frac{B - \mu}{\sigma}$$

della distribuzione normale standardizzata



❑ i punti di flesso della curva coincidono con $x = \mu \pm \sigma$

❑ nell'intervallo $(\mu - \sigma) \div (\mu + \sigma)$ cade il 68.26% delle misure

❑ nell'intervallo $(\mu - 2\sigma) \div (\mu + 2\sigma)$ cade il 95.44% delle misure

❑ nell'intervallo $(\mu - 3\sigma) \div (\mu + 3\sigma)$ cade il 99.74% delle misure

In breve ...

□ la distribuzione delle medie è di tipo normale

□ poichè $\sigma_m = \frac{\sigma_x}{\sqrt{n}}$ per ridurre la variabilità delle misure eseguo sempre la media di n misure consecutive

□ presa una media m_i qualsiasi posso affermare che essa cade:

- ♦ nell'intervallo $(\mu - \sigma) \div (\mu + \sigma)$ con una propabilità del 68.26%
- ♦ nell'intervallo $(\mu - 2\sigma) \div (\mu + 2\sigma)$ con una propabilità del 95.44%
- ♦ nell'intervallo $(\mu - 3\sigma) \div (\mu + 3\sigma)$ con una propabilità del 99.74%

Stima di σ_m

❑ vale la relazione $\sigma_m = \frac{\sigma}{\sqrt{n}}$

❑ essendo σ_x incognita posso stimare σ_m con la varianza campionaria S_m

❑ per S_m vale la relazione $S_m = \frac{S}{\sqrt{n}}$

❑ al variare dei campioni su cui calcolo la media ottengo valori diversi di S_m

❑ S_m è una (variabile) statistica che stima σ_m

❑ dalla statistica sappiamo che, indicando con ν i gradi di libertà:

$$\sigma^2(S_m) \cong \frac{\sigma_m^2}{2 \cdot \nu} = \frac{\sigma_m^2}{2 \cdot (n-1)}$$

❑ la stima dell'errore standard della media è migliore per campioni di elevata numerosità

continua ...

- ❑ se ripeto le misure con lo stesso processo, nelle stesse condizioni, anche in tempi diversi posso combinare i risultati di tutte le misure per ottenere una stima attendibile dell'errore standard della media

- ❑ a) per ciascuna sessione di misura possiamo calcolare:

$$S_{m,i}^2 = \frac{\sum (m_i - \bar{m}_i)^2}{r_i - 1} = \frac{\sum d_i^2}{r_i - 1}$$

$$\sum d_i^2 = S_{m,i}^2 \cdot (r_i - 1)$$

- ❑ b) combinando i dati di tutte le sessioni: $\hat{\sigma}_m^2 = \frac{\sum (r_i - 1) \cdot S_{mi}^2}{\sum (r_i - 1)}$

- ❑ per un processo di misura, il monitoraggio di $\hat{\sigma}_m^2$ può essere utile per il controllo di qualità

Medie con diversa numerosità

□ si dispone di p medie, ciascuna ottenuta da n_i misure singole

□ la media complessiva vale: $\tilde{m} = \frac{\sum n_i \cdot m_i}{\sum_p n_i}$

□ per ciascuna popolazione delle medie si ha:

$$\sigma_{mi}^2 = \frac{\sigma^2}{n_i} \quad \Rightarrow \quad n_i = \frac{\sigma^2}{\sigma_{mi}^2}$$

□ posto $w_i \stackrel{def}{=} \frac{1}{\sigma_{mi}^2} \quad \Rightarrow \quad n_i = \sigma^2 \cdot w_i$

continua ...

□ ricordando che σ è da considerare costante:

$$\tilde{m} = \frac{\sum_p n_i \cdot m_i}{\sum_p n_i} \quad \Rightarrow \quad \tilde{m} = \frac{\sum_p w_i \sigma^2 \cdot m_i}{\sum_p w_i \sigma^2} = \frac{\sum_p w_i \cdot m_i}{\sum_p w_i} \quad w_i \stackrel{def}{=} \frac{1}{\sigma_{mi}^2}$$

- le medie calcolate a partire da campioni di numerosità diversa possono essere combinate per stimare la media della popolazione delle misure possibili pur di attribuire a ciascuna media il giusto peso.
- nel computo peseranno di più le medie derivate da campioni più numerosi e quindi estratte da una distribuzione con varianza più piccola