

# LEZIONI DI STATISTICA MEDICA

Prof. SIMONE ACCORDINI

**Lezione n.12**  
- Test statistico



Sezione di Epidemiologia & Statistica Medica  
Università degli Studi di Verona

## Confronto della media tra due popolazioni

Esempio:

*Cockburn et al (1980) riportano una ricerca clinica per la prevenzione dell'ipocalcemia infantile, nella quale donne in gravidanza che ricevevano un supplemento di vitamina D venivano messe a confronto con donne non trattate.*

**Calcemia del bambino misurata 6 giorni dopo la nascita:**

	Numero di pazienti	Media (mg /100 ml)	DS (mg /100 ml)
<b>Vitamina D (D<sub>1</sub>)</b>	233	<b>9.36</b>	1.15
<b>Controllo (D<sub>0</sub>)</b>	394	<b>9.01</b>	1.33

La differenza osservata è dovuta al caso oppure alla vitamina D?



## TEST STATISTICO

- una delle tecniche inferenziali più utilizzate in medicina
- utile in situazioni nelle quali si è interessati a prendere **decisioni tra due o più alternative possibili sulla base delle osservazioni campionarie**

esempi:

- valutare l'efficacia di un nuovo farmaco rispetto ad un farmaco standard
- valutare se il trattamento chirurgico di un particolare tumore in una data fase allunga la vita dei pazienti rispetto al trattamento chemioterapico
- valutare se l'esposizione ad una determinata sostanza chimica è responsabile di un eccesso di tumori

In tali situazioni, la valutazione dell'alternativa migliore è finalizzata a **decidere quale intervento operare sulla realtà** (scelta del farmaco, tipo di terapia, tipo di intervento preventivo).



## Test statistico per il confronto della media tra due popolazioni

**GRUPPO NON ESPOSTO ( $D_0$ ):**

$x_{01}, x_{02}, \dots, x_{0n_0} \rightarrow n_0$  determinazioni **indipendenti** della v.c.  $X_0 \sim \text{Norm}(\mu_0, \sigma_0)$

**GRUPPO ESPOSTO ( $D_1$ ):**

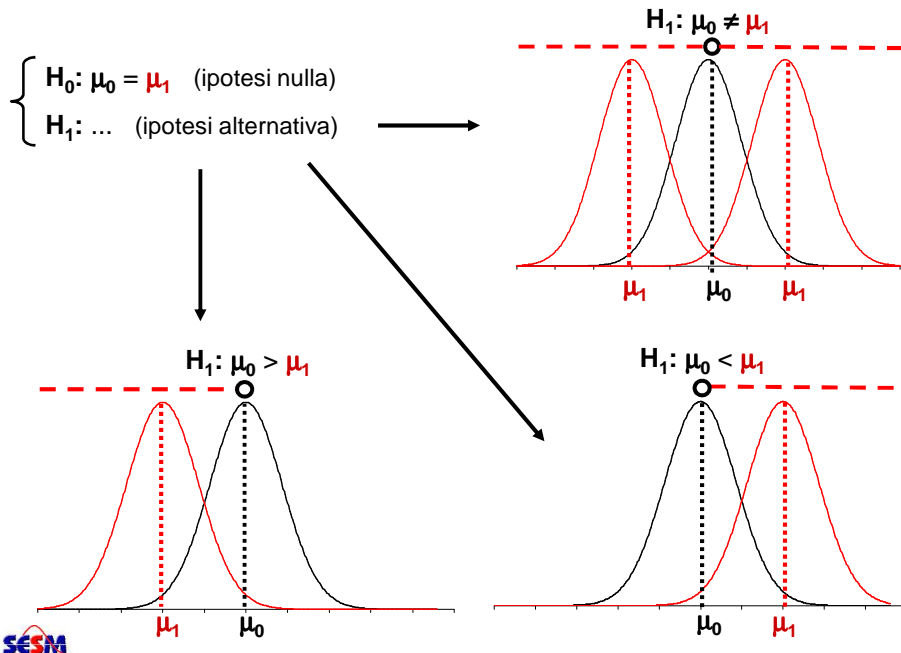
$x_{11}, x_{12}, \dots, x_{1n_1} \rightarrow n_1$  determinazioni **indipendenti** della v.c.  $X_1 \sim \text{Norm}(\mu_1, \sigma_1)$



**I° STEP: definire il sistema di ipotesi da verificare**

$H_0: \mu_0 = \mu_1$  (ipotesi nulla)

$H_1: \dots$  (ipotesi alternativa)



$H_0: \mu_0 = \mu_1$  (ipotesi nulla)

$H_1: \mu_0 \neq \mu_1$  (ipotesi alternativa)

L'ipotesi nulla ( $H_0$ ) è l'ipotesi che spiega la **differenza osservata tra le medie come dovute al caso**

⇒ variazioni casuali

L'ipotesi alternativa ( $H_1$ ) è l'ipotesi che spiega la **differenza osservata tra le medie come dovuta alla variabile in studio** (trattamento, fattori di rischio, ...)

⇒ variazioni sistematiche

**II° STEP: definire la statistica test**

$$\begin{cases} H_0: \mu_0 = \mu_1 & (\text{ipotesi nulla}) \\ H_1: \mu_0 \neq \mu_1 & (\text{ipotesi alternativa}) \end{cases}$$

differenza tra le medie = stima dell'effetto

$$Z = \frac{\bar{X}_1 - \bar{X}_0}{ES[\bar{X}_1 - \bar{X}_0]}$$

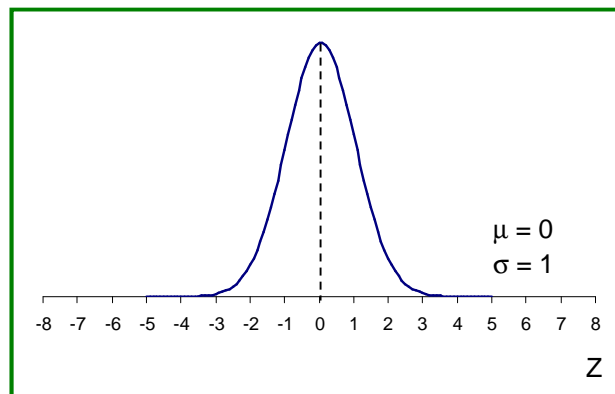
errore standard della differenza = misura della precisione della stima



$$Z = \frac{\bar{X}_1 - \bar{X}_0}{ES[\bar{X}_1 - \bar{X}_0]} = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\sigma_1^2/n_1 + \sigma_0^2/n_0}}$$

$\sigma_1^2$  e  $\sigma_0^2$  note

Se è vera l'ipotesi nulla ( $H_0$ ):



Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D <sub>1</sub> )	233	9.36	1.15
Controllo (D <sub>0</sub> )	394	9.01	1.33

$$\left\{ \begin{array}{l} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{array} \right. \quad \leftarrow \text{I}^\circ \text{ STEP}$$

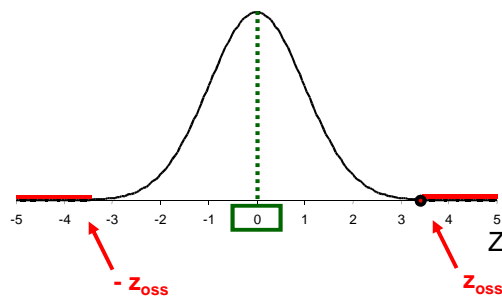
$$z_{\text{oss}} = \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{\sigma_1^2/n_1 + \sigma_0^2/n_0}} = \frac{9.36 - 9.01}{0.1008} = 3.47 \quad \leftarrow \text{II}^\circ \text{ STEP}$$



III° STEP: calcolare il p-value

misura continua di quanto i dati supportano l'ipotesi nulla (H<sub>0</sub>),  
ovvero di quanto i dati sono plausibili sotto l'ipotesi nulla (H<sub>0</sub>)

$$\left\{ \begin{array}{l} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{array} \right.$$



$$p\text{-value} = 2 \text{ prob}(Z \geq z_{\text{oss}} \mid H_0 \text{ è vera})$$



Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D <sub>1</sub> )	233	9.36	1.15
Controllo (D <sub>0</sub> )	394	9.01	1.33

$$\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$$

I° STEP

$$z_{\text{oss}} = \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{\sigma_1^2/n_1 + \sigma_0^2/n_0}} = \frac{9.36 - 9.01}{0.1008} = 3.47$$

II° STEP

$$p\text{-value} = 2 \text{ prob} (Z \geq 3.47 \mid H_0 \text{ è vera}) = 0.0005$$

III° STEP

i dati non sembrano supportare l'ipotesi nulla (H<sub>0</sub>)

→ la differenza osservata sembra essere dovuta alla vitamina D



## SIGNIFICATIVITA' STATISTICA

Confronto tra il valore osservato del p-value e il cut-off di significatività (0.05):

**p-value ≥ 0.05** → la differenza osservata **NON è statisticamente significativa**

⇒ **NON rifiuto l'ipotesi nulla (H<sub>0</sub>)**

(la differenza osservata sembra essere dovuta al caso)

**p-value < 0.05** → la differenza osservata è **statisticamente significativa**

⇒ **rifiuto l'ipotesi nulla (H<sub>0</sub>)**

(la differenza osservata sembra essere dovuta ad altri fattori: trattamento, fattori di rischio, ...)



Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
<i>Vitamina D (D<sub>1</sub>)</i>	233	9.36	1.15
<i>Controllo (D<sub>0</sub>)</i>	394	9.01	1.33

p-value = **0.0005 < 0.05**

la differenza osservata è **statisticamente significativa**

⇒ **rifiuto l'ipotesi nulla (H<sub>0</sub>)**

(la differenza osservata sembra essere dovuta alla vitamina D)



## RELAZIONE TRA INTERVALLO DI CONFIDENZA E TEST STATISTICO

Esempio (continua):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
<i>Vitamina D (D<sub>1</sub>)</i>	233	9.36	1.15
<i>Controllo (D<sub>0</sub>)</i>	394	9.01	1.33

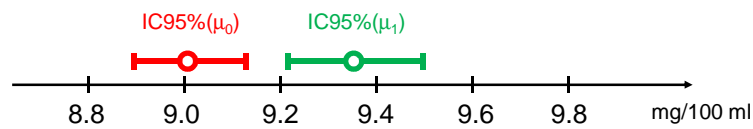
$$\text{IC95\% } (\mu_0) = \bar{x}_0 \pm 1.96 \sigma_0 / \sqrt{n_0} = 9.01 \pm 1.96 * 1.33 / \sqrt{394}$$

$$\text{IC95\% } (\mu_1) = \bar{x}_1 \pm 1.96 \sigma_1 / \sqrt{n_1} = 9.36 \pm 1.96 * 1.15 / \sqrt{233}$$



Esempio (calcemia):

	controllo (n = 394)	vitamina D (n = 233)	p-value
media [IC95%]	9.01 [8.88, 9.14]	9.36 [9.21, 9.51]	0.0005 (<0.05)



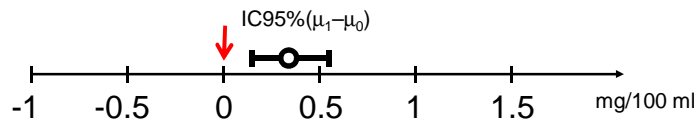
gli intervalli di confidenza non sono sovrapposti  
 ⇒ la differenza osservata è statisticamente significativa



Esempio (calcemia):

differenza delle medie [IC95%]	0.35 [0.15, 0.55]
-----------------------------------	----------------------

mg/100 ml

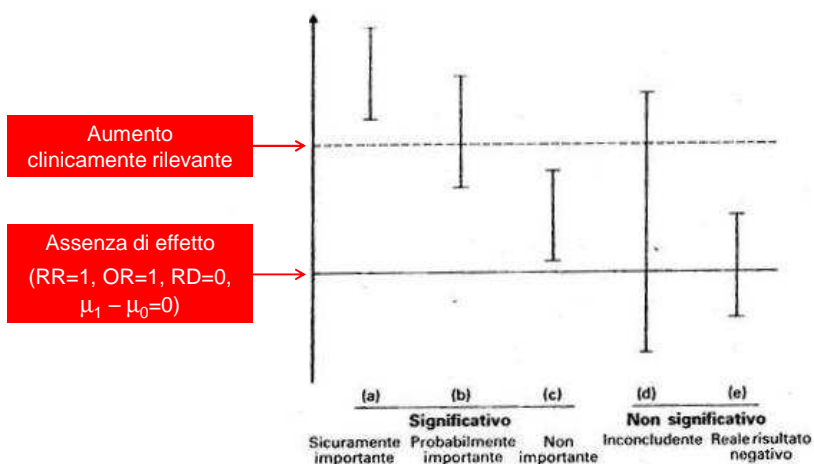


l'intervallo di confidenza della differenza delle medie  
 non contiene il valore nullo (0 mg/100ml)  
 ⇒ la differenza osservata è statisticamente significativa  
 ma NON è clinicamente rilevante





## SIGNIFICATIVITÀ STATISTICA vs RILEVANZA CLINICA



### Esempio:



Numero di cardiopatie ischemiche (CHD) in funzione del tipo di personalità (Western Collaborative Group Study):

	CHD	NO CHD	
Tipo A	178	1.411	1589
Tipo B	79	1.486	1565
	257	2.897	3154

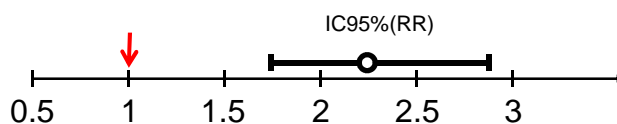
$$RR = \frac{11.2}{5.0} = 2.22$$

$$IC_{95\%}(RR) = [1.72, 2.87]$$

tipo A: competitivo, apprensivo

tipo B: rilassato e non competitivo

} Coorte di individui (34-59 anni)  
seguiti per un periodo di 8 anni



l'intervallo di confidenza del RR non contiene il valore nullo (1)

⇒ l'aumento osservato dell'incidenza di CHD è statisticamente significativo