NOTE DALLE LEZIONI
DI
STATISTICA MEDICA
ED ESERCIZI

I METODI PER IL CONFRONTO DI MEDIE
(*Campioni indipendenti)*

IL PROBLEMA

Sono stati rilevati i dati relativi alla frequenza cardiaca (misurata in battiti al minuto) in un gruppo di soggetti con angina e in un gruppo di soggetti con infarto.

Anginosi	Infartuati
81	61
65	75
77	78
87	80
95	68
89	65
103	68
89	69
78	70
83	79
91	61

Si vuole verificare se la frequenza cardiaca e differente nei due gruppi

Prima di continuare……..

CALCOLATE
LA MEDIA E LA VARIANZA
DEI DUE GRUPPI

VERIFICA DI IPOTESI:
CONFRONTO DI DUE MEDIE PER CAMPIONI INDIPENDENTI
VARIANZE INCOGNITE

	Popolazione	
μ_1 σ_1		μ_2 σ_2

	Campione	
$n_1=11$ $\bar{x}_1=85{,}27$ $S_1=10{,}14$		$n_2=11$ $\bar{x}_2=70{,}36$ $S_2=6{,}81$

Dati: frequenza cardiaca in battiti al minuto degli 11 soggetti dei due campioni.
1. anginosi; 2. infartuati

Assunzioni:

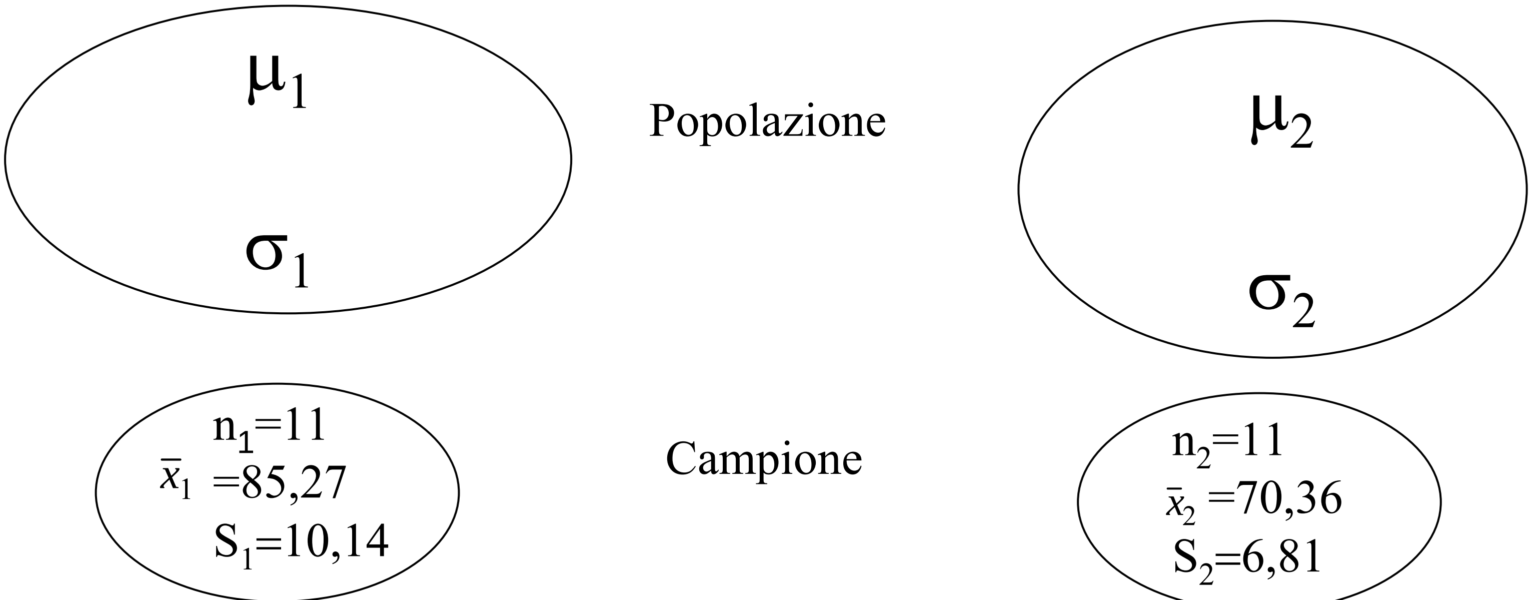
☞

☞

☞ Indica le assunzioni prima di andare avanti….

☞

VERIFICA DI IPOTESI:
CONFRONTO DI DUE MEDIE PER CAMPIONI INDIPENDENTI
VARIANZE INCOGNITE



<u>Dati</u>: frequenza cardiaca in battiti al minuto degli 11 soggetti dei due campioni.
1. anginosi; 2. infartuati

<u>Assunzioni</u>:

- ☞ campioni estratti da popolazioni con distribuzione di Gauss;
- ☞ campioni indipendenti
- ☞ varianza delle popolazioni non nota;
- ☞ varianza delle popolazioni omogenee.

Coincidono con quelle indicate? Se si procedi…altrimenti rivedi la teoria….

Ipotesi

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 \neq \mu_2 \\ & \mu_1 > \mu_2 \\ & \mu_1 < \mu_2 \end{cases}$$

$\mu_1 \neq \mu_2$ → Ipotesi alternativa bidirezionale

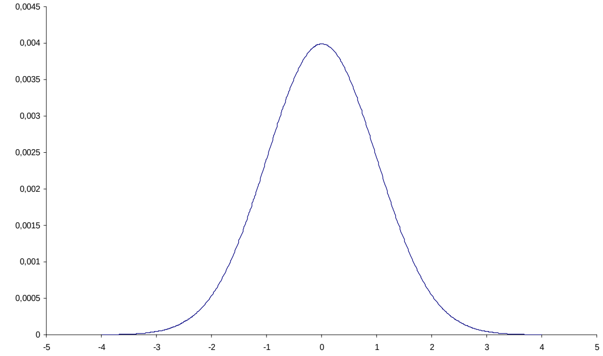
$\mu_1 > \mu_2$, $\mu_1 < \mu_2$ → Ipotesi alternativa unidirezionale

Statistica test

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

Distribuzione della statistica test

Distribuzione t-Student con g.l. n_1+n_2-2 della varianza comune S^2_p.

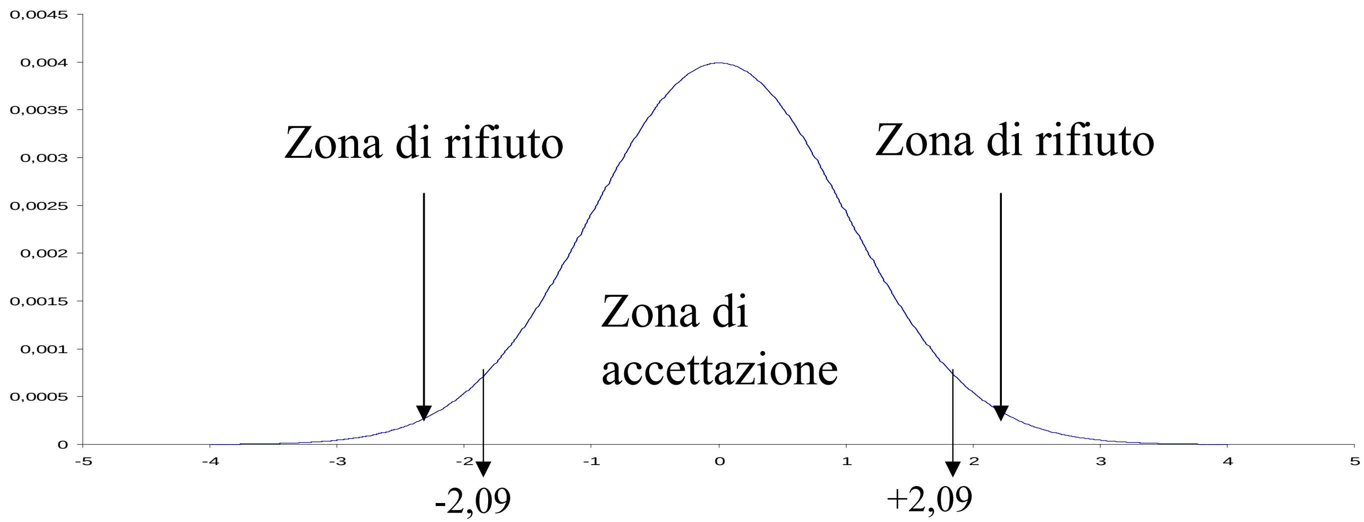


Si ricorda che :

$$S_p^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

Regola di decisione

Fisso α, la probabilità di rifiutare l'ipotesi nulla quando è vera, α=0,05. Supponendo di aver formulato un ipotesi bidirezionale dovremo suddividere α nei due lati della curva (α/2 nelle code). Sapendo che la distribuzione della statistica è di t-Student i limiti si ricercheranno nelle apposite tavole in corrispondenza di α e dei gradi di libertà n_1+n_2-2 e saranno uguali ed opposti (nel nostro esempio i limiti sono ±2,09).



CALCOLATE :

✓LA VARIANZA COMUNE

✓ LA STATISTICA TEST

Poi verificate il risultato….

Calcolo della statistica test: Sapendo che S^2_p =74,635 con g.l.=20

$$t_{calc} = \frac{85{,}27 - 70{,}36}{\sqrt{\frac{74{,}635}{11} + \frac{74{,}635}{11}}} = 4{,}05$$

<u>Decisione statistica</u>: rifiuto l'ipotesi nulla dato che 4,05>2,09

<u>Decisione clinica</u>: il valore medio dei battiti minuto nelle due popolazioni è significativamente diverso

Il confronto tra campioni indipendenti prevede il calcolo della varianza comune S^2_p che è possibile calcolare solo se le varianze sono omogenee . Pertanto è necessario verificare l'omogeneità delle varianze:

IPOTESI

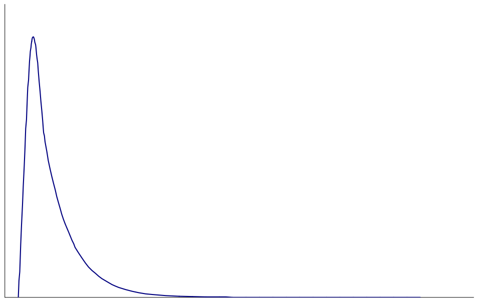
$$\begin{cases} \mathbf{H_0}: \quad \sigma^2_1 = \sigma^2_2 \\ \mathbf{H_1}: \quad \sigma^2_1 \neq \sigma^2_2 \end{cases}$$

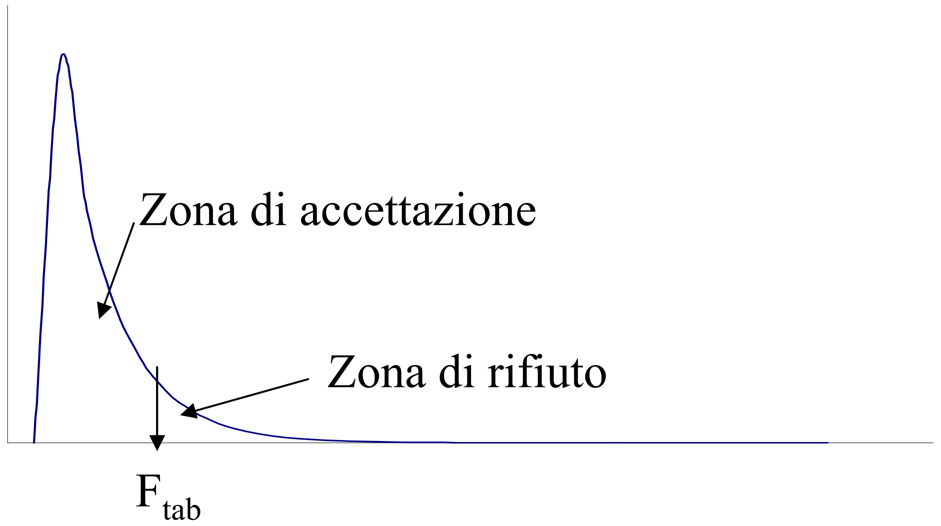
Statistica test

$\mathbf{F_{calc}=S_1^2/S_2^2}$
(si raccomanda di porre a numeratore la varianza maggiore)

Distribuzione della statistica test

Distribuzione F-Fisher che dipende dai g.l. del numeratore e del denominatore.





Regola di decisione

Fisso α= 0,05, la probabilità di rifiutare l'ipotesi nulla quando è vera. Si individueranno in questo modo i limiti della zona di rifiuto (coda); sapendo che la distribuzione della statistica è di F-Fisher i limiti si cercheranno nelle apposite tavole in corrispondenza di α e dei gradi di libertà n_1-1 (g.l.=10) del numeratore e n_2-1 (g.l.=10) del denominatore. Nel nostro esempio F_{tab}=2,98

Calcolo della statistica test:

$$F = 102{,}82 \,/\, 46{,}45 = 2{,}21$$

Decisione statistica

non rifiuto l'ipotesi nulla dato che 2,21<2,98.

Le due varianze si possono considerare omogenee e quindi posso calcolare S^2_p

Decisione clinica: le due varianze si possono considerare uguali nel limite della variabilità biologica e dell'errore casuale.

Nel caso in cui il test F abbia portato a rifiutare l'ipotesi nulla, allora le due varianze non si possono considerare uguali.
Il test per il confronto tra due medie diventa allora :

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

I gradi di libertà di questa statistica che ha ancora una distribuzione t-Student sono determinati dalla formula di Dixon e Massey:

$$g.l. = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2}}$$

SE NON FOSSE STATO POSSIBILE IPOTIZZARE

PER

LA VARIABILE "FREQUENZA CARDIACA"

UNA DISTRIBUZIONE DI GAUSS

COME AVRESTE DOVUTO EFFETTUARE

LA VERIFICA DELLE IPOTESI?

Riprendi i dati dell'esercizio e rispondi alla domanda prima di proseguire….

CONFRONTO TRA CAMPIONI INDIPENDENTI
TEST DELLA SOMMA DEI RANGHI
(WILCOXON)

Dati: frequenza cardiaca misurata in battiti al minuto degli 11 soggetti con angina e degli 11 soggetti con infarto.

Assunzioni:
- ☞ campioni indipendenti
- ☞ non si possono fare assunzioni sulla distribuzione.

Anginosi	Infartuati	Campione combinato	
81	61	61i	61i
65	75	65i	65a
77	78	68i	68i
87	80	69i	70i
95	68	75i	77a
89	65	78i	78a
103	68	79i	80i
89	69	81a	83a
78	70	87a	89a
83	79	89a	91a
91	61	95a	103a

Campione combinato	Assegnazione dei ranghi	
61i	1	1,5
61i	2	1,5
65i	3	3,5
65a	4	3,5
68i	5	5,5
68i	6	5,5
69i	7	7
70i	8	8
75i	9	9
77a	10	10
78i	11	11,5
78a	12	11,5
79i	13	13
80i	14	14
81a	15	15
83a	16	16
87a	17	17
89a	18	18,5
89a	19	18,5
91a	20	20
95a	21	21
103a	22	22

Si noti che a valori uguali si assegna rango pari alla media dei ranghi che ciascun valore avrebbe se fossero diversi

Campione combinato	Assegnazione dei ranghi
61i	1,5
61i	1,5
65i	3,5
65a	3,5
68i	5,5
68i	5,5
69i	7
70i	8
75i	9
77a	10
78i	11,5
78a	11,5
79i	13
80i	14
81a	15
83a	16
87a	17
89a	18,5
89a	18,5
91a	20
95a	21
103a	22

Somma ranghi degli anginosi
$\Sigma Ra=173$

Somma ranghi degli infartuati
$\Sigma Ri=80$

Cerco sulle tavole di Wilcoxon il valore tabulato in corrispondenza della numerosità dei due campioni e del livello di significatività.
Confronto il valore del campione di numerosità più piccola con l'intervallo tabulato.
Se il valore calcolato cade fuori dall'intervallo tabulato allora rifiuto H_0, altrimenti accetterò H_0.
Nel nostro caso i valori tabulati sono 96-157, per cui rifiuto H_0.