

PROGRAMMA DEL CORSO DI STATISTICA MEDICA A.A. 2021/22

Richiamo sulle basi razionali di inferenza statistica, stima e verifica di ipotesi su medie (un campione, campioni indipendenti, campioni appaiati), test non parametrici

Analisi della varianza ad una via

Confronti multipli

Modello Lineare Generale

Analisi della varianza a due e più vie e per misure ripetute

Analisi della varianza non parametrica

Regressione lineare semplice

Stima dei parametri del modello

Inferenza sul coefficiente di regressione: analisi della varianza della regressione

Regressione multipla

Regressione logistica (cenni)

Correlazione parametrica e non parametrica

Lettura critica di articoli scientifici.

Utilizzo di software statistici disponibili on line e interpretazione degli output

Testi consigliati:

W.W.DANIEL: Biostatistica. EdiSES

MARTIN BLAND **Statistica Medica** APOGEO

M.M.TRIOLA&M.F.TRIOLA Statistica per le discipline biosanitarie. Pearson-Addison Wesley

R.J.FREUND & W.J.WILSON, Metodi Statistici. PICCIN /

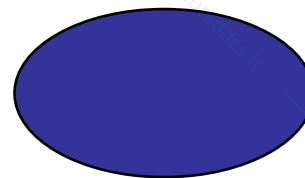
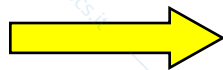
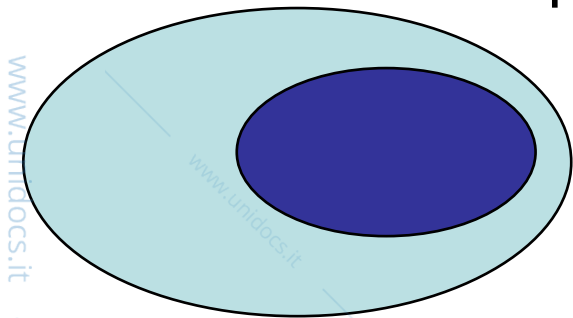
Obiettivi del corso

Alla fine del corso lo studente deve essere in grado di

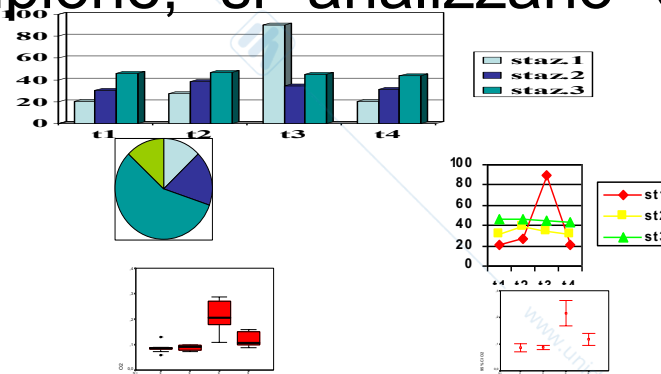
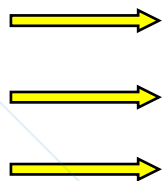
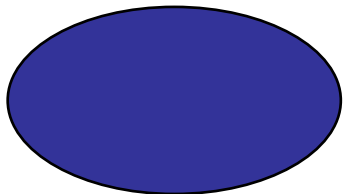
- Leggere in maniera critica un articolo scientifico
- Pianificare e condurre uno studio statistico
- Individuare opportunamente i metodi statistici da utilizzare per verificare le ipotesi di ricerca
- Eseguire l'analisi statistica utilizzando i software statistici disponibili online

Fasi elaborazione

Si estrae un campione dalla popolazione

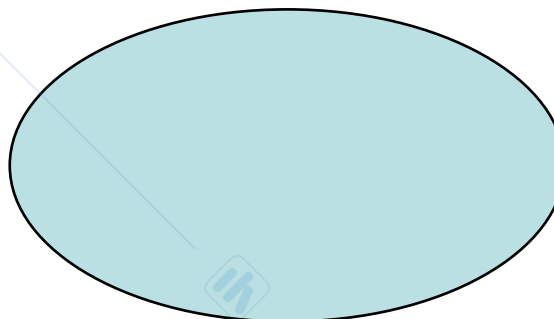
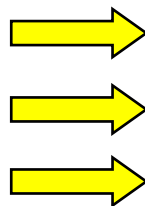
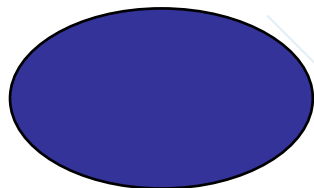


Si rilevano i dati nel campione, si analizzano e si descrivono



**Statistica
descrittiva**

Si estendono i risultati del campione alla popolazione



**Statistica
inferenziale**

Inferenza statistica

- Nell'estendere i risultati dal campione alla popolazione si passa da una condizione di certezza ad una di incertezza
- I risultati ottenuti nel campione sono solo una variante di tutti i possibili risultati degli infiniti campioni che si possono estrarre da una popolazione
- Occorre quantificare la probabilità che ogni deduzione basata sui dati sia corretta

Stima dei parametri

- **Stima puntuale:** si calcola un singolo valore numerico per stimare il corrispondente parametro. Es.: una media, una proporzione, una deviazione standard.
- **Stima intervallare:** si calcola un intervallo di valori che, con un certo grado di probabilità, conterrà il parametro da stimare

$\mu = ?$
 $\sigma = ?$

$\bar{x}_1 = 12$
 $S_1 = 3$

$\bar{x}_2 = 14$
 $S_2 = 3$

$\bar{x}_3 = 11$
 $S_3 = 2$

$\bar{x}_4 = 14$
 $S_4 = 3.5$

$\bar{x}_1 = 12$
 $S_1 = 3$

$\bar{x}_{150} = 12$
 $S_{150} = 3$

$\bar{x}_6 = 13$
 $S_6 = 1,8$

$\bar{x}_7 = 18$
 $S_7 = 2,4$

$\bar{x}_5 = 10$
 $S_5 = 1$

$\bar{x}_{200} = 17$
 $S_{200} = 4$

$\bar{x}_n = 16$
 $S_n = 2$

Quale di questi campioni descrive meglio la popolazione?

Nessun campione mi dà la certezza dei valori della popolazione, ma mi è utile conoscere la probabilità con cui si possono verificare i risultati campionari per prendere una decisione

Stima puntuale

- Consistono nel calcolo dei valori caratteristici di una distribuzione nel campione
- Non danno informazioni circa l'attendibilità della stima
- Non tengono conto della variabilità campionaria

$\mu = ?$
 $\sigma = ?$

$\bar{x}_1 = 12$
 $S_1 = 3$

$\bar{x}_2 = 14$
 $S_2 = 3$

$\bar{x}_3 = 11$
 $S_3 = 2$

$\bar{x}_4 = 14$
 $S_4 = 3.5$

$\bar{x}_1 = 12$
 $S_1 = 3$

$\bar{x}_{150} = 12$
 $S_{150} = 3$

$\bar{x}_6 = 13$
 $S_6 = 1,8$

$\bar{x}_7 = 18$
 $S_7 = 2,4$

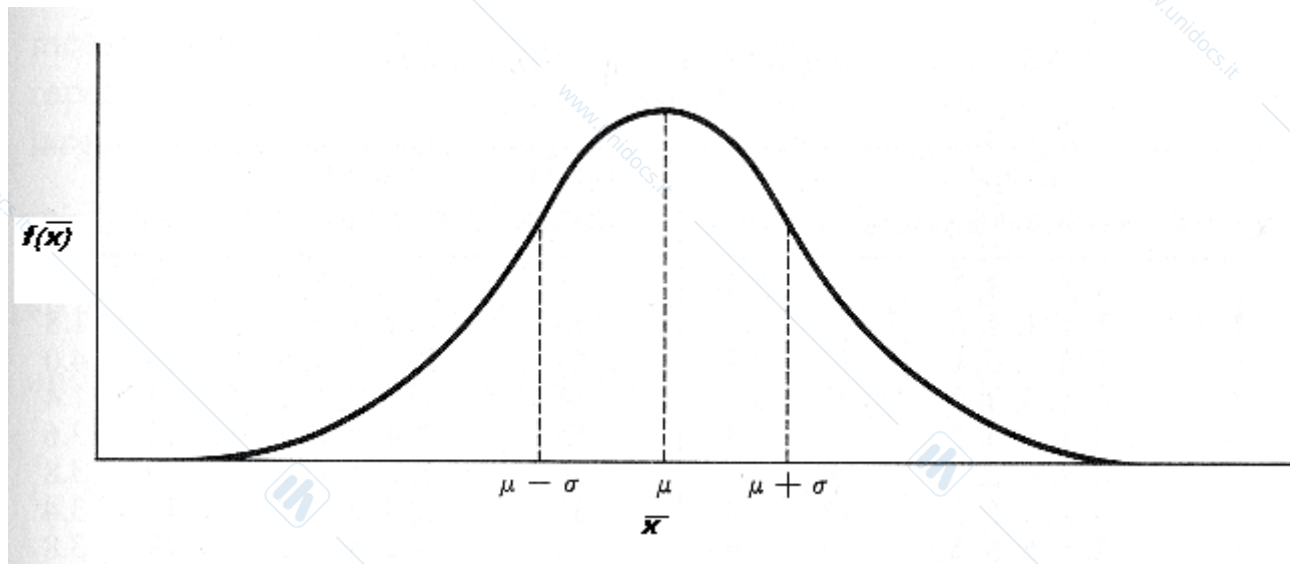
$\bar{x}_5 = 10$
 $S_5 = 1$

$\bar{x}_{200} = 17$
 $S_{200} = 4$

$\bar{x}_n = 16$
 $S_n = 2$

Le stime puntuali variando da campione a campione sono delle variabili e pertanto avranno una propria distribuzione

- Per qualsiasi x si ha
$$\bar{x} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$
- con media pari alla media della popolazione μ
- con varianza pari alla varianza della popolazione, σ^2 , diviso la numerosità del campione.
- $\frac{\sigma}{\sqrt{n}}$ è la radice della varianza campionaria nonché l'errore standard della media campionaria



Stima intervallare

- Le stime di intervallo forniscono informazioni sia sul valore numerico del parametro incognito che sul grado di attendibilità della stima
- occorre determinare, a partire dalle informazioni campionarie, due limiti entro i quali, con una probabilità $1-\alpha$, sarà contenuto il parametro
- $1-\alpha = P(L_1 \leq \theta \leq L_2)$ con $0 \leq \alpha \leq 1$
- con L_1 e L_2 dipendenti dalle informazioni campionarie e dalla distribuzione campionaria del parametro
- $1-\alpha$ livello di confidenza

- Intervallo di confidenza per una media
- Si parte dalla variabile $\bar{x} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$
- E dalla sua standardizzata

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$1 - \alpha = P\left(L_1 \leq \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}} \leq L_2\right)$$

.....

$$1 - \alpha = P\left(\bar{x} - L \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + L \frac{\sigma}{\sqrt{n}}\right)$$

$$1 - \alpha = P\left(\bar{x} - z \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z \frac{\sigma}{\sqrt{n}}\right)$$

- per $1-\alpha=0,95$ (cioè $\alpha = 0,05$) z è pari a 1,96
- per $1-\alpha=0,99$ (cioè $\alpha = 0,01$) z è pari a 2,58

- Intervallo di confidenza per una media con varianza incognita:

- Non si può utilizzare σ ma la sua stima s
- Per cui la distribuzione della statistica non è più una gauss standard ma una t-Student

- Si procede come prima e si giunge a:

$$1 - \alpha = P\left(\bar{x} - t \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t \frac{s}{\sqrt{n}}\right)$$

Intervallo di confidenza per una media

$$1 - \alpha = P\left(\bar{x} - t \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t \frac{s}{\sqrt{n}}\right)$$

Il valore di t si determina in corrispondenza di $\alpha=0.05$ o 0.01 e dei gradi di libertà $n-1$

Verifica di ipotesi

I passi per la verifica delle ipotesi sono i seguenti:

1. Si formulano delle ipotesi sulla popolazione:

- Il diametro medio dei noduli tiroidei in una popolazione di pazienti con gozzo plurinodulare è pari a 1.5 (cm)
- L'età media dei lavoratori edili è di 40 anni
- Il livello di polveri rilevato in 4 zone della città è diverso
- Il rischio di infortuni sul lavoro è maggiore in piccoli cantieri rispetto a grandi cantieri

2. Si valuta la natura dei dati a disposizione nonché la distribuzione teorica sottostante la variabile in studio effettuando una opportuna valutazione delle assunzioni necessarie per procedere con una verifica di tipo parametrico oppure non parametrica.

3. Si traducono tali ipotesi in ipotesi statistiche:

1. H_0 : ipotesi nulla (ipotesi di uguaglianza)
2. H_1 : ipotesi alternativa (ipotesi di diversità)

esempi:

$$H_0: \mu = 1.5$$

$$H_1: \mu > 1.5$$

$$H_0: \mu = 40$$

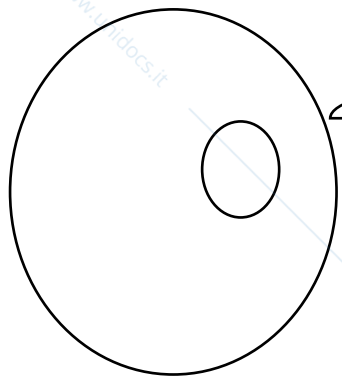
$$H_1: \mu > 40$$

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1: \mu_r \neq \mu_s$$

4. Si effettuano nel campione i calcoli necessari per la verifica delle ipotesi formulate:

- Si calcola il diametro medio dei noduli
- Si calcola l'età media delle donne operate al seno
- Si calcola la dimensione media del tumore
- Si valutano le differenze nelle misure dei due operatori
-
- Si descrive il campione in relazione alle ipotesi che si vogliono verificare



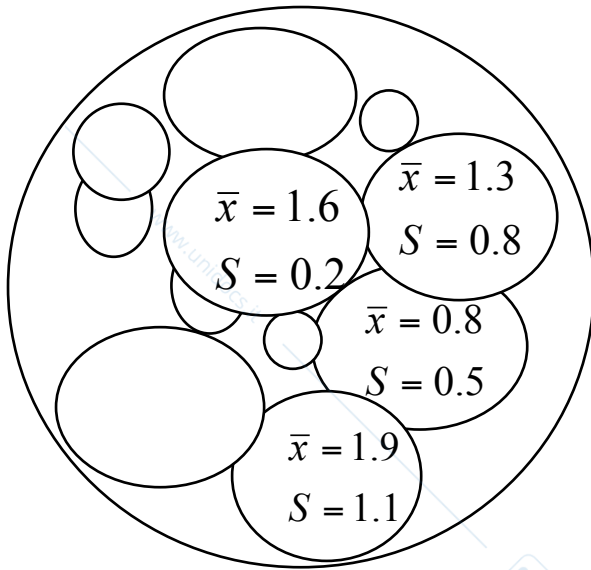
$$\bar{x} = 2$$

$$S = 0.2$$

diametro medio dei
noduli nel campione

4. **Si conduce il TEST STATISTICO appropriato che, tenendo conto che il risultato ottenuto dal campione estratto è uno dei tanti possibili risultati che si potevano ottenere da tutti i possibili campioni, consente di prendere la decisione:**

- **rifiutare H_0** ossia dire che il risultato ottenuto dal campione è significativamente diverso da quello ipotizzato in H_0 , ossia che non è un caso aver avuto un risultato così diverso; **il test è significativo**
- **non rifiutare H_0** ossia dire che non vi è evidenza di una differenza significativa da quanto ipotizzato in H_0 , l'eventuale differenza osservata è da attribuire al caso o a fluttuazioni campionarie; **il test non è significativo**



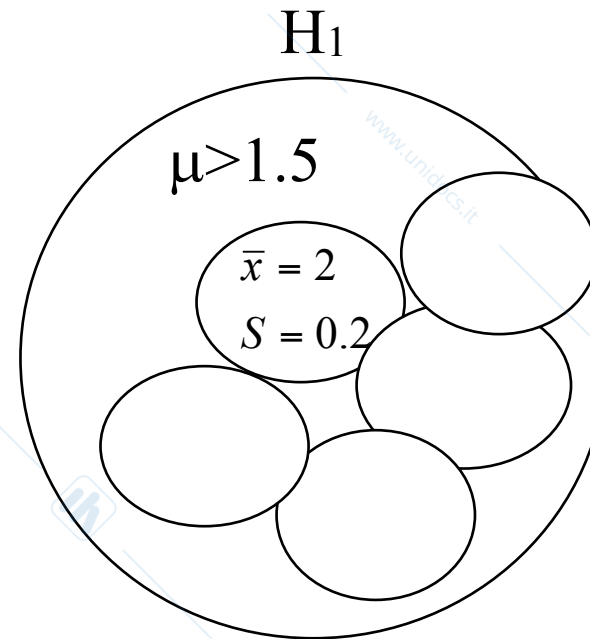
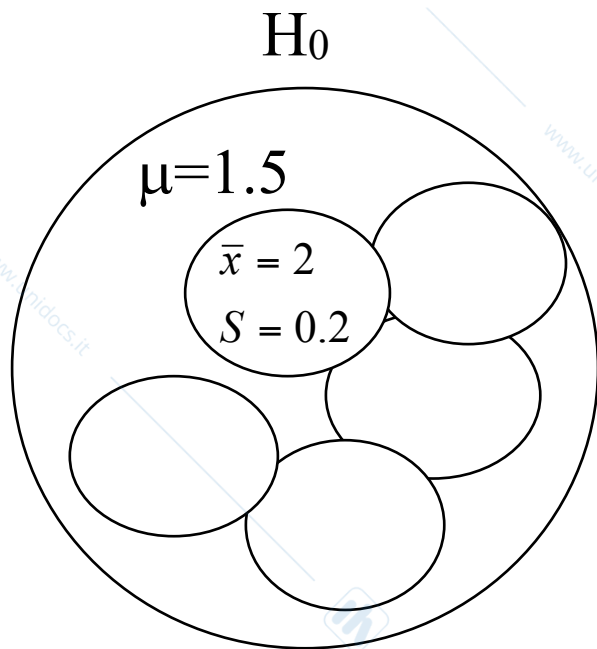
Da una popolazione si possono estrarre infiniti campioni.

Ogni ipotetico campione ha dei valori caratteristici (es. media, varianza) differenti.

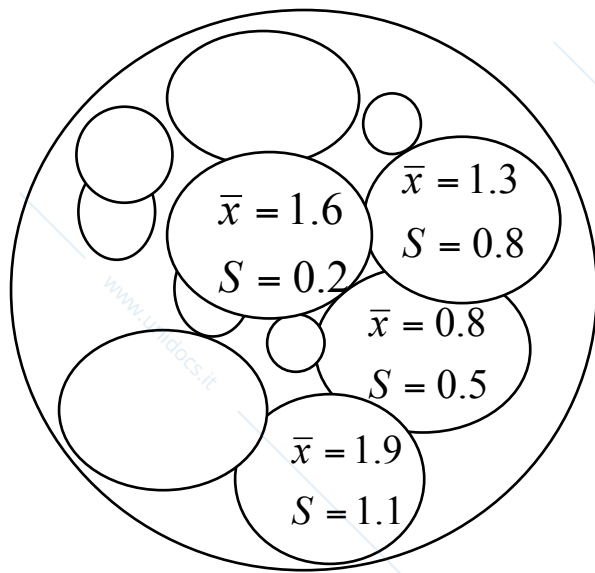
Media e varianza sono variabili con una determinata distribuzione.

Per la media:

$$\bar{x} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$



Prendere una decisione significa capire se il risultato ottenuto dal campione è rappresentativo di una popolazione per la quale sia vera l'ipotesi nulla o di una popolazione per la quale sia vera l'ipotesi alternativa, ovviamente tale decisione è di tipo probabilistico perché consideriamo un solo campione e non tutti i possibili campioni o l'intera popolazione, e la decisione presa può essere affetta da errori.



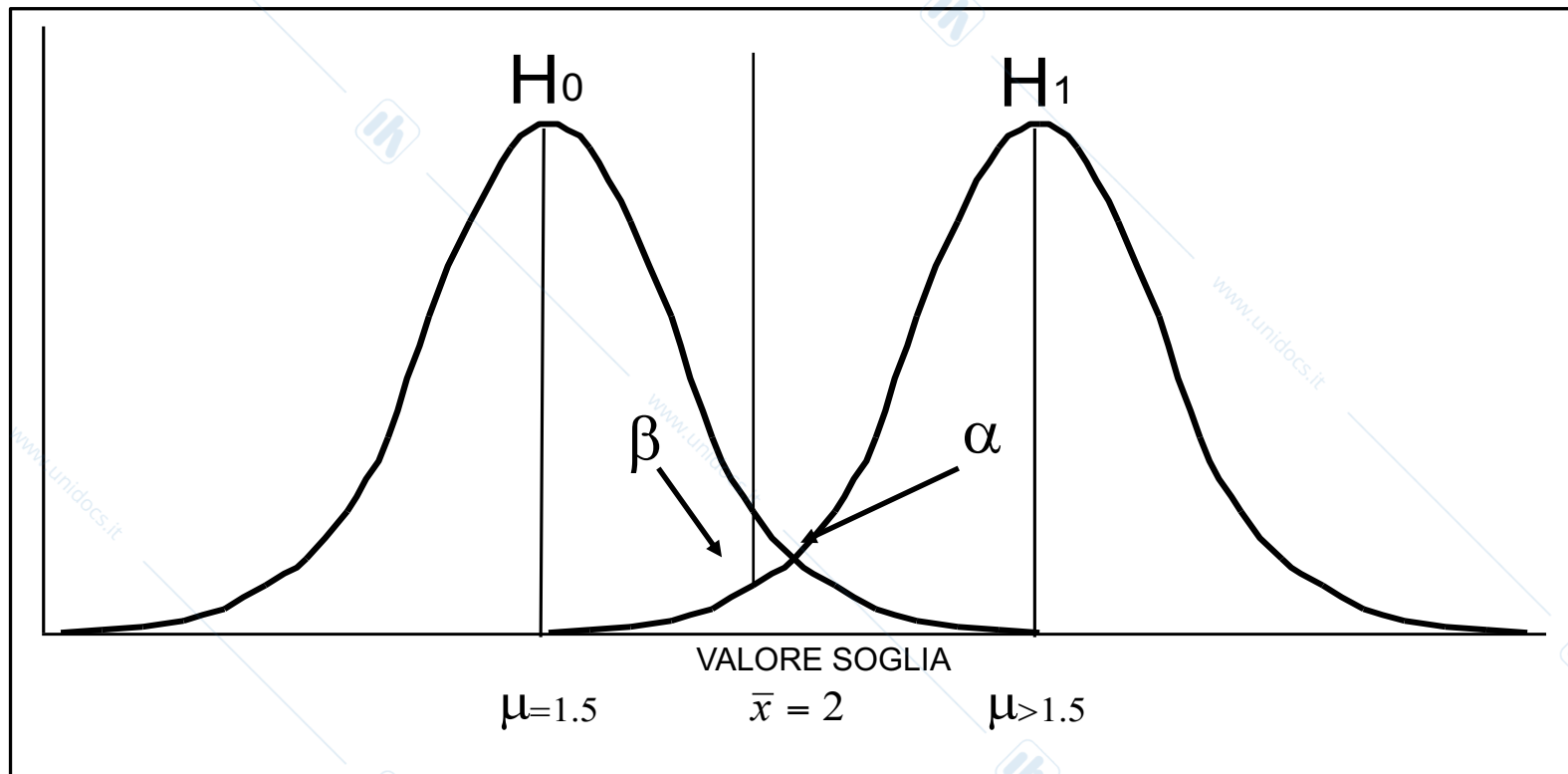
$$\bar{x} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

La decisione statistica deve tener conto della variabilità dei campioni, da qui la necessità di costruire una **statistica test** che, sfruttando le caratteristiche della sua distribuzione di probabilità, consenta di decidere in favore di H_0 o di H_1 sapendo di poter commettere due errori:

α : probabilità di rifiutare H_0 quando è vera (**errore I tipo**)

β : probabilità di accettare H_0 quando è falsa (**errore II tipo**)

Esempio diametro medio dei noduli: partendo dal risultato campionario e dalla distribuzione della media campionaria, ci chiediamo se 2 è più probabile che sia ottenuta da una popolazione con $\mu=1.5$ oppure con $\mu>1.5$



- A questo punto per verificare l'ipotesi:

$$H_0: \mu=1.5 \quad H_1: \mu>1.5$$

occorre:

- costruire una statistica test
- stabilire un valore soglia in corrispondenza di un determinato α (0.05, 0.01, 0.001 ecc.)
- prendere una decisione

- Nel caso di verifica su una media, ipotizzando di conoscere la deviazione standard, la statistica test da utilizzare è:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- posto $\alpha=0.05$ il valore soglia è 1.645
- la regola di decisione sarà che se z calcolato >1.645 si rifiuta H_0

- Nel nostro esempio di un campione di 500 noduli con diametro medio=2, ipotizzando di conoscere la deviazione standard $\sigma=0.4$, la statistica test da utilizzare è:

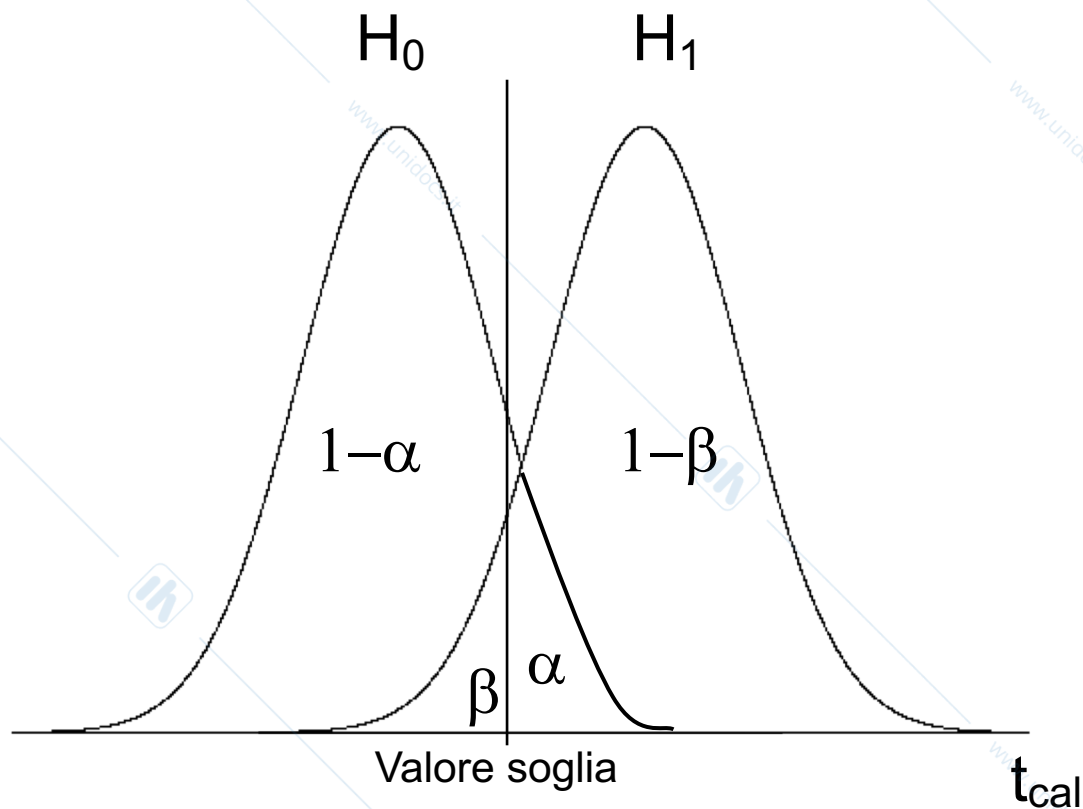
$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{2 - 1.5}{0.4 / \sqrt{500}} = 27.95$$

- Posto $\alpha=0.05$, il valore soglia è 1.645
- Essendo $z=27.95 > 1.645$ si rifiuta H_0 , il test è significativo
- La media della popolazione è significativamente diversa da 1.5 (>1.5)

- Nel caso più concreto di deviazione standard incognita, sfruttando la stima $s=0.2$ la statistica test da utilizzare è:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{2 - 1.5}{\frac{0.2}{\sqrt{500}}} = 55.9$$

- Posto $\alpha=0.05$, e g.l.=499 il valore soglia t è 1.677
- Essendo $t=55.9 > 1.677$ si rifiuta H_0 , il test è significativo
- Il diametro medio dei tumori è significativamente $>$ di 1.5



α : probabilità di rifiutare H_0 quando è vera (**errore I tipo**)

β : probabilità di accettare H_0 quando è falsa (**errore II tipo**)

$1-\alpha$: probabilità di accettare H_0 quando è vera

$1-\beta$: probabilità di rifiutare H_0 quando è falsa
(**potenza del test**)

Test di verifica su medie: un gruppo

- Si formulano le ipotesi

$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

oppure $H_1: \mu < \mu_0$

$$H_1: \mu \neq \mu_0$$

- Si fanno delle assunzioni circa:

- La distribuzione (di Gauss?)
- La varianza (nota?)

- Si costruisce la statistica test

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- Si effettuano i calcoli e si prende la decisione: se $t_{\text{calcolato}} > t_{\text{tabulato}}$ si rifiuta H_0 , il test è significativo

“Confrontiamo due o più gruppi”

- Test di verifica su medie: due gruppi.
- Test di verifica per misure ripetute (due osservazioni ripetute)
- Test di verifica su medie: più gruppi (analisi della varianza).
- Test non parametrici alternativi

- **Quando si vogliono confrontare due o più gruppi occorre valutare se:**
 1. i campioni sono indipendenti o dipendenti
 2. la distribuzione della variabile è di Gauss
 3. altre condizioni

1. Campioni indipendenti o dipendenti?

- Campioni indipendenti: si tratta di unità distinte classificate in
 - diverse condizioni sperimentali
 - diverse patologie
 - diverse terapie
 - diverso sesso
 - diverse classi di età
 - diversi luoghi

- **Campioni dipendenti: si tratta delle stesse unità misurate**
 - in diverse condizioni sperimentali
 - in diversi tempi (prima e dopo, t_0 - t_6 - t_{12} ...)
 - con diversi strumenti
 - da diversi operatori

2. Verifica della normalità

- Esplorazione grafica (box-plot, steam&leaf, Q-Qplot ecc.
- Verifica normalità mediante opportuni test:
 - Kolmogorov-Smirnov
 - Shapiro-Wilks
 - Cramer-Von Mises
 - Anderson-Darling

**Ricordate che qualsiasi test utilizzate
l'ipotesi nulla è l'ipotesi di normalità**

- Se la distribuzione è di Gauss, o è possibile effettuare una trasformazione “normalizzante” o comunque, se il campione è sufficientemente grande, si può procedere con un test parametrico, altrimenti sarà preferibile optare per un test non parametrico.

Altre condizioni da verificare

- Quando si fanno dei confronti fra gruppi (due o più gruppi) è necessario verificare preliminarmente se le varianze sono omogenee oppure no e tenerne conto nel test.

Test T-Student per campioni indipendenti

- L'ipotesi da verificare è

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2 \quad \text{oppure}$$

$$H_1: \mu_1 < \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

- Il test da utilizzare è:

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

Che si distribuisce come un t-Student con $n_1 + n_2 - 2$ gradi di libertà

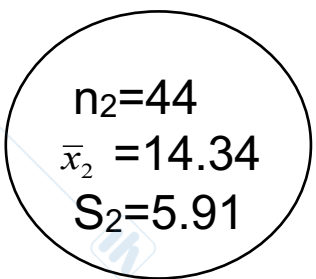
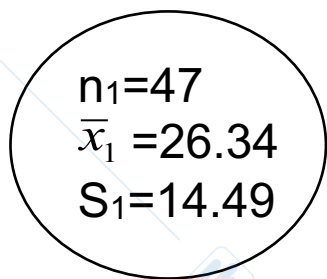
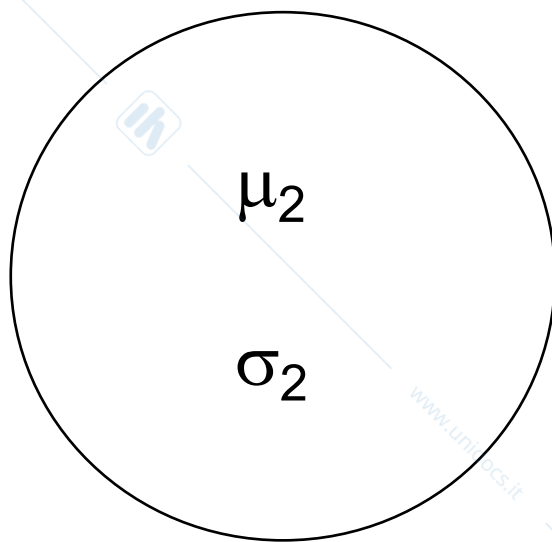
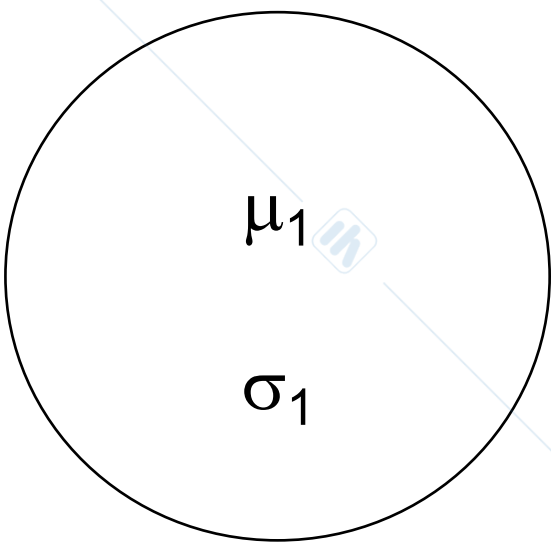
con
$$S_p^2 = \frac{S_1^2(n_1 - 1) + S_2^2(n_2 - 1)}{n_1 + n_2 - 2} =$$

Quando le due varianze sono uguali, altrimenti S_p^2 è sostituito dalle due varianze singole e cambiano i g.l.

- posto $\alpha=0.05$ e g.l.= n_1+n_2-2 , si determina il valore soglia (t_{tab})
- la regola di decisione sarà che se t calcolato $>t_{tab}$ si rifiuta H_0
- in questo caso $p<0.05$, il test è significativo

Esempio 1: diametro noduli tiroidei

Si vuole verificare se in media il diametro di noduli singoli sia uguale al diametro di noduli multipli



1. Si estrae un campione dalla popolazione di noduli singoli ed uno dalla popolazione di noduli multipli
2. Si tratta di due campioni indipendenti
3. Si verifica la normalità del diametro
4. Si stimano i parametri media e dev. Standard
5. L'ipotesi da verificare è

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

- Il test da utilizzare è:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}} = \frac{(26.34 - 14.34) - 0}{\sqrt{\frac{125.39}{47} + \frac{125.39}{44}}} = 5.11$$

- Che si distribuisce come un t-Student con 89 gradi di libertà
- posto $\alpha=0.05$ e g.l.= 89, si determina il valore soglia ($t_{\text{tab}}=1.96$)
- essendo t calcolato $>t_{\text{tab}}$ si rifiuta H_0
- in questo caso $p<0.05$, il test è significativo
- Il diametro medio dei noduli è significativamente diverso nelle due popolazioni

- In realtà in questo caso le varianze sarebbero risultate diverse al test di omogeneità e quindi avremmo dovuto utilizzare un test t-Student per varianze non omogenee.
- I software forniscono i risultati di entrambi i test oltre che il test F di Fisher per la verifica dell' omogeneità delle varianze

CONFRONTO DIAMETRO NODULI PATOLOGICI VS NON PATOLOGICI

La procedura TTEST

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Variabile	ist2	N	Media		Media LC superiore	Dev std LC inferiore
			LC inferiore	Media		
diam	0	3205	16.654	17.021	17.388	10.346
diam	1	107	14.129	15.972	17.815	8.4781
diam	Diff (1-2)		-0.988	1.0489	3.0855	10.321

Variabile	ist2	N	Dev std		Err std	Minimo	Massimo
			LC superiore	Dev std			
diam	0		10.866	0.1872	3	80	
diam	1		11.111	0.9297	4	54	
diam	Diff (1-2)		10.83	1.0387			

Variabile	Metodo	Varianze	DF	Test T	
				Valore t	Pr > t
diam	Pooled	Equal	3310	1.01	0.3126
diam	Satterthwaite	Unequal	115	1.11	0.2710

Uguaglianza di varianze

Variabile	Metodo	Num DF	Den DF	Uguaglianza di varianze		NON SI DI UGU VARIAN
				Valore F	Pr > F	
diam	Folded F	3204	106	1.21	0.1907	

$H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 \neq \mu_2$

Variabile	alone	N	Media LC inferiore	Media	Media LC superiore	Dev std LC inferiore	Dev s
diam	N	2128	14.33	14.747	15.163	9.5167	9.80
diam	Y	1257	20.217	20.818	21.419	10.454	10.1
diam	Diff (1-2)		-6.783	-6.071	-5.359	9.9716	10.4

Statistiche

Variabile	alone	Dev std LC superiore	Err std	Minimo	Massimo
diam	N	10.106	0.2125	3	80
diam	Y	11.305	0.3064	4	72
diam	Diff (1-2)	10.458	0.3632		

Test T

Variabile	Metodo	Varianze	DF	Valore t	Pr > t	
diam	Pooled	Equal	3383	-16.72	<.0001	SI LE
diam	Satterthwaite	Unequal	2424	-16.28	<.0001	

Uguaglianza di varianze

Variabile	Metodo	Num DF	Den DF	Valore F	Pr > F	SI RIFIUTA UGUAGLIANZA DI VARIANZE
diam	Folded F	1256	2127	1.23	<.0001	

variabile	$H_1: \mu_1 \neq \mu_2$	N	LC inferiore	Media inferiore	Media	LC superiore	Media superiore	Dev std inferiore	Dev std
diam	N	3034		16.182	16.547		16.913	10.028	10.281
diam	Y	359		19.645	20.939		22.233	11.617	12.467
diam	Diff (1-2)			-5.544	-4.391		-3.239	10.288	10.533

Statistiche

Variabile	microcal	Dev std LC superiore	Err std	Minimo	Massimo
diam	N	10.546	0.1866	3	80
diam	Y	13.452	0.658	4	70
diam	Diff (1-2)	10.79	0.5879		

Test T

Variabile	Metodo	Varianze	DF	Valore t	Pr > t
diam	Pooled	Equal	3391	-7.47	<.0001
diam	Satterthwaite	Unequal	418	-6.42	<.0001

Uguaglianza di varianze

Variabile	Metodo	Num DF	Den DF	Valore F	Pr > F
diam	Folded F	358	3033	1.47	<.0001

Test non parametrico per campioni indipendenti: test di Wilcoxon

- L'ipotesi da verificare è

H_0 : i due gruppi sono uguali

H_1 : i due gruppi sono diversi

- Come si procede?

- si crea un campione combinato ponendo le osservazioni in ordine non decrescente
- si assegna il rango a tutte le osservazioni
- si calcola la somma dei ranghi distinta per i due campioni
- si confronta la somma dei ranghi del campione più piccolo con l'intervallo tabulato
- Si rifiuta H_0 se il valore è esterno

- **Esistono delle varianti:**
 - Test di U Mann Withney
 - approssimazione di z sulla somma dei ranghi
 - ed altri sempre basati sui ranghi

Esempio 1: diametro dei noduli

1. Si estrae un campione dalla popolazione di noduli singoli ed uno dalla popolazione di noduli multipli
2. Si tratta di due campioni indipendenti
3. Si verifica la “non normalità” del diametro
4. Si crea il campione combinato, ponendo in ordine crescente i dati relativi ai due gruppi
5. Si assegnano i ranghi
6. Si calcola la somma dei ranghi del campione più piccolo e si confronta con l'intervallo tabulato
7. Se il valore è interno si accetta H_0 .

gruppo	diametro	rango
2	6	1,5
2	6	1,5
2	7	3,5
2	7	3,5
2	8	5,5
2	8	5,5
2	9	8,0
2	9	8,0
2	9	8,0
1	10	11,5
1	10	11,5
1	10	11,5
1	10	11,5
1	11	15,0
2	11	15,0
2	11	15,0
1	12	18,0
1	12	18,0
2	12	18,0
2	13	21,0
2	13	21,0
2	13	21,0
2	.	.
2	.	.

$$\Sigma R_1 = 2756.5$$

$$\Sigma R_2 = 1429.5$$

Si confronta la somma del campione più piccolo con l'intervallo
In questo caso si rifiuta H_0 .

ist2	N	Somma degli score	Expected Under H0	Dev std sotto H0	Score medio
0	3205	5319422.0	5309082.50	9718.70488	1659.72605
1	107	166906.0	177245.50	9718.70488	1559.86916

Sono stati utilizzati score medi per i legami.

Test a due campioni di Wilcoxon

Statistica	166906.0000
Approssimazione normale	
Z	-1.0638
Pr unilaterale < Z	0.1437
Pr bilaterale Pr > Z	$p = 0.2874$
Approssimazione t	
Pr unilaterale < Z	0.1437
Pr bilaterale Pr > Z	0.2875

NON SI RIFIUTA H0
I CAMPIONI SONO UC

Z include una correzione di continuità di 0.5.

Test di Kruskal-Wallis

Chi-quadrato	1.1318
DF	1
Pr > Chi-quadrato	0.2874

Score di Wilcoxon (somme per ranghi) per la variabile diam
Classificata dalla variabile alone

alone	N	Somma degli score	Expected Under H0	Dev std sotto H0	Score medio
N	2128	3058068.0	3602704.0	27439.6501	1437.06203
Y	1257	2672737.0	2128101.0	27439.6501	2126.28242

Sono stati utilizzati score medi per i legami.

Test a due campioni di Wilcoxon

Statistica 2672737.0000

Approssimazione normale

Z 19.8485

Pr unilaterale > Z <.0001

Pr bilaterale Pr > |Z| p <.0001

SI RIFUTA H0: I DUE CAMPIO DIVERSI

Approssimazione t

Pr unilaterale > Z <.0001

Pr bilaterale Pr > |Z| <.0001

Z include una correzione di continuità di 0.5.

Test di Kruskal-Wallis

Chi-quadrato 393.9631

DF 1

Pr > Chi-quadrato <.0001



**Score di Wilcoxon (somme per ranghi) per la variabile diam
Classificata dalla variabile microcal**

microcal	N	Somma degli score	Expected Under H0	Dev std sotto H0	Score medio
Y	359	731703.0	609223.0	17530.5286	2038.16992
N	3034	5026218.0	5148698.0	17530.5286	1656.63085

Sono stati utilizzati score medi per i legami.

Test a due campioni di Wilcoxon

Statistica	731703.0000
Approssimazione normale	
Z	6.9866
Pr unilaterale > Z	<.0001
Pr bilaterale Pr > Z	<.0001
Approssimazione t	
Pr unilaterale > Z	<.0001
Pr bilaterale Pr > Z	<.0001

Z include una correzione di continuità di 0.5.

Test di Kruskal-Wallis

Chi-quadrato	48.8135
DF	1
Pr > Chi-quadrato	<.0001

Il caso di due misure ripetute

- Come prima valutiamo se:
 1. i campioni sono indipendenti o dipendenti
 2. la distribuzione della variabile è di Gauss

- **Campioni dipendenti: si tratta delle stesse unità misurate due volte**
 - in due condizioni sperimentali
 - in diversi tempi (prima e dopo)
 - con due strumenti
 - da due operatori
- **Trattandosi di stessi individui ha senso valutare le differenze dei valori**
 - Tra le due condizioni sperimentali
 - Tra prima e dopo
 - Tra i due strumenti
 -

- Ciò che vogliamo verificare è se le due situazioni sono sovrapponibili o meno
- Pertanto lavoreremo sulle differenze e ci chiederemo se in media le due serie di dati sono uguali oppure no

Test T-Student per campioni dipendenti

- L'ipotesi da verificare è

$$H_0: \mu_d = 0$$

$$H_1: \mu_d > 0 \quad \text{oppure} \quad H_1: \mu_d < 0$$

$$H_1: \mu_d \neq 0$$

- Il test da utilizzare è:

$$t = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}}$$

Che si distribuisce come un t-Student con n-1 gradi di libertà

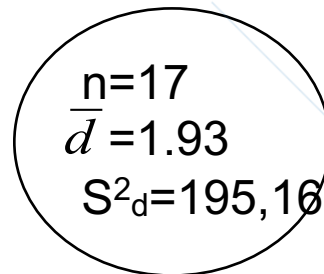
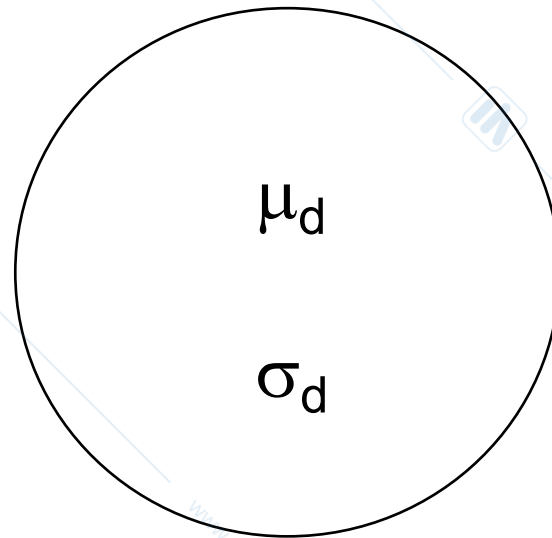
con
$$\bar{d} = \frac{\sum d_i}{n}$$

$$S_d^2 = \frac{\sum (d_i - \bar{d})^2}{n - 1}$$

- posto $\alpha=0.05$ e g.l.= $n-1$, si determina il valore soglia (t_{tab})
- la regola di decisione sarà che se t calcolato $>t_{\text{tab}}$ si rifiuta H_0
- in questo caso $p<0.05$, il test è significativo

Esempio 1: diametro nodulo

Si vuole verificare se le misure del diametro del nodulo effettuate da due operatori diversi sono le stesse oppure no



1. Si estrae un campione di soggetti da una popolazione e si fa misurare il diametro dei noduli da due operatori
2. Si tratta di due campioni dipendenti o appaiati
3. Si effettuano le differenze (op.1-op.2)
4. Si verifica la normalità della differenza dei diametri
5. Si stimano i parametri media e dev. Standard della differenza
6. L'ipotesi da verificare è
 $H_0: \mu_d = 0$
 $H_1: \mu_d \neq 0$

- Il test da utilizzare è:

$$t = \frac{\bar{d} - \mu_d}{\sqrt{\frac{S_d^2}{n}}} = \frac{1.93 - 0}{\sqrt{\frac{195.16}{17}}} = 0.571$$

- Che si distribuisce come un t-Student con 16 gradi di libertà
- posto $\alpha=0.05$ e g.l.= 16, si determina il valore soglia ($t_{\text{tab}}=2.12$)
- essendo $|t \text{ calcolato}| < |t_{\text{tab}}|$ si accetta H_0
- in questo caso $p>0.05$, il test non è significativo
- la misura del diametro del nodulo effettuata dai due operatori è in media la stessa.

Test non parametrico per campioni appaiati: test dei ranghi con segno di Wilcoxon

- L'ipotesi da verificare è

H_0 : i due gruppi sono sovrapponibili

H_1 : i due gruppi non sono sovrapponibili

- Come si procede?

- Si effettuano le differenze tra le coppie di valori
- Si eliminano le differenze nulle (riducendo la dim del campione)
- si assegna il rango alle differenze private del segno
- si calcola la somma dei ranghi assegnati alle differenze positive
- si calcola la somma dei ranghi assegnati alle differenze negative
- si confrontano le somme dei ranghi ottenute con i valori tabulati (dipende dalle tavole)
- Si rifiuta H_0 se le somme sono esterne

Esempio 1: diametro noduli

1. Si estrae un campione di soggetti da una popolazione e si fa misurare il diametro dei noduli da due operatori
2. Si tratta di due campioni dipendenti o appaiati
3. Si effettuano le differenze (op.1-op.2)
4. Si verifica la non normalità delle differenze
5. Si assegnano i ranghi alle differenze senza segno
6. Si effettua la somma dei ranghi assegnati a differenze positive ed a quelle negative
7. si confrontano entrambe le somme con i valori tabulati
8. se le somme cadono nell' intervallo si accetta H_0

$$\Sigma \text{ranghi } + = 82.5$$

$$\Sigma \text{ranghi } - = 48$$

in questo caso si accetta H_0

Verifica su più gruppi

- Occorre valutare se:
 1. i campioni sono indipendenti
 2. la distribuzione della variabile è di Gauss
 3. I campioni hanno varianze omogenee (test di Bartlett o altro)

Verifica su più gruppi indipendenti

- I gruppi da porre a confronto possono essere costituiti da
 - Osservazioni campionarie provenienti da popolazioni differenti (ad es. diverse patologie, diverse sedi, diverso grado istologico)
 - Dati risultanti da un esperimento controllato che prevede diversi trattamenti (piano completamente randomizzato)

• In entrambi i casi si può ipotizzare una variabile (ad es. durata degenza per patologia) che può dipendere da un fattore (diversi ospedali oppure diversi protocolli terapeutici) con diversi livelli

• In entrambi i casi il procedimento da scegliere è quello dell'analisi della varianza ad una via (ad 1 solo fattore)

• In entrambi i casi l'ipotesi da verificare è

- $H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$
- $H_1: \mu_r \neq \mu_s$

Domande ricorrenti

1. Perché non posso fare il confronto delle medie a due a due?
2. Perché per fare un confronto fra le medie di più gruppi devo fare una analisi sulle varianze?

Risposta:

1. Perché il livello reale di significatività complessiva (experiment-wise) non sarebbe quello specificato per ciascun set (comparison-wise) ma sarebbe considerevolmente maggiore.
2. Perché dalla valutazione della variabilità metto in evidenza le differenze tra i gruppi

2. Partiamo dalle ipotesi

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_r \neq \mu_s$$

H_0 indica che i gruppi sono uguali ovvero che non c'è "variabilità" tra i gruppi (between)

ma se non c'è variabilità fra i gruppi le osservazioni sono tutte uguali? Impossibile c'è sempre una variabilità biologica (within)

L'analisi della varianza ci consente di valutare se, date le misure rilevate su diversi gruppi, prevale la variabilità tra gruppi (dovuta al fattore) o la variabilità entro i gruppi (variabilità biologica)

gruppi

Osservaz.	1	2	...	i	...	k
1	y_{11}	y_{12}	...	y_{1i}	...	y_{1k}
2	y_{21}	y_{22}	...	y_{2i}	...	y_{2k}
...
j	y_{j1}	y_{j2}	...	y_{ji}	...	y_{jk}
...
n_i	y_{ni1}	y_{ni2}	...	y_{nii}	...	y_{nik}

ospedali

osservazione	A	B	...	i	...	G
1	3	2	...	4	...	2
2	5	3	...	4	...	3
...
j	2	4	...	5	...	4
...
10	2	2	...	3	...	3

Costruiamo la statistica test

Partendo dall'ipotesi nulla di uguaglianza consideriamo una media totale ed una variabilità totale che in termini di devianza è:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

Che esprime la variabilità del fenomeno.

Si può ipotizzare che a tale variabilità concorrano due fonti di variabilità:

una variabilità dovuta al fattore (ospedale, laboratorio, patologia, sede, terapia...)

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2$$

Una variabilità biologica (variabilità d' errore, variabilità biologica)

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)$$

Si dimostra che la variabilità totale è data dalla somma delle due variabilità

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}) = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i) + \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2$$

Con g. l.

N-1

N-k

k-1

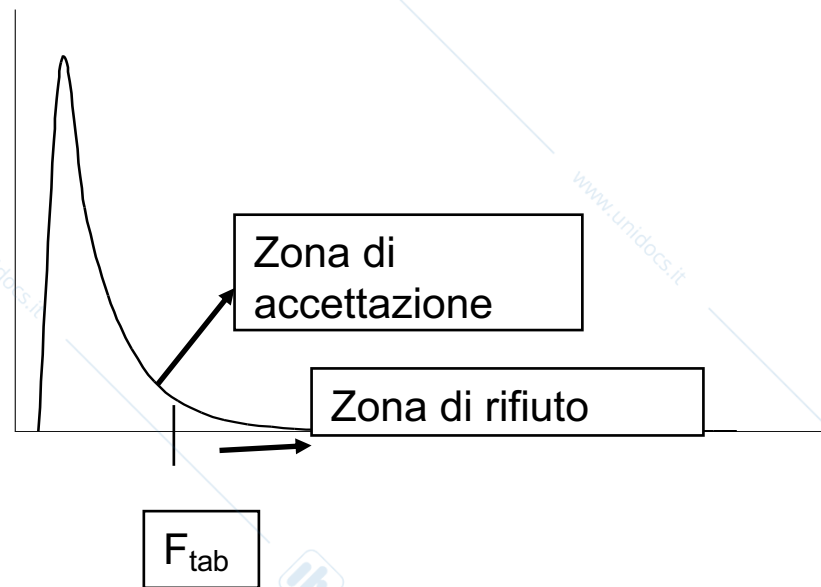
A questo punto occorre valutare se prevale la differenza tra i gruppi (le medie sono differenti) o quella entro i gruppi.

A tal fine si confrontano le due varianze costruendo la statistica test

$$F = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2 / k - 1}{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 / N - k} = \frac{\text{varianza tra gruppi}}{\text{varianza entro gruppi}}$$

Che sotto H_0 si distribuisce come una F di Fisher con $k-1$ ed $N-k$ gradi di libertà

- posto $\alpha=0.05$
- in corrispondenza dei gradi di libertà del numeratore e del denominatore si determina un valore tabulato che delimita la zona di rifiuto.
- la regola di decisione sarà che se F calcolato $>F_{\text{tab}}$ si rifiuta H_0
- in questo caso $p<0.05$, il test è significativo



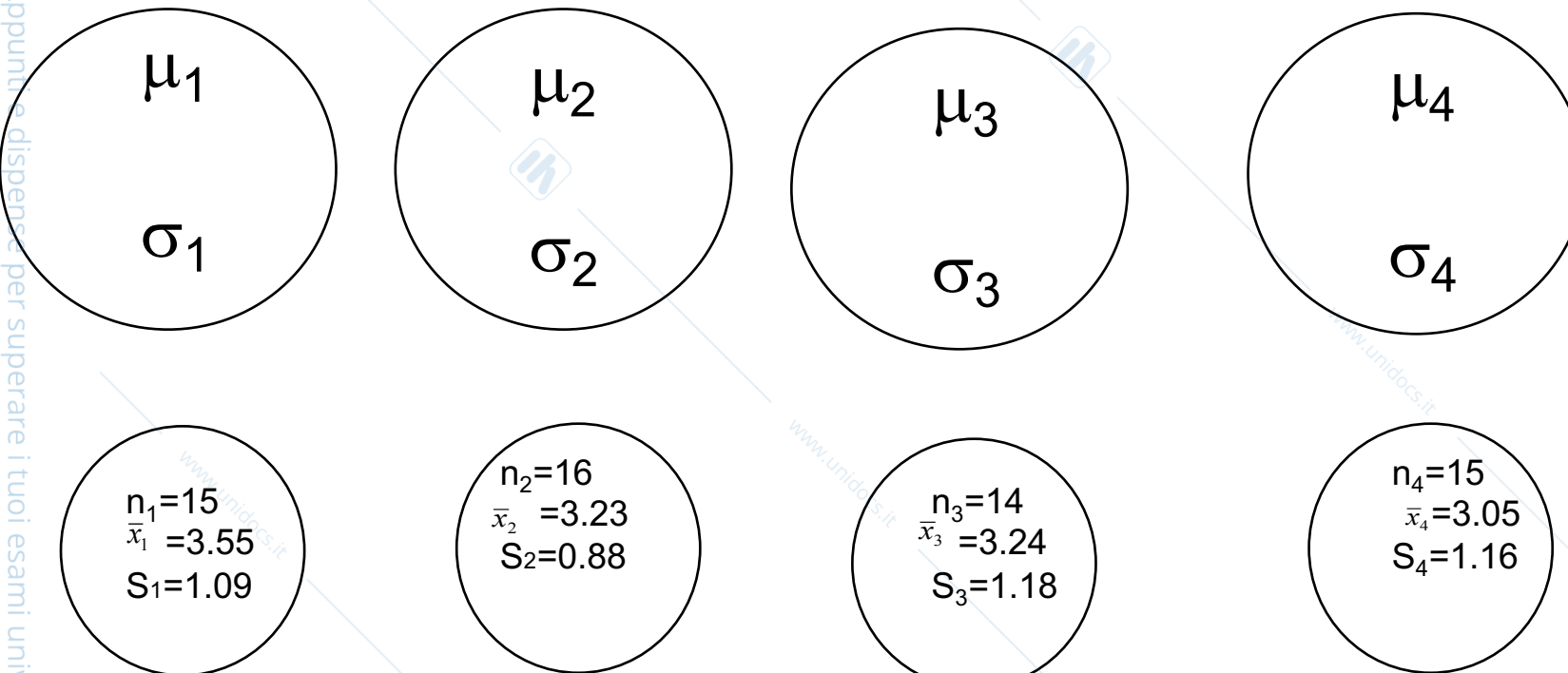
ANOVA

Sorgenti di variazione	devianze	g.l.	varianze	F
Tra gruppi	$\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2$	K-1	$\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2}{k-1}$	$\frac{\text{var tra gruppi}}{\text{var entro gruppi}}$
Entro gruppi	$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	N-k	$\frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{N-k}$	
Totale	$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_i - \bar{y})^2$	N-1		

Se $F_{\text{calc}} > F_{\text{tab}}$ si rifiuta l'ipotesi nulla

Esempio 1: durata degenza

Si vuole verificare se la durata della degenza per una stessa patologia sia uguale in 4 diversi ospedali



1. Si estrae un campione dalla popolazione di ricoverati per la patologia in ogni ospedale
2. Si tratta di quattro campioni indipendenti
3. Si verifica la normalità della distribuzione della durata del ricovero
4. Si verifica l'omogeneità delle 4 varianze
5. Si procede con la scomposizione della varianza (ANOVA)
6. Si costruisce la statistica F
7. Si prende la decisione

- l'ipotesi da verificare è

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$$H_1: \mu_r \neq \mu_s$$

- Si calcolano le devianze

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2 = 1.94$$

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 = 65.19$$

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = 67.13$$

• Si costruisce la tabella ANOVA

Sorgenti di variazione	devianze	g.l.	varianze	F
Tra gruppi	1.94	3	0.65	
Entro gruppi	65.19	56	1.16	0.55
Totale	67.13	59		

Essendo $F=0.55 < F_{0.05,3,56}=2.76$

Si accetta l'ipotesi nulla di uguaglianza delle medie

Group	N (count)	Mean	Std. Dev.
1:	15	3.55	1.09
2:	16	3.23	0.88
3:	14	3.24	1.18
4:	15	3.05	1.16
5:			
6:			
7:			
8:			

Compute

ANOVA Table...

Source of Variation	Sum of Squares	d.f.	Variance	F	p
Between Groups:	1.9398	3	0.6466	0.5554	0.6466
Within Groups:	65.1890	56	1.1641		
Total:	67.1288	59			

Coming soon -- post-hoc tests to determine which specific groups are significantly different from each other.

Sample Size(i)	15	16	14	15						
Sample Mean(i)	3.55	3.23	3.24	3.05						
Sample Variance(i)	1.19	0.77	1.39	1.34						

CALCULATE: TEST FOR EQUALITY OF MEANS		CLEAR	
Variation Between	0.6465911	Variation Within	1.1614286
F-Statistic	0.5567205	P-Value	0.64629
Conclusion			
Little or no real evidences against the null hypothesis			<input type="button" value="▲"/> <input type="button" value="▼"/>

For Technical Details, Back to:
[Statistical Thinking for Decision Making](#)

Kindly [email](#) your comments to:
[Professor Hossein Arsham](#)

Cosa succede quando la F è significativa?

- Indica semplicemente che le medie non sono tutte uguali
- Come individuare i livelli del fattore che si differenziano?

- Si effettuano delle comparazioni multiple
- Si distinguono due tipi:
 - **Comparazioni prepianificate:** sono generate a priori sull' esperimento
 - **Comparazioni a posteriori (post-hoc):** scaturiscono dal risultato dell' analisi
- Quando possibile è sempre preferibile utilizzare le prime anziché le seconde.
- I contrasti pianificati consentono una ulteriore scomposizione della variabilità tra gruppi (rientrano nel modello)

- I post-hoc test, tuttavia, sono molto utilizzati e ne esistono di vario tipo
- Tutti hanno come obiettivo il compromesso tra controllare l' experiment-wise e l' eccessiva conservatività (quando non si rifiuta H_0)
- I più noti sono:
 - LSD di Fisher :calcola la differenza minima significativa(Least Significant Difference) e confronta tutte le differenze con la LSD. Basso controllo dell' experiment-wise.

- La procedura di Tukey che si basa sul campo di variazione studentizzato
- Test del campo di variazione multiplo di Duncan, uno dei più utilizzati.
- La procedura di Scheffè che permette un numero qualsiasi di contrasti ma è il più conservativo

La procedura preliminare di analisi della varianza pur non essendo sempre necessaria è **ASSOLUTAMENTE CONSIGLIATA** per evitare di avere risultati apparentemente contraddittori. Meglio ancora i contrasti pianificati!

Test non parametrico per campioni indipendenti: test di Kruskal Wallis

- L'ipotesi da verificare è
H₀: i gruppi sono uguali
H₁: i gruppi non sono tutti uguali

• Come si procede?

- si crea un campione combinato ponendo le osservazioni in ordine non decrescente
- si assegna il rango a tutte le osservazioni
- si calcola la somma dei ranghi distinta per i K campioni (T_i)

- Si calcola

$$H = \frac{12}{n(n+1)} \sum \frac{T_i^2}{n_i} - 3(n+1)$$

- Si rifiuta H₀ se $H > \chi^2$ con k-1 g.l.

Modello lineare generale

- Il confronto fra le medie può essere espresso con un modello del tipo

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

dove

y_{ij} = j-mo valore campionario osservato nella i-ma popolazione

μ = livello di base (media globale)

τ_i = parametro che misura l'effetto attribuibile al trattamento i-mo

ε_{ij} = deviazione del j-mo valore osservato dalla media della i-ma popolazione

pertanto l'ipotesi di uguaglianza delle medie (piano completamente randomizzato) coincide con quella sull'effetto del trattamento che diventa:

$H_0: \tau_i = 0$ per tutte le i

$H_1: \tau_i \neq 0$ per alcune i

Analisi della varianza a due vie

Consente di analizzare quanto due fattori influiscono simultaneamente sulla variabile di interesse

Il modello in questo caso può esprimersi come:

$$y_{ijl} = \mu + \tau_i + \beta_j + \tau\beta_{ij} + \varepsilon_{ijl} \quad \text{con} \quad \begin{array}{l} i=1 \dots p; j=1 \dots q; \\ l=1 \dots n \end{array}$$

- dove

y_{ijl} = l-mo valore campionario osservato nella i-ma popolazione trattata con il j-esimo trattamento

μ = livello di base (media globale)

τ_i = parametro che misura l'effetto attribuibile al livello i-mo del fattore T

β_j = parametro che misura l'effetto attribuibile al livello j-mo del fattore B

$\tau\beta_{ij}$ = effetto dell'interazione dei livelli T_i e B_j (effetto simultaneo)

ε_{ijl} = deviazione del l-mo valore osservato dalla media del i-mo livello del fattore T e del j-mo livello del fattore B

Con questo disegno è possibile verificare le ipotesi

H_0 : le medie dei livelli del fattore T sono uguali

H_0 : le medie dei livelli del fattore B sono uguali

H_0 : le medie delle diverse interazioni dei livelli di T e B sono uguali

Esempio: riduzione emoglobina in funzione di due fattori: tipo di carcinoma e terapia

cervicale		Prostata		Colon-retto	
Terapia 1	Terapia 2	Terapia 1	Terapia 2	Terapia 1	Terapia 2
1.7	2.3	2.7	0.6	-0.3	1.6
-0.2	1.2	1.6	1.7	1.9	-2.2
1.7	-0.6	2.5	0.8	1.7	1.9
2.3	1.3	0.5	1.7	0.5	-1.6
2.7	-1.1	2.6	1.4	2.1	0.8
.
.

come prima bisogna procedere alla scomposizione della devianza totale

questa volta le fonti di variazione sono:

- i livelli del fattore T
- i livelli del fattore B
- le combinazioni dei livelli dell'interazione
- l'errore

occorrerà calcolare le rispettive varianze e le statistiche F per verificare le varie ipotesi

ANOVA

Sorgenti di variazione	devianze	g.l.	varianze	F	p
trattamenti	5.27	1	5.27	4.11	0.049*
tipo di carcinoma	9.11	2	9.11	3.55	0.038*
trat*carc	0.92	2	0.92	0.36	0.702
errore	53.88	42	53.88		
totale	69.18	47			

INTERPRETAZIONE OUTPUT

- Si verifica preliminarmente l'interazione che in questo caso non è significativa
- Si valutano i fattori principali (main effect): risulta significativo l'effetto del trattamento e l'effetto del tipo di carcinoma
- Conclusioni: la riduzione dell'emoglobina dipende sia dal trattamento che dal tipo di carcinoma, non si evidenzia un effetto combinato trattamento*tipo di carcinoma.

Si vuole verificare se la variazione nella composizione minerale ossea dopo 6 mesi di terapia sia significativamente diversa nei 3 gruppi (Controllo, Dieta, Terapia Ormonale)

Fattori tra soggetti

		N
gruppo	C	35
	D	22
	TO	36

Test degli effetti fra soggetti

Variabile dipendente: moctdif

Sorgente	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
Modello corretto	,003 ^a	2	,002	,176	,839
Intercetta	,020	1	,020	2,154	,146
gruppo	,003	2	,002	,176	,839
Errore	,815	90	,009		
Totale	,836	93			
Totale corretto	,818	92			

Fattori tra soggetti

		N
gruppo	C	35
	D	22
	TO	36

Test degli effetti fra soggetti

Variabile dipendente: moccdif

Sorgente	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
Modello corretto	,010 ^a	2	,005	,495	,611
Intercetta	,000	1	,000	,026	,873
gruppo	,010	2	,005	,495	,611
Errore	,924	90	,010		
Totale	,934	93			
Totale corretto	,934	92			

Si vuole verificare se la variazione nella composizione minerale ossea dopo 6 mesi di terapia dipenda dal gruppo (Controllo, Dieta, Terapia Ormonale)

Fattori tra soggetti

		N
gruppo	C	35
	D	22
	TO	36
tipo	c	17
	s	76

Test degli effetti fra soggetti

Variabile dipendente: moccdif

Sorgente	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
Modello corretto	,085 ^a	5	,017	1,745	,133
Intercetta	,003	1	,003	,313	,577
gruppo	,037	2	,018	1,880	,159
tipo	,013	1	,013	1,338	,251
gruppo * tipo	,051	2	,026	2,623	,078
Errore	,849	87	,010		
Totale	,934	93			
Totale corretto	,934	92			

Considerando le variabili mocc t0 e mocc t6 si costruisce un modello per misure ripetute (fattore1). I risultati sono equivalenti a quelli ottenuti con la variabile differenza.

Sorgente		Somma dei quadrati	Tipo	df	Media dei quadrati	F	Sig.
		III					
fattore1	Assumendo la sfericità	,002		1	,002	,313	,577
	Greenhouse-Geisser	,002		1,000	,002	,313	,577
	Huynh-Feldt	,002		1,000	,002	,313	,577
	Limite inferiore	,002		1,000	,002	,313	,577
fattore1 * gruppo	Assumendo la sfericità	,018		2	,009	1,880	,159
	Greenhouse-Geisser	,018		2,000	,009	1,880	,159
	Huynh-Feldt	,018		2,000	,009	1,880	,159
	Limite inferiore	,018		2,000	,009	1,880	,159
fattore1 * tipo	Assumendo la sfericità	,007		1	,007	1,338	,251
	Greenhouse-Geisser	,007		1,000	,007	1,338	,251
	Huynh-Feldt	,007		1,000	,007	1,338	,251
	Limite inferiore	,007		1,000	,007	1,338	,251
fattore1 * gruppo * tipo	Assumendo la sfericità	,026		2	,013	2,623	,078
	Greenhouse-Geisser	,026		2,000	,013	2,623	,078
	Huynh-Feldt	,026		2,000	,013	2,623	,078
	Limite inferiore	,026		2,000	,013	2,623	,078
Errore(fattore1)	Assumendo la sfericità	,424		87	,005		
	Greenhouse-Geisser	,424		87,000	,005		
	Huynh-Feldt	,424		87,000	,005		
	Limite inferiore	,424		87,000	,005		

Fattori tra soggetti

		N
gruppo	C	35
	D	22
	TO	36
fumo	ex	6
	no	74
	si	13

Test degli effetti fra soggetti

Variabile dipendente: moccdif

Sorgente	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.
Modello corretto	,066 ^a	7	,009	,926	,491
Intercetta	,005	1	,005	,539	,465
gruppo	,025	2	,012	1,224	,299
fumo	,008	2	,004	,410	,665
gruppo * fumo	,048	3	,016	1,553	,207
Errore	,868	85	,010		
Totale	,934	93			
Totale corretto	,934	92			

Relazione tra variabili quantitative

- Si distinguono due metodi
 1. La regressione
 2. La correlazione

La regressione si utilizza quando si può individuare una variabile dipendente ed una o più variabili indipendenti

La correlazione si utilizza quando si vuole valutare l'interazione tra le variabili

La regressione

- La regressione è un metodo statistico che consente di analizzare la relazione tra due o più variabili affinché una variabile possa essere predetta o spiegata attraverso le altre
- La relazione può avere svariate forme e l'obiettivo è quello di spiegare la relazione tra le variabili attraverso una funzione: il modello di regressione

La regressione lineare semplice

- In questo caso la relazione viene descritta attraverso una retta con un' unica variabile esplicativa
- Il modello di regressione lineare semplice ha la forma

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Con x = variabile indipendente

y = variabile dipendente

β_0 = intercetta (valore di y quando $x=0$)

β_1 = coefficiente di regressione (pendenza della retta)

ε = termine d' errore

β_0 =intercetta (valore di y quando $x=0$) è un parametro spesso privo di significato pratico

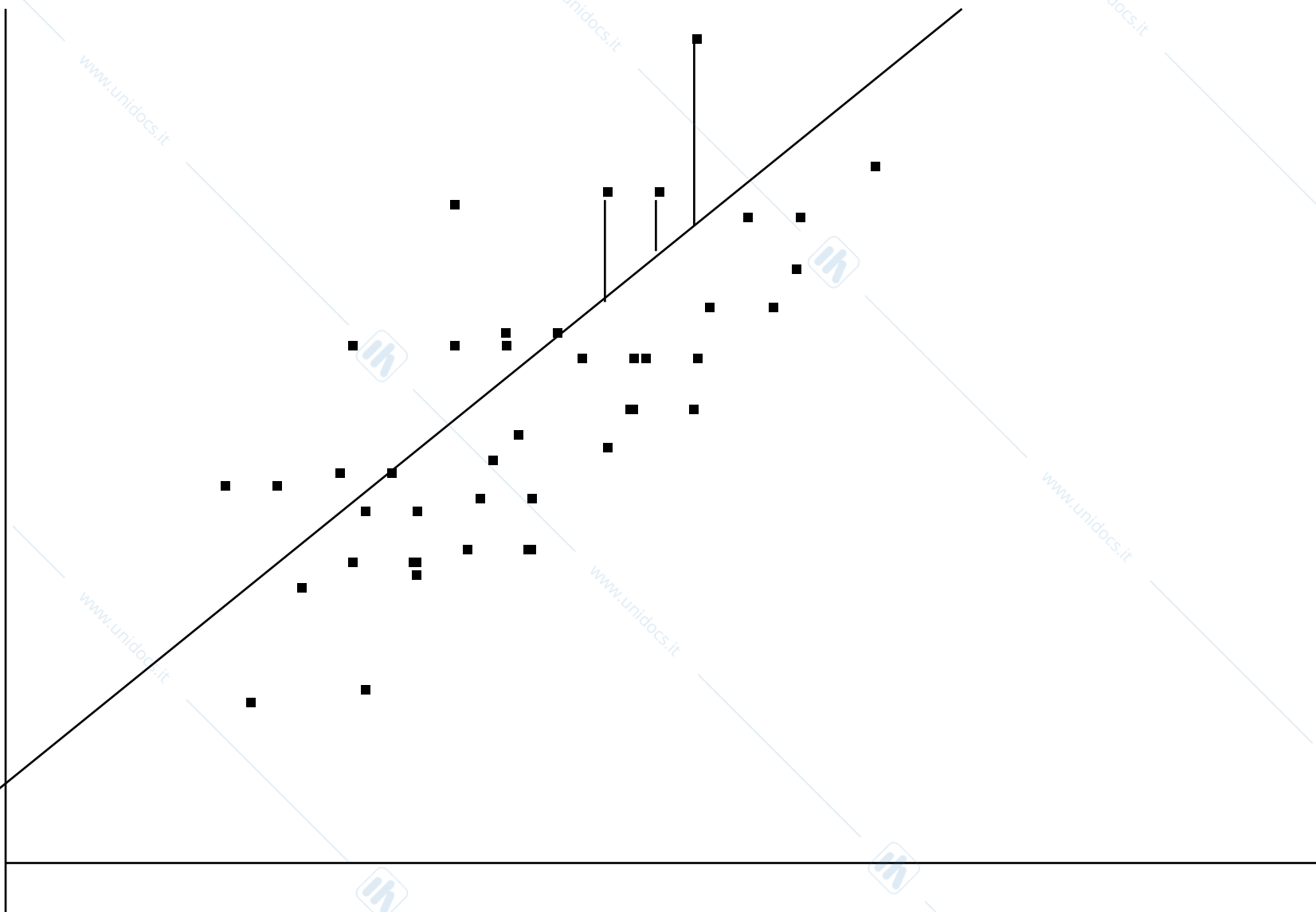
β_1 =coefficiente di regressione
rappresenta la pendenza della retta, ovvero
esprime **quanto varia in media y in funzione di variazioni unitarie di x**

assunzioni

- Il modello lineare è appropriato
- I termini di errore sono indipendenti
- Per ogni x esiste una sottopopolazione di valori y distribuite secondo una Gauss
- Le varianze delle sottopopolazioni sono tutte uguali
- Le medie delle sottopopolazioni di y giacciono tutte sulla stessa retta
- I valori di y sono statisticamente indipendenti

I passi per l'analisi della regressione

- Verificare le assunzioni sottostanti in base ai dati disponibili
- Stimare i parametri del modello
- Effettuare inferenza statistica sui parametri



Stima dei parametri del modello

- L'obiettivo è individuare la retta, detta di regressione, che interpoli tutte le y e che risulti il miglior adattamento ai dati
- Per raggiungere tale obiettivo è necessario individuare il valore dei parametri della regressione β_0 e β_1
- Il metodo generalmente utilizzato è il

METODO DEI MINIMI QUADRATI

che consiste nell'individuare la retta che rende minima la somma delle distanze al quadrato di ciascun punto y_i dai punti della retta di regressione stessa

$$\sum [y - (\beta_0 + \beta_1 x)]^2 = \sum \varepsilon^2 = \text{minimo}$$

- La metodologia matematica che permette di individuare i valori di β_0 e β_1 che soddisfino i minimi quadrati consiste nel determinare la derivata per β_0 e β_1
- Si ottengono le stime dei due parametri della regressione come segue

$$\hat{\beta}_1 = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

- L'equazione della retta diventa pertanto:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

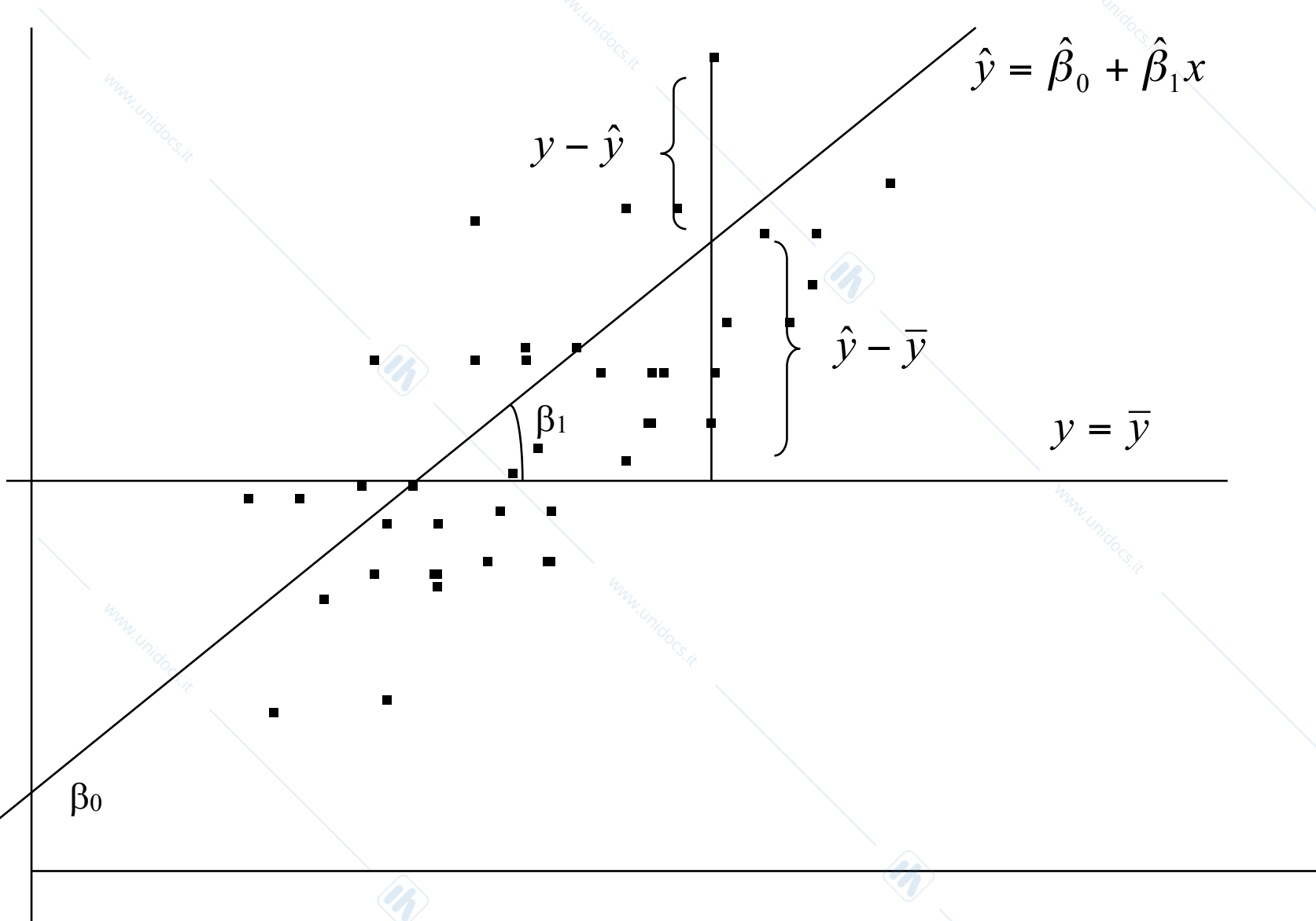
Inferenza sulla regressione

- Occorre verificare se le stime ottenute sui parametri del modello sono significative
- In particolare le ipotesi da verificare sono:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Verifica da un punto di vista grafico



Costruiamo la statistica test

Partendo dall'ipotesi nulla di $\beta_1=0$ e quindi di assenza di influenza di x su y , consideriamo una media totale ed una variabilità totale della y (a prescindere la x) che in termini di devianza è:

$$\sum (y_i - \bar{y})^2$$

Che esprime la variabilità del fenomeno.

Si può ipotizzare che a tale variabilità concorrano due fonti di variabilità:

una variabilità dovuta alla relazione della y con la x

$$\sum (\hat{y}_i - \bar{y})^2$$

Anche detta somma dei quadrati spiegata o devianza di regressione

Una variabilità non spiegata (dalla relazione) ovvero devianza d' errore o residua

$$\sum (y_i - \hat{y})^2$$

Si dimostra che la variabilità totale è data dalla somma delle due variabilità

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y})^2$$

Con g. l.

N-1

1

N-2

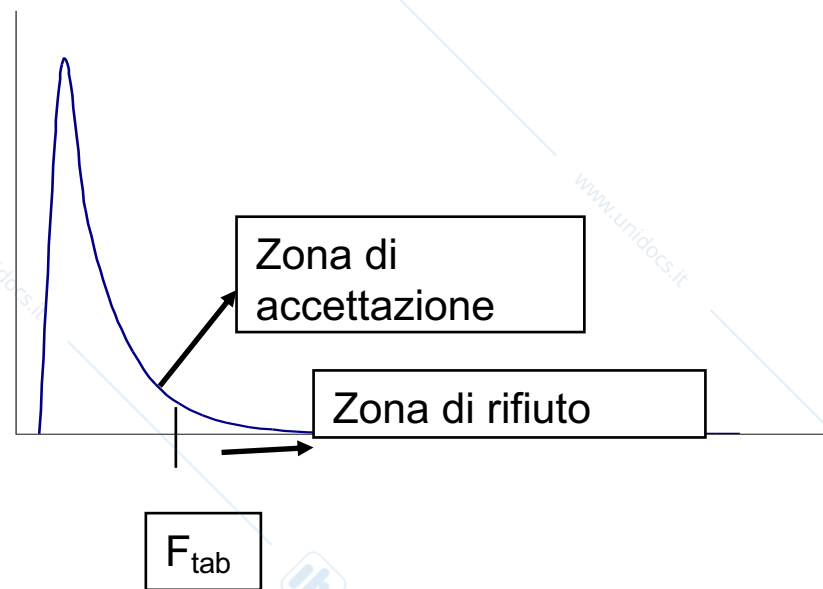
A questo punto occorre valutare se prevale la variabilità spiegata o la variabilità residua

A tal fine si confrontano le due varianze costruendo la statistica test

$$F = \frac{\sum (\hat{y}_i - \bar{y})^2 / 1}{\sum (y_i - \hat{y}_i)^2 / (N - 2)} = \frac{\text{varianza spiegata}}{\text{varianza residua}}$$

Che sotto H_0 si distribuisce come una F di Fisher con 1 ed $N-2$ gradi di libertà

- posto $\alpha=0.05$
- in corrispondenza dei gradi di libertà del numeratore e del denominatore si determina un valore tabulato che delimita la zona di rifiuto.
- la regola di decisione sarà che se F calcolato $>F_{\text{tab}}$ si rifiuta H_0
- in questo caso il test è significativo



ANOVA

Sorgenti di variazione	devianze	g.l.	varianze	F
regressione	$\sum (\hat{y}_i - \bar{y})^2$	1	$\frac{\sum (\hat{y}_i - \bar{y})^2}{1}$	$\frac{\text{varianza spiegata}}{\text{varianza residua}}$
residui	$\sum (y_i - \hat{y}_i)^2$	N-2	$\frac{\sum (y_i - \hat{y}_i)^2}{(n-2)}$	
Totale	$\sum (y_i - \bar{y})^2$	N-1		

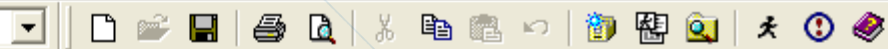
Se $F_{\text{calc}} > F_{\text{tab}}$ si rifiuta l'ipotesi nulla

Bontà del modello

- Il coefficiente di determinazione indica quanta parte delle nostre osservazioni sono spiegati dal modello
- quanti dati cadono sulla retta stimata
- Quanto il modello determinato possa considerarsi adeguato

$$R^2 = \frac{devregr}{devtot} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

$$0 < R^2 < 1$$



Output - (Senza nome)

The SAS System 14:44 Thursday, May 5, 2005 237
 The REG Procedure
 Model: MODEL1
 Dependent Variable: osteopo

Analysis of Variance

Origine	DF	Sum of Squares	Mean Square	Valore F	Pr > F
Model	1	0.00093656	0.00093656	0.00	0.9970
Error	160	10419	65.12113		
Corrected Total	161	10419			

Root MSE	8.06977	R-Square	0.0000
Dependent Mean	10.28735	Adj R-Sq	-0.0062
Coeff Var	78.44362		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	Valore t	Pr > t
Intercept	1	10.26122	6.91812	1.48	0.1400
eta	1	0.00047041	0.12404	0.00	0.9970

Modello:
 Osteopontina = 10.26 + 0.00047 * età

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Riepilogo del modello

Modello	R	R- quadrato	R- quadrato corretto	Deviazion e standard Errore della
1	,205 ^a	,042	,031	,1100100

a. Predittori: (Costante), anni

Anova^a

Modello		dei quadrati	df	Media dei quadrati	F	Sig.
1	Regressi	,048	1	,048	3,979	,049 ^b
	Residuo	1,101	91	,012		
	Totale	1,149	92			

a. Variabile dipendente: MOCT

b. Predittori: (Costante), anni

Coefficienti^a

Modello		Coefficienti non standardizzati		Coefficie nti standardi zzati	t	Sig.
		B	Deviazion e standard Errore	Beta		
1	(Costante)	1,000	,156		6,423	,000
	anni	-,006	,003	-,205	-1,995	,049

a. Variabile dipendente: MOCT

Riepilogo del modello

Modello	R	R- quadrato	R- quadrato corretto	Deviazione standard Errore della stima
1	,194 ^a	,037	,027	,1086315

a. Predittori: (Costante), mesi

Anova^a

Modello		dei quadrati	df	Media dei quadrati	F	Sig.
1	Regress	,040	1	,040	3,428	,067 ^b
	Residuo	1,038	88	,012		
	Totale	1,079	89			

a. Variabile dipendente: MOCT

b. Predittori: (Costante), mesi

Coefficienti^a

Modello		Coefficienti non standardizzati		Coefficienti standardizz ati	t	Sig.
		B	Deviazion e standard Errore	Beta		
1	(Costant	,708	,017		41,970	,000
	mesi	,000	,000	-,194	-1,851	,067

a. Variabile dipendente: MOCT

Riepilogo del modello

Modello	R	R- quadrato	R- quadrato corretto	Deviazione standard Errore della stima
1	,282 ^a	,079	,069	,1078384

a. Predittori: (Costante), BMI

Anova^a

Modello		dei quadrati	df	Media dei quadrati	F	Sig.
1	Regressi	,091	1	,091	7,842	,006 ^b
	Residuo	1,058	91	,012		
	Totale	1,149	92			

a. Variabile dipendente: MOCT

b. Predittori: (Costante), BMI

Coefficienti^a

Modello		Coefficients non standardizzati		Coefficients standardizzati	t	Sig.
		B	Deviazio ne standard Errore	Beta		
1	(Costant	,499	,069		7,184	,000
	BMI	,007	,003	,282	2,800	,006

a. Variabile dipendente: MOCT

Riepilogo del modello

Modello	R	R- quadrato	R- quadrato corretto	Deviazione standard Errore della stima
1	,332 ^a	,110	,079	,1056666

a. Predittori: (Costante), BMI, mesi, anni

Anova^a

Modello		dei quadrati	df	Media dei quadrati	F	Sig.
1	Regressio	,119	3	,040	3,543	,018 ^b
	Residuo	,960	86	,011		
	Totale	1,079	89			

a. Variabile dipendente: MOCT

b. Predittori: (Costante), BMI, mesi, anni

Coefficienti^a

Modello		Coefficienti non standardizzati		Coefficienti standardizzati		
		B	Deviazion e standard Errore	Beta	t	Sig.
1	(Costante)	,657	,212		3,099	,003
	anni	-,002	,004	-,072	-,620	,537
	mesi	,000	,000	-,169	-1,480	,142
	BMI	,006	,003	,246	2,343	,021

a. Variabile dipendente: MOCT

Modello lineare generale

- Il procedimento di analisi della varianza della regressione è del tutto sovrapponibile a quello del confronto fra più medie
- Infatti anche il confronto fra le medie può essere espresso con un modello

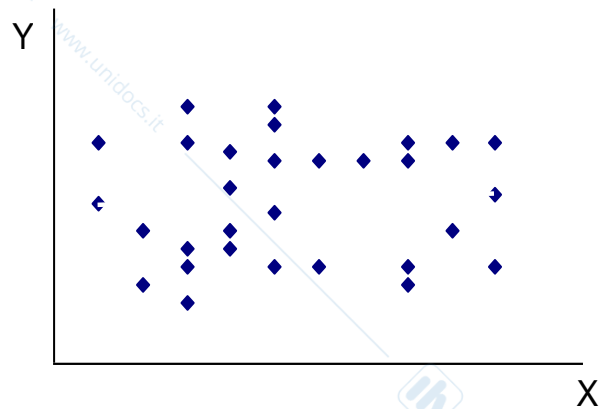
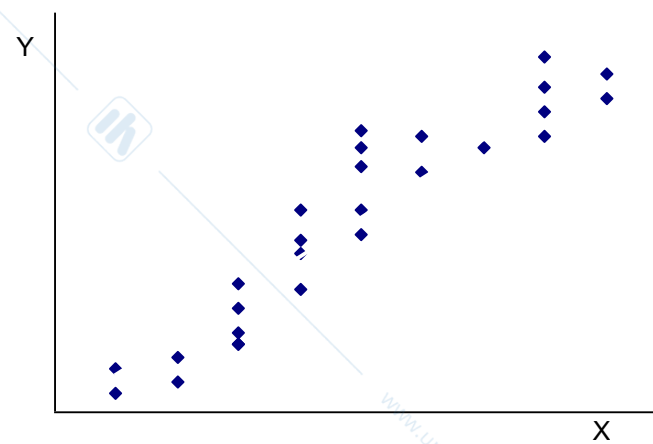
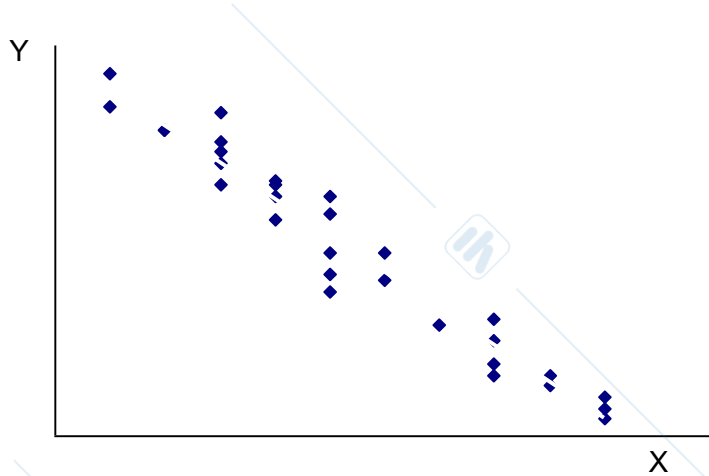
$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

- analisi della varianza e regressione utilizzano la stessa tipologia di modello che rientra nella grande famiglia dei modelli lineari generali

La correlazione

- La correlazione pone sullo stesso piano le variabili senza distinguere una variabile indipendente ed una dipendente
- consente di:
 - valutare il grado di interazione o reciproca influenza tra due variabili
 - valutare il grado di associazione di due variabili che potrebbero essere influenzate entrambe da una causa esterna

- La relazione esistente tra due variabili può essere analizzata graficamente ponendo i dati osservati in un diagramma a dispersione o scatterplot

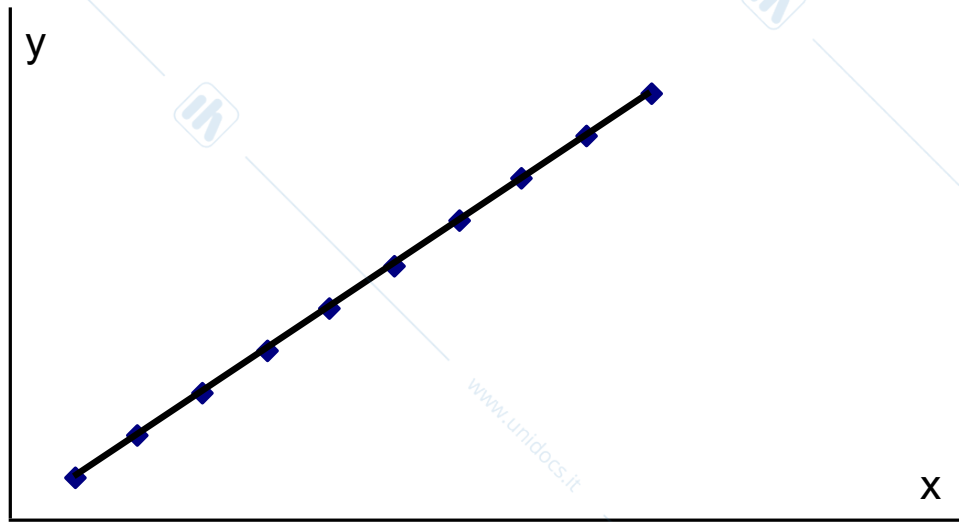


- Il grado di associazione lineare tra le due variabili è dato dal coefficiente di correlazione di Pearson

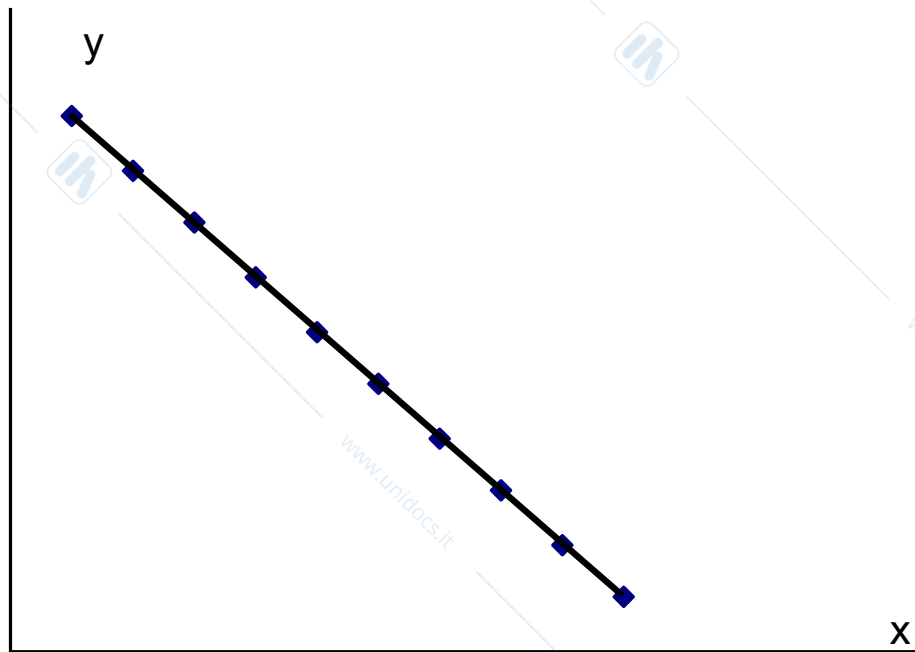
$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Con $-1 \leq r \leq +1$

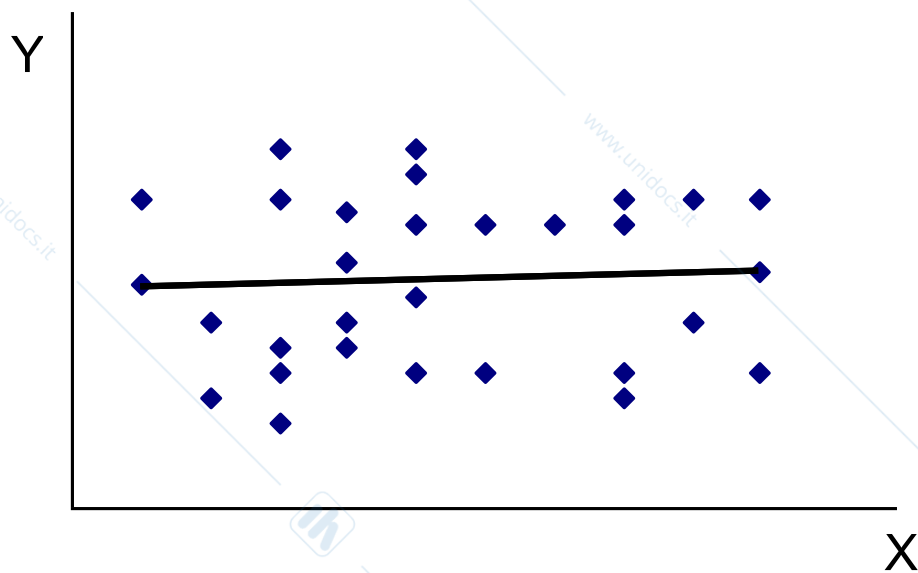
- $r = +1$: massima correlazione con proporzionalità diretta tra le due variabili, al crescere della X cresce anche la Y



- $r = -1$: massima correlazione con proporzionalità inversa tra le due variabili, al crescere della X decresce la Y e viceversa



- $r = 0$: vuol dire che non esiste correlazione tra le due variabili ovvero non esiste una relazione lineare ma si potrebbe pensare ad un'altra forma di legame (parabola, esponenziale, ...)
- Se si può assumere che le due variabili seguano una distribuzione normale bivariata allora l'assenza di correlazione significa assenza di qualsiasi legame e quindi indipendenza



Test di significatività su r

- Il valore di r è una stima campionaria del coefficiente di correlazione ρ della popolazione.
- occorre eseguire un test di verifica della significatività del coefficiente r campionario.
- si assume che le due variabili seguano una distribuzione normale bivariata

- L'ipotesi da verificare è
 $H_0: \rho = 0$ (ipotesi di indipendenza)
 $H_1: \rho \neq 0$

- Il test da utilizzare è

$$T = r \sqrt{\frac{n-2}{1-r^2}}$$

- Che sotto H_0 ha una distribuzione t-Student con $n-2$ gradi di libertà
- Se t calcolato $>$ t tabulato si rifiuta H_0
- In questo caso il test è significativo

- Esempio: valutiamo la relazione esistente tra il BMI e la massa grassa (GRKG) in un campione di 10 soggetti

Calcoliamo r

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} =$$
$$= \frac{\sum xy - \sum x \sum y / n}{\sqrt{(\sum x^2 - (\sum x)^2 / n) (\sum y^2 - (\sum y)^2 / n)}}$$

N	X BMI	Y GRKG	x ²	y ²	xy
1	24,7	17,2	610,1	297,05	425,7
2	18,6	6,2	346	39,038	116,2
3	22,0	13,9	484	192,79	305,5
4	24,0	19,2	576	369,45	461,3
5	24,6	17,3	605,2	298,6	425,1
6	31,2	29,5	973,4	868,54	919,5
7	32,5	36,5	1056	1330,8	1186
8	27,9	24,0	778,4	577,15	670,3
9	29,2	26,7	852,6	712,84	779,6
10	29,4	22,0	864,4	482,99	646,1
Σ	264,1	212,52	7146	5169,2	5935

$$r = \frac{5935 - 264.1 \cdot 212.52/10}{\sqrt{(7146 - 264.1^2/10)(5169.2 - 212.52^2/10)}} = 0,96$$

Verifichiamo l'ipotesi

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

$$T = r \sqrt{\frac{n-2}{1-r^2}} = 0,96 \sqrt{\frac{10-2}{1-0,96^2}} = 9.7$$

Essendo $t_{cal} > t_{tab}$ si rifiuta H_0

Nella popolazione considerata c'è relazione tra BMI e massa grassa

Test non parametrico: correlazione di spearman

- quando non è possibile fare assunzioni sulla distribuzione delle variabili si segue un procedimento non parametrico:
- Si assegnano i ranghi alle due variabili
- Si calcolano per ogni osservazione le differenze dei ranghi
- Si effettua la somma delle differenze al quadrato

il coefficiente di correlazione da usare è :

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

- con $-1 \leq r_s \leq +1$
- dove d_i sono le differenze dei ranghi attribuiti ai valori delle due variabili

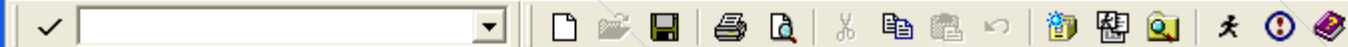
Correlazione non parametrica

N	x BMI	y GRKG	Rx	Ry	d	d ²
1	24,7	17,2	5	3	2,0	4
2	18,6	6,2	1	1	0,0	0
3	22,0	13,9	2	2	0,0	0
4	24,0	19,2	3	5	-2,0	4
5	24,6	17,3	4	4	0,0	0
6	31,2	29,5	9	9	0,0	0
7	32,5	36,5	10	10	0,0	0
8	27,9	24,0	6	7	-1,0	1
9	29,2	26,7	7	8	-1,0	1
10	29,4	22,0	8	6	2,0	4
						14

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} =$$

$$= 1 - \frac{6 \cdot 14}{10(10^2 - 1)} = 0.915$$

Essendo r_s calcolato $>$ r_s tabulato = 0.648 si rifiuta H_0



Output - (Senza nome)

The SAS System

14:44 Thursday, May 5, 2005 86

La procedura CORR

2 Variabili: osteop Tscore

Statistiche semplici

Variabile	N	Media	Dev std	Mediana	Minimo	Massimo
osteop	31	3.95484	1.50662	3.80000	1.10000	7.50000
Tscore	31	-2.22903	0.91413	-2.10000	-4.00000	-0.50000

Coefficienti di correlazione di Pearson, N = 31
 Prob > |r| rispetto a H0: Rho=0

	osteop	Tscore
osteop	1.00000	0.08615 0.6449
Tscore	0.08615 0.6449	1.00000

Coefficienti di correlazione di Spearman, N = 31
 Prob > |r| rispetto a H0: Rho=0

	osteop	Tscore
osteop	1.00000	0.11179 0.5494
Tscore	0.11179 0.5494	1.00000

La procedura CORR

3 Variabili: OPG RANKL osteopo

Statistiche semplici

variabile	N	Media	Dev std	Mediana	Minimo	Massimo
OPG	15	17.94350	26.32295	5.33900	0.20000	100.00000
RANKL	15	431.44500	827.36444	76.76274	0.01000	2470
osteopo	15	22.15155	25.92078	9.93640	0	100.00000

Coefficienti di correlazione di Pearson, N = 15
 Prob > |r| rispetto a H0: Rho=0

	OPG	RANKL	osteopo
OPG	1.00000	0.10009 0.7226	0.97230 <.0001
RANKL	0.10009 0.7226	1.00000	0.26171 0.3461
osteopo	0.97230 <.0001	0.26171 0.3461	1.00000

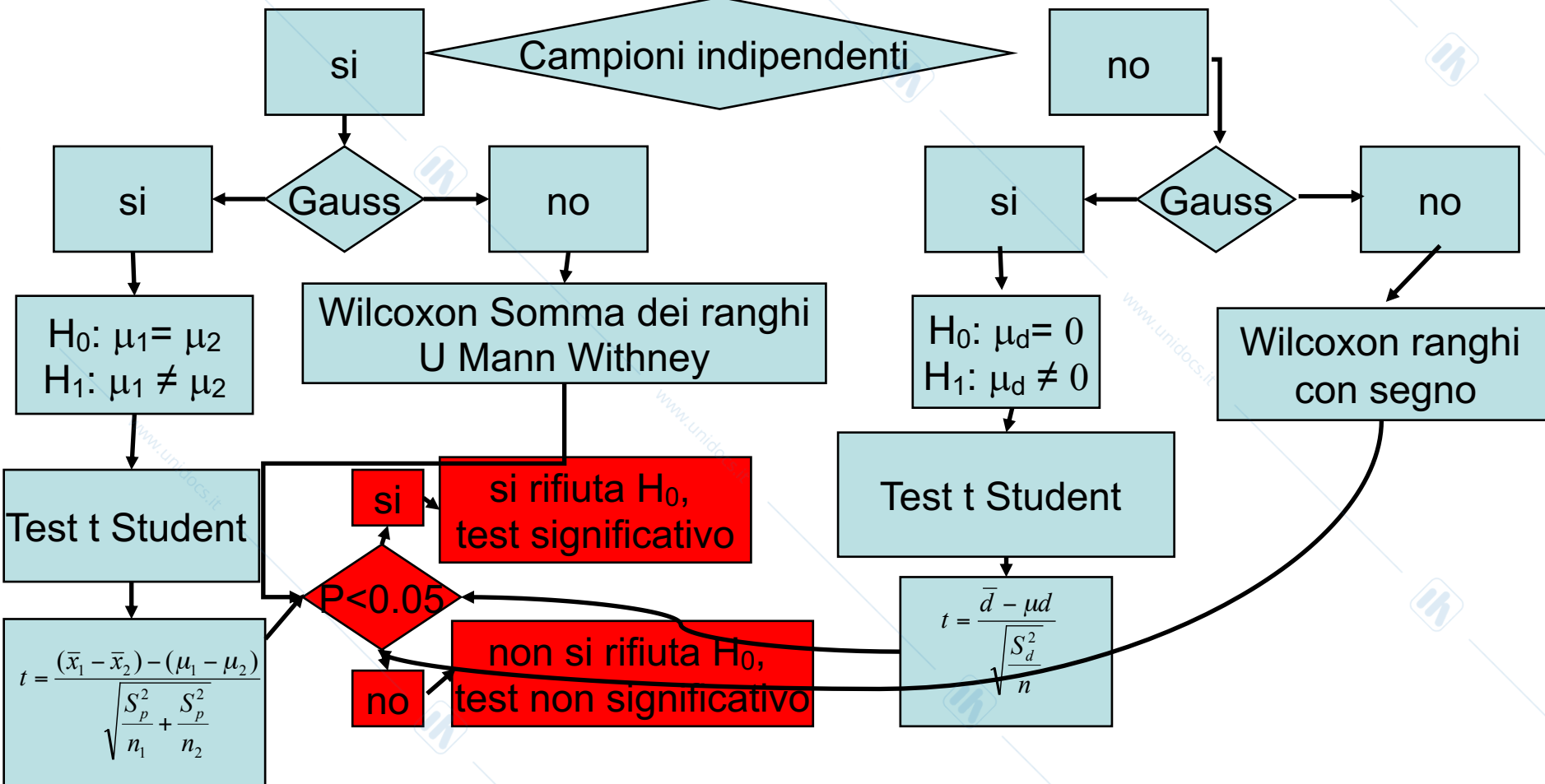
Coefficienti di correlazione di Spearman, N = 15
 Prob > |r| rispetto a H0: Rho=0

	OPG	RANKL	osteopo
OPG	1.00000	0.39068 0.1499	0.93834 <.0001
RANKL	0.39068 0.1499	1.00000	0.54171 0.0370
osteopo	0.93834 <.0001	0.54171 0.0370	1.00000

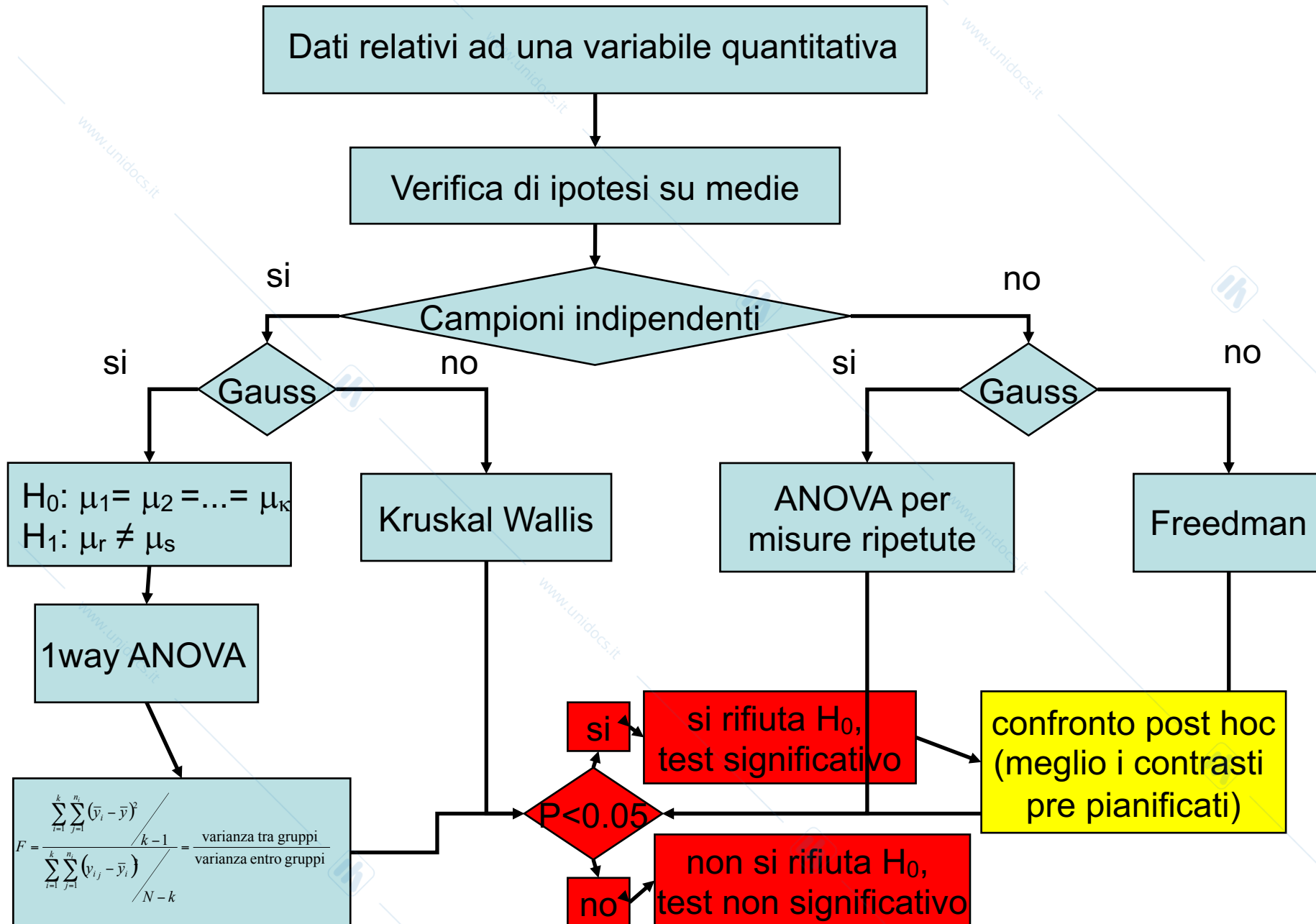
Confronto fra 2 gruppi

Dati relativi ad una variabile quantitativa

Verifica di ipotesi su medie



Confronto fra più gruppi



$$F = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (v_{ij} - \bar{y}_i)^2 / (N-k)} = \frac{\text{varianza tra gruppi}}{\text{varianza entro gruppi}}$$

Relazione tra variabili



Concordanza

A volte è necessario valutare non solo il grado di accordo che si osserva fra due test diagnostici, ma anche quello fra due (o più) operatori che interpretano l'esito di uno stesso test (es. radiografie, elettrocardiogramma, auscultazione cardiaca ecc.), oppure fra due letture effettuate da uno stesso operatore in tempi diversi (prima e dopo l'acquisizione di altre informazioni).

Concordanza

Test A

	+	-
Test B	a	b
	c	d

Concordanza tra 2 test

$$\frac{a + d}{a + b + c + d} = \frac{a + d}{N}$$

Tale misura di concordanza però non esclude la concordanza dovuta al caso

Concordanza dovuta al caso

Uno studente ed un radiologo esperto esaminano separatamente 100 radiografie del torace per la diagnosi di polmonite ottenendo i seguenti risultati

		studente		tot
		+	-	
radiologo	+	0	6	6
	-	0	94	94
		0	100	100

Per lo studente le ecografie sono tutte negative, il radiologo ne identifica 6 positive

Concordanza dovuta al caso

Calcolando la concordanza si ottiene

$$\frac{a+d}{N} = \frac{0+94}{100} = 0.94 = 94\%$$

studente

		+	-	tot
radiologo	+	0	6	6
	-	0	94	94
		0	100	100

Può ritenersi possibile una concordanza così alta?

La concordanza del 94% non è reale poiché una quota della concordanza osservata è dovuta al caso

Concordanza dovuta al caso

Procediamo con un esempio: nella tabella sono riportate le valutazioni di 100 TC fatte da 2 radiologi esperti

		Radiologo 1		
		+	-	tot
Radiologo 2	+	60	10	70
	-	5	25	30
		65	35	100

L'accordo osservato è pari a $\frac{60 + 25}{100} = 0.85 = 85\%$

Concordanza dovuta al caso

Radiologo 1

	+	-	tot
Radiologo 2	60	10	70
	5	25	30
	65	35	100

Radiologo 2

Il radiologo 1 ha espresso
 $65/100=0.65$ diagnosi positive
 $35/100=0.35$ diagnosi negative

Il radiologo 2 ha espresso
 $70/100=0.70$ diagnosi positive
 $30/100=0.30$ diagnosi negative

La probabilità dovuta al caso di avere per entrambi
una diagnosi positiva $\Rightarrow 0.65 \cdot 0.70 = 0.455$
una diagnosi negativa $\Rightarrow 0.35 \cdot 0.30 = 0.105$

La probabilità che i due radiologi abbiano espresso la stessa diagnosi
è pari a $\Rightarrow 0.455 + 0.105 = 0.560 = 56\%$

Questo è l'accordo dovuto al caso

Di conseguenza la probabilità che i due radiologi abbiano espresso la stessa
diagnosi non per caso è pari a $1 - 0.560 = 0.44 = 44\%$

K di Cohen

Disponendo delle quantità

Accordo osservato = 0.85

Accordo dovuto al caso (atteso) = 0.56

Accordo non dovuto al caso = 0.44

Si può calcolare il K di Cohen con la formula

$$K = \frac{\text{accordo osservato} - \text{accordo per caso}}{\text{accordo non per caso}} = \frac{0.85 - 0.56}{0.44} = 0.659$$

K di Cohen

Calcolando i valori attesi di concordanza è possibile dalla tabella

		Radiologo 1		
		+	-	tot
Radiologo 2	+	60 (45.5)	10	70
	-	5	25 (10.5)	30
		65	35	100

Calcolare direttamente il K di Cohen

$$K = \frac{O - E}{N - E} = \frac{(60 + 25) - (45.5 + 10.5)}{100 - (45.5 + 10.5)} = \frac{85 - 56}{100 - 56} = 0.659$$

Interpretazione del K di Cohen

$k \leq 0$ concordanza scarsissima

$k = 0.01-0.20$ concordanza scarsa

$k = 0.21-0.40$ concordanza discreta

$k = 0.41-0.60$ concordanza moderata

$k = 0.61-0.80$ concordanza buona

$k = 0.81-1.00$ concordanza ottima

Landis & Koch

Interpretazione del K di Cohen

$k = 0.00-0.40$ concordanza marginale

$k = 0.41-0.75$ concordanza buona

$k = 0.75-1.00$ concordanza eccellente

Fleiss

Studi statistici

- **Studi osservazionali**

Si osservano diversi aspetti di una situazione esistente e si interpretano i dati per dare una spiegazione sull'origine della situazione osservata

- **Studi sperimentali**

Si compie un'azione (somministrazione di un farmaco) e si osservano e si interpretano gli effetti di tale azione

Studi sperimentali

- Possono essere condotti
 - in vitro
 - su animali
 - su volontari
 - su pazienti
 - su soggetti sani
- Gli studi di trattamenti condotti su soggetti umani si dicono **studi clinici**

Come verificare se un nuovo trattamento (sperimentale) è più efficace di quello in uso?

- Confrontare i risultati ottenuti con il trattamento sperimentale con i risultati ottenuti precedentemente con il trattamento standard recuperati da eventuali registri?
- Confrontare i nostri pazienti (trattamento sperimentale) con pazienti di altro ospedale (trattamento standard)?

- Somministrare il trattamento sperimentale a pazienti che lo accettano volontariamente e il trattamento tradizionale ai pazienti che non lo accettano?
- Allocare i pazienti al trattamento sperimentale o a quello tradizionale in modo alternato?
- Allocare i pazienti al trattamento sperimentale o a quello tradizionale seguendo un procedimento casuale?

Tavola dei numeri casuali

01 19223 95034 05756 28713 96409 12531 42544 82853
02 73676 47150 99400 01927 27754 42648 82425 36290
03 45467 71709 77558 00095 32863 29485 82226 90056
04 52711 38889 93074 60227 40011 85848 48767 52573
05 95592 94007 69971 91481 60779 53791 17297 59335

06 68417 35013 15529 72765 85089 57067 50211 47487
07 82739 57890 20807 47511 81676 55300 94383 14893
08 60940 72024 17868 24943 61790 90656 87964 18883
09 36009 19365 15412 39638 85453 46816 83485 41979
10 38448 48789 18338 24697 39364 42006 76688 08708

11 81486 69487 60513 09297 00412 71238 27649 39950
12 59636 88804 04634 71197 19352 73089 84898 45785
13 62568 70206 40325 03699 71080 22553 11486 11776
14 45149 32992 75730 66280 03819 56202 02938 70915
15 61041 77684 94322 24709 73698 14526 31893 32592

16 14459 26056 31424 80371 65103 62253 50490 61181
17 38167 98532 62183 70632 23417 26185 41448 75532
18 73190 32533 04470 29669 84407 90785 65956 86382
19 95857 07118 87664 92099 58806 66979 98624 84826
20 35476 55972 39421 65850 04266 35435 43742 11937

21 71487 09984 29077 14863 61683 47052 62224 51025
22 13873 81598 95052 90908 73592 75186 87136 95761
23 54580 81507 27102 56027 55892 33063 41842 81868
24 71035 09001 43367 49497 72719 96758 27611 91596
2596746 12149 37823 71868 18442 35119 62103 39244

26 96927 19931 36809 74192 77567 88741 48409 41903
27 43909 99477 25330 64359 40085 16925 85117 36071
28 15689 14227 06565 14374 13352 49367 81982 87209
29 36759 58984 68288 22913 18638 54303 00795 08727
30 69051 81717 40951 78453 4004 89872 87201 97245

31 05007 16632 81194 14873 04197 85576 45195 96565
32 68732 55259 84292 08796 43165 93739 31685 97150
33 45740 41807 65561 33302 07051 93623 18132 09547
34 27816 78416 18329 21337 35213 37741 04312 68508
35 66925 55658 39100 78458 11206 19876 87151 31260

36 08421 44753 77377 28744 75592 08563 79140 92454
37 53645 66812 61421 47836 12609 15373 98481 14592
38 66831 68908 40772 21558 47781 33586 79177 06928
39 55588 99404 70708 41098 43563 56934 48394 51719
40 12975 13258 13048 45144 72321 81940 00360 02428

41 96767 35964 23822 96012 94591 65194 50842 53372
42 72829 50232 97892 63408 77919 44575 24870 04178
43 88565 42628 17797 49376 61762 16953 88604 12724
44 62964 88145 83083 69453 46109 59505 69680 00900
45 19687 12633 57857 95806 09931 02150 43163 58636

46 37609 59057 66967 83401 60705 02384 90597 93600
47 54973 86278 88737 74351 47500 84552 19909 67181
48 00694 05977 19664 65441 20903 62371 22725 53340
49 71546 05233 53946 68743 72460 27601 45403 88692
50 07511 88915 41267 16853 84569 79367 32337 03316

Effetto placebo, cieco e doppio cieco

Essere a conoscenza del trattamento può influire sulla risposta al trattamento

Soluzioni:

- Il paziente non conosce il trattamento (cieco)
- Il paziente e il ricercatore non conoscono il trattamento (doppio cieco)

Studi osservazionali

- Uno studio osservazionale è uno studio in cui i soggetti e le variabili non sono in alcun modo manipolate ma vengono semplicemente osservate al fine di studiare la relazione tra una caratteristica ed un evento
- La caratteristica può essere
 - un trattamento
 - un fattore demografico (età, sesso...)
 - un fattore comportamentale (n. sigarette, ore attività fisica....)
 - un fattore ambientale (esposizione sostanze tossiche, polline...)
 - dati di laboratorio, fattori genetici...

- L'evento può essere:

- l'insorgenza o la diagnosi della malattia o la sua recrudescenza
- la guarigione
- la presenza di sintomi
- il verificarsi di effetti collaterali
- la morte
-

- Una caratteristica capace di modificare l'insorgenza o il decorso di una malattia in maniera prevedibile e riproducibile si definisce **fattore prognostico**
- Se il fattore prognostico agisce in maniera positiva è detto **fattore di protezione**
- Se il fattore prognostico agisce in maniera negativa è detto **fattore di rischio**

Tipologie di studi osservazionali

- Studi trasversali
- Studi prospettici (studi di coorte)
- Studi retrospettivi (studi caso-controllo)

—

Studi trasversali o di sorveglianza

- sono studi fatti su una popolazione in un determinato momento
- consentono di valutare la prevalenza di fattori prognostici (di rischio o di protezione), di casi di una malattia, ma non consentono di stabilire delle relazioni di “causa-effetto”

Studi prospettici (studi di coorte)

- Si selezionano due campioni casuali di soggetti: un campione costituito da soggetti che presentano il fattore di rischio e un altro campione costituito da soggetti che non presentano il fattore di rischio
- I soggetti sono seguiti nel tempo e si registra il numero di soggetti in ogni campione che, alla fine dell'osservazione, sarà classificato in ciascuna delle categorie della variabile risposta

Studio prospettico

popolazione a rischio e popolazione non a rischio

coorte degli esposti seguita nel tempo

coorte dei non esposti seguita nel tempo

evento presente (malati)

evento assente (non malati)

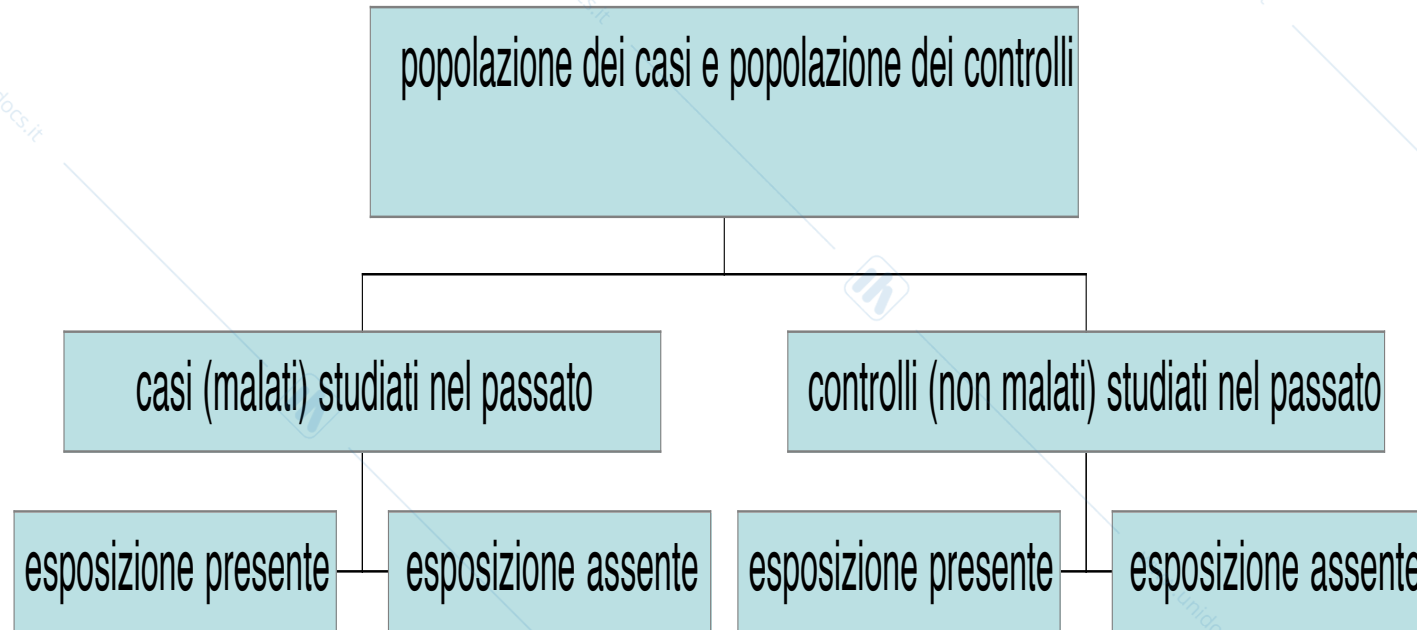
evento presente (malati)

evento assente (non malati)

• **Studi retrospettivi (studi caso-controllo):**

- Si selezionano due campioni casuali di soggetti tra le categorie della variabile risposta: un campione costituito da soggetti che presentano la caratteristica in studio (casi) e un altro campione costituito da soggetti che non presentano la caratteristica in studio (controlli)
- Lo studioso guarda retrospettivamente ai soggetti valutando chi presenta o presentava e chi non presenta o non presentava il fattore di rischio

Studio retrospettivo



In entrambi i casi i dati possono essere riassunti in una tabella di contingenza

2×2 che fornisce informazioni sul numero di soggetti con o senza il fattore di rischio insieme al numero di soggetti con o senza la caratteristica in studio (ad. es. la malattia) ed alle frequenze relative a tutte le combinazioni delle categorie delle due variabili.

Studio prospettico

classificazione di un campione di soggetti secondo la malattia e il fattore di rischio

Fattore di rischio	Malattia presente	Malattia assente	totale
presente	a	b	a+b
assente	c	d	c+d
totale	a+c	b+d	N

Rischio Relativo (RR)

- Si calcola in uno studio prospettico
- Il rischio relativo è il rapporto tra il rischio di contrarre la malattia tra i soggetti che presentano il fattore di rischio, e il rischio di contrarre la malattia tra i soggetti che non presentano il fattore di rischio

- Rischio di contrarre la malattia tra i soggetti con il fattore di rischio: $a/(a+b)$
- Rischio di contrarre la malattia tra i soggetti senza il fattore di rischio: $c/(c+d)$
- Rischio Relativo:

$$RR = \frac{a/(a+b)}{c/(c+d)}$$

- $RR=1$ uguale incidenza tra esposti e non esposti
- $RR>1$ esposti hanno maggiore incidenza
- $RR<1$ esposti minore incidenza)
- Si costruisce un intervallo di confidenza per il Rischio Relativo

Studio retrospettivo

casi e controlli classificati in relazione al fattore di rischio

Fattore di rischio	Casi	Controlli	totale
presente	a	b	a+b
assente	c	d	c+d
totale	a+c	b+d	N

Odds Ratio (OR)

- Si calcola in uno studio retrospettivo
- L'Odds di un evento è il rapporto tra la probabilità che esso si verifichi e la probabilità che esso non si verifichi
- L'Odds di essere un caso rispetto ad essere un controllo tra i soggetti con il fattore di rischio è $[a/(a+b)]/[b/(a+b)] = a/b$
- L'Odds di essere un caso rispetto ad essere un controllo tra i soggetti senza il fattore di rischio è $[c/(c+d)]/[d/(c+d)] = c/d$

- Odds ratio:

$$OR = \frac{a/b}{c/d} = \frac{ad}{bc}$$

- OR=1 l'esposizione non è associata all'evento
- OR>1 l'esposizione è associata all'evento (è un fattore di rischio)
- OR<1 l'esposizione è associata all'evento (è un fattore di protezione)
- Si costruisce un intervallo di confidenza per l'Odds Ratio

TEST DIAGNOSTICO

Un *test di diagnostico* è una procedura o tecnica che si basa :
su un **criterio obiettivo**, piuttosto che su un **giudizio soggettivo**.

Il test diagnostico definisce un “**valore soglia**” della misurazione di
una variabile biologica rispetto al quale i pazienti sono classificati

- come **positivi (+)**
- o come **negativi (-)**.

Una *diagnosi clinica* è un processo che si basa
sulla valutazione di **test diagnostici, sintomi, segni ed esami
di laboratorio.**
oltre che sul **giudizio soggettivo** [Occhio clinico].

TEST DIAGNOSTICO

test diagnostico è una qualunque procedura utile all'identificazione di uno stato di malattia.

Esempi : misura e valutazione di ...

Glicemia ⇒ diabete

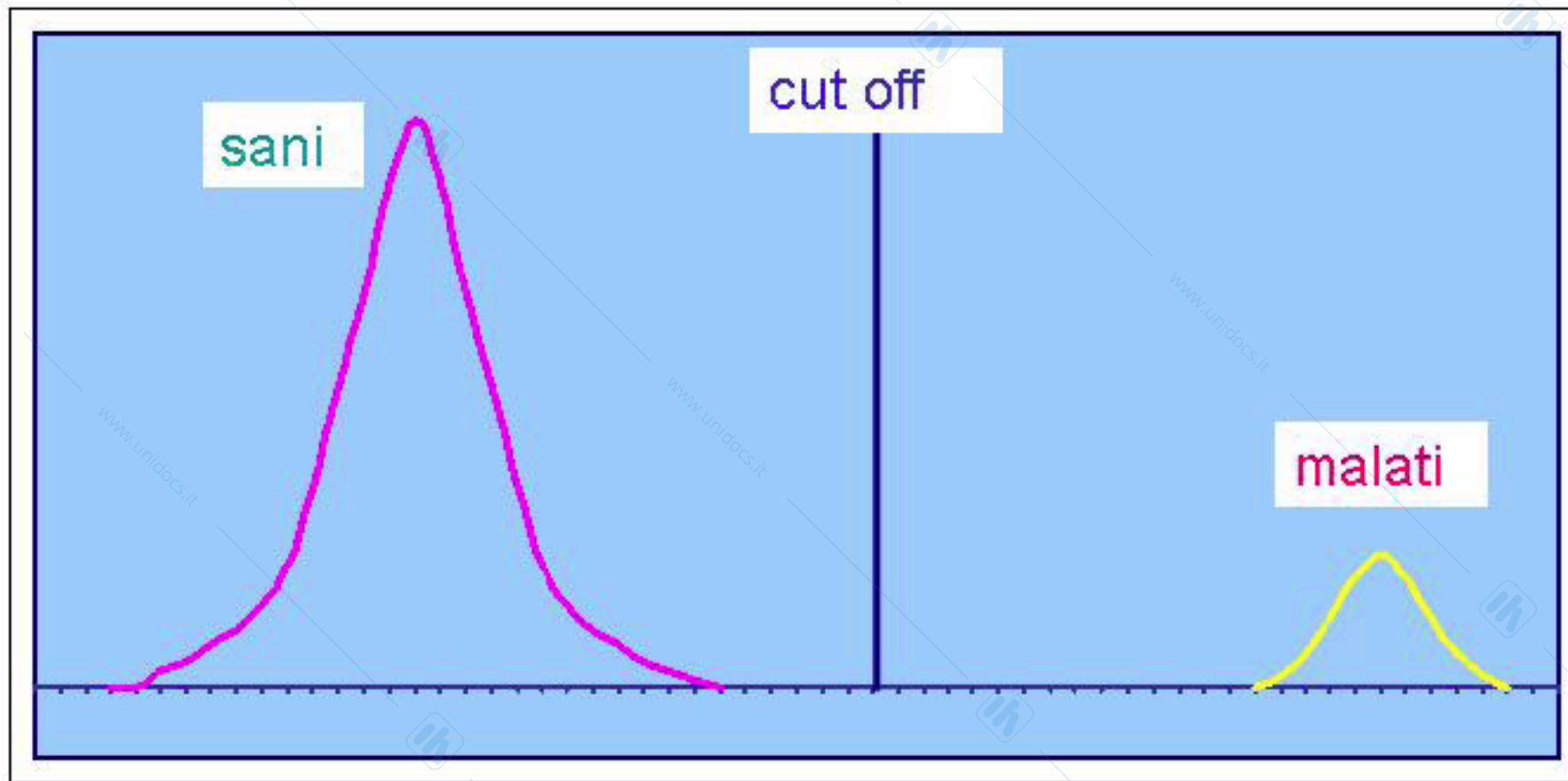
SGOT e SGPT ⇒ malattie epatiche

Proteinuria ⇒ malattie renali

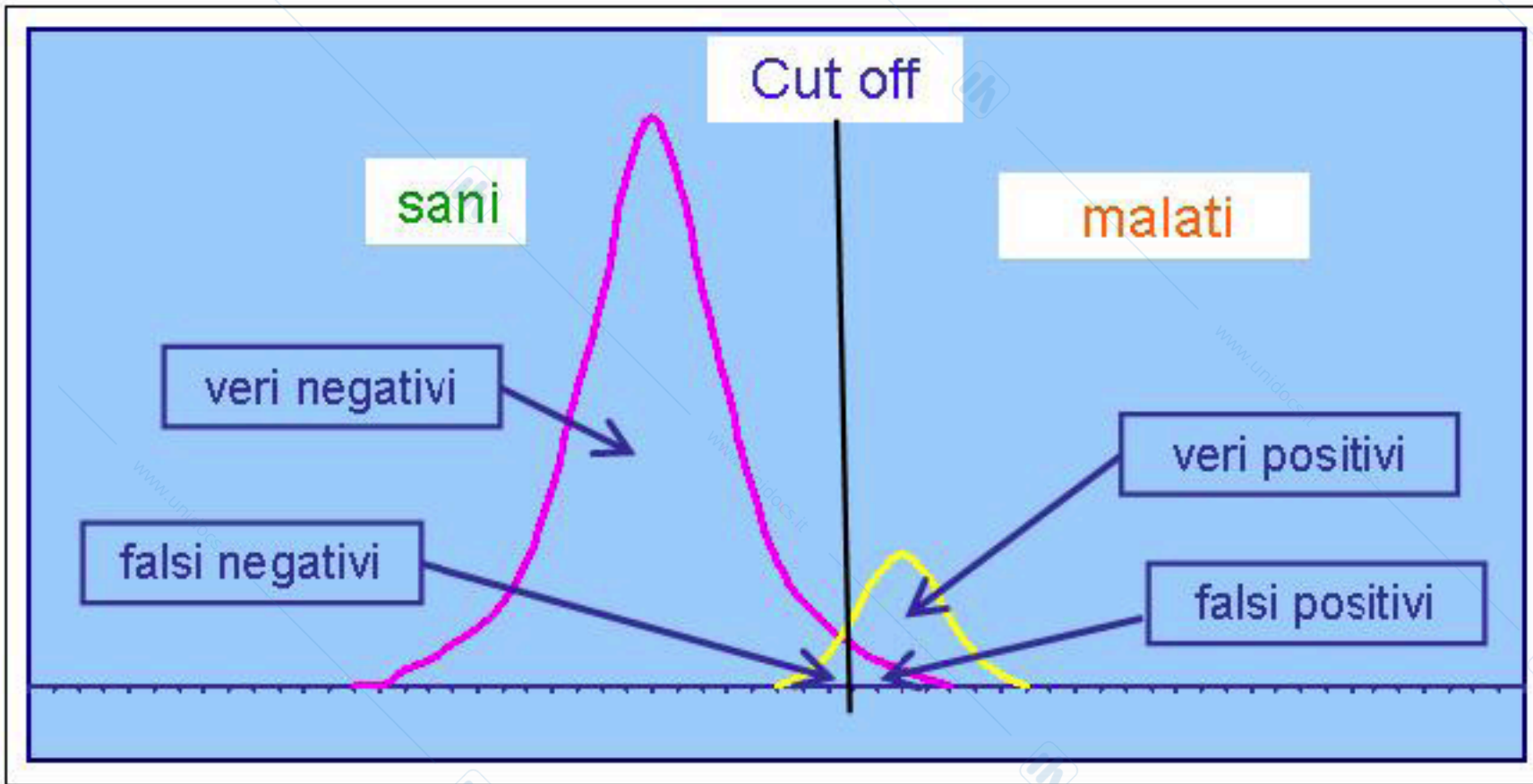
L'esito di un test è positivo se induce a sospettare la presenza della malattia.

L'esito di un test è negativo se induce ad escludere la presenza della malattia.

La situazione ideale



La situazione reale



un buon test diagnostico

tende a fornire esiti positivi in soggetti che presentano la malattia.

Si consideri l'insieme dei **soggetti che hanno la malattia M** , e si supponga di sottoporli al test ...

Se il test risulta

positivo (T+)

⇒ **Veri positivi (VP)**

Se il test risulta

negativo (T-)

⇒ **Falsi negativi (FN).**

La probabilità che un test diagnostico ha di dare **esiti positivi (T+) nei malati ($M+$)** prende nome di **sensibilità (S_n)**.

Sensibilità = VP / malati

La sensibilità di un test è la sua capacità di identificare i soggetti che realmente presentano la malattia

un buon test diagnostico

tende a fornire esiti negativi in soggetti che non hanno la malattia

Si consideri l'insieme dei soggetti che non hanno la malattia M, e si supponga di sottoporli al test ...

Se il test risulta

positivo (T+) ⇒ **Falsi positivi (FP).**

Se il test risulta

negativo (T-) ⇒ **Veri negativi (VN).**

La probabilità che un test diagnostico ha di dare **esiti negativi (T-)** nei non malati (M-) prende nome di **specificità (Sp)**.

Specificità = VN/sani

La specificità di un test è la sua capacità di identificare i soggetti che realmente non presentano la malattia

sensibilità e specificità sono

- ✓ caratteristiche interne e proprie di un test diagnostico, poiché ciascuna è riferita ad un insieme omogeneo (malati o sani);
- ✓ caratteristiche misurabili da la frequenza relativa di esiti positivi o negativi su campioni di pazienti affetti da malattia o di soggetti sani;
- ✓ comprese tra 0 e 1: esse infatti esprimono valori di probabilità;
- ✓ raramente entrambi uguali a 1.

MALATTIA ED ESITI DEL TEST

Gli individui sottoposti a test diagnostico, possono essere classificati come veri negativi, falsi positivi, falsi negativi e veri positivi in funzione dell'esito del test e della presenza della malattia.

	T-	T+	<i>Totale</i>
M-	VN	FP	Sani
M+	FN	VP	Malati
<i>Totale</i>	Negativi	Positivi	Popolazione

Il rapporto **malati/popolazione** è detto **prevalenza** di malattia.

Nella pratica clinica

noi non sappiamo se il soggetto è malato
né conosciamo il risultato dell' esame
“gold standard” (ad es. biopsia)

il risultato del test è tutto ciò che abbiamo
perciò la domanda vera è: qual' è la
probabilità che il test ci dia una diagnosi
corretta?

Un esempio

	Doppler +	Doppler -	Totale
angiografia +	90	10	100
angiografia -	50	850	900
Totale	140	860	1000

Sensibilità=90% (83%-95%)

Specificità=94% (93%-96%)

Ma qual è la probabilità che il doppler dica il vero?

A questa domanda non danno risposta né sensibilità né specificità

VALORE PREDITTIVO DI UN TEST

Il **valore predittivo** di un esito **positivo** al test (VPP) è la probabilità della **presenza** della **malattia in** un soggetto con **esito positivo**:

$$VPP = \frac{VP}{positivi}$$

Il **valore predittivo** di un esito **negativo** al test (VPN) è la probabilità dell'**assenza** della **malattia in** un soggetto con **esito negativo**:

$$VPN = \frac{VN}{negativi}$$

ACCURATEZZA DI UN TEST

L'**ACCURATEZZA** di un test è la
quantità di diagnosi corrette

$$= (VN + VP) / \text{POPOLAZIONE}$$

Dall' esempio

	Doppler +	Doppler -	Totale
angiografia +	90	10	100
angiografia -	50	850	900
Totale	140	860	1000

Sensibilità=90% (83%-95%)

Specificità=94% (93%-96%)

VPP=64% (56%-72%)

VPN=99% (98%-99%) ACCURATEZZA=(90+850)/1000*100=94%

ATTENZIONE ALLA PREVALENZA

I VP di un test nella pratica clinica dipendono in maniera sostanziale dalla prevalenza della patologia nella popolazione studiata, e questa può essere anche molto differente da quella osservata negli studi che hanno calcolato originariamente i VP

esempio

Se eseguiamo un ecodoppler carotideo su tutti i presenti in questa aula, la probabilità pre esame (o prevalenza) di stenosi di interesse chirurgico sarà certamente molto bassa

Ma se facciamo la stessa cosa in una corsia dove sono ricoverati pazienti con recente ictus ischemico non lacunare, che non fibrillano, la prevalenza sarà molto alta

Lo stesso esame, fatto dallo stesso operatore, avrà VP molto diversi in queste due circostanze

Sensibilità 90% specificità 94% ma diversa prevalenza

ANGIOGRAFIA		DOPPLER		
		+	-	
+	45	5	50	
-	57	893	950	
		102	898	1000

Prevalenza 5%

VPP=44%

VPN=99%

ANGIOGRAFIA		DOPPLER		
		+	-	
+	675	75	750	
-	15	235	250	
		690	310	1000

Prevalenza 75%

VPP=98%

VPN=76%

VPP e VPN in funzione della prevalenza a parità di sensibilità e specificità

prevalenza	VPP	VPN	realità	test		totale
				positivo	negativo	
50%	84%	81%	malato	400	100	500
			sano	75	425	50
			totale	475	525	1000
20%	57%	94%	malato	160	40	200
			sano	120	680	800
			totale	280	720	1000
10%	37%	97%	malato	80	20	100
			sano	135	765	900
			totale	215	785	1000
1%	5%	99%	malato	8	2	10
			sano	148	842	990
			totale	156	844	1000

Dal pre test al post test

Conoscendo sensibilità e specificità (probabilità pre test) e la prevalenza si possono calcolare il VPP e il VPN (probabilità post test)

$$VPP = \frac{sens * prev}{sens * prev + (1 - spec) * (1 - prev)}$$

$$VPN = \frac{spec * (1 - prev)}{spec * (1 - prev) + (1 - sens) * prev}$$

In sintesi

- L'**accuratezza** dipende dalla prevalenza della malattia nella popolazione. Due test con la stessa accuratezza possono essere molto diversi, perchè la proporzione di veri positivi e veri negativi può essere molto diversa.
- La **sensibilità e la specificità** sono indici largamente accettati per giudicare la performance di un test. Hanno il vantaggio di essere indipendenti dalla prevalenza della malattia nella popolazione.
- Il **valore predittivo di un test positivo e negativo** dipendono dalla prevalenza della malattia nella popolazione, per cui i dati ricavati su una popolazione possono non essere applicabili ad un gruppo diverso.

ancora un esempio

- Supponiamo una popolazione di maschi asintomatici con età compresa fra i 50 e i 70 anni, con storia familiare di carcinoma prostatico (PCa), con esame rettale positivo e da sottoporre a biopsia prostatica.
- Sottoponiamo tale popolazione all' esame dei livelli di PSA al fine di identificare i soggetti che potrebbero essere affetti da PCa.
- I risultati del test vengono riassunti nella seguente Tabella:

ancora un esempio

	T-	T+	TOTALE
M-	2576	1732	4308
M+	275	936	1211
TOTALE	2851	2668	5519

Risultati dicotomizzati sulla base del valore soglia più diffusamente utilizzato di 4 ng/ml

ancora un esempio

- Sensibilità= $936/1211=0,90$
- Specificità= $2576/4308=0,60$
- VP+= $936/2668=0,35$
- VP-= $2576/2851=0,90$

Un altro modo per valutare il test diagnostico

- **Rapporto di verosimiglianza per esito positivo (positive likelihood ratio, LR+)**

Rapporto tra la proporzione di test positivi nei malati e la proporzione di test positivi nei sani:

$$LR_+ = \frac{P(T^+ / M^+)}{P(T^+ / M^-)} = \frac{\textit{sens}}{1 - \textit{spec}}$$

Un altro modo per valutare il test diagnostico

- **Rapporto di verosimiglianza per esito negativo (negative likelihood ratio, LR-)**

Rapporto tra la proporzione di test negativi nei malati e la proporzione di test negativi nei sani:

$$LR_{-} = \frac{P(T^{-} / M^{+})}{P(T^{-} / M^{-})} = \frac{1 - \textit{sens}}{\textit{spec}}$$

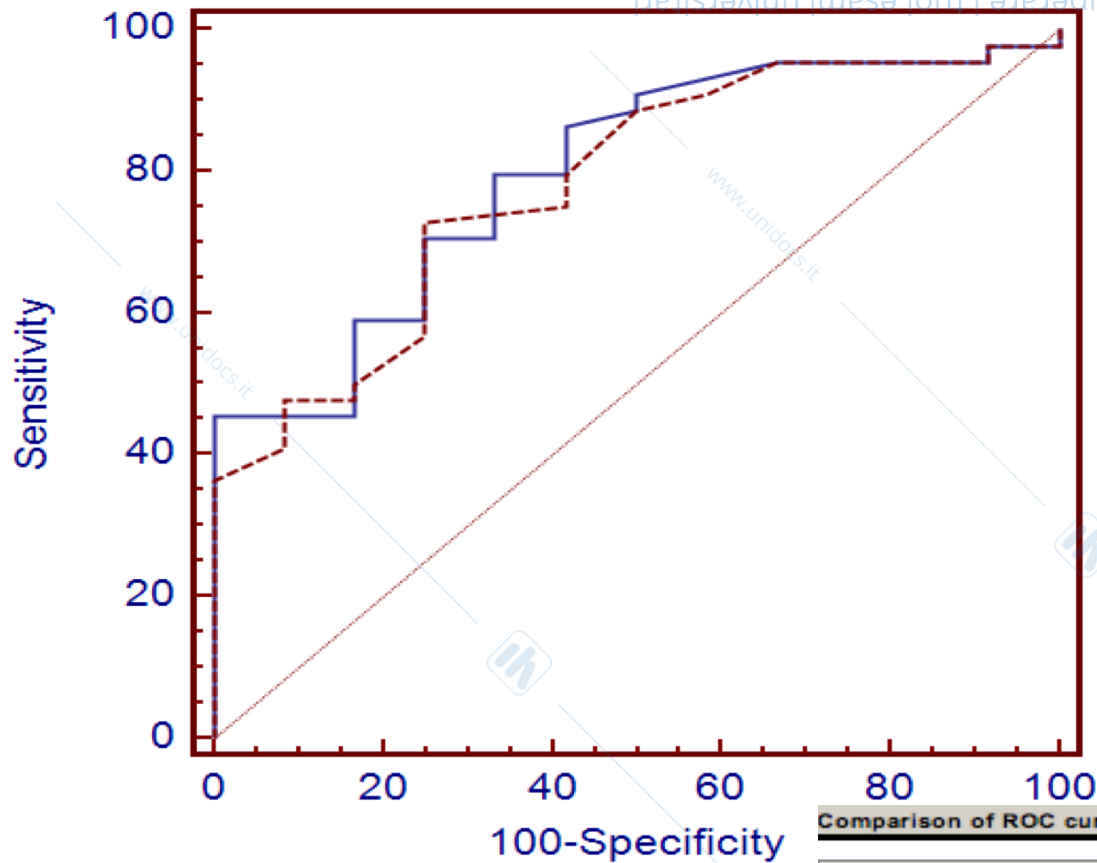
Cut-off per LR+ ed LR-

- Quando LR+ è maggiore di 5, un test positivo consente di confermare la presenza della malattia con un buon livello di confidenza
- Quando LR- è minore di 0,2, un test negativo consente di escludere la malattia con un buon livello di confidenza

Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994 Mar 2;271(9):703-7.

Curva ROC

- ROC: Receiver Operating Characteristic
- La Curva ROC è la rappresentazione grafica della sensibilità (asse Y) rispetto alla specificità (asse X: 1-specificità))
- Per ogni valore della variabile in esame valutata rispetto ad un criterio diagnostico di riferimento (Gold Standard) si calcola la specificità e la sensibilità, la curva si crea unendo tali punti
- AUC: Area Under the Curve è l'area sotto la curva



— SUVmax 2
 - - - SUVmean 2

Comparison of ROC curves			
Variable 1	SUVmax_2		
Variable 2	SUVmean_2		
Classification variable	T01		
Positive group	T01		
T01	= 1		
Sample size	44		
Negative group	T01		
T01	= 0		
Sample size	12		
	AUC	SE	95% CI
SUVmax_2	0,798	0,063	0,670 to 0,894
SUVmean_2	0,782	0,066	0,652 to 0,881
Pairwise comparison of ROC curves			
SUVmax_2 - SUVmean_2			
Difference between areas	0,016		
Standard error	0,012		
95% Confidence interval	-0,008 to 0,040		
Significance level	P = 0,192		

Interpretazione di AUC

- 0.5-0.7 bassa capacità discriminante
 - 0.7-0.9 media-buona capacità discriminante
 - > 0.9 ottima capacità discriminante
-
- Il cut off viene identificato nel punto più in alto a sinistra e cioè quello corrispondente alla migliore combinazione di sensibilità e specificità

Rappresentazione grafica del cut-off

