

Si ricorda nell'effettuare un test di specificare ogni volta le ipotesi nulla e alternativa.

Es 1 (tutti). Si considerino i seguenti dati

-1.3 3.0 0.7 -0.1 2.7 -0.2 2.2 1.5 1.3 1.4

- 1) Si costruisca l'istogramma suddividendo in 3 classi [-1.5, 0[, [0,1.5[, [1.5,3]. Se ne faccia il grafico.
- 2) Si stimino la media pesata e la varianza pesata
- 3) Si ipotizzi che i dati derivino da una distribuzione gaussiana, si faccia il grafico della densità di probabilità dalla quale i dati sono ottenuti

In questo caso è rilevante che la media pesata siano calcolata pesando i valori centrali degli intervalli. Inoltre nel grafico della densità di probabilità devono essere tarate opportunamente le ascisse.

Es 2 (tutti)

Si vuole testare se due dispositivi per la produzione di sfere in acciaio, producano sfere delle stesse dimensioni. Si confrontano due serie di misure effettuate su 24 sfere, 12 prodotte dal dispositivo A e 12 dal dispositivo B. Si può dire che le due macchine funzionino nello stesso modo? Si ipotizzino i dati derivanti da distribuzioni gaussiane. ($\alpha = 0.05$). Si descrivano le ipotesi che devono essere garantite per l'esecuzione corretta del test in oggetto.

A 0.265 0.265 0.266 0.267 0.267 0.265 0.267 0.267 0.265 0.268 0.268 0.265
 B 0.264 0.265 0.264 0.266 0.267 0.268 0.264 0.265 0.265 0.267 0.268 0.269

*Un t-test per dati indipendenti e varianze incognite.
 I gradi di libertà sono 22 e l'ipotesi è bilatera (t critico= 2.074)
 Il valore di t trovato è 0.405.*

Es 3.1

L'Analisi di Sholl è utilizzata per caratterizzare le ramificazioni dendritiche di neuroni. In particolare, da questa analisi viene deciso di valutare il numero medio di ramificazioni al variare della distanza dal soma. Due algoritmi di segmentazione di immagini di microscopia confocale vengono utilizzati per stimare tale grandezza su un campione di 20 neuroni estratti casualmente dalla stessa fetta di tessuto. 10 di questi campioni sono analizzati con un algoritmo, gli altri 10 con l'altro. Nelle ipotesi che le caratteristiche dei due gruppi siano omogenee, si vuole testare se i due algoritmi hanno prestazioni analoghe ($\alpha = 0.05$). (n.d.r. questa modalità di verifica delle prestazioni dei due algoritmi non è consigliata).

Alg1	
40	62
45	44
50	42
33	40
36	39

Alg2	
38	55
46	51
47	40
30	29
37	37

*In questo caso i campioni sono indipendenti. Vista la non certa e difficilmente verificabile non gaussianità è opportuno usare un test di Mann Whitney.
 Si può utilizzare l'approssimazione gaussiana, per cui il valore medio dei ranghi nell'ipotesi nulla è pari a $\mu=10*21/2=105$ e la dev std è pari a $\sigma=13.23$. Visto che la somma dei ranghi è 111 (o 99). Si ottiene $z=0.453$ (o -0.453). Tale valore è inferiore a 1.96 (superiore a -1.96).*

Es 3.2

L'Analisi di Sholl è utilizzata per caratterizzare le ramificazioni dendritiche di neuroni. In particolare, da questa analisi viene deciso di valutare il numero medio di ramificazioni al variare della distanza dal soma. Due algoritmi di segmentazione di immagini di microscopia confocale vengono utilizzati per stimare tale grandezza su un campione di 10 neuroni. Si vuole testare se i due algoritmi hanno prestazioni analoghe ($\alpha = 0.05$).

Neurone	Alg1	Alg2
1	40	38
2	45	46
3	50	47
4	33	30
5	36	37

Neurone	Alg1	Alg2
6	62	55
7	44	51
8	42	40
9	40	29
10	39	37

In questo caso i dati non sono indipendenti. Gli algoritmi sono applicati agli stessi neuroni. Il valore della statistica è pari a 32 (sommati sia positivi che negativi). È possibile usare sia l'approssimazione che la tabella esatta (preferibile visto che i valori sono tabulati). Si ottiene che 39 corrisponde al valore critico per $p=0.048$. Quindi non si può rifiutare l'ipotesi nulla.

Es 4 (tutti)

Tramite fMRI si misura l'ampiezza della risposta nel cingolo anteriore ad tre diversi stimoli a diversi livelli di arousal (L1,L2,L3). Si includono nello studio 18 soggetti che vengono assegnati in modo casuale ai tre compiti. Le misure delle risposte sono determinate in variazione percentuale rispetto alla linea base e sono di seguito riportate. Nell'ipotesi di distribuzione gaussiana delle risposte, si dica se esiste una differenza nelle risposte ai tre stimoli ($\alpha = 0.05$). Nel caso esista, eseguire dei test per ottenere maggiori informazioni sull'origine di queste differenze. Nel caso non esista, si dica come in generale queste informazioni si sarebbero potute ottenere nel caso di differenze significative.

L1	L2	L3
0.2	-0.4	1.1
-0.1	1.2	0.7
0.3	0.8	0.4
0.5	0.3	1.5
1.1	-0.4	1.7
-0.4	0.4	2.1

L'analisi della varianza porta a questi valori

Analisi varianza: ad un fattore

RIEPILOGO

Gruppi	Conteggio	Somma	Media	Varianza
Colonna 1	6	1.6	0.266666667	0.266666667
Colonna 2	6	1.9	0.316666667	0.409666667
Colonna 3	6	7.5	1.25	0.407

ANALISI VARIANZA

Origine della variazione	SQ	gdl	MQ	F	Valore di significatività	F crit
Tra gruppi	3.681111111	2	1.840555556	5.096923077	0.020463311	3.682320344
In gruppi	5.416666667	15	0.361111111			
Totale	9.097777778	17				

Successivamente si eseguono dei t test a coppie (post hoc). Per la successiva analisi post hoc è importante:

- 1) Dividere la significatività per il numero di test multipli $0.05/3=0.016667$ (visto che il test è a due code

si trova 0.00833)

2) Inoltre, i gdl sono $3 \cdot (6-1) = 15$ quindi il valore più vicino per il t critico è 2.602.

Es. 5 (tutti)

Si confrontano due algoritmi (A1 e A2) per l'analisi delle immagini radiologiche. Si vuole verificare se i due approcci abbiano prestazioni differenti o meno ($\alpha = 0.05$). Si testano i due algoritmi su 334 immagini di soggetti malati. L'algoritmo A1 rileva 138 casi, mentre l'algoritmo A2 ne rileva 136. 84 casi sono stati rilevati da entrambi, mentre 144 risultano negativi ad entrambi gli algoritmi.

Mc Nemar

		A2		
		Pos.	Neg.	
A1	Pos.	84	54	138
	Neg..	52	144	196
		136	198	334

Si considerano solo gli elementi che hanno fornito valori discordanti e se ne calcola la frequenza attesa nell'ipotesi nulla (53)

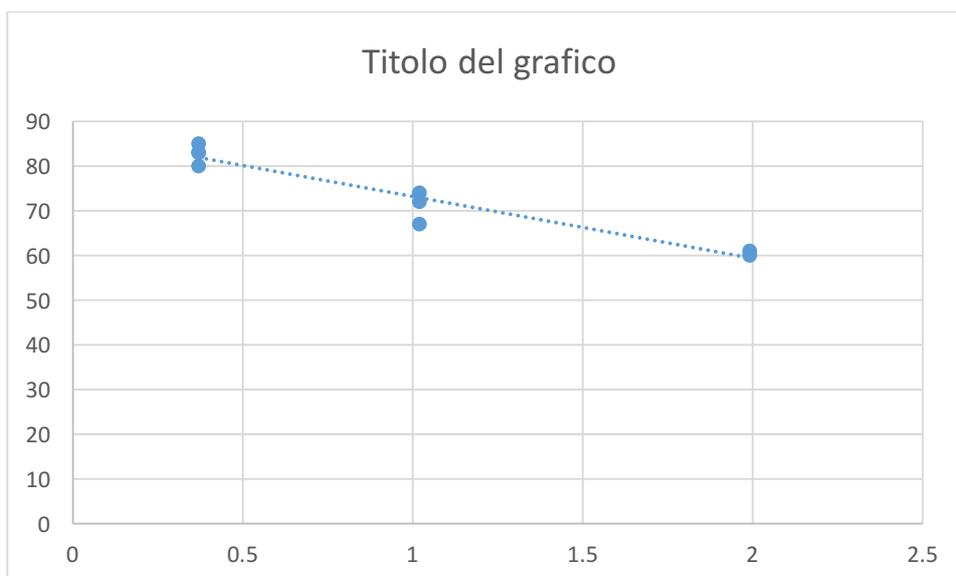
$$W = \frac{(|54 - 53| - 0.5)^2}{53} + \frac{(|52 - 53| - 0.5)^2}{53} = 0.009$$

va confrontato con il valore critico del χ^2 con 1 gdl che è pari a 3.841

Es 6 (2016-2017)

Nella tabella seguente sono mostrate le percentuali di Radon rilasciate durante una doccia, in funzione del diametro degli fori per il passaggio dell'acqua. Si vuole verificare se esista una relazione significativa tra le due grandezze ($\alpha = 0.05$). Si faccia il grafico dei dati che rappresenti tale relazione.

Diam. Foro (mm)	% di Radon rilasciata			
0.37	80	83	83	85
1.02	67	72	74	
1.99	60	61		



OUTPUT RIEPILOGO

<i>Statistica della regressione</i>	
R multiplo	0.961905173
R al quadrato	0.925261562
R al quadrato corretto	0.914584642
Errore standard	2.804942699
Osservazioni	9

ANALISI VARIANZA

	<i>gdl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>Significatività F</i>
Regressione	1	681.8149641	681.8149641	86.6599714	3.42273E-05
Residuo	7	55.07392481	7.867703544		
Totale	8	736.8888889			

	<i>Coefficienti</i>	<i>Errore standard</i>	<i>Stat t</i>	<i>Valore di significatività</i>
Intercetta	87.05778509	1.695683309	51.34082799	2.78581E-10
Variabile X 1	-13.91080585	1.494317968	-9.309133762	3.42273E-05

D1 (tutti)

Si descriva la rilevanza dei test per la verifica della normalità dei dati. Si discutano i dettagli applicativi di un test per la verifica della gaussianità dei dati.

D2 (tutti)

Discutere il significato dell'operazione di inferenza sulla media e suoi utilizzi.

D3 (2016-2017)

Si descrivano le finalità dell'anova a due fattori di classificazione. Si introducano le procedure per il suo utilizzo. Si discuta il significato di interazione tra fattori e in quale caso sia possibile studiarlo.

Importante: per analizzare l'interazione abbiamo bisogno di repliche per ogni coppia di fattori.

D4 (2016-2017)

Si discutano le ipotesi che devono essere soddisfatte affinché un modello di regressione sia valido. Si mostri il significato del coefficiente di correlazione ed il suo legame con i parametri e/o le statistiche del modello. Si possono utilizzare anche dei grafici esplicativi.

Nota importante: pensare alla differenza tra validità e significatività.