

AMSB 28/1/2019 AA1718 e precedenti

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

Es 1 (4 punti) Si consideri una variabile aleatoria caratterizzata da una distribuzione t di Student con 5 gradi di libertà. Si ipotizzi di ottenere N valori campionari di tale misura.

- 1) Discutere la forma dell'istogramma di tali misure all'aumentare di N
- 2) Discutere la forma dell'istogramma di tali misure all'aumentare dei gradi di libertà
- 3) Consideriamo il valore medio di N misure derivate da tale distribuzione con 5 gradi di libertà. Discutere come può essere descritto dal punto di vista statistico tale parametro al variare del numero N di misure.

- 1) *Non bisogna confondere il numero di misure con i gdl della variabile t. Se aumentiamo i campioni l'istogramma diventa più fedele alla forma della densità di probabilità t di student a 5 gdl.*
- 2) *In questo caso l'istogramma assomiglia a quello di una variabile gaussiana*
- 3) *Il valore medio è una variabile aleatoria e all'aumentare del numero di misure che vengono mediate, la sua distribuzione assomiglierà sempre di più a quella di una gaussiana*

Es 2 (3 punti) Vorremmo sapere se le persone di 40 anni siano in media più pesanti di quelle di 30. Si eseguono quindi delle misure su soggetti appartenenti ai due gruppi. Si considerano i dati a distribuzione gaussiana. Si indichino ipotesi nulla e alternativa.

30 anni	77	65	73	58	63	49	51	82	103	69					
40 anni	102	73	56	55	83	72	88	70	81	85	44	71	62	78	75

Si commenti se questo esperimento è da considerarsi valido per verificare l'ipotesi posta e se indichino eventuali modifiche al piano sperimentale.

Si usa un t test. Le varianze sono simili.

L'ipotesi alternativa è monolatera.

La statistica vale 0.6458 e il valore critico pari a (per alfa 0.05) è 1.714. Si accetta l'ipotesi nulla. Non era stato data la significatività: lo studente poteva sceglierla in autonomia. È ovvio che se stiamo facendo un test dobbiamo scegliere un valore per la significatività

Es 3 (3 punti) Si vuole verificare se trattamenti del dolore, dopo danni muscolo-scheletrici, siano equivalenti o meno. Si riportano i risultati ottenuti su 300 soggetti. Si utilizzi una significatività $\alpha = 0.01$.

	Farmaco A	Farmaco B	Farmaco C
Miglioramento Significativo	58	81	61
Leggero Miglioramento	42	19	39

	Farmaco A	Farmaco B	Farmaco C	
Miglioramento Significativo	58	81	61	200
Leggero Miglioramento	42	19	39	100
	100	100	100	300

Si usa un test del chi quadro. Le frequenze attese sono calcolate come se i farmaci non spiegassero le differenze osservate.

	Farmaco A	Farmaco B	Farmaco C	
Miglioramento Significativo	$100 \cdot 200 / 300 = 66.7$	$100 \cdot 200 / 300 = 66.7$	$100 \cdot 200 / 300 = 66.7$	200
Leggero Miglioramento	$100 \cdot 100 / 300 = 33.3$	$100 \cdot 100 / 300 = 33.3$	$100 \cdot 100 / 300 = 33.3$	100
	100	100	100	300

Chi-quadro=14.08

Il valore critico è pari a 9.21. Quindi possiamo rifiutare l'ipotesi nulla di uguaglianza degli effetti dei farmaci

Es 4 (3 punti) Ci interessa avere la stima del tempo di vita di un dispositivo. Vengono effettuate le seguenti misure (in ore)

1052 1271 836 962 1019 1051 512 1027 1219 1040

Sapendo che la deviazione standard del tempo di vita di tale dispositivo sia nota e pari a 80 ore, i trovi l'intervallo di confidenza al 95% di tale parametro.

Mantenendo tale percentuale, si forniscano indicazioni precise per fare in modo che tale intervallo sia ampio 50 ore.

Per intervallo di confidenza, in questo caso, si usa la statistica z

L'intervallo è pari a (949.32, 1048.5)

Per il secondo punto si trova che il numero di misure è pari circa a 39. Infatti,

$$\frac{2 \cdot 1.96 \cdot 80}{\sqrt{n}} = 50$$

Es 5 (3 punti)

Nella tabella è mostrata la densità di cellule nervose in due sezioni diverse dell'intestino di 6 individui. La distribuzione della densità segue una distribuzione t di Student. È possibile dire che le due sezioni differiscano in termini di densità di cellule nervose? Si usi una significatività pari a $\alpha = 0.05$.

Soggetto	Reg. A	Reg. B
1	50.6	38.0
2	39.2	39.2
3	35.2	23.2
4	17.0	19.0
5	11.2	6.6
6	14.2	16.4

Si deve usare un test non parametrico per dati appaiati. Il test dei ranghi con segno (Wilcoxon è adatto). Vediamo le differenze e ne facciamo i ranghi. Una differenza è zero, quindi i gradi di libertà si riducono a 5. La statistica è pari a 9 (o -9) e il valore critico pari a 15 (o -15). Si accetta l'ipotesi nulla di uguaglianza dei valori nelle due sezioni.

Es 6 (3 punti) In un modello di regressione, che tiene conto di 7 coppie di misure si trova che la somma dei quadrati relativa alla regressione è pari a 199.84. La somma totale dei quadrati dell'errore è pari a 205.71. I valori dell'intercetta e della pendenza della regressione sono pari a -0.0083 e 0.2438 rispettivamente. Si dica se tale modello sia significativo o meno ($\alpha = 0.05$). Si discuta se e come potrebbe essere determinato l'intervallo di confidenza per la pendenza del modello.

Nel testo c'è una ambiguità non voluta: 205.71 può essere considerato SSTO o SSE (la frase "somma totale dei quadrati dell'errore" è ambigua). Di conseguenza entrambe le interpretazioni sono state considerate valide. (N.B. nel primo caso $SSE=SSTO-SSR$)

Qui di seguito si risolve il secondo caso ($SSE=205.71$)

$$MSR=SSR/1=199.84$$

$$MSE=SSE/5=205.71/5=41.14$$

$$F=199.84/41.14=4.8576$$

Il valore critico per $F(1,5)=6.61$. Il modello di regressione non è significativo.

L'intervallo di confidenza per la pendenza del modello può essere ottenuta a partire dalla sua stima, utilizzando la deviazione standard della stima e tenendo in mente che è una variabile t di Student con N-2 gradi di libertà.

D1 (2 punti) (solo per coloro che hanno seguito nel AA 2016_2017)

Si descrivano le ipotesi che devono essere verificate per la validità del test anova.

D1 (2 punti) (solo per coloro che hanno seguito nel AA 2017_2018)

Enunciare il teorema di Bayes e il teorema delle probabilità totali. Discutere con un esempio l'utilizzo di uno o di entrambi i teoremi.

D2 (3 punti)

Si indichi la procedura per determinare i valori critici nel caso di un test di Mann Whitney. Si faccia l'esempio nel caso di un test applicato a due insiemi di misure con numerosità pari a 2 e 8 rispettivamente.

Non viene chiesto come si esegue un test di Mann Whitney ma come si determinano i valori critici.

D3 (3 punti)

Si consideri una variabile aleatoria discreta che assuma valori pari a -1, 0, 1 e 2, e solo questi, con probabilità rispettivamente pari a 0.3, 0.1, 0.3 e x.

Fare il grafico della densità di probabilità e della funzione di distribuzione di tale variabile.

Stimarne il valore medio e il momento centrale del terzo ordine.

D4 (3 punti)

Si consideri una variabile aleatoria continua y che sia funzione di una seconda variabile continua x: $y=g(x)$

Si introduca la metodologia per la stima della densità di probabilità di y a partire da quella di x.