AMSB 28/1/2019 AA1819

Es 1 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 Si vuole verificare se un farmaco provochi o meno nausea, rispetto all'uso di un placebo. Nella tabella seguente si riportano i risultati su 565 soggetti. Si usi una significatività pari a $\alpha = 0.05$.

	Farmaco	**Placebo**
Nausea	34	13
No Nausea	256	262

	Farmaco	***Placebo***	
Nausea	*34*	*13*	*47*
No Nausea	*256*	*262*	*518*
	290	*275*	*565*

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

	Farmaco	***Placebo***	
Nausea	*260*47/518=24.1*	*275*47/565=22.9*	*47*
No Nausea	*290*518/565=265.9*	*275*518/565=252.1*	*518*
	290	*275*	*565*

Chi-quadro=9.1
Il valore critico è pari a 3.841. Quindi possiamo rifiutare l'ipotesi nulla di uguaglianza degli effetti tra farmaco e placebo.

Es 3 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 4 Un panettiere vuole verificare se tenendo la finestra aperta del suo laboratorio, l'odore del pane permetterà, come lui ritiene, di aumentare le vendite. Sceglie casualmente 10 giorni nei quali terrà la finestra aperta e 10 giorni nei quali la terrà chiusa e prende nota delle vendite. È possibile dire che la sua ipotesi è corretta? Si usi una significatività pari a $\alpha = 0.05$.

Vendite finestra aperta
202.0 204.5 207.0 215.5 190.8 215.6 208.8 187.8 204.1 185.7

Vendite finestra chiusa
193.5 192.2 199.4 177.6 205.4 200.6 181.8 169.2 172.2 192.8

Si specifichino ipotesi nulla e alternativa e si considerino i dati a distribuzione gaussiana.

Si può usare un t test per dati indipendenti. Le varianze sono simili (si deve effettuare il test di Fisher). Il test fornisce un valore pari a t=2.63 a fronte di un valore critico pari a 1.734 (ipotesi alternativa monolatera). I gdl sono 18

Es 5 Nella seguente tabella vengono mostrati il contenuto di clorofilla ottenute da misure effettuate su 4 varietà di carciofi. Si può dire che il contenuto di clorofilla dipenda dalla varietà? In caso affermativo si verifichi l'origine di una eventuale differenza. Si usi una significatività pari a $\alpha = 0.05$.

	Numerosità	**Media**	**Dev. Stand.**
Varietà 1	5	0.3	0.12
Varietà 2	5	0.24	0.089
Varietà 3	4	0.41	0.04
Varietà 4	6	0.33	0.054

Test ANOVA
*SSentro=4*0.12^2+4*0.089^2+3*0.04^2+5*0.054^2=0.1087*
SStra=5(0.3-0.316)^2+5*(0.24-0.316)^2+4*(0.41-0.316)^2+6*(0.33-0.316)^2=0.0667*

S2_entro=SSentro/(20-4)=0.0068
S2_tra=SStra/(4-1)=0.0222

F=3.27

Il valore critico per F è 3.24. Si deve rifiutare l'ipotesi nulla. Si deve porre attenzione al fatto che la statistica è molto vicina alla soglia della non significatività.
Si devono poi eseguire dei t test a coppie: si deve dividere la significatività per il numero contemporanei di test (6), usare i gdl pari a 16 e l'errore standard pari a $\sqrt{S_{entro}^2\left(\frac{1}{n1}+\frac{1}{n2}\right)}$ *con n1 e n2 la numerosità delle varietà che si vogliono confrontare. Il valore critico dei test è pari quindi a 2.921 (il migliore che otteniamo dalla tabella a disposizione, quello esatto sarebbe 3.01).*

Es 6 Si faccia il grafico della massa di probabilità di una binomiale con p=0.25 e numero di prove pari a 4.
Si calcoli la probabilità di avere

1) 2 successi su 4 prove
2) più di 2 successi

Consideriamo un mazzo da 52 carte da gioco, contenenti 4 assi e prendiamone a caso 13. Ogni asso avrà la probabilità pari a 0.25 di essere estratto. Consideriamo adesso la probabilità che nell'estrazione di 13 carte abbiamo due assi.
Si discuta se questa probabilità è uguale a quella calcolata al punto 1.

1) *0.2109*
2) *0.0508*

La probabilità richiesta nell'ultimo punto è $\frac{C_{4,2}C_{48,11}}{C_{52,13}} = 0.2135$

D1 In un esperimento di fisica il numero di particelle al minuto su un rilevatore è pari a 15.
Si forniscano le seguenti probabilità:

1) Probabilità che, osservando il fenomeno per 1 minuto, si abbiano due eventi nei primi 10 secondi.
2) Probabilità che si abbia un numero di eventi uguale o inferiore a 2
3) La probabilità di avere 3 eventi negli ultimi 15 secondi

1) *Noi siamo interessati a contare gli eventi in 10 secondi. Quindi* $\Lambda = \frac{15}{6} = 2.5$. *N.B. 6 perché 10 secondi sono 1/6 di un minuto.*

La probabilità che cerchiamo è $e^{-2.5}\frac{2.5^2}{2!} = 0.2565$

2) $\sum_{i=0}^{2} e^{-2.5}\frac{2.5^i}{i!} = 0.5438$
3) *Non conta il fatto che siano gli ultimi 15 secondi o i primi 10, conta solo il tempo che stiamo considerando*

$$\Lambda = \frac{15}{4} = 3.75$$
$$e^{-3.75}\frac{3.75^3}{3!} = 0.2067$$

D2 Si discuta il significato di stimatore polarizzato e consistente. Si indichi uno stimatore consistente e se ne dia la dimostrazione.

D3
Si facciano i grafici della densità di probabilità e della funzione di distribuzione di una variabile aleatoria continua uniformemente distribuita tra -3 e 2
Si calcoli il momento centrale del secondo ordine di tale variabile.
Si calcoli la probabilità che tale variabile assuma valori tra -1 e 0

La variabile aleatoria continua è descritta negli appunti: in questo caso la densità di probabilità tra -3 e 2 vale 0.2. la Funzione di distribuzione vale 0 per x minore di 3 e cresce linearmente fino al valore 1 in x=2 e rimane da quel punto costante.
Si ricorda che il momento centrale del secondo ordine è la varianza. La probabilità al terzo punto è 0.2.

D4
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.