

Biostatistica - 4 Luglio 2017

Nome:

Cognome:

Matricola:

Esercizio 1

Un laboratorio di analisi biochimiche ha condotto uno studio per valutare i valori di riferimento della concentrazione di emoglobina nel sangue. In particolare si voleva valutare se vi fossero differenze per diverse classi di età e per sesso dei soggetti.

Per tale studio vengono scelti 3 campioni di soggetti, ciascuno con numerosità 7:

uomini under 50	donne under 50	uomini over 50
14.1	12.9	13.5
17.3	15.5	14.5
15.8	14.3	13.2
16.5	13.7	12.8
15.6	15.1	13.6
17.2	13.2	14.3
14.8	12.4	14.9

a) Considerando i campioni come realizzazione di una variabile aleatoria caratterizzata da una funzione densità di probabilità risultante dalla differenza tra una Gaussiana con media μ_1 e una distribuzione Z , verificare, con un opportuno test statistico, se i campioni sono realizzazione della medesima variabile aleatoria. In caso negativo, indicare le coppie di campioni statisticamente differenti;

b) Ripetere il punto a) considerando i campioni come appartenenti ad una distribuzione t-Student a 3 g.d.l.,

c) Ripetere il punto a) considerando i campioni come appartenenti ad una distribuzione t-Student a 300 g.d.l.,

d) Ripetere il punto a) considerando i campioni come appartenenti ad una distribuzione completamente caratterizzata dal momento del primo ordine, dal momento centrale del 2° ordine, dal momento del 3° e 4° ordine, e momento centrale del 5° ordine.

(10 punti)

Esercizio 2

I valori di Pressione Arteriosa Sistolica (SAP), in soggetti sani, sono completamente caratterizzati da una funzione densità di probabilità Gaussiana con mediana pari a 120 mmHg e varianza 100 mmHg.

a) Calcolare la probabilità che i soggetti sani abbiano un valore di pressione arteriosa sistolica maggiore di 140 mmHg;

b) Verificare se un campione di 18 pazienti con patologie muscolari, su cui è stato misurato un valore medio di SAP pari a 125 mmHg con una varianza pari a 121 mmHg, possano essere inclusi nella popolazione di soggetti sani per quanto riguarda i valori SAP;

c) Nel campione di dati indicati al punto precedente, indicare un valore o un range di valori riguardo la SAP media della popolazione di pazienti con patologie muscolari

d) Considerando come ignota la varianza della SAP nella popolazione dei sani (e media nota), cosa si può affermare sulla probabilità che i soggetti sani abbiano un valore di pressione arteriosa sistolica maggiore, tutt'al più uguale a 150 mmHg?

(7 punti)

Esercizio 3

Una statistica nazionale indica che nel 2005 il 75% dei bambini di età compresa tra i 2 e i 17 anni si è sottoposto ad almeno un intervento odontoiatrico. Si vuol sapere se la richiesta di interventi è simile in bambini che vivono a Pisa. Per questo, sono stati intervistati 125 bambini dello stesso range di età, evidenziando che 64 di essi hanno subito interventi odontoiatrici nel 2005.

E' possibile affermare che esiste una similarità statistica tra gli interventi a Pisa e quelli nazionali?
(4 punti)

Esercizio 4

Considerando le seguente realizzazione di un vettore aleatorio:

$$V = [4.3, 1.2, 7.4, 8.3, 3.5, 2.6, 9.1, 5.5, 6.8, 0.7, 9.3, 3.4, 8.2, 1.6, 0.3]$$

- Dimostrare che la Gaussianità della popolazione da cui è stato estratto il campione mediante il test di Shapiro-Wilk
- In base all'esito dei punti precedenti, descrivere il campione mediante indici descrittivi opportuni, giustificando la scelta
- Dopo aver elencato le opportune ipotesi sulla variabile aleatoria in oggetto, descrivere il campione mediante opportuni grafici di frequenze relative, assolute e cumulative
- Considerando il campione come realizzazione di una variabile aleatoria Gaussiana, fare inferenza sulla varianza della popolazione con una probabilità del 99%
- Considerando il campione: K calcolare un coefficiente di correlazione non-parametrico e un coefficiente parametrico. Di quest'ultimo verificare al 95% l'ipotesi di avere variabili incorrelate:

$$K = [2.2 \ 0.8 \ 3.9 \ 4.1 \ 1.9 \ 1.5 \ 4.9 \ 2.8 \ 3.7 \ 0.5 \ 4.9 \ 2.2 \ 4.2 \ 1.1 \ 0.2]$$

(8 punti)

Esercizio 5

Dopo aver dato la definizione formale di probabilità:

a) calcolare la probabilità che, tra 10 medici chirurghi identificati come M_1, M_2, \dots, M_{10} , per un delicato intervento un vengano scelti esattamente M_2, M_4 , e M_7 , in modo totalmente casuale (ogni medico ha la stessa probabilità di essere scelto)

b) ricalcolare la probabilità richiesta al punto a) sapendo che saranno scelti solo medici identificati con pedice pari (quindi tra $M_2, M_4, M_6, M_8, M_{10}$)

(4 punti)

Alcuni chiarimenti per la presentazione dell'elaborato:

- Sarà corretto solo quanto è riportato a penna. Di questa, è ammesso un solo colore: nero o blu.
- Non sono ammessi strumenti per la cancellazione di quanto scritto (es. bianchetto). Ciò non esclude la possibilità di cancellare del testo che si ritiene errato mediante una linea sul testo stesso.
- La lingua ufficiale di questo esame è l'Italiano. Per questo, non saranno considerate risposte date in altre lingue (es. Inglese), malgrado queste possano essere corrette.
- Il riferimento al numero di ogni esercizio deve essere chiaramente indicato prima dello svolgimento di quest'ultimo per essere considerato valido.
- Gli esercizi presentati senza svolgimento o formule o esaustive giustificazioni verranno considerati con punteggio nullo anche se è presente il risultato corretto.
- Le tabelle e l'eventuale formulario utilizzati per lo svolgimento dell'esame devono essere consegnati insieme alla traccia e allo svolgimento dello stesso.