

Esame di Statistica del 20 giugno 2025

Tempo a disposizione 1h 40m.

Le domande vero/falso valgono 1.5 punti, risposte errate nelle domande vero/falso pesano in negativo sulla valutazione. Risposte errate alle altre domande di questa pagina non pesano in negativo.

Il testo deve essere consegnato (con nome e cognome) per consentire la correzione.

Nome e cognome	Matricola	1
----------------	-----------	---

☐ V ☐ F Un intervallo di confidenza per la media è tanto più piccolo quanto meno numeroso è il campione (a parità di tutto il resto).

☐ V ☐ F La pendenza della retta di regressione tra due variabili è positiva se e solo se la correlazione tra le due è positiva.

☐ V ☐ F A parità di tutto il resto, una regione di accettazione per un test sulla media è tanto più grande quanto più piccolo è il campione.

☐ V ☐ F Un campione, per essere rappresentativo, dev'essere tanto più grande quanto più grande è la popolazione.

☐ V ☐ F Se la varianza di un insieme di osservazioni è 0, primo e terzo quartile coincidono

☐ V ☐ F La media può essere calcolata per qualunque tipo di variabile

☐ V ☐ F I quantili empirici sono le osservazioni ordinate.

☐ V ☐ F La probabilità di un evento impossibile è 1.

☐ V ☐ F La deviazione standard è la radice quadrata della media.

☐ V ☐ F La media aritmetica di una distribuzione simmetrica coincide sempre con la mediana.

1 (5 punti) Il ritardo con cui un treno arriva al binario, espresso in secondi, è una variabile aleatoria distribuita secondo un'esponenziale di parametro 0.0111 (la funzione di ripartizione di un'esponenziale di parametro λ è $F(x) = 1 - e^{-\lambda x}$).

a. Qual è la probabilità di attendere più di 207 secondi?

b. Su 50 arrivi, quante volte in media il treno ritarda più di 207 secondi?

2 (5 punti) La tabella a doppia entrata sottostante rappresenta la distribuzione di frequenze assolute congiunte della variabile quantitativa X con valori interi da 0 a 5 e la variabile Y qualitativa con valori A e B.

	0	1	2	3	4	5
A	2	2	3	3	2	5
B	3	7	4	4	6	8

a. Si calcolino le mediane di X condizionate ai possibili valori di Y .

b. Si calcolino media e varianza marginali di X .

3 (5 punti) Un operatore telefonico propone a 617 clienti un cambiamento nell'offerta. Di questi 352 accettano l'offerta.

a. Si fornisca un intervallo di confidenza al 98% per la probabilità che la proposta venga accettata.

b. Si verifichi l'ipotesi $H_0 : \pi = 0.611$ fornendo il valore p .

Fine I parte, II parte sul retro del foglio

Si ricorda che **rispondendo solo alla prima parte il voto massimo è 21** e che i quesiti della seconda parte vengono corretti solo se si è ottenuto almeno 18 nella prima parte.

ATTENZIONE:

per ottenere il massimo dei voti, è sufficiente svolgere tre dei quattro quesiti in questa pagina. Risposte errate possono pesare in negativo.

4

Uno studio ha raccolto dati su 19 studenti, registrando le ore medie di studio giornaliere durante la preparazione di un esame e il voto ottenuto.

Ore studio	Voto
5.8	33.8
9.1	37.3
7.4	31.8
4.4	26.8
9.8	34.1
6.3	32.6
8.1	32.5
7.7	33
9.4	33.9
8.3	33.9
6.4	31.5
6.7	31.9
6	30.4
7.7	31.2
7.8	32.7
3.8	23.7
2.6	23.6
8.7	33.8
4	24.5

- Costruire una rappresentazione grafica delle due variabili, e commentare la forma e il tipo della relazione.
- Calcolare un indice per valutare l'associazione tra le ore di studio ed il voto, e commentarne il significato. Nel caso in cui tale indice fosse risultato molto vicino a 0, come avremmo interpretato il risultato?

5

In una popolazione, la prevalenza della malattia M è pari a 8%. Un test diagnostico per questa malattia presenta le seguenti caratteristiche:

- se una persona è effettivamente malata, il test risulta positivo nel 90% dei casi; - se una persona è effettivamente sana, il test risulta negativo nel 92% dei casi.

- Calcolare la probabilità che una persona scelta a caso nella popolazione abbia un risultato positivo al test.
- Calcolare la probabilità che una persona con test positivo sia effettivamente malata. Commentare il risultato ottenuto.