

Esame di Statistica del 5 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 La varianza campionaria corretta è sempre maggiore di quella non corretta.
- ☐ V ☐ F Le altezze dei rettangoli che costituiscono un istogramma sono le frequenze relative.
- ☐ V ☐ F Al crescere della numerosità campionaria (tutto il resto rimanendo costante), la regione di rifiuto diventa più grande.
- ☐ V ☐ F Se $P(A|B) = 0.1$ e $P(A|\bar{B}) = 0.6$ allora $0.1 \leq P(A) \leq 0.6$
- ☐ V ☐ F Sia Φ la funzione di ripartizione della normale standard, allora $\Phi(-x) = \Phi(1-x)$
- ☐ V ☐ F Un intervallo di confidenza è tanto più ampio quanto maggiore è il livello di confidenza (a parità di tutto il resto).
- ☐ V ☐ F La correlazione è sempre compresa tra -1 e +1.
- ☐ V ☐ F Il valore medio può essere inferiore alla mediana.
- ☐ V ☐ F La distribuzione di una variabile categoriale può essere rappresentata graficamente da un istogramma
- ☐ V ☐ F La probabilità è una misura di incertezza.

1 (5 punti) Un'azienda di e-commerce si avvale per le consegne di un'azienda terza. Le ultime 88 consegne sono state completate mediamente in 2.03 giorni (varianza campionaria corretta 0.242).

- Si ottenga un intervallo di confidenza al 95% per il tempo medio di consegna in giorni.
- Si dica quante consegne ci si attende impieghino più di 2.5 giorni tra le prossime 73

2 (5 punti) Siano X_i , $i = 1, 2$ variabili aleatorie indipendenti e identicamente distribuite secondo una $\mathcal{N}(-6, \sigma^2)$, si sa inoltre che $P(X_i < -7.27) = 0.4$.

- Si ricavi σ^2 .
- Si ricavi il quantile 0.9 della distribuzione di $Y = -2.3X_1 + 4.6X_2$.

3 (5 punti) Si consideri la seguente distribuzione di frequenza per classi

Classe	[0, 1[	[1, 2[	[2, 4[	[4, 5[	[5, 7[	[7, 9[	[9, 11[	[11, 15[	[15, 20[	[20, 30]
Freq	0.13	0.16	0.26	0.04	0.06	0.02	0.14	0.13	0.02	0.04

- Si ottenga la distribuzione di frequenza per le classi [0, 5[, [5, 11[, [11, 30].
- Si ottengano mediana e quartili.

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

Un'azienda che imbottiglia acqua minerale vuole verificare se il processo produttivo garantisce che, in media, il volume imbottigliato sia pari a 1 litro. Un campione casuale di 20 bottiglie ha fornito i seguenti volumi (in litri):

1.006	0.997
0.997	0.996
1.005	0.999
1.002	0.996
0.996	0.992
0.999	1.007
1.004	1.01
0.998	1.001
1.004	1.001
1.001	0.999

- Indicare la media e la mediana del campione. Formulare l'ipotesi nulla e alternativa per un test sulla media. Indicare le assunzioni necessarie per la validità del test.
- Calcolare la statistica del test e confrontarla con il valore critico per un livello di significatività $\alpha = 0.05$. Commentare il risultato ottenuto. Indicare anche il livello di significatività osservato (o valore p o p -value).

5 Un'indagine su 210 adulti ha raccolto informazioni su due variabili: se la persona è fumatore (sì/no) e se pratica attività sportiva regolare (sì/no). I risultati sono i seguenti: 64 persone sono fumatori e fanno sport; 38 persone sono fumatori e non fanno sport; 51 persone non fumano e fanno sport; 57 persone non fumano e non fanno sport.

- Calcolare un indice di associazione tra le due variabili.
- Verificare, al livello di significatività $\alpha = 0.05$, se c'è una differenza significativa tra la probabilità che una persona che fa sport fumi e la probabilità che una persona che non fa sport fumi.