

Esame di Statistica del 10 luglio 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 Se $P(A \cap B) = 0$ allora o $P(A) = 0$ o $P(B) = 0$.
- ☐ V ☐ F L'intervallo interquartile comprende il 75% dei dati.
- ☐ V ☐ F È sempre vero che $E(\sqrt{X}) = \sqrt{E(X)}$
- ☐ V ☐ F Qualunque sia la variabile aleatoria X si ha $P(X \geq E(X)) = 0.5$
- ☐ V ☐ F La media non può essere inferiore al minimo delle osservazioni.
- ☐ V ☐ F Se X è binomiale con dimensione n e probabilità π , e n è grande, allora approssimativamente $P(X \leq n\pi + z\sqrt{n\pi(1-\pi)}) = \Phi(z)$
- ☐ V ☐ F Se una popolazione è composta di metà maschi e metà femmine, un campione, per essere rappresentativo, deve contenere (esattamente) metà maschi e metà femmine.
- ☐ V ☐ F La lunghezza di un intervallo di confidenza per una proporzione cresce al crescere del livello di confidenza.
- ☐ V ☐ F Se si moltiplicano delle osservazioni per 2 la loro varianza (diversa da zero) quadruplica.
- ☐ V ☐ F Per conoscere il valore di un parametro bisognerebbe osservare l'intera popolazione.

1 (5 punti) Si sono rilevati i redditi di un campione di 119 individui. Il reddito medio è risultato pari a 20.7 migliaia di euro, la media dei quadrati pari a 433.

- Si fornisca una regione di rifiuto al 2% per l'ipotesi nulla $H_0 : \mu = 21.1$ dove μ è il reddito medio della popolazione.
- Se la media della popolazione è 21.4, qual è la probabilità che il test rifiuti?

2 (5 punti) Una tratta extraurbana è percorsa da un pullman che è guidato da uno di due autisti: Maria e Claudia. Maria arriva in ritardo con probabilità 0.19, Claudia con probabilità 0.27. La metà delle giornate l'autista è Maria.

- Qual è la probabilità che il pullman arrivi in ritardo.
- Se si sa che le ultime due corse l'autista era lo stesso e il pullman è arrivato entrambe le volte in orario, qual è la probabilità che fosse guidato da Claudia?

3 (5 punti) Si considerino le seguenti frequenze relative cumulate.

	X	-2	3	4	5	8
Freq. cumulate		0.591	0.713	0.826	0.896	1.000

- Si ottengano media e varianza;
- Si disegni il boxplot.

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

Si consideri una variabile casuale discreta X , che rappresenta il numero di chiamate ricevute da un centralino in un'ora. La sua distribuzione di probabilità è la seguente:

x	$P(X = x)$
0	0.29
1	0.15
2	0.16
3	?
4	0.19

- Completa la tabella trovando la probabilità mancante. Rappresenta graficamente la funzione di ripartizione empirica (ECDF) della variabile X , e determina il primo quartile, rappresentandolo sul grafico.
- Si osservano 6 ore indipendenti. Sia Y il numero di ore in cui si ricevono almeno 2 chiamate.
 - Specifica la distribuzione della variabile Y .
 - Calcola la probabilità che $Y = 3$.
 - Calcola la probabilità che $2 \leq Y \leq 4$.

5

Un dirigente afferma che, grazie a un nuovo metodo di formazione, il punteggio medio ottenuto dai dipendenti in un test di abilità è superiore a 70.

Sono stati raccolti i seguenti punteggi (su 100) da un campione casuale di 8 dipendenti:

71, 70, 72, 72, 75, 75, 71, 75

- Si vuole verificare se i dati forniscono sufficiente evidenza a supporto dell'affermazione del dirigente.
 - Formula le ipotesi statistiche appropriate.
 - Esegui il test al livello di significatività del 5%.
 - Interpreta il risultato ottenuto.
- Se si volesse costruire un intervallo di confidenza per la media con ampiezza pari a 4 punti, quante osservazioni sarebbero necessarie mantenendo lo stesso livello di confidenza e la stessa varianza campionaria?