

Esame di Statistica del 31 maggio 2023

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	Bonus	1
----------------	-----------	-------	---

- V F La differenza interquartile è sempre maggiore o uguale a 0.
- V F L'intervallo di confidenza per la media al 95% contiene sempre il primo e il terzo quartile.
- V F La media aritmetica dei logaritmi è sempre maggiore del logaritmo della media aritmetica.
- V F Se $P(A|B) = 0.3$ e $P(A|\bar{B}) = 0.8$ allora $0.3 \leq P(A) \leq 0.8$
- V F La distribuzione binomiale è simmetrica rispetto alla media.
- V F Per ottenere una stima affidabile di una media, il campione dev'essere almeno pari al 10% della popolazione.
- V F Se X è normale di media -1 e varianza 1 e Y è normale di media -2 e varianza 1 e $c < -1$, allora $P(X \leq c) < P(Y \leq c)$.
- V F L'intervallo $[0, 1]$ è un intervallo di confidenza al livello del 100% per la proporzione.
- V F Se $X \sim \text{Binom}(11, 0.5)$ e Y è la percentuale di successi, la probabilità che Y sia (esattamente) pari al 50% è zero.
- V F La mediana non cambia se si aggiunge una costante a tutte le osservazioni.

1 Per stimare il peso medio delle confezioni di biscotti prodotte da un macchinario si pesano $n = 1000$ confezioni, la media aritmetica dei pesi delle confezioni osservate è 97 grammi, mentre la media dei quadrati dei è 9418.3.

- (2) Si fornisca un intervallo di confidenza al 90% per il parametro "peso medio delle confezioni prodotte dal macchinario".
- (3) Il peso nominale delle confezioni è 100 grammi, si fornisca il valore p per l'ipotesi nulla che la media è non inferiore a 100.

2 Sia X una variabile aleatoria con distribuzione binomiale con probabilità 0.6 e dimensione 661, si trovino

- (1) $E(X)$;
- (2) $P(371.4 < X < 421.8)$;
- (2) $P(371.4 < X < 421.8 | X > 371.4)$

3 Si consideri la seguente distribuzione di frequenze relative.

y	0	1	2	3	4
Freq	0.1	0.2	0.2	0.2	0.3

- (3) Si rappresenti la funzione di ripartizione empirica;
- (2) Si ottenga lo scarto interquartile;

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 Viene effettuato uno studio su studenti universitari afferenti a due diversi dipartimenti dell'Università di Trieste, Economia e Scienze Politiche, per verificare se esista una differenza nelle modalità di fruizione delle lezioni: esclusivamente in presenza, usufruendo delle registrazioni, o entrambe. I dati sono riassunti nella seguente tabella:

	In Presenza	Remoto	Entrambi	Totale
Economia	?	91	58	?
Scienze Politiche	?	?	692	3595
Totale	2309	?	?	4023

- Si verifichi attraverso un opportuno indice il grado di associazione tra le due variabili in esame, e si commenti il risultato.
- Si rappresentino graficamente le due variabili analizzate su di una stessa rappresentazione grafica.

5 Su di un campione casuale di 16 giorni del mese di maggio, l'intervallo di confidenza relativo alle temperature minime a Trieste è il seguente: (16.8; 22.4), con un livello di confidenza $1 - \alpha$ pari a 95%. Supponendo che le temperature minime si distribuiscano come una Normale e che la varianza sia ignota:

- Si trovi la media campionaria, la deviazione standard campionaria. Cosa si può dire della moda?
- Nuovi studi scientifici rivelano che, per tali dati sulle temperature minime, la varianza si può considerare nota e pari a 10. Quanto ampio deve essere il campione dei giorni osservati per ottenere un intervallo di confidenza al 98% non più ampio di 3?