

Esame di Statistica del 14 giugno 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
----------------	-----------	-------	---

La distribuzione di una variabile qualitativa ordinata può essere rappresentata graficamente da un istogramma.

Data una variabile X si ha $P(X \geq E(X)) = 0.5$

La regione di accettazione in una verifica d'ipotesi per la media è centrata sulla media campionaria.

Date due v.a. X e Y qualunque, è sempre vero che $P(X + Y \leq x + y) \leq P(X \leq x \cap Y \leq y)$.

Se la densità di frequenza della classe $[2; 6]$ è 0.2, la frequenza relativa della classe è 0.05.

Un campione costruito selezionando a caso n individui dalla popolazione in modo che ciascuno abbia la stessa probabilità di essere incluso è un campione rappresentativo

La stima è una caratteristica della popolazione.

Al crescere della numerosità campionaria (tutto il resto rimanendo costante), la regione di confidenza al 90% diventa più grande.

La probabilità che esca più di 10 lanciando due dadi a 6 facce è $2/12$.

È sempre vero che $P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$.

1 Il gestore di un sito di commercio elettronico vuole stimare la probabilità che un visitatore effettui un acquisto. Si sa che degli ultimi $n = 95$ visitatori 49 hanno effettuato un acquisto.

a. (2) Si fornisca un intervallo di confidenza al 95% per la probabilità cercata.

b. (3) Usando come probabilità la stima ottenuta al punto precedente, qual è la probabilità che ci siano più di 67 acquisti se ci sono 121 visitatori?

2 Di due eventi A e B è noto che $P(A \cap B) = 0.1443$ e $P(A \cup B) = 0.5852$ e che $P(A) = 0.39$. Si ricavino

a. (2) $P(B)$;

b. (3) $P(A|B)$;

3 Si osservano, per alcuni individui, se siano stati al cinema nell'ultimo mese (Sì/No) e il numero di libri letti nell'ultimo mese (0, 1-2 o più di 2). Le osservazioni sono riportate di seguito.

$(S, 0); (N, 0); (N, 0); (S, 0); (N, 0); (N, 1 - 2); (S, 1 - 2); (S, 3+); (N, 3+)$

a. (1) Si riassumano i dati in una tabella a due vie.

b. (2) Si ottenga la frequenza relativa condizionata delle persone che leggono almeno un libro tra chi non va al cinema.

c. (2) Si ottengano le frequenze attese se vi è indipendenza tra le due variabili.

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

I seguenti dati riguardano il tempo impiegato nel completamento di un compito di statistica, in minuti, registrato su di un campione casuale di 16 studenti:

63.00	68.00	112.00	85.00	98.00	97.00	88.00	113.00
68.00	99.00	66.00	103.00	81.00	66.00	93.00	80.00

- Si testi l'ipotesi nulla che il tempo medio è uguale a 100 contro l'ipotesi alternativa che sia minore di 100, chiarendo le ipotesi alla base.
- Si interpreti il risultato, anche sulla base del livello di significatività scelto. Cosa sarebbe potuto succedere al test se il livello di significatività scelto fosse stato più elevato?

5

Su di un campione casuale di 7 individui, le due variabili "ore di studio giornaliera" e "ore di sport giornaliera" vengono registrate:

	1	2	3	4	5	6	7
Ore di studio	4.50	2.20	0.30	1.10	5.30	1.50	6.10
Ore di sport	2.50	2.20	2.00	1.40	2.60	0.40	3.40

- Si rappresenti la relazione tra le due variabili graficamente e si calcoli un opportuno indice.
- Si calcoli il coefficiente angolare della retta di regressione e si discuta il risultato.