Esame 13/07/22

Quesito 1. Quando il modello di regressione lineare semplice è corretto, i residui standardizzati:

☒ hanno media nulla

☐ devono mostrare una forte regolarità prima crescente e poi decrescente all'aumentare del livello di ingresso

☐ hanno media diversa da zero

Quesito 2. La mediana campionaria di un insieme di dati

☒ dipende direttamente solo da uno o due valori in centro alla distribuzione e non risente dei dati estremi

☐ fa uso di tutti i valori ed è influenzata in maniera sensibile da valori eccezionalmente alti o bassi

☐ ne descrive la variabilità

Quesito 3. Si estraggono a caso due palline da un'urna che ne contiene 6 di bianche e 5 di nere. Qual è la probabilità che le due estratte siano una bianca e una nera?

☒ $\frac{6}{11}$

☐ $\frac{33}{55}$

☐ $\frac{3}{11}$

6B 5N 2 estrazioni

$$P(1B\,1N) = \frac{6 \cdot 5}{\binom{11}{2}} = \frac{30}{\frac{11!}{2!\,9!}} = \frac{30}{\frac{11 \cdot 10}{2}} = \frac{30}{55} = \frac{6}{11}$$

Quesito 4. Data una popolazione normale di media incognita μ e varianza nota σ^2 e definiti con n il numero di campioni estratti dalla popolazione e $\bar{x}$ la media campionaria, quando affermiamo che, con il 95% di confidenza, la media vera della distribuzione appartiene all'intervallo $(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}})$, stiamo parlando di

☒ intervallo di confidenza al 95% per μ

☐ intervallo di confidenza al 99% per μ

☐ probabilità che μ appartenga a quell'intervallo

Quesito 5. Se in un box-plot il rettangolo ha area nulla

☒ la differenza interquartile (IQR) è pari a zero

☐ la distribuzione dei valori dei campioni è asimmetrica

☐ questa situazione non si può verificare

Quesito 6. La distribuzione di probabilità di una variabile aleatoria binomiale (dove p = probabilità di successo in una prova, n = numero di prove) è asimmetrica se:

☒ il parametro p è diverso da 0.5

☐ il parametro p è uguale a 0.5

☐ non tutte le prove vengono effettuate nella stesse condizioni

Quesito 7. La resa di un certo investimento (in migliaia di dollari) è una variabile aleatoria X con distribuzione di probabilità $P\{X = -1\} = 0.7$, $P\{X = 4\} = 0.2$, $P\{X = 8\} = 0.1$. Quanto vale la varianza della resa dell'investimento?

☒ 9.49

☐ 10.3

☐ 0.849

$\bar{x} = -1 \cdot 0.7 + 4 \cdot 0,2 + 8 \cdot 0,1 = 0,9$

$\sigma^2 = E[x^2] - \bar{x}^2$

$E[x^2] = (-1)^2 \cdot 0.7 + 4^2 \cdot 0,2 + 8^2 \cdot 0,1 = 10,3$

$\sigma^2 = 10,3 - 0,9^2 = 9,49$

Quesito 8. Due eventi E e F si dicono indipendenti se vale

☒ $P(E \cap F) = P(E)P(F)$

☐ $P(E \mid F) \neq P(E)$

☐ $E \cap F = \emptyset$

Quesito 9. Quale delle seguenti proprietà riguardanti la distribuzione t di Student con n gradi di libertà (t_n) è vera?

☒ Ciascuna curva t_n è caratterizzata da "code" più spesse (più pesanti), ad indicare una variabilità maggiore rispetto alla gaussiana

☐ Al crescere di n, la densità di t_n non convergerà mai a quella della normale standard

☐ La densità delle distribuzioni t, a differenza di quella normale, non è simmetrica rispetto all'asse di ascissa 0

Quesito 10. La stima puntuale di un parametro di interesse è

☒ il valore assunto dallo stimatore in corrispondenza di un particolare campione

☐ una statistica

☐ una variabile aleatoria

Esercizio 1 [5 punti]

La probabilità che un soggetto abbia un'infezione virale è pari a 0.001. La diagnosi dell'infezione è effettuata mediante un test clinico che ha le seguenti caratteristiche: la probabilità che un soggetto infetto risulti positivo al test è 0.8, mentre la probabilità che un soggetto non infetto non risulti positivo al test è 0.84.
a) Qual è la probabilità che il test risulti positivo? (usare QUATTRO cifre decimali dopo la virgola)
b) Qual è la probabilità che un soggetto sia infetto dato che è risultato positivo al test? (usare TRE cifre decimali dopo la virgola)
Indichiamo con I l'evento "soggetto infetto" e con T l'evento "test positivo".

$P(I)=0{,}001 \qquad P(T|I)=0{,}8$

$P(S)=1-P(I) \qquad P(T|S)=1-0{,}84$

a)

$P(T)=P(T|I)\cdot P(I)+P(T|S)\cdot P(S)=$

$=0{,}8\cdot 0{,}001+(1-0{,}84)\cdot(1-0{,}001)=0{,}1606$

b)

$$P(I|T)=\frac{P(T|I)\cdot P(I)}{P(T)}=\frac{0{,}8\cdot 0{,}001}{0{,}1606}=0{,}005$$

Esercizio 2 [4 punti]

Un'indagine effettuata su un campione di 50 famiglie ha dato il seguente risultato:

N	f_{ass}	**Frequenza cumulativa**
0	5	**5**
1	13	**18**
2	15	**33**
3	10	**43**
4	4	**47**
5	2	**49**
6	1	**50**

N = numero figli per famiglia; f_{ass}= frequenza assoluta

Con queste informazioni a tua disposizione, calcolare:
a) il numero medio di figli per famiglia (usare UNA cifra decimale dopo la virgola)
b) la distanza interquartile
c) la moda campionaria
d) la deviazione standard campionaria (usare TRE cifre decimali dopo la virgola)

a)

$$\mu=\frac{0\cdot 5+1\cdot 13+2\cdot 15+3\cdot 10+4\cdot 4+5\cdot 2+6\cdot 1}{5+13+15+10+4+2+1}=2{,}1$$

b) $Q_1=\frac{n+1}{4}=\frac{51}{4}=12{,}75 \qquad Q_3=\frac{3(n+1)}{4}=38{,}25$

$\Downarrow \qquad \Downarrow$

1 Figlio $\qquad$ 3 Figli

$IQR = Q_3 - Q_1 = 3 - 1 = 2$

c) moda $= 2$

d) $\sigma^2 = \frac{1}{n-1} \sum \left((x - \bar{x})^2 \cdot f_{ass} \right) =$

$= \frac{1}{49} \cdot \left((0-2,1)^2 \cdot 5 + (1-2,1)^2 \cdot 13 + (2-2,1)^2 \cdot 15 + (3-2,1)^2 \cdot 10 + (4-2,1)^2 \cdot 4 + (5-2,1)^2 \cdot 2 + (6-2,1)^2 \cdot 1 \right) =$

$= 1,888$

$\sigma = \sqrt{1,888} = 1,374$