Esame 04/02/2022

Quesito 1. Dato un modello di regressione lineare semplice $Y = \alpha + \beta x$. Il metodo dei minimi quadrati consiste nello scegliere come stimatori dei coefficienti di regressione α e β i due valori di A e B che minimizzano:

☒ la somma del quadrato degli scarti tra dati osservati e dati teorici

☐ la somma degli scarti tra dati teorici e dati osservati

☐ la variabilità dei dati teorici

La prima é l'unica che nomina i quadrati

Quesito 2. Sia $X_1, X_2, ...X_n$ un campione estratto da una popolazione normale di media incognita μ e varianza nota σ^2, con n il numero di campioni estratti dalla popolazione e $\bar{x}$ la media campionaria. L'intervallo di confidenza bilaterale per la media μ è costruito aggiungendo e sottraendo alla media campionaria un margine d'errore $z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$. Il margine d'errore può essere ridotto se:

☒ si aumenta la dimensione del campione

☐ si aumenta il livello di confidenza $(1 - \alpha)$

☐ si diminuisce l'ampiezza del campione

Quesito 3. Abbiamo due campioni provenienti da una distribuzione normale relativi alla velocità di esecuzione di uno script con modelli differenti di processore. Ipotizzando che i due campioni abbiamo varianza simile, vogliamo testare, tramite test di Student a due code, se il tempo di esecuzione dipende dal modello usato (H_0 = le medie delle due distribuzioni sono uguali). Fissando una soglia di significatività $\alpha = 1\%$, otteniamo un intervallo di confidenza $[-2.4, +2.4]$ e un valore della statistica $t = 2.8$. Quale delle seguenti affermazioni è corretta:

☒ rifiutiamo l'ipotesi nulla H_0 sapendo che possiamo commettere un errore di tipo I (falso positivo) nell'1% dei casi

☐ il test suggerisce di accettare l'ipotesi nulla H_0

☐ sappiamo di avere l'1% di probabilità di commettere un errore di tipo II (falso negativo)

$t > 2.4 \rightarrow$ Rifiuto H_0

Quesito 4. La funzione di ripartizione F (cumulative distribution function - CDF) di una variabile aleatoria X è definita, per ogni numero reale x, tramite:

☒ $F(x) := P(X \leq x)$

☐ $F(x) := \int_B f(x)dx$

☐ $F(x) := P(X = x)$

Quesito 5. 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

☐ 0.9

Già fatto

Quesito 6. Sia $X_1, X_2, ...X_n$ un campione estratto da una popolazione normale di media incognita μ e varianza nota σ^2. Vogliamo produrre un intervallo di confidenza bilaterale per la media μ. Se è necessario ridurre l'ampiezza dell'intervallo come possiamo procedere?

☒ si diminuisce la varianza del campione

☐ si diminuisce l'ampiezza del campione

☐ si aumenta la varianza del campione

Quesito 7. Dato un modello di regressione lineare semplice $Y = \alpha + \beta x$. Quale delle seguenti risposte indica che il modello di regressione usato interpreta bene i dati?

☒ coefficiente di determinazione $R^2 = 0.98$

☐ coefficiente di correlazione $r = 0.55$

☐ coefficiente di determinazione $R^2 = 0.6$

Quesito 8. Sia $X_1, X_2, ..., X_n$ un campione proveniente da una popolazione di media μ e varianza σ^2. Se la distribuzione della popolazione è normale, allora la media campionaria $\bar{X}$

☒ sarà a sua volta normale indipendentemente dall'ampiezza del campione

☐ non sarà mai normale per qualsiasi n

☐ sarà a sua volta normale solo per $n > 30$

Quesito 9. Una commissione di 5 elementi deve essere selezionata da un gruppo di 6 uomini e 9 donne. Se la scelta viene fatta a caso, che probabilità vi è che vengano presi 3 uomini e 2 donne?

☐ $\frac{240}{1001}$

☐ 2

☐ $\frac{1001}{240}$

$$\binom{15}{5} = \frac{15!}{5!\,10!} = 3003$$

$$\frac{\binom{6}{3}\cdot\binom{9}{2}}{3003} = \frac{\frac{6!}{3!\,3!}\cdot\frac{9!}{2!\,7!}}{3003} = \frac{20\cdot 36}{3003} = \frac{720}{3003} = \frac{240}{1001}$$

Quesito 10. Se X e Y sono due variabili aleatorie indipendenti, allora

☒ $Cov(X, Y) = 0$

☐ $E[XY] \neq E[X]E[Y]$

☐ $Var(X) = Var(X + Y)$

Esercizio 1 [5 punti]

Tra i partecipanti ad un concorso per giovani compositori il 50% suona il pianoforte, il 30% suona il violino e il 20% la chitarra. Partecipano ad un concorso per la prima volta il 10% dei pianisti, il 33% dei violinisti e il 10% dei chitarristi. Applicando i concetti di probabilità condizionata e il teorema di Bayes, rispondere alle seguenti domande. Indichiamo con A l'evento "Aspiranti compositori alla prima esperienza", con B l'evento "Pianisti", con C l'evento "Violinisti" e con D l'evento "Chitarristi".
1. Qual è la percentuale di aspiranti compositori alla prima esperienza? (usare due cifre decimali dopo la virgola)
2. Sapendo che ad esibirsi per primo sarà un compositore alla prima esperienza, qual è la probabilità che sia un chitarrista? (usare due cifre decimali dopo la virgola)

$P(\text{Piano}) = 0.5$ $\quad P(N \mid \text{Piano}) = 0.1$

$P(\text{Violino}) = 0.3$ $\quad P(N \mid \text{Violino}) = 0.33$

$P(\text{Chitarra}) = 0.2$ $\quad P(N \mid \text{Chitarra}) = 0.1$

1) $P(N) = 0.5 \cdot 0.1 + 0.3 \cdot 0.33 + 0.2 \cdot 0.1 = 0.17$

2) $P(\text{Chitarra} \mid N) = \dfrac{P(N \mid \text{Chitarra}) P(\text{Chitarra})}{P(N)} = \dfrac{0.1 \cdot 0.2}{0.17} = 0.12$

Esercizio 2 [4 punti]

Uno studio sulla relazione tra età e varie funzioni visive (come acuità e percezione della profondità) ha riportato le seguenti osservazioni sull'area della lamina sclerale (mm^2) provenienti dal nervo ottico umano ("Morphometry of Nerve Fiber Bundle Pores in the Optic Nerve Head of the Human," Experimental Eye Research, 1988: 559–568): 2.75 2.62 2.74 3.85 2.34 2.74 3.93 4.21 3.88 4.33 3.46 4.52 2.43 3.65 2.78 3.56 3.01
Calcolare le media (a), la mediana (b), la deviazione standard (c) e la varianza (d) campionarie dei dati (~~usare una cifra decimale dopo la virgola~~).

$\mu = \dfrac{\sum x_i}{n} = 3{,}341$

$M = 3.46$

$\sigma^2 = 0{,}501$

$\sigma = \sqrt{0{,}501} = 0{,}708$