Esercizi set 10

1.1

Il modello di regressione lineare è un modello statistico che permette di determinare un'equazione per descrivere la relazione tra

- ☐ una variabile di ingresso e una o più variabili di risposta;
- ☐ diverse variabili di ingresso e diverse variabili di risposta;
- ☒ una variabile di risposta e una o più variabili di ingresso;
- ☐ tutte le affermazioni precedenti sono corrette.

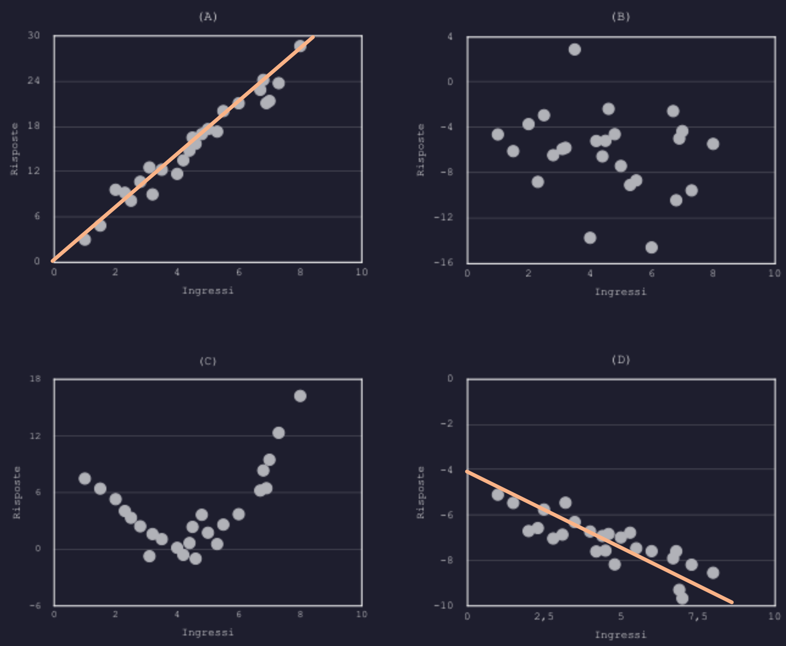
1.2

In un contesto di regressione lineare, quale delle seguenti affermazioni *non* corrisponde ad una delle ipotesi che si fanno sull'errore casuale ξ?

- ☒ La media dell'errore casuale vale uno.
- ☐ La varianza dell'errore casuale è la stessa per tutti i valori della variabile di ingresso.
- ☐ Gli errori casuali relativi a diversi valori della variabile di ingresso sono tra loro indipendenti.
- ☐ L'errore casuale ha distribuzione normale.

1.3

In un contesto di inferenza sui coefficienti di regressione, quali dei seguenti diagrammi di dispersione sono plausibilmente compatibili con il non rifiuto dell'ipotesi nulla $H_0 : \beta_1 = 0$?



- ☐ A e D
- ☒ B e C
- ☐ A, B e D
- ☐ Solo B

1.4

Il coefficiente di determinazione aiuta a capire

- ☐ il valore della risposta ottenuto per un certo valore della variabile di ingresso;
- ☐ il valore della variabile di ingresso necessario per ottenere un certo valore della risposta;
- ☒ la plausibilità di una relazione lineare tra la variabile di ingresso e la risposta;
- ☐ nessuna delle affermazioni precedenti è vera.

1.5

Se tra due variabili c'è una forte dipendenza lineare, il valore del coefficiente di determinazione deve essere

- ☒ vicino a 1;
- ☐ uguale a 0.5;
- ☐ molto vicino a 0;
- ☐ non ci sono abbastanza informazioni per poter rispondere.

1.6

Se il valore del coefficiente di determinazione è vicino a 1, allora le osservazioni sono raggruppate più vicino a

- ☐ la media campionaria degli ingressi;
- ☐ la media campionaria delle risposte;
- ☒ la retta di regressione lineare;
- ☐ l'origine.

1.7

La relazione tra il numero di birre consumate (x) e il livello di alcol presente nel sangue (Y) è stata studiata utilizzando un modello di regressione lineare basato su un campione di 16 studenti universitari. È stata ottenuta la retta di regressione $y = -0.0127 + 0.0180x$.
Per poter guidare, il limite legale per il livello di alcol nel sangue è 0.08. Se Paperino ha bevuto 5 birre, il modello prevede che il suo livello di alcol nel sangue è

- ☐ 0.09 unità sopra il limite legale;
- ☒ 0.0027 unità sotto il limite legale;
- ☐ 0.0027 unità sopra il limite legale;
- ☐ 0.0073 unità sopra il limite legale.

$y = -0.0127 + 0.0180 \cdot 5 = 0.0773$

$0,08 - 0,0773 = 0,0027$

1.8

In un modello di regressione lineare, se la varianza residua e la varianza spiegata valgono rispettivamente 200 e 300, il coefficiente di determinazione risulta

- ☐ 0.6667;
- ☒ 0.6000;
- ☐ 0.4000;
- ☐ 1.5000.

$SS_R = 200 \quad SS_E = 300 \qquad S_{YY} = 200+300 = 500$

$$\frac{SS_E}{S_{YY}} = \frac{300}{500} = 0{,}6$$

1.9

In un contesto di regressione lineare, è stato osservato un campione di dati $(x_1, y_1), \dots, (x_{10}, y_{10})$, per cui risultano $\overline{x} = 9$, $\overline{y} = 17$, $\sum_{i=1}^{10} x_i y_i = 3373$, $\sum_{i=1}^{10} x_i^2 = 1775$, $\widehat{s}_{YY} = 1434$ e $\widehat{ss}_R = 505.98$.

1. La stima dei minimi quadrati per β_1 (coefficiente angolare della retta di regressione) è

- ☒ 1.91;
- ☐ 2.98;
- ☐ −2.98;
- ☐ −1.91.

$$\beta_1 = \frac{\sum_1^n x_i y_i - \bar{x} \sum_1^n y_i}{\sum_1^n x_i^2 - n\bar{x}^2} = \frac{3373 - 9 \cdot 17 \cdot 10}{1775 - 10 \cdot 81} = 1{,}91$$

2. La stima dei minimi quadrati per β_0 (intercetta della retta di regressione) è

- ☐ 0.24;
- ☐ 0.19;
- ☒ −0.19;
- ☐ −0.24.

$$\beta_0 = \bar{Y} - \beta_1 \bar{x} = 17 - 1{,}91 \cdot 9 = -0{,}19$$

3. La varianza spiegata vale

- ☐ 1434; $= S_{YY}$
- ☐ 505.98;
- ☐ 50.598;
- ☒ 928.02.

$SS_E = S_{YY} - SS_R$

$= 1434 - 505{,}98$

$= 928{,}02$

4. Il coefficiente di determinazione vale

- ☒ 0.6471;
- ☐ 0.3547;
- ☐ 0;
- ☐ 1.

$$1 - \frac{\widehat{SS}_R}{\widehat{S}_{YY}} = 1 - \frac{505{,}98}{1434} = 1 - 0{,}3528 = 0{,}6471$$

1.10

Quando un modello di regressione lineare ha un buon adattamento con il campione di dati osservati, il suo diagramma dei residui

- ☐ mostra una forte regolarità: deve avere una tendenza lineare;
- ☐ mostra una forte regolarità: i residui devono crescere in valore assoluto al crescere della variabile di ingresso;
- ☒ non deve mostrare alcuna regolarità;
- ☐ deve mostrare valori concentrati attorno all'origine.