

Regressione lineare: esercizi

9.3 Gli antropologi cercano spesso di ricavare informazioni sul passato utilizzando resti umani rinvenuti in antiche sepolture. Ad esempio, attraverso il rinvenimento di un femore, prevedono quanto alto fosse l'individuo che ne era il possessore. L'equazione che impiegano per lo scopo è $\hat{y} = 61.4 + 2.4x$, dove \hat{y} è l'altezza prevista e x è la lunghezza del femore (entrambe le grandezze sono espresse in centimetri).⁶

- Identifica l'intercetta e il coefficiente angolare dell'equazione (pendenza). Interpreta il coefficiente angolare.
- Un femore ritrovato in un certo sito ha una lunghezza di 50 cm. Quale è l'altezza prevista per la persona che aveva quel femore?

9.4 La OECD (*Organization for Economic Cooperation and Development*) è un'organizzazione formata da 20 nazioni ad avanzato livello di industrializzazione. Per queste nazioni,⁷ l'equazione di previsione che mette in relazione $y =$ tasso di povertà infantile nel 2000 con $x =$ spesa sociale in percentuale del PIL (GDP) è $\hat{y} = 22 - 1.3x$. I valori y vanno dal 2.8% della Finlandia al 21.9% degli USA. I valori x vanno dal 2% degli USA al 16% della Danimarca.

- Interpreta l'intercetta e il coefficiente angolare.
- Trova il tasso di povertà previsto per gli USA e per la Danimarca.
- La correlazione è -0.79 . Interpreta.

9.8 L'addetto all'ufficio immatricolazioni di un college USA impiega una regressione per approssimare la relazione fra $y =$ GPA del college e $x =$ GPA della scuola superiore (entrambi misurati su una scala a quattro punti).

- Quale equazione è più realistica: $\hat{y} = 0.5 + 7.0x$, o $\hat{y} = 0.5 + 0.7x$? Perché?
- Supponi che l'equazione di previsione sia $\hat{y} = x$. Identifica l'intercetta y e interpreta l'inclinazione.

9.16 Per l'indagine sugli studenti del Problema 1.11, la correlazione campionaria fra $y =$ ideologia politica (con punteggi da 1 a 7, dove i valori più alti rappresentano maggiore conservatorismo) e $x =$ numero di volte che viene letto un quotidiano durante la settimana è pari a $r = -0.066$.

- a. Ritieni che la correlazione sia forte o debole?
- b. Interpreta il quadrato della correlazione.
- c. Quando y viene previsto utilizzando $x =$ religiosità (quanto spesso partecipi a funzioni religiose, 0, 1, 2, 3) la correlazione campionaria è risultata essere pari a $r = 0.580$. Quale di queste due variabili esplicative sembra avere una maggiore relazione lineare con y ? Spiega.

9.18 Per gli studenti che hanno seguito l'insegnamento di Statistica "101" al Lake Wobegon College nel Minnesota sia $x =$ voto all'esame intermedio che $y =$ voto all'esame finale avevano media = 75 e deviazione standard = 10.

- a. L'equazione di previsione è $\hat{y} = 30 + 0.60x$. Trova il voto finale previsto per gli studenti che hanno ottenuto un voto all'esame intermedio (i) pari a 100, (ii) pari a 50. Nota che il voto finale previsto regredisce verso la media rispetto al voto dell'esame intermedio.
- b. Mostra che la correlazione è pari a 0.60.
- c. Se $\hat{y} = x$, mostra che $r = 1.0$.
- d. Se $\hat{y} = 75$ (cioè, il coefficiente angolare = 0), mostra che $r = 0.0$.