Soluciones del examen - Estadística descriptiva

Curso: Universidad 1 · Duración: 90 minutos · /20

Solución del ejercicio 1 (4 puntos)

Criterios. Se valora el cálculo de las medidas de posición y dispersión y la aplicación del criterio de Tukey.

a. $\bar{x} = 5$, $Me = 5$, moda 5. La suma es 50, así que $\bar{x} = \frac{50}{10} = 5$. La mediana es la media de los datos 5.º y 6.º, $\frac{5+5}{2}$. El 5 aparece tres veces.
b. $\sigma^2 = \frac{298}{10} - 25 = 4{,}80$, $\sigma \approx 2{,}19$. Se usa $\sum x_i^2 = 4 + 9 + 9 + 16 + 25 + 25 + 25 + 36 + 49 + 100 = 298$.
c. $IQR = 3$, límite superior $6 + 4{,}5 = 10{,}5$: el 10 no es atípico. Las vallas de Tukey son $Q_1 - 1{,}5 \cdot IQR = -1{,}5$ y $Q_3 + 1{,}5 \cdot IQR = 10{,}5$. Como $10 \leq 10{,}5$, el dato no es atípico, aunque está cerca del límite.

Solución del ejercicio 2 (4 puntos)

Criterios. Se valora usar las marcas de clase y la interpolación correcta para la mediana.

a. $\bar{x} = \frac{1100}{50} = 22{,}00$ años; $\sigma \approx 9{,}00$ años. Marcas de clase: 5, 15, 25 y 35. $\sum m_i f_i = 25 + 225 + 500 + 350 = 1100$ y $\sum m_i^2 f_i = 125 + 3375 + 12500 + 12250 = 28250$, luego $\sigma^2 = \frac{28250}{50} - 22^2 = 565 - 484 = 81$ y $\sigma = 9$.
b. 25 datos: intervalo $[20, 30)$; $Me = 20 + \frac{25-20}{20} \cdot 10 = 22{,}5$ años. Las frecuencias acumuladas son 5, 20, 40 y 50; la clase mediana es la primera que supera 25. Dentro de ella faltan 5 de sus 20 datos, es decir, un cuarto de la amplitud de 10 años.

Solución del ejercicio 3 (5 puntos)

Criterios. Se valora la covarianza, la correlación, la recta y la predicción con unidades.

a. $\bar{x} = 3$, $\bar{y} = 41{,}00$, $\sigma_{xy} = -9{,}00$, $r \approx -0{,}996$: relación lineal negativa muy fuerte. Se usa $\sum xy = 50 + 90 + 126 + 144 + 160 = 570$, así que $\sigma_{xy} = \frac{570}{5} - 3 \cdot 41 = -9$. Además $\sigma_x = \sqrt{2}$ y $\sigma_y = \sqrt{40{,}8} \approx 6{,}39$.
b. Pendiente $-4{,}50$: $y = 54{,}50 - 4{,}50x$. La pendiente es $\frac{\sigma_{xy}}{\sigma_x^2} = \frac{-9}{2}$ y la recta pasa por $(3, 41)$: $y - 41 = -4{,}5(x - 3)$. Cada año el coche pierde unos 450 €.
c. $y(6) \approx 27{,}50$, unos 2750 €. Es una ligera extrapolación (los datos llegan a 5 años), aceptable porque $|r|$ es casi 1.

Solución del ejercicio 4 (4 puntos)

Criterios. Se valora el cálculo de los cuartiles y de las vallas de Tukey.

a. $Q_1 = 6$, $Me = 7{,}5$, $Q_3 = 10$. Con 10 datos, la mediana es la media del 5.º y el 6.º, $\frac{7+8}{2}$. La mitad inferior es 3, 5, 6, 6, 7, con mediana 6; la superior, 8, 9, 10, 11, 25, con mediana 10.
b. $IQR = 4$; límites 0 y 16: el 25 es atípico. Las vallas son $6 - 1{,}5 \cdot 4 = 0$ y $10 + 1{,}5 \cdot 4 = 16$. Solo el 25 queda fuera; en el diagrama de caja se marca aparte y el bigote superior llega hasta 11.

Solución del ejercicio 5 (3 puntos)

Criterios. Se valora el coeficiente de variación y la conclusión razonada.

a. $CV_A = \frac{300}{1800} \approx 0{,}167$ y $CV_B = \frac{375}{2500} = 0{,}15$. El CV es el cociente entre la desviación típica y la media, y no tiene unidades. En porcentaje: 16,7 % y 15 %.
b. La B, porque su dispersión relativa es menor. Aunque B tiene mayor desviación típica en euros, en relación con su media sus salarios están menos dispersos. Por eso se compara el CV y no σ.