Soluciones del examen - Estadística bidimensional

Curso: 1 Bachillerato Ccss · Duración: 55 minutos · /20

Solución del ejercicio 1 (4 puntos)

a. Fila «andando»: $45 + 30 = 75$. Fila «no andando»: $30 + 45 = 75$. Columna «deporte»: $45 + 30 = 75$. Columna «no deporte»: $30 + 45 = 75$. Total 150.

b. $\dfrac{75}{150} = 0{,}5$. La mitad de los alumnos practica deporte con regularidad.

c. Entre los que van andando: $\dfrac{45}{75} = 0{,}6$. Entre los que no van andando: $\dfrac{30}{75} = 0{,}4$.

d. Si fueran independientes, el porcentaje de deportistas sería el mismo en los dos grupos (50 %). Como es 60 % y 40 %, hay cierta dependencia, aunque moderada. Formalmente, $P(D \mid A) = 0{,}6 \neq P(D) = 0{,}5$.

Solución del ejercicio 2 (4 puntos)

a. $n = 5$; $\sum x = 15$, $\sum y = 25$, $\sum x^2 = 55$, $\sum y^2 = 155$, $\sum xy = 58$. Medias: $\bar{x} = 3$, $\bar{y} = 5$. Varianzas: $\sigma_x^2 = 2$ y $\sigma_y^2 = 6$. Covarianza: $\sigma_{xy} = -3{,}400$. Desviaciones: $\sigma_x = 1{,}414$, $\sigma_y = 2{,}449$. Se obtienen con $\sigma_x^2 = \dfrac{55}{5} - 3^2 = 2$ y $\sigma_y^2 = \dfrac{155}{5} - 5^2 = 6$.

b. . Aquí  . El signo negativo indica que, al crecer x, y tiende a disminuir.

c. $r = \dfrac{-3{,}400}{1{,}414 \cdot 2{,}449} = -0{,}9815$. Correlación negativa muy fuerte: si x aumenta, y disminuye casi linealmente. Como $|r|$ está muy cerca de 1, los puntos están casi alineados sobre una recta de pendiente negativa.

Solución del ejercicio 3 (4 puntos)

a. $n = 5$; $\sum x = 30$, $\sum y = 140$, $\sum x^2 = 220$, $\sum y^2 = 4398$, $\sum xy = 978$. Medias: $\bar{x} = 6$, $\bar{y} = 28$. Varianzas: $\sigma_x^2 = 8$ y $\sigma_y^2 = 95{,}600$. Covarianza: $\sigma_{xy} = 27{,}600$. Desviaciones: $\sigma_x = 2{,}828$, $\sigma_y = 9{,}778$. Entonces $b = \dfrac{27{,}60}{8} = 3{,}450$ y $a = 28 - 3{,}450 \cdot 6 = 7{,}300$. Recta: $y = 3{,}450\,x + 7{,}300$.

b. $y(7) = 3{,}450 \cdot 7 + 7{,}300 = 31{,}45$, unos 31 clientes.

c. $r = 0{,}9980$ (casi 1) y $x = 7$ está dentro del intervalo de datos (2 a 10): es una interpolación y la estimación es fiable. El coeficiente de determinación, $r^2 \approx 0{,}996$, indica que la recta explica casi toda la variación.

Solución del ejercicio 4 (4 puntos)

a. Correlación positiva fuerte: los meses con más helados tienden a tener más insolaciones.
b. $r^2 = 0{,}81$: el 81 % de la variación de las insolaciones se asocia linealmente con las ventas de helados.
c. No. La correlación no implica causalidad: ambas variables dependen de una tercera, la temperatura (más calor, más helados y más insolaciones).
d. No: 5 grados está fuera del rango de datos (20 a 35); sería extrapolar y el comportamiento puede ser distinto.

Solución del ejercicio 5 (4 puntos)

a. $n = 6$; $\sum x = 90$, $\sum y = 910$, $\sum x^2 = 1420$, $\sum y^2 = 144900$, $\sum xy = 12960$. Medias: $\bar{x} = 15$, $\bar{y} = 151{,}667$. Varianzas: $\sigma_x^2 = 11{,}667$ y $\sigma_y^2 = 1147{,}222$. Covarianza: $\sigma_{xy} = -115$. Desviaciones: $\sigma_x = 3{,}416$, $\sigma_y = 33{,}871$.
b. $r = \dfrac{-115}{3{,}416 \cdot 33{,}871} = -0{,}9940$: correlación negativa muy fuerte (si sube el precio, bajan las ventas).
c. $b = \dfrac{-115}{11{,}67} = -9{,}857$ y $a = 151{,}67 - (-9{,}857) \cdot 15 = 299{,}52$. Recta: $y = -9{,}857\,x + 299{,}52$.
d. $y(15) = -9{,}857 \cdot 15 + 299{,}52 = 151{,}7$ camisetas; ingresos $= 15 \cdot 151{,}7 = 2275$ €.
e. $r^2 = 0{,}9881$: el 98,8 % de la variación de las ventas se explica por el precio. El modelo lineal es muy adecuado en el rango de 10 a 20 euros.