# Soluciones - Estadística unidimensional

Curso: 1 Bachillerato Ccss

## Solución del ejercicio 1

a. La variable es «número de hijos»: es cuantitativa discreta.
b. Frecuencias relativas $h_i = f_i/40$: 0: 0,150, 1: 0,350, 2: 0,300, 3: 0,125, 4: 0,075. Porcentajes: 0: 15,0 %, 1: 35,0 %, 2: 30,0 %, 3: 12,5 %, 4: 7,5 %. Acumuladas $F_i$: 0: 6, 1: 20, 2: 32, 3: 37, 4: 40.
c. Como máximo 2 hijos: $F = 32$ familias, es decir, $\dfrac{32}{40} = 80{,}0\,\%$.
d. Al menos 3 hijos: $5 + 3 = 8$ familias, es decir, $\dfrac{8}{40} = 20\,\%$.

## Solución del ejercicio 2

a. $\bar{x} = \dfrac{\sum x_i f_i}{n} = \dfrac{0\cdot 6 + 1\cdot 14 + 2\cdot 12 + 3\cdot 5 + 4\cdot 3}{40} = \dfrac{65}{40} = 1{,}625$ hijos.
b. $n = 40$: la mediana ocupa las posiciones 20 y 21. Las frecuencias acumuladas son 6, 20, 32, 37, 40. El dato 20 es un 1 y el 21 es un 2, luego la mediana es $\dfrac{1+2}{2} = 1{,}5$.
c. La moda es el valor con mayor frecuencia: 1 hijo (14 familias).
d. La media (1,62) es algo mayor que la mediana (1,5) y que la moda (1) porque hay algunas familias con 3 y 4 hijos que tiran de la media hacia arriba. La distribución es ligeramente asimétrica hacia la derecha.

## Solución del ejercicio 3

a. $\sum x_i^2 f_i = 0\cdot 6 + 1\cdot 14 + 4\cdot 12 + 9\cdot 5 + 16\cdot 3 = 14 + 48 + 45 + 48 = 155$.
b. $\sigma^2 = \dfrac{\sum x_i^2 f_i}{n} - \bar{x}^2 = \dfrac{155}{40} - 1{,}62^2 = 3{,}875 - 2{,}6406 = 1{,}2344$.
c. $\sigma = \sqrt{1{,}2344} = 1{,}111$ hijos.
d. $CV = \dfrac{\sigma}{\bar{x}} = \dfrac{1{,}111}{1{,}62} = 0{,}684$, es decir, el 68,4 %. Una dispersión relativa alta (por encima del 50 %): los datos están bastante dispersos respecto de la media.

## Solución del ejercicio 4

a. Marcas de clase: 5, 15, 25, 35, 45 minutos (punto medio de cada intervalo).

b. $\sum m_i f_i = 990$ y $\sum m_i^2 f_i = 29400$. $\bar{x} = \dfrac{990}{40} = 24{,}75$ min.
$\sigma^2 = \dfrac{29400}{40} - 24{,}75^2 = 122{,}44$, así que $\sigma = 11{,}07$ min.

c. Clase modal: la de mayor frecuencia, $[20, 30)$ con 15 clientes. Clase mediana: $n/2 = 20$; las frecuencias acumuladas son 4, 13, 28, 36, 40; la primera que supera 20 es $[20, 30)$.

d. $Me = L + \dfrac{n/2 - F_{\text{ant}}}{f} \cdot a = 20 + \dfrac{20 - 13}{15} \cdot 10 = 20 + 4{,}67 = 24{,}67$ min.

## Solución del ejercicio 5

a. Con 11 datos ordenados, la mediana es el sexto dato: $Me = 9$. La mitad inferior (5 datos: 3, 5, 6, 7, 8) tiene mediana 6 y la superior (10, 12, 13, 15, 20) tiene mediana 13. Entonces $Q_1 = 6$ y $Q_3 = 13$.

b. Recorrido: $20 - 3 = 17$. Recorrido intercuartílico: $RI = Q_3 - Q_1 = 13 - 6 = 7$.

c. Límites: $6 - 1{,}5 \cdot 7 = -4{,}5$ y $13 + 1{,}5 \cdot 7 = 23{,}5$. Todos los datos están dentro: no hay atípicos.

d. La caja va de 6 a 13 con una línea en 9; los bigotes llegan a 3 y a 20. La caja es algo más larga a la derecha de la mediana (9 a 13) que a la izquierda (6 a 9): ligera asimetría positiva.

## Solución del ejercicio 6

a. $CV_A = \dfrac{8}{50} = 0{,}16$, es decir, el 16 %.

b. $CV_B = \dfrac{1{,}8}{9} = 0{,}20$, es decir, el 20 %.

c. La edad en A, porque su coeficiente de variación (16 %) es menor: está relativamente menos dispersa que la antigüedad de B (20 %).

d. Porque las variables tienen distintas unidades y distinto tamaño medio (edad y antigüedad). El CV es adimensional y expresa la dispersión relativa a la media.

## Solución del ejercicio 7

a. $n = 30$. $\sum x_i f_i = 202$, luego $\bar{x} = \dfrac{202}{30} = 6{,}733$.

b. Hay $n = 30$ datos: la mediana está entre los puestos 15 y 16. Las frecuencias acumuladas son 1, 3, 7, 14, 20, 25, 28, 30. Los puestos 15 y 16 caen en la nota 7: $Me = 7$. La moda es 6 (la de mayor frecuencia, 7 alumnos).

c. $\sum x_i^2 f_i = 1450$, $\sigma^2 = \dfrac{1450}{30} - 6{,}733^2 = 2{,}996$ y $\sigma = 1{,}731$.

d. Suspensos: $1 + 2 = 3$. Aprobados: $30 - 3 = 27$, es decir, el $90\,\%$. Notable o más:

$6+5+3+2=16$ alumnos, el $53{,}3\,\%$.

## Solución del ejercicio 8

a. Marcas de clase: 1, 3, 5, 7, 9. $\sum m_i f_i = 416$. $\bar{x} = \dfrac{416}{100} = 4{,}16$ min.

b. $\sum m_i^2 f_i = 2260$. $\sigma^2 = \dfrac{2260}{100} - 4{,}16^2 = 5{,}294$ y $\sigma = 2{,}301$ min.

c. $n/2 = 50$. Acumuladas: 18, 52, 80, 92, 100. La clase mediana es $[2, 4)$.
$Me = 2 + \dfrac{50-18}{34} \cdot 2 = 2 + 1{,}88 = 3{,}88$ min.

d. Espera menos de 4 min: $18 + 34 = 52$ viajeros, es decir, el 52 %.

## Solución del ejercicio 9

a. $\bar{x} = \dfrac{1200+1250+1300+1300+1350+1400+5400}{7} = \dfrac{13200}{7} = 1885{,}71$ €. La mediana es el cuarto dato: $Me = 1300$ €.

b. La moda es 1 300 € (aparece dos veces).

c. Sin el gerente: $\dfrac{7800}{6} = 1300$ €.

d. La mediana (o la moda), porque no se deja arrastrar por el salario extremo. La media de 1 885,71 € no representa a ningún trabajador: la mayoría gana menos de 1 400 €.

## Solución del ejercicio 10

a. Con 10 datos, la mediana es la media de los puestos 5 y 6:
$Me = \dfrac{640+700}{2} = 670$ €.

b. Mitad inferior: 450, 520, 560, 600, 640, con mediana 560; mitad superior: 700, 750, 820, 900, 1 100, con mediana 820. Entonces $Q_1 = 560$ € y $Q_3 = 820$ €.

c. $RI = 820 - 560 = 260$ €. La mitad central de los alquileres está entre 560 y 820 €.

d. 800 € está entre la mediana (670) y $Q_3$ (820): pertenece al tercer cuarto, entre el 50 % y el 75 % de los alquileres más baratos.

## Solución del ejercicio 11

a. $y = 1{,}03\,x + 50$.

b. $\bar{y} = 1{,}03 \cdot 1400 + 50 = 1442 + 50 = 1492$ €.

c. La desviación típica se multiplica por $|a|$ y no cambia con la suma:

$\sigma_y = 1{,}03 \cdot 200 = 206$ €.

d. Antes: $CV = \dfrac{200}{1400} = 14{,}29\,\%$. Ahora: $CV = \dfrac{206}{1492} = 13{,}81\,\%$. Disminuye un poco, porque al sumar 50 € a todos la dispersión relativa baja.

## Solución del ejercicio 12

a. (1) Cualitativa nominal. (2) Cuantitativa discreta. (3) Cuantitativa continua. (4) Cualitativa ordinal (hay un orden natural).
b. (1) Diagrama de sectores o de barras. (2) Diagrama de barras. (3) Histograma (o de caja). (4) Diagrama de barras (respetando el orden).
c. La media requiere valores numéricos con los que operar; el destino es una categoría y la suma de destinos no tiene sentido. Para el destino solo se puede dar la moda.
d. La frecuencia (absoluta o relativa) o la densidad de frecuencia; con amplitudes iguales, la altura de cada barra es proporcional a la frecuencia del intervalo.

## Solución del ejercicio 13

a. Marcas: 2,25; 2,75; 3,25; 3,75; 4,25. $\sum m_i f_i = 322$ y $\bar{x} = 3{,}220$ kg. $\sum m_i^2 f_i = 1059{,}250$; $\sigma^2 = 10{,}5925 - 3{,}220^2 = 0{,}2241$ y $\sigma = 0{,}473$ kg.
b. $\bar{x} \pm \sigma = 3{,}220 \pm 0{,}473$, es decir, $(2{,}75;\ 3{,}69)$. Suponiendo que los bebés se reparten de forma uniforme en cada clase: la mitad de la clase $[2{,}5; 3{,}0)$ (12 bebés), la clase $[3{,}0; 3{,}5)$ completa (46) y la parte de $[3{,}5; 4{,}0)$ hasta 3,69 ( $\dfrac{0{,}19}{0{,}5} \cdot 18 \approx 6{,}8$). Total aproximado: $64{,}8$ bebés, el 65 %.
c. $\bar{x} \pm 2\sigma = (2{,}27;\ 4{,}17)$: contiene a casi todos: $\dfrac{0{,}23}{0{,}5} \cdot 6 + 24 + 46 + 18 + \dfrac{0{,}17}{0{,}5} \cdot 6 \approx 2{,}8 + 24 + 46 + 18 + 2{,}0 = 92{,}8$ bebés, es decir, alrededor del 93 %.
d. Menos de 2,5 kg: los de la primera clase, $6\,\%$ (6 de 100).

## Solución del ejercicio 14

a. Grupo A ($n = 20$): $\bar{x} = \dfrac{128}{20} = 6{,}40$; $\sigma^2 = \dfrac{834}{20} - 6{,}40^2 = 0{,}7400$; $\sigma = 0{,}860$. Grupo B ($n = 20$): $\bar{x} = \dfrac{129}{20} = 6{,}45$; $\sigma^2 = \dfrac{877}{20} - 6{,}45^2 = 2{,}2475$; $\sigma = 1{,}499$.
b. $CV_A = \dfrac{0{,}860}{6{,}40} = 13{,}4\,\%$ y $CV_B = \dfrac{1{,}499}{6{,}45} = 23{,}2\,\%$.
c. Las medias son 6,40 (A) y 6,45 (B): son casi iguales. El grupo A es más homogéneo, porque su CV es menor (13,4 % frente a 23,2 %).
d. Grupo A: puestos 10 y 11 caen en la nota 6, así que $Me = 6$. Grupo B: puestos 10

y 11 caen en 6 y 7, así que $Me = 6{,}5$.

## Solución del ejercicio 15

a. Marcas: 50, 150, 250, 350, 450. Acumuladas: 6, 20, 37, 46, 50. Relativas: 0,12; 0,28; 0,34; 0,18; 0,08.

b. $\sum m_i f_i = 11600$ y $\bar{x} = \dfrac{11600}{50} = 232$ kWh. $\sum m_i^2 f_i = 3305000$, $\sigma^2 = \dfrac{3305000}{50} - 232^2 = 12276$, $\sigma = 110{,}80$ kWh.

c. $n/2 = 25$: clase $[200, 300)$, $Me = 200 + \dfrac{25 - 20}{17} \cdot 100 = 229{,}4$. $n/4 = 12{,}5$: clase $[100, 200)$, $Q_1 = 100 + \dfrac{12{,}5 - 6}{14} \cdot 100 = 146{,}4$. $3n/4 = 37{,}5$: clase $[300, 400)$, $Q_3 = 300 + \dfrac{37{,}5 - 37}{9} \cdot 100 = 305{,}6$ kWh.

d. $CV = \dfrac{110{,}80}{232} = 47{,}8\,\%$: dispersión elevada, los consumos varían mucho de unos hogares a otros.

e. Tarifa reducida: hasta $Q_1 \approx 146$ kWh. Tarifa especial para el 25 % que más consume: desde $Q_3 \approx 306$ kWh.

## Solución del ejercicio 16

a. Burgos: $\dfrac{202}{12} = 16{,}83$ °C. Málaga: $\dfrac{228}{12} = 19$ °C.

b. Burgos: $\sigma = 6{,}78$ °C. Málaga: $\sigma = 4{,}30$ °C. (Se calcula : Burgos $329{,}33 - 16{,}83^2$; Málaga $379{,}50 - 19^2$.)

c. Burgos: $27 - 8 = 19$ °C. Málaga: $26 - 14 = 12$ °C.

d. Burgos: $CV = \dfrac{6{,}78}{16{,}83} = 40{,}3\,\%$. Málaga: $CV = \dfrac{4{,}30}{19} = 22{,}6\,\%$. Málaga es más uniforme (menor dispersión absoluta y relativa).

## Solución del ejercicio 17

a. Datos ordenados: 22, 23, 24, 24, 25, 26, 26, 27, 28, 95. $\bar{x} = \dfrac{320}{10} = 32$ clientes. Mediana: media de los puestos 5 y 6 $= \dfrac{25 + 26}{2} = 25{,}5$.

b. $\sum x^2 = 14680$; $\sigma^2 = 1468 - 32^2 = 444$; $\sigma = 21{,}07$.

c. Sin el 95: $\bar{x} = \dfrac{225}{9} = 25$. Mediana (9 datos): el quinto $= 25$. $\sigma = 1{,}83$.

d. La media pasa de 32 a 25 (cambia 7 clientes), la mediana de 25,5 a 25 (casi

nada) y la desviación típica de 21,1 a 1,8 (se reduce muchísimo). La mediana es la medida más resistente a un valor atípico; la media y la desviación típica son muy sensibles.

## Solución del ejercicio 18

a. Marcas: 15, 25, 35, 45, 55, 65. $\sum m_i f_i = 1560$, $\bar{x} = \dfrac{1560}{40} = 39$ años.
$\sum m_i^2 f_i = 67000$, $\sigma^2 = \dfrac{67000}{40} - 39^2 = 154$ y $\sigma = 12{,}41$ años.

b. $n/2 = 20$. Acumuladas: 2, 10, 22, 32, 38, 40. Clase mediana $[30, 40)$:
$Me = 30 + \dfrac{20 - 10}{12} \cdot 10 = 38{,}33$ años. Clase modal: $[30, 40)$ (frecuencia 12).

c. $Q_1$: $n/4 = 10$, que coincide con la acumulada de $[10, 20)$ y $[20, 30)$:
$Q_1 = 20 + \dfrac{10 - 2}{8} \cdot 10 = 30$ años. $Q_3$: $3n/4 = 30$ en $[40, 50)$:
$Q_3 = 40 + \dfrac{30 - 22}{10} \cdot 10 = 48$ años. La mitad central está entre 30 y 48 años.

d. Menos de 30 años: $2 + 8 = 10$ socios, es decir, el 25 %. Con 50 años o más: $6 + 2 = 8$ socios.

e. El 10 % de 40 son 4 socios. El percentil 90 está en $0{,}9 \cdot 40 = 36$: clase $[50, 60)$:
$P_{90} = 50 + \dfrac{36 - 32}{6} \cdot 10 = 56{,}7$ años.

## Solución del ejercicio 19

a. Marcas: 65, 75, 85, 95, 105, 115. $\sum m_i f_i = 26875$; $\bar{x} = \dfrac{26875}{295} = 91{,}10$ km/h.
$\sum m_i^2 f_i = 2485375$; $\sigma^2 = \dfrac{2485375}{295} - 91{,}10^2 = 125{,}48$; $\sigma = 11{,}20$ km/h.

b. $n/2 = 147{,}5$. Acumuladas: 10, 45, 130, 240, 280, 295. La primera clase cuya acumulada supera 147,5 es $[90, 100)$, luego es la clase mediana:
$Me = 90 + \dfrac{147{,}5 - 130}{110} \cdot 10 = 91{,}59$ km/h.

c. Superan 90 km/h los de las clases $[90, 100)$, $[100, 110)$ y $[110, 120)$:
$110 + 40 + 15 = 165$ vehículos, el $\dfrac{165}{295} = 55{,}9\,\%$ (suponiendo que el 90 está en el límite de clase).

d. $0{,}9 \cdot 295 = 265{,}5$: clase $[100, 110)$ (acumuladas 240 y 280).
$P_{90} = 100 + \dfrac{265{,}5 - 240}{40} \cdot 10 = 106{,}4$ km/h.

e. $CV = \dfrac{11{,}20}{91{,}10} = 12{,}3\,\%$: es bajo (menos del 15 %), las velocidades son bastante homogéneas, aunque casi 6 de cada 10 conductores supera el límite.

## Solución del ejercicio 20

a. $\bar{x} = \dfrac{12 \cdot 6 + 18 \cdot 7}{30} = \dfrac{72 + 126}{30} = 6{,}6$.
b. Suma total $= 5 \cdot 8 = 40$, luego el quinto es $40 - (6 + 6 + 7 + 12) = 9$. Datos ordenados: 6, 6, 7, 9, 12; mediana 7 y moda 6, como pedía el enunciado.
c. Suma inicial: $6 \cdot 10 = 60$. Suma final: $7 \cdot 11 = 77$. Nuevo dato: $77 - 60 = 17$.
d. Nueva media: $-2 \cdot 4 + 5 = -3$. Nueva desviación típica: $|-2| \cdot 3 = 6$ (el signo no importa y la suma no cambia la dispersión).