EJERCICIOS ESTADISTICA DESCRIPTIVA DATOS NO AGRUPADOS

EJERCICIOS EXAMEN JUNIO 2020 26-28/06/2020 MODELO B

1. Se recoge información para una variable discreta con un rango de valores 10 a 25 y se recoge en la siguiente tabla de frecuencias:

Variable estadística	Frecuencia absoluta n_i	N_i Frecuencia absoluta acumulada
10	2	2
13	4	6
16	n_3 10	16
19	15	31
22	6	37

$\sum n_i = n = 37$

$2+4+n_3+15+6=37$

$n_3 = 10$

$\sum n_i = 37$ $\quad n = 37$

¿Cuál es el valor de la tabla que falta?

A. 6
B. 10
C. 16
D. 21

2. Las edades de los 554 empleados de una determinada empresa son las que aparecen en la siguiente tabla:

Edad	Frecuencia absoluta n_i
[15 - 25)	22
[25 - 35)	70
[35 - 45)	121
[45 - 55)	157
[55 - 65)	184

F_i

$n = 554$

0'167 menores de 35 : $22+70=92$

$$\frac{92}{554} = 0'16606 \simeq 0'167$$

¿Cuál es la proporción de trabajadores menores de 35 años?

A. 0.272
B. 0.167
C. 0.380
D. 0.126

3. Una empresa realiza una encuesta sobre la facilidad de navegación por su sitio web a 200 usuarios elegidos al azar obteniendo los siguientes resultados:

Navegación	Nº usuarios
Excelente	102
Sobresaliente	58
Buena	30
Mala	10

variable cualitativa ordinal → media no se puede calcular

¿Cuál es la media de esta distribución de datos?

A. 102

B. 160

C. Excelente

D. No se puede calcular

4. Una empresa recoge los datos del número de errores que mensualmente tienen los productos que fabrican. Se tienen los siguientes valores a lo largo de 8 meses:

12	24	16	18	14	10	15	20

A. Los datos presentan una mediana en el valor ~~18~~. ✗

B. Los datos incluyen un valor atípico que es la observación con valor "24". ✗

C. La distribución no tiene moda por lo que no podemos calcular su asimetría. ✗

D. Ninguna de las anteriores es correcta.

10, 12, 14, 15, 16, 18, 20, 24 $n=8$

$Q_1 = 13$ $Q_3 = 19$

$Me = \frac{15+16}{2} = 15'5$

$n=8$ par $Me = \frac{x_{n/2} + x_{n/2+1}}{2} = \frac{x_4+x_5}{2} = 15'5$

$LS = Q_3 + 1'5(Q_3 - Q_1)$

Pos Q_1 $\frac{n+1}{4} = \frac{8+1}{2} = 2'25$ $Q_1 = \frac{x_2+x_3}{2} = \frac{12+14}{2} = 13$

Pos Q_3 $3\left(\frac{n+1}{4}\right) = 6'75$ $Q_3 = \frac{x_6+x_7}{2} = \frac{18+20}{2} = 19$

$LS = 19 + 1'5(19-13) = 28$ $24 ? < 28$

↳ no atípico

5. Los sueldos mensuales (en miles de €) de 25 empresarios del sector de la construcción en el año 2019 fueron:

Salario (miles €)	Nº empresarios (n_i)
3.7	3
3.9	8
4.3	4
4.7	6
4.9	4

Varianza (S^2)	$S^2 = \frac{\sum (x_i - \bar{X})^2 \cdot n_i}{N}$ Fórmula alternativa: $S^2 = \frac{\sum x_i^2 \cdot n_i}{N} - \bar{X}^2 = \overline{X^2} - \bar{X}^2$ Datos agrupados utilizamos m_i y en lugar del valor x_i

$\sum n_i = 25$ $n = 25$

Un resumen por medio de estadísticos descriptivos para esta variable es:

A. El salario medio de los trabajadores es de 107.3 euros con una varianza de 0.1903 euros^2

B. El salario medio de los trabajadores es de 4.29 euros con una desviación típica de 0.036 euros^~~2~~

C. El salario medio de los trabajadores es de 107.3 euros con una varianza de 18.61 euros^2

D. El salario medio de los trabajadores es de 4.29 euros con una desviación típica de 0.4363 euros

$$\bar{x} = \frac{\sum x_i \cdot n_i}{n} = \frac{3'7\cdot 3 + 3'9\cdot 8 + 4'3\cdot 4 + 4'7\cdot 6 + 4'9\cdot 4}{25} = 4'29 \text{ miles €}$$

$$S^2 = \frac{\sum x_i^2 \cdot n_i}{n} - \bar{x}^2 = \frac{3'7^2\cdot 3 + 3'9^2\cdot 8 + 4'3^2\cdot 4 + 4'7^2\cdot 6 + 4'9^2\cdot 4}{25} - 4'29^2 =$$

$$= 18'61 - 4'29^2 = 0'2075$$

$$S = \sqrt{0'2075} = 0'455$$

6. Los empleados que forman parte de la población activa de un país son 3 millones de los cuales: 1 millón en el sector primario, 1.75 millones en el sector secundario y 1.25 millones en el sector servicios. ¿Cuál es la mejor forma de representar esta información?

A. Gráfico de sectores

V. cualitativa nominal

B. Polígono acumulativo de frecuencias

C. Histograma

D. Diagrama de dispersión

TUACADEMIAFACIL

EJERCICIO 1 EXAMEN SEPTIEMBRE 2024 (6-11/09/2024) MODELO B

1. En un hospital se ha creado un índice para controlar la cantidad que se provee de un nuevo fármaco a los pacientes con fallos renales. Los valores de dicho índice oscilan entre 0 (menor nivel de suministro del fármaco) y 7 (nivel más elevado de suministro del fármaco). El fármaco se está aplicando a todos los pacientes que lo requieren. A continuación se muestra la distribución de frecuencias:

Nivel de suministro del fármaco	0	1	2	3	4	5	x_i
Número de pacientes que reciben la dosis	3	1	4	5	6	7	n_i

a) Calcula la media, la mediana y la moda del nivel de suministro del fármaco. ¿Se tratan de estadísticos o de parámetros?, ¿por qué? (1 punto).

b) ¿Qué ocurriría con la media y con la varianza si se aumentara la dosis en uno a todos los pacientes? Razona tu respuesta (0,5 puntos).

c) Calcula el rango intercuartílico en base a los datos de la tabla. ¿Por qué presenta un valor menor que el rango? (0,5 puntos).

d) Calcula el coeficiente de asimetría de Fisher e interpreta el resultado obtenido (0,5 puntos).

e) Si el coeficiente de curtosis de esta distribución es de 0,78674, obtén conclusiones sobre el grado de apuntamiento de esta distribución (0,5 puntos).

x_i	n_i	N_i	$x_i \cdot n_i$
0	3	3	
1	1	4	
2	4	8	
3	5	13	
4	6	19	
5	7	26	

$N = \sum n_i = 26$

0, 0, 0, 1, 2, 2, 2, 2, 3, ... , 3, 4, 4, ...

(13)(14)

a) Calcula la media, la mediana y la moda del nivel de suministro del fármaco. ¿Se tratan de estadísticos o de parámetros?, ¿por qué? (1 punto).

- Media: $\bar{x} = \frac{\sum x_i \cdot n_i}{N} =$

$$= \frac{0\cdot 3 + 1\cdot 1 + 2\cdot 4 + 3\cdot 5 + 4\cdot 6 + 5\cdot 7}{26} = 3'192$$

- Mediana: $N = 26$ par

$$Me = \frac{x_{(N/2)} + x_{(N/2+1)}}{2} = \frac{x_{13} + x_{14}}{2} = \frac{3+4}{2} = 3'5$$

- Moda: $Mo = 5$

Son parámetros porque analizamos a todos los paciente que toman el fármaco.

b) ¿Qué ocurriría con la media y con la varianza si se aumentara la dosis en uno a todos los pacientes? Razona tu respuesta (0,5 puntos).

$X \sim$ dosis

$X' = X + 1 \longrightarrow$ cambio de origen

$\bar{x}' = \bar{x} + 1 = 3'192 + 1 = 4'192$

$S^2_{x'} = S^2_x$

Aumentar la dosis es un cambio de origen, la media se ve afectada en 1 unidad de aumento, pero la varianza no cambia ya que no le afectan los cambios de origen.

c) Calcula el rango intercuartílico en base a los datos de la tabla. ¿Por qué presenta un valor menor que el rango? (0,5 puntos).

Cuartiles (Q_i) con las fórmulas que se dan se tiene la posición	Posición si N impar: $Q_1: \frac{N+1}{4}$ $Q_2;\ M_e = \frac{N+1}{2}$ $Q_3: 3 \cdot \frac{N+1}{4}$	Posición si N par: $Q_1: \frac{N}{4}$ $Q_2: \frac{N}{2}$ $Q_3: 3 \cdot \frac{N}{4}$
Medidas de dispersión		
Rango	$X_{max} - X_{min}$	
Rango intercuartílico (RIQ)	$RIQ = Q_3 - Q_1$	

- Rango $= X_{max} - X_{min} =$

$$= 5 - 0 = 5$$

- $RIQ = Q_3 - Q_1$

Posición Q_1 : $\frac{N}{4} = \frac{26}{4} = 6'5 \longrightarrow Q_1 = \frac{X_6 + X_7}{2} = \frac{2+2}{2} = 2$

Posición Q_3 : $3\frac{N}{4} = 3 \cdot 6'5 = 19'5 \longrightarrow Q_3$

$$Q_3 = \frac{X_{19} + X_{20}}{2} = \frac{4+5}{2} = 4'5$$

x_i	n_i	N_i
0	3	3
1	1	4
2	4	8
3	5	13
4	6	19
5	7	26

$N = \sum n_i = 26$

$$\left.\begin{array}{l} RIQ = 4'5 - 2 = 2'5 \\ \text{Rango} = 5 \end{array}\right\} RIQ < \text{Rango}$$

$RIQ <$ Rango tenemos menos dispersión en los datos centrales de la distribución. El 50% de los datos centrales están más concentrados que los de los extremos.

d) Calcula el coeficiente de asimetría de Fisher e interpreta el resultado obtenido (0,5 puntos).

$$Fisher = \frac{\sum(x_i - \bar{X})^3}{N \cdot S^3}$$

Asimetría ? $\left.\begin{array}{l}\bar{x} = 3'192 \\ Me = 3'5\end{array}\right\}$ $\bar{x} < Me$ Asimetría negativa

$$S_x^2 = \frac{\sum x_i^2 \cdot n_i}{N} - \bar{x}^2 = \frac{0^2 \cdot 3 + 1^2 \cdot 1 + 2^2 \cdot 4 + 3^2 \cdot 5 + 4^2 \cdot 6 + 5^2 \cdot 7}{26} - 3'192^2 = 2'618$$

$$S_x = \sqrt{2'618} = 1'618$$

$$C.Fisher = \frac{(0-3'192)^3 + (1-3'192)^3 + (2-3'192)^3 + (3-3'192)^3 + (4-3'192)^3 + (5-3'192)^3}{26 \cdot 1'618^3} =$$

$$= \frac{-38'31}{26 \cdot 1'618^3} = -0'347 < 0 \longrightarrow \text{Asimetría negativa}$$

e) Si el coeficiente de curtosis de esta distribución es de 0,78674, obtén conclusiones sobre el grado de apuntamiento de esta distribución (0,5 puntos).

$$C.\ curtosis = 0'78674 > 0 \longrightarrow \text{Leptocúrtica}$$

La distribución tiene más apuntamiento que la Normal.

ANALISIS UNIDIMENSIONAL

Medidas de centralización

Media ($\bar{X}$)		$\bar{X} = \frac{\sum X_i \cdot n_i}{N}$ Datos agrupados: $\bar{X} = \frac{\sum c_i \cdot n_i}{N}$ donde c_i es la marca de clase	
Mediana (Me)	Posición: $\frac{N}{2}$ si N es par (en este caso tomaremos los dos datos intermedios para hacer un promedio) $\frac{N+1}{2}$ si N es impar	Valor de Me para datos agrupados: n par: $Me = L_I + \frac{\frac{n}{2} - N_{i-1}}{n_i} \cdot amplitud$ n impar: $Me = L_I + \frac{\frac{n+1}{2} - N_{i-1}}{n_i} \cdot amplitud$	
Moda (Mo)		Valor de Mo par datos agrupados: $M_o = L_I + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot amplitud$ donde $d_i = \frac{n_i}{amplitud}$	
Cuartiles (Q_i) con las fórmulas que se dan se tiene la posición		Posición si N impar: $Q_1 : \frac{N+1}{4}$ $Q_2;\ M_e = \frac{N+1}{2}$ $Q_3 : 3 \cdot \frac{N+1}{4}$	Posición si N par: $Q_1 : \frac{N}{4}$ $Q_2 : \frac{N}{2}$ $Q_3 : 3 \cdot \frac{N}{4}$

Medidas de dispersión

Rango	$X_{max} - X_{min}$	
Rango intercuartílico (RIQ)	$RIQ = Q_3 - Q_1$	
Varianza (S^2)	$S^2 = \frac{\sum (x_i - \bar{X})^2 \cdot n_i}{N}$ Fórmula alternativa: $S^2 = \frac{\sum x_i^2 \cdot n_i}{N} - \bar{X}^2 = \overline{X^2} - \bar{X}^2$ Datos agrupados utilizamos m_i y en lugar del valor x_i	
Desviación típica (S)	$S = \sqrt{S^2}$	
Coeficiente de variación (CV)	$CV = \frac{S}{\bar{X}}$ con $\bar{X} \neq 0$	
Coeficientes de asimetría	$Pearson$ $= \frac{\bar{X} - M_o}{S}$	$Fisher$ $= \frac{\sum (x_i - \bar{X})^3}{N \cdot S^3}$