ESTADÍSTICA NOVENO

MEDIDAS DE POSICIÓN DATOS AGRUPADOS

Las medidas de posición son estadísticas que se utilizan para resumir la ubicación de los datos dentro de un conjunto de datos.

Cuando trabajamos con datos agrupados, es decir, cuando los datos se presentan en intervalos o clases en lugar de valores individuales, las medidas de posición se calculan de manera ligeramente diferente a como se haría con datos no agrupados.

Las principales medidas de posición utilizadas con datos agrupados son:

Cuartiles (Q o K), deciles (D) y percentiles (P)

CUARTILES:

Los cuartiles son los tres valores de la variable que dividen a un conjunto de datos ordenados en cuatro partes iguales.

Se representan con la letra Q.

$Q_1 = 25\%$
$Q_2 = 50\%$ o también mediana.
$Q_3 = 75\%$
$Q_4 = 100\%$

Su fórmula:

$$Q_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i$$

Li = límite inferior.
Fi - 1 = Frecuencia absoluta acumulada anterior.
ai = amplitud de los intervalos.
fi = frecuencia absoluta.
K = valor del cuartil a hallar.

Lo primero que se debe hacer es hallar la posición del cuartil con la fórmula:

$$\text{Posición} = \frac{K.N}{4}$$

Luego se tiene la tabla de frecuencias, con marca de clase y frecuencia absoluta acumulada.
Se aplica la fórmula.

EJEMPLO:

Datos	f.a.	F.A.A
50 - 60	8	8
60 - 70	10	18
70 - 80	16	34
80 - 90	14	48
90 - 100	10	58
100 - 110	5	63
110 - 120	2	65
Total	**65**	

P $= \frac{1 \times N}{4}$ 1 es el Q1 o Q2 o Q3
4 cuartiles

Q1 $= \frac{1 \times 65}{4} = 16{,}25$

$$Q_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i$$

k = Q1
N = total datos
Fi = frecuencia acumulada anterior al intervalo
fi = frecuencia absoluta del intervalo
ai = amplitud del intervalo

Q1 $= 60 + \frac{\frac{1 * 65}{4} - 8}{10} \times 10$

Q1 $= 60 + \frac{16{,}25 - 8}{10} \times 10$

Q1 $= 60 + \frac{8{,}25}{10} \times 10$

Q1 $= 60 + 0{,}825 \times 10$

Q1 $= 60 + 0{,}825 \times 10$

Q1 $= 60 + 8{,}25$

Q1 $= 68{,}825$

DECILES:

Los deciles son los 10 valores de la variable que dividen a un conjunto de datos ordenados en diez partes iguales.

Se representan con la letra D.

$D_1 = 10\%$ $D^5 = 50\%$ $D_9 = 90\%$
$D_2 = 20\%$ $D_6 = 60\%$ $D_{10} = 100\%$
$D_3 = 30\%$ $D_7 = 70\%$
$D_4 = 40\%$ $D_8 = 80\%$

Su fórmula:

$$D_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i$$

Li = límite inferior.
Fi - 1 = Frecuencia absoluta acumulada anterior.

ai = amplitud de los intervalos.
fi = frecuencia absoluta.
K = valor del decil a hallar.

Lo primero que se debe hacer es hallar la posición del decil con la fórmula:

$$\text{Posición} = \frac{K.N}{4}$$

Luego se tiene la tabla de frecuencias, con marca de clase y frecuencia absoluta acumulada.
Se aplica la fórmula.

PERCENTI

Los percentiles son los 100 valores de la variable que dividen a un conjunto de datos ordenados en cien partes iguales.

Se representan con la letra P.

P_{1} = 1%	P^{50} = 50%	P_{94} = 94%
P_{2} = 2%	P_{68} = 68%	P_{100} = 100%
P_{3} = 3%	P_{76} = 76%	
P_{40} = 40%	P_{83} = 83%	

Su fórmula:

$$P_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i$$

Li = límite inferior.
Fi - 1 = Frecuencia absoluta acumulada anterior.
ai = amplitud de los intervalos.
fi = frecuencia absoluta.
K = valor del percentil a hallar.

Lo primero que se debe hacer es hallar la posición del percentil con la fórmula:

$$\text{Posición} = \frac{K.N}{4}$$

Luego se tiene la tabla de frecuencias, con marca de clase y frecuencia absoluta acumulada.
Se aplica la fórmula.

EJEMPLO:

Se tienen las edades de un grupo de estudiantes de un colegio, con los datos se desea conocer el valor de algunas medidas de posición así: Q1- Q3 – D6 – D9 – P15 – P36:

Datos	f.a.
10 - 14.	15
14 - 18.	9
18 - 22.	23
22 - 26.	18
26 - 30.	8
30 - 34.	12
34 - 38.	21
38 - 42.	36
42 - 46.	14
Total	156

Luego se completa la tabla:

Datos	f.a.	FAA
10 - 14.	15	15
14 - 18.	9	24
18 - 22.	23	47
22 - 26.	18	65
26 - 30.	8	73
30 - 34.	12	85
34 - 38.	21	106
38 - 42.	36	142
42 - 46.	14	156
Total	156	

	Datos	f.a.	FAA
	10 - 14.	15	15
P15	14 - 18.	9	24
Q1	18 - 22.	23	47
P36	22 - 26.	18	65
	26 - 30.	8	73
	30 - 34.	12	85
D6	34 - 38.	21	106
Q3 / D9	38 - 42.	36	142
	42 - 46.	14	156
	Total	156	

Halla la posición: Ubica en la tabla:

Q1=			
P =	1 * 156	156	39
	4	4	

Q3 =			
P =	3 * 156	468	117
	4	4	

D6 =			
P =	6 * 156	936	93,6
	10	10	

D9 =			
P =	9 * 156	1404	140,4
	10	10	

P15=			
P =	15 * 156	2340	23,4
	100	100	

P36=			
P =	36 * 156	5616	56,16
	100	100	

Despeja la fórmula con la posición encontrada y ubicada en la tabla:

Q1 =	18	.+.	39	.-.	24	*4
				23		
Q1 =	18	.+.	15			*4
Q3 =	38	.+.	0,305556	*4		
Q3 =	38	.+.	1,222222			
Q3 =	39,22222	39				

Q3 =	38	.+.	117	.-.	106	*4
				36		
Q3 =	38	.+.	11			*4
Q1 =	18	.+.	0,652174	*4		
Q1 =	18	.+.	2,608696			
Q1 =	20,60869565	21				

D9 =	38	.+.	140	.-.	106	*4
				36		
D9 =	38	.+.	34			*4
			30			
D9 =	38	.+.	1,133333	*4		
D9 =	38	.+.	4,533333			
D9 =	42,53333	43				

D6 =	34	.+.	94	.-.	85	*4
				21		
D6 =	34	.+.	9			*4
			21			
D6 =	34	.+.	0,428571	*4		
D6 =	34	.+.	1,714286			
D6 =	35,71428571	36				

P15	14	.+.	23	.-.	15	*4
				9		
P15	14	.+.	8			*4
			9			
P15	14	.+.	0,888889	*4		
P15	14	.+.	3,555556			
P15	17,55555556	18				

P36 =	22	.+.	56	.-.	47	*4
				18		
P36 =	22	.+.	9			*4
			30			
P36 =	22	.+.	0,3	*4		
P36 =	22	.+.	1,2			
P36 =	23,2	23				

EJERCICIOS

1. Un estudio sobre la cantidad de horas que duermen los estudiantes de una universidad ha recopilado los siguientes datos agrupados:

Horas de sueño	f.a.
5 - 6	10
6 - 7	25
7 - 8	35
8 - 9	20
9 - 10	10

Elabora la tabla.
Halla el Q2.
Halla el Q3.
Halla el D7.
Halla el D9.
Halla el P55.
Halla el P98.

MEDIDAS DE DISPERSIÓN PARA DATOS AGRUPADOS

Las medidas de dispersión son importantes porque proporcionan información sobre la homogeneidad o heterogeneidad de un conjunto de datos.

Un conjunto de datos con poca dispersión tendrá valores que están más cerca del valor central, mientras que un conjunto de datos con mucha dispersión tendrá valores que están más dispersos alrededor del valor central.

Algunas de las medidas de dispersión más comunes incluyen.

VARIANZA:

La varianza es una medida de dispersión que indica qué tan dispersos están los valores en relación con la media en un conjunto de datos.

Cuando los datos están agrupados en intervalos o clases, la varianza de datos agrupados se calcula sumando el cuadrado de las desviaciones de cada punto medio de clase ponderadas por sus frecuencias, y luego dividiendo por el número total de observaciones.

Fórmula:

$$\boxed{S^2 = \frac{\Sigma(x_i - \bar{x})^2 f_i}{n-1}}$$

Para hallar la varianza primero debe encontrar la media aritmética o promedio de los intervalos y completar la tabla.

ESTUDIANTES	f.a.
10 - 15.	5
15 - 20	9
20 - 25	12
25 - 30	15
30 - 35	11
35 - 40	8
Total	60

DESVIACIÓN TÍPICA O ESTANDAR:

Es la raíz cuadrada de la varianza. Proporciona una medida de dispersión en las mismas unidades que los datos, lo que la hace más interpretable.

Fórmula:

$$\boxed{\text{De} = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 f_i}{n-1}}}$$

Si ya tienes el valor de la varianza, únicamente la sacas a ese valor el de la raíz cuadrada, pero si no conoces el valor de la varianza debes despejar toda la fórmula.

EJEMPLO:

Con la tabla anterior encuentra la varianza y la desviación:

Se completa la tabla: recuerda la fórmula de la media.

ESTUDIANTES	f.a.	X	X * f.a.	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 * f$
10 - 15.	5	12,50	62,50	182,25	911,25
15 - 20	9	17,50	157,50	72,25	650,25
20 - 25	12	22,50	270,00	12,25	147,00
25 - 30	15	27,50	412,50	2,25	33,75
30 - 35	11	32,50	357,50	42,25	464,75
35 - 40	8	37,50	300,00	132,25	1 058,00
Total	60		1.560,00		3 265,00

$$\bar{x} = \sum \frac{x_i \cdot f_i}{n} = \frac{1.560,0}{60,0} \quad 26,0$$

Con ese dato se completa la tabla y se despeja la fórmula:

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2 f_i}{n-1} = \frac{3.265,0}{60 - 1} \quad \frac{3.265,0}{59,0} \quad 55,3$$

Varianza.

Ahora se saca la desviación:

De = $\sqrt{\frac{\Sigma(x_i - \bar{x})^2 f_i}{n-1}}$ = $\sqrt{55{,}3}$ = 7,43 Desviación.

EJERCICIOS

1. Con los datos de la siguiente tabla halla la varianza y la desviación de ellos.

Datos	f.a
10 - 30	15
30 - 50	23
50 - 70	9
70 - 90	12
90 - 110	25
Total	84

Recuerda completar la tabla para que puedas despejar las fórmulas con mayor facilidad.

PREGUNTAS TIPO SABER

1. Teniendo en cuenta la siguiente fórmula, escoge la opción en la cual se la utiliza.

Posición = % / 100 x N

A. Hallar el valor mínimo y máximo de un conjunto de datos.
B. Hallar valor de los cuartiles.
C. Hallar valor de los deciles y percentiles.
D. Hallar la posición de cuartiles, deciles y percentiles y ubicarlos para saber su valor en un grupo de datos.

2. ¿Cuál de las siguientes afirmaciones es correcta sobre la varianza y la desviación de datos agrupados?

A. La varianza y la desviación de datos agrupados se calculan de la misma manera que para datos no agrupados.
B. La varianza se calcula sumando los cuadrados de las desviaciones de cada punto de datos de la media, dividido por el número total de observaciones.
C. La desviación estándar de datos agrupados se calcula tomando la raíz cuadrada de la media de los cuadrados de las desviaciones de cada punto de datos de la media.
D. La varianza de datos agrupados se calcula multiplicando las frecuencias de cada clase por las desviaciones cuadradas de los puntos medios de las clases y dividiendo por el total de observaciones.

3. ¿Cuál de las siguientes medidas proporciona una estimación de cuánto varían los valores en relación con la media en un conjunto de datos agrupados?

A. Rango.
B. Mediana.
C. Desviación.
D. Moda.

4. Si se tiene un conjunto de datos agrupados en intervalos de clase y se quiere calcular la varianza. ¿Qué fórmula se utilizaría?

a) $\sigma = \sqrt{\frac{\sum f_i \cdot (X_i - \overline{X})^2}{N}}$

b) $\sigma = \frac{\sum f_i \cdot (X_i - \overline{X})^2}{N}$

c) $\sigma^2 = \frac{\sum f_i \cdot (X_i - \overline{X})^2}{N}$

d) $\sigma^2 = \sqrt{\frac{\sum f_i \cdot (X_i - \overline{X})^2}{N}}$