MEDIDAS DE DISPERSION

Una presentación de tu profesor Gonzalo H. Hueitra

ANTES DE COMENZAR

- Para ayudarte a resolver tus dudas he creado una clase en **Google Classroom**.
- Descarga la App y únete a la clase usando el siguiente código

Google Classroom

nltnzry

- Iré subiendo material de apoyo y podrás contactarme en caso de que necesites ayuda.

MEDIDAS DE DISPERCIÓN

Las **medidas de dispersión** sirven para determinar si los datos se concentran en torno a la media o si están muy dispersos. Es decir, indican la **variabilidad de los datos.**

Las más conocidas son: el **rango**, la **desviación media**, la **varianza** y la **desviación estándar**.

El **Rango** corresponde a la diferencia entre el mayor y el menor de los datos de la distribución. Indica cuán dispersos están los datos de los extremos.

$$R = x_{max} - x_{min}$$

Por ejemplo. Las notas obtenidas por un estudiante en una asignatura son las siguientes: 2,0 ; 3,9 ; 5,0 ; 5,9 ; 6,2. Obtener el rango.

$$\boldsymbol{R = 6,2 - 2,0 = 4,2}$$

Lo que indica que las notas son bastantes dispersas, ya que la amplitud entre ambos valores es "grande".

RANGO

La **media aritmética** (o promedio) de una distribución estadística permite hacernos una idea de sus características. Sin embargo, dos países pueden tener igual renta (o sueldo) promedio, pero desiguales distribuciones. Por ejemplo uno de ellos puede tener sectores muy adinerados y otros muy pobres, mientras que en el otro la distribución sea más equitativa. Por esto, se hace necesario conocer otras características de estas. Para ello se debe estudiar la **desviación media**, la **varianza** y la **desviación media.**

La **desviación de una variable x con respecto a su media aritmética** es la diferencia entre los valores de la variable (x_i) y la media aritmética ($\bar{X}$). Es decir cada dato lleva asociado una desviación, pero al obtener la sumatoria de todas ellas resulta cero.

$$D = x_i - \bar{X}$$

DESVIACIÓN

La **desviación media de un conjunto de datos** ($D_{\bar{X}}$) es la media aritmética de los valores absolutos de las desviaciones de cada dato respecto a la media ($\bar{X}$). Estos es,

Para **datos no agrupados** se tiene:

$$D_{\bar{X}} = \frac{\sum|x_i - \bar{X}|}{n} = \frac{|x_1 - \bar{X}| + |x_2 - \bar{X}| + |x_3 - \bar{X}| + \cdots + |x_n - \bar{X}|}{n}$$

Para **datos agrupados** se tiene:

$$D_{\bar{X}} = \frac{\sum(|x_i - \bar{X}| \cdot f_i)}{n} = \frac{|x_1 - \bar{X}| \cdot f_1 + |x_2 - \bar{X}| \cdot f_2 + |x_3 - \bar{X}| \cdot f_3 + \cdots + |x_n - \bar{X}| \cdot f_n}{n}$$

Donde,
x_i: son los valores de la variable.
f_i: es la frecuencia absoluta de cada valor de la variable.
n: es el número total de datos.

DESVIACION MEDIA

Por ejemplo. Las notas obtenidas por un estudiante en una asignatura son las siguientes: 2,0 ; 3,9 ; 5,0 ; 5,9 ; 6,2. Obtener la desviación media.

Paso 1: Para calcular la desviación media, primero se debe calcular la media aritmética.

$$\bar{X} = \frac{2{,}0 + 3{,}9 + 5{,}0 + 5{,}9 + 6{,}2}{5} = \frac{23}{5} = 4{,}6$$

Paso 2: Con esta media aritmética se pueden obtener las desviaciones para cada nota:

Notas	x	2,0	3,9	5,0	5,9	6,2
Desviación	$x - \bar{X}$	$-2{,}6$	$-0{,}7$	0,4	1,3	1,6

Estos valores indican el mayor o menor alejamiento de las notas respecto a $\bar{X}$

Paso 3: Se verifica que la suma de las desviaciones respecto al promedio:

$$(-2{,}6) + (-0{,}7) + 0{,}4 + 1{,}3 + 1{,}6 = 0$$

DESVIACION MEDIA

Paso 4: Se calcula la **desviación media** de la siguiente manera:

$$D_{\bar{X}} = \frac{|2{,}0-4{,}6|+|3{,}9-4{,}6|+|5{,}0-4{,}6|+|5{,}9-4{,}6|+|6{,}2-4{,}6|}{5}$$
$$= \frac{2{,}6+0{,}7+0{,}4+1{,}3+1{,}6}{5} = \frac{6{,}6}{5} = 1{,}32$$

La **desviación media** permite determinar en cuánto varían, en promedio, los datos de una distribución con respecto a la media aritmética.

DESVIACION MEDIA

La desviación media permite determinar en cuánto varían, en promedio, los datos de una distribución con respecto a la media aritmética. Para cuantificar ese grado de dispersión se estudian la **varianza** y la **desviación estándar**.

La **varianza** (σ^2 (corresponde a la media aritmética de los cuadrados de las desviaciones. Se expresa en unidades cuadradas.

Su fórmula está para **datos no agrupados** dada por:

$$\sigma^2 = \frac{\sum (x_i - \bar{X})^2}{n} = \frac{(x_1 - \bar{X})^2 + (x_2 - \bar{X})^2 + \cdots + (x_n - \bar{X})^2}{n}$$

Para **datos agrupados**, se tiene:

$$\sigma^2 = \frac{\sum [(x_i - \bar{X})^2 \cdot f_i]}{n} = \frac{(x_1 - \bar{X})^2 \cdot f_1 + (x_2 - \bar{X})^2 \cdot f_2 + \cdots + (x_n - \bar{X})^2 \cdot f_n}{n}$$

VARIANZA

Por ejemplo. Las notas obtenidas por un estudiante en una asignatura son las siguientes: $2{,}0\ ;\ 3{,}9\ ;\ 5{,}0\ ;\ 5{,}9\ ;\ 6{,}2$. Obtener la varianza y la desviación estándar, si se sabe que el promedio es $\bar{X} = 4{,}6$.

Paso 1: Para calcular la media de los cuadrados de las diferencias entre cada dato y el promedio. Se obtiene así la **varianza** (σ^2):

$$\sigma^2 = \frac{(2{,}0 - 4{,}6)^2 + (3{,}9 - 4{,}6)^2 + (5{,}0 - 4{,}6)^2 + (5{,}9 - 4{,}6)^2 + (6{,}2 - 4{,}6)^2}{5}$$

$$= \frac{(-2{,}6)^2 + (-0{,}7)^2 + (0{,}4)^2 + (1{,}3)^2 + (1{,}6)^2}{5} = \frac{6{,}76 + 0{,}49 + 0{,}16 + 1{,}69 + 2{,}56}{5}$$

$$= \frac{11{,}66}{5} = 2{,}332$$

VARIANZA

Paso 2: Calculando la raíz cuadrada del valor anterior se obtiene la **desviación estándar** (σ):

$$\sigma = \sqrt{2{,}332} \approx 1{,}53 \left.\vphantom{\sqrt{2}}\right\} \text{Grado de dispersión respecto a } \bar{X}$$

La **desviación estándar** (σ) se obtiene extrayendo la raíz cuadrada de la varianza. Se expresa en la misma unidad que la variable, por lo que nos puede dar una idea más cercana de lo disperso que es el conjunto.

Mientras menos sea la desviación estándar, los datos son más homogéneos, es decir, a menor dispersión mayor homogeneidad, y viceversa.

DESVIACION ESTANDAR

ACTIVIDADES

Estas actividades están pensadas para que sean desarrolladas en su cuaderno. **Recuerda** que estas actividades deben ser entregadas al retorno de clases para su evaluación.

1. Ordena los siguientes datos de menor a mayor. Luego calcular el rango de los datos.

$$3,22 \;\; ; \;\; 2,93 \;\; ; \;\; 3,01 \;\; ; \;\; 4,48 \;\; ; \;\; 5,06 \;\; ; \;\; 4,31 \;\; ; \;\; 2,98 \;\; ; \;\; 3,07$$

2. Si la estatura de la estudiante más alta de un curso es 1,72 m y la estatura del menor es 1,48 m, calcular el rango de la estatura de este curso.

3. Calcular la desviación estándar de los siguientes datos:

$$3 \;\; ; \;\; 5 \;\; ; \;\; 6 \;\; ; \;\; 7 \;\; ; \;\; 10 \;\; ; \;\; 12 \;\; ; \;\; 15 \;\; ; \;\; 18$$

4. Un grupo de deportistas obtuvo las siguientes marcas en atletismo (expresada en horas): $2,60$; $2,40$; $2,30$; $2,00$; $2,10$. Calcular e interpretar la desviación media de los datos.

5. Calcula la desviación media para la cantidad de dólares que tiene un grupo de turistas. Para ello completa la siguiente tabla.

Dólares que tienen los turistas

Cantidad de dólares	Marca de clase x	Desviación $x - \bar{X}$	Frecuencia f	Desviación $\lvert x_i - \bar{X} \rvert \cdot f$
$[150 - 250[$			7	
$[250 - 350[$			5	
$[350 - 450[$			13	
$[450 - 550[$			5	
$[550 - 650[$			2	
	Total		**32**	

Marca de clase. Se calcula como la suma del límite inferior (menor valor) y el límite superior (mayor valor) del intervalo, dividido en 2

$$D_{\bar{X}} = \frac{\sum \lvert x_i - \bar{X} \rvert \cdot f_i}{n} =$$

6. **Analiza la información de las tablas. Luego, responde.**
Las tablas muestran las notas de dos cursos diferentes obtenidas en una misma prueba de matemática.

III Medio A

Notas	Frecuencia
7,0	3
6,7	5
6,3	4
6,0	8
4,0	8
3,4	2
3,0	4
Total	**34**

III Medio B

Notas	Frecuencia
6,0	2
5,5	5
5,3	9
5,2	5
5,1	3
5,0	10
Total	**32**

A. ¿Cuál es la desviación estándar de los cursos?
B. ¿Qué curso tiene mejor rendimiento? Fundamenta

7. Completa la tabla. Luego, responde

Edad a la que contrajo matrimonio

Edad	x	f	$(x-\bar{X})^2$	$(x-\bar{X})^2 \cdot f$
$[16, 20[$		2		
$[20, 24[$		8		
$[24, 28[$		8		
$[28, 32[$		18		
$[32, 36[$		20		
$[36, 40[$		18		
$[40, 44[$		15		
$[44, 48[$		8		
$[48, 52[$		3		
Total		**100**		

A. ¿Cuál es la varianza de las edades?

B. ¿Cuál es la desviación estándar de las edades?

C. ¿Los datos están distribuidos de forma homogénea? Justifica,

Nos vemos pronto!

¿Alguien tiene alguna pregunta?
Escríbeme a
hola@profegonzalo.cl