

SESIÓN 7

VARIANZA Y DESVIACIÓN ESTÁNDAR EN DISTRIBUCIÓN DE DATOS AGRUPADOS

I. CONTENIDOS:

1. Varianza y desviación estándar en datos agrupados.
2. Coeficiente de variabilidad.

II. OBJETIVOS:

Al término de la Sesión, el alumno:

- Comprenderá el concepto de desviación a partir de la medida en un conjunto de datos agrupados

III. PROBLEMATIZACIÓN:

Comenta las preguntas con tu Asesor y selecciona las ideas más significativas.

- Si dos distribuciones tienen las mismas medidas de tendencia central y de dispersión, ¿El análisis de una de ellas servirá también para la otra?
- ¿Crees que además de la mediana habrá otras formas útiles de dividir una distribución de datos no agrupados?

IV. TEXTO INFORMATIVO-FORMATIVO:

1.1. Varianza y desviación estándar en datos agrupados

Para calcular la varianza y la desviación estándar en una distribución de datos agrupados se deben desarrollar en la tabla de distribución de frecuencias los siguientes campos:

Intervalos de clase, frecuencia, marca de clase, producto de la frecuencia por la marca de clase, marca de clase elevada al cuadrado, y producto de la frecuencia por el cuadrado de la marca de clase.

$$S^2 = \frac{\sum fX^2 - \frac{(\sum fX)^2}{N}}{N}$$

La varianza se calcula con la fórmula:

La desviación estándar con: $S = \sqrt{S^2}$

Ejemplo 1: Determinar la desviación estándar en la siguiente distribución, que trata sobre un grupo de empleados en un negocio clasificándolos por grupos de edad.

EDAD	EMPLEADOS
17 – 23	7
24 – 30	15
31 – 37	8
38 - 44	7
45 – 51	3
52 – 58	4
59 - 65	2

Primero debemos construir la tabla con los campos que se señalaron al inicio de esta clase:

Intervalos de clase	f	X	fX	X ²	fX ²
17 – 23	7	20	140	400	2800
24 – 30	15	27	405	729	10935
31 – 37	8	34	272	1156	9248
38 - 44	7	41	287	1681	11767
45 – 51	3	48	144	2304	6912
52 – 58	4	55	220	3025	12100
59 - 65	2	62	124	3844	7688
	46		1592		61450

Para calcular la varianza se utiliza la fórmula:

$$S^2 = \frac{\sum fX^2 - \frac{(\sum fX)^2}{\sum f}}{\sum f}$$

Sustituyendo:

$$S^2 = \frac{61450 - \frac{(1592)^2}{46}}{46} = 138.107$$

Para calcular la desviación estándar se utiliza:

$$S = \sqrt{S^2}$$

Sustituyendo:

$$S = \sqrt{138.107} = 11.75$$

Las edades de los empleados del negocio se desvían de la media, en promedio, 11.75 años

2.1. Coeficiente de variabilidad

Para comparar una distribución con otra, en cuanto a su dispersión, existe otra medida conocida como el coeficiente de variabilidad. Es el cociente entre la desviación estándar y la media de una distribución:

$$CV = \frac{S}{\bar{X}}$$

Se acostumbra expresar en porcentaje, para esto debemos multiplicar por 100 el cociente anterior. Entre más bajo es el coeficiente de variabilidad, menos dispersos son los datos de la distribución.

Para el ejemplo anterior el coeficiente de variabilidad se calcula así:

Primero se calcula la media $1592 \div 46 = 34.60$

Luego, sustituyendo en la fórmula:

$$CV = \frac{11.75}{34.60} = 0.3396$$

El coeficiente de variabilidad de la distribución es de 33.96%

V. ESTRATEGIAS CENTRADAS EN EL APRENDIZAJE:

A. Agrupa los siguientes datos y con las distribuciones de frecuencia, calcula e interpreta la varianza, desviación estándar y el coeficiente de variabilidad.

a) Tiempo (en semanas) que han requerido 30 personas para conseguir empleo.

12	25	9	24	5	11
15	23	7	20	19	7
5	4	3	1	16	5
12	0	6	0	3	21
18	17	13	8	2	16

b) Juguetes defectuosos por lote, de una fábrica.

0	2	3	19	4	1	4	8	5	7	0	18
3	19	12	13	8	5	7	17	9	8	6	21
16	6	4	9	15	6	9	16	4	3	9	15
24	15	16	7	29	9	8	17	3	16	12	13
5	12	21	6	16	13	5	21	16	2	5	9
6	8	0	2	23	12	2	6	15	21	7	27
7	3	4	1	5	21	7	5	11	5	8	5