Medidas de dispersión

Las **medidas de dispersión** nos informan sobre cuánto se alejan del centro los valores de la distribución.

Las **medidas de dispersión** son:

Rango o recorrido

El **rango** es la **diferencia** entre el **mayor** y el **menor** de los **datos** de una distribución estadística.

Desviación media

La **desviación media** es la **media aritmética** de los **valores absolutos** de las **desviaciones** respecto a la **media**.

Varianza

La **varianza** es la **media aritmética** del **cuadrado de las desviaciones** respecto a la **media**.

Desviación típica

La **desviación típica** es la **raíz cuadrada** de la **varianza**.

Vamos a ver cómo se calcula cada una de estas medidas de dispersión, pincha sobre cada uno de sus nombres para ver su explicación

- **Desviación media**
- **Varianza**
- **Desviación típica**

También tienes información sobre el cálculo del **coeficiente de variación**.

Desviación media

La **desviación media** es la **media aritmética** de los **valores absolutos de las desviaciones respecto a la media**. La **desviación respecto a la media** es la **diferencia** entre cada **valor** de la variable estadística y la **media aritmética**.

$$D_i = x_i - \bar{x}$$

La **desviación media** se representa por $D_{\bar{x}}$

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \ldots + |x_n - \bar{x}|}{N}$$

$$D_{\bar{x}} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{N}$$

Ejemplo

Calcular la **desviación media** de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

$$\bar{x} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$D_{\bar{x}} = \frac{|9-9| + |3-9| + |8-9| + |8-9| + |9-9| + |8-9| + |9-9| + |18-9|}{8} = 2.25$$

Desviación media para datos agrupados

Si los datos vienen agrupados en una **tabla de frecuencias**, la expresión de la **desviación media** es:

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}|f_1 + |x_2 - \bar{x}|f_2 + \ldots + |x_n - \bar{x}|f_n}{N}$$

$$D_{\bar{x}} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|f_i}{N}$$

Ejemplo

Calcular la **desviación media** de la distribución:

	x_i	f_i	$x_i \cdot f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot f_i$
[10, 15)	**12.5**	**3**	**37.5**	**9.286**	**27.858**
[15, 20)	**17.5**	**5**	**87.5**	**4.286**	**21.43**
[20, 25)	**22.5**	**7**	**157.5**	**0.714**	**4.998**
[25, 30)	**27.5**	**4**	**110**	**5.714**	**22.856**
[30, 35)	**32.5**	**2**	**65**	**10.174**	**21.428**

		21	**457.5**		**98.57**

$$\bar{x} = \frac{457.5}{21} = 21.786$$

$$D_{\bar{x}} = \frac{98.57}{21} = 4.69$$

Varianza

La **varianza** es la **media aritmética del cuadrado de las desviaciones respecto a la media** de una distribución estadística.

La varianza se representa por σ^2.

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{N} \quad \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}$$

Varianza para datos agrupados

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \ldots + (x_n - \bar{x})^2 f_n}{N}$$

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 f_i}{N}$$

Para simplificar el **cálculo de la varianza** vamos o utilizar las siguientes expresiones que son equivalentes a las anteriores.

$$\sigma^2 = \frac{x_1^2 + x_2^2 + \ldots + x_n^2}{N} - \bar{x}^2 \quad \sigma^2 = \sum_{i=1}^{n}\frac{x_i^2}{N} - \bar{x}^2$$

Varianza para datos agrupados

$$\sigma^2 = \frac{x_1^2 f_1 + x_2^2 f_2 + \ldots + x_n^2 f_n}{N} - \bar{x}^2 \quad \sigma^2 = \sum_{i=1}^{n}\frac{x_i^2 f_i}{N} - \bar{x}^2$$

Ejercicios de varianza

Calcular la varianza de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

$$\bar{x} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$\sigma^2 = \frac{(9-9)^2 + (3-9)^2 + (8-9)^2 + (8-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2 + (18-9)^2}{8} = 15$$

Calcular la varianza de la distribución de la tabla:

	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[10, 20)	15	1	15	225
[20, 30)	25	8	200	5000
[30,40)	35	10	350	12 250
[40, 50)	45	9	405	18 225
[50, 60	55	8	440	24 200

[60,70)	**65**	**4**	**260**	**16 900**
[70, 80)	**75**	**2**	**150**	**11 250**
		42	**1 820**	**88 050**

$$\bar{x} = \frac{1820}{42} = 43.33$$

$$\sigma^2 = \frac{88050}{42} - 43.33^2 = 218.94$$

Propiedades de la varianza

1 La **varianza** será siempre un **valor positivo o cero**, en el caso de que las puntuaciones sean iguales.

2 Si a todos los **valores** de la variable se les **suma** un **número** la **varianza no varía**.

3 Si todos los **valores** de la variable se **multiplican** por un **número** la **varianza** queda **multiplicada** por el **cuadrado** de dicho **número**.

4 Si tenemos varias distribuciones con la misma **media** y conocemos sus respectivas **varianzas** se puede calcular la **varianza total**.

Si todas las muestras tienen el mismo tamaño:

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2}{n}$$

Si las muestras tienen distinto tamaño:

$$\sigma^2 = \frac{k_1 \cdot \sigma_1^2 + k_2 \cdot \sigma_2^2 + \ldots + k_n \cdot \sigma_n^2}{k_1 + k_2 + \ldots + k_n}$$

Observaciones sobre la varianza

1 La **varianza**, al igual que la media, es un índice muy sensible a las puntuaciones extremas.

2 En los casos que **no se pueda hallar la media** tampoco será posible hallar la **varianza**.

3 La **varianza** no viene expresada en las mismas unidades que los datos, ya que las desviaciones están elevadas al cuadrado.

Desviación típica

La **desviación típica** es la **raíz cuadrada de la varianza**.

Es decir, la raíz cuadrada de la media de los cuadrados de las puntuaciones de desviación.

La **desviación típica** se representa por **σ**.

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{N}} \qquad \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}}$$

Desviación típica para datos agrupados

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 f_1 + (x_2 - \bar{x})^2 f_2 + \ldots + (x_n - \bar{x})^2 f_n}{N}}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2 f_i}{N}}$$

Para simplificar el cálculo vamos o utilizar las siguientes expresiones que son equivalentes a las anteriores.

$$\sigma = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{N} - \bar{x}^2} \quad \sigma = \sqrt{\sum_{i=1}^{n} \frac{x_i^2 f_i}{N} - \bar{x}^2}$$

Desviación típica para datos agrupados

$$\sigma = \sqrt{\frac{x_1^2 f + x_2^2 f_2 + \ldots + x_n^2 f_n}{N} - \bar{x}^2} \quad \sigma = \sqrt{\sum_{i=1}^{n} \frac{x_i^2}{N} - \bar{x}^2}$$

Ejercicios de desviación típica

Calcular la **desviación típica** de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

$$\bar{x} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$\sigma^2 = \sqrt{\frac{(9-9)^2 + (3-9)^2 + (8-9)^2 + (8-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2 + (18-9)^2}{8}} = 3.87$$

Calcular la desviación típica de la distribución de la tabla:

	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[10, 20)	15	1	15	225
[20, 30)	25	8	200	5000
[30,40)	35	10	350	12 250
[40, 50)	45	9	405	18 225
[50, 60)	55	8	440	24 200
[60,70)	65	4	260	16 900
[70, 80)	75	2	150	11 250
		42	1 820	88 050

$$\bar{x} = \frac{1820}{42} = 43.33$$

$$\sigma = \sqrt{\frac{88050}{42} - 43.33^2} = 14.797$$

Propiedades de la desviación típica

1 La **desviación típica** será siempre un **valor positivo o cero**, en el caso de que las puntuaciones sean iguales.

2 Si a todos los **valores** de la variable se les **suma** un **número** la **desviación típica no varía**.

3 Si todos los **valores** de la variable se **multiplican** por un **número** la **desviación típica** queda **multiplicada** por dicho **número**.

4 Si tenemos varias distribuciones con la misma **media** y conocemos sus respectivas **desviaciones típicas** se puede calcular la **desviación típica total**.

Si todas las muestras tienen el mismo tamaño:

$$\sigma = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}{n}}$$

Si las muestras tienen distinto tamaño:

$$\sigma = \sqrt{\frac{k_1 \cdot \sigma_1^2 + k_2 \cdot \sigma_2^2 + \dots + k_n \cdot \sigma_n^2}{k_1 + k_2 + \dots + k_n}}$$

Observaciones sobre la desviación típica

1 La **desviación típica**, al igual que la media y la varianza, es un índice muy sensible a las puntuaciones extremas.

2 En los casos que **no se pueda hallar la media** tampoco será posible hallar la **desviación típica**.

3 Cuanta más pequeña sea la **desviación típica** mayor será la **concentración de datos** alrededor de la **media**.

Coeficiente de variación

El **coeficiente de variación** es la relación entre la **desviación típica** de una muestra y su **media**.

$$C.V = \frac{\sigma}{\bar{x}}$$

El **coeficiente de variación** se suele expresar en **porcentajes**:

$$C.V = \frac{\sigma}{\bar{x}} \cdot 100$$

El **coeficiente de variación** permite comparar las **dispersiones** de dos distribuciones distintas, siempre que sus **medias** sean **positivas**.

Se calcula para cada una de las distribuciones y los valores que se obtienen se comparan entre sí.

La **mayor dispersión** corresponderá al valor del **coeficiente de variación mayor**.

Ejercicio

Una distribución tiene x = 140 y σ = 28.28 y otra x = 150 y σ = 25. ¿Cuál de las dos presenta mayor dispersión?

$$C.V_1 = \frac{28.28}{140} \cdot 100 = 20.2\%$$

$$C.V_2 = \frac{24}{150} \cdot 100 = 16\%$$

La primera distribución presenta mayor dispersión.