

Medidas de dispersión

Consideraremos diversas medidas de dispersión. El rango se sustenta en los valores máximo y mínimo del conjunto de datos, es decir, sólo se consideran dos valores. La desviación media, la varianza y la desviación estándar se basan en desviaciones de la media aritmética.

Rango La medida más simple de dispersión es el rango. Representa la diferencia entre los valores máximo y mínimo de un conjunto de datos.

En forma de ecuación: $Rango = x_{m\acute{a}x} - x_{m\acute{i}n}$

Es decir $Rango = valor_{m\acute{a}ximo} - valor_{m\acute{i}nimo}$

El rango se emplea mucho en aplicaciones de control de procesos estadísticos (CPE), debido a que resulta fácil de calcular y entender.

Desviación media.

Un problema que presenta el rango estriba en que parte de dos valores, el más alto y el más bajo, es decir, no los toma en cuenta a todos. La desviación media sí lo hace; mide la cantidad media respecto de la cual los valores de una población o muestra varían. Expresado en forma de definición:

DESVIACIÓN MEDIA.

Media aritmética de los valores absolutos de las desviaciones con respecto a la media aritmética.

En el caso de una muestra, la desviación media, designada ***DM***, se calcula mediante la fórmula:

$$DM = \frac{\sum |x - \bar{x}|}{n}$$

en donde:

x es el valor de cada observación. ***x̄*** es la media aritmética de los valores.

n es el número de observaciones en la muestra.

| | indica el valor absoluto.

¿Por qué ignorar los signos de las desviaciones de la media? De no hacerlo, las desviaciones positivas y negativas se compensarían con exactitud unas a otras y la desviación media siempre sería cero. **Dicha medida (cero) resultaría un estadístico sin utilidad.**

La siguiente tabla muestra el número de recién nacidos en el hospital Rovirosa y hospital de los Ángeles entre las 4 y las 5 de la tarde, de una muestra de 5 días el mes pasado.

20	40	50	60	80
----	----	----	----	----

20	49	50	51	80
----	----	----	----	----

Determine la media, la mediana, el rango y la desviación media de cada hospital. Compare las similitudes y diferencias.

Calculamos la media del hospital Rovirosa.

$$\bar{x} = \frac{20+40+50+60+80}{5} = \frac{250}{5} = 50$$

$$\bar{x} = 50$$

Calculamos la media del hospital Los ángeles.

$$\bar{x} = \frac{20+49+50+51+80}{5} = \frac{250}{5} = 50$$

$$\bar{x} = 50$$

Las medias son iguales.

Calculamos la mediana de ambos hospitales.

20	40	50	60	80
----	----	----	----	----

20	49	50	51	80
----	----	----	----	----

Las medianas son iguales.

Ahora calculemos el rango de ambos hospitales.

20	40	50	60	80
----	----	----	----	----

20	49	50	51	80
----	----	----	----	----

$$\text{Rango} = x_{\text{máx}} - x_{\text{mín}}$$

$$\text{Rango} = 80 - 20 = 60$$

Ambos hospitales tienen el mismo rango.

Ahora calculemos la desviación media y empleemos la formula.

20	40	50	60	80
----	----	----	----	----

$$DM = \frac{\sum|x-\bar{x}|}{n}$$

$$DM = \frac{|20-50|+|40-50|+|50-50|+|60-50|+|80-50|}{5}$$

$$DM = \frac{|-30|+|-10|+|0|+|10|+|30|}{5}$$

$$DM = \frac{30+10+10+30}{5}$$

$$DM = \frac{80}{5}$$

$$DM = 16$$

20	49	50	51	80
----	----	----	----	----

$$DM = \frac{\sum|x-\bar{x}|}{n}$$

$$DM = \frac{|20-50|+|49-50|+|50-50|+|51-50|+|80-50|}{5}$$

$$DM = \frac{|-30|+|-1|+|0|+|1|+|30|}{5}$$

$$DM = \frac{30+1+1+30}{5}$$

$$DM = \frac{62}{5}$$

$$DM = 12.4$$

La desviación media son distintos, por tanto, el hospital Rovirosa está más cerca de la media, es decir está más concentrado las demandas.

La desviación media posee dos ventajas.

Primero, incluye todos los valores de los cálculos. Recuerde que el rango sólo incluye los valores máximo y mínimo. Segundo, es fácil de definir: es la cantidad promedio que los valores se desvían de la media. Sin embargo, su inconveniente es el empleo de valores absolutos. Por lo general, es difícil trabajar con valores absolutos, así que la desviación media no se emplea con tanta frecuencia como otras medidas de dispersión, como la desviación estándar.

Actividades de aprendizajes.

Los pesos de los contenedores enviados a Irlanda son (en miles de libras):

95, 103, 105, 110, 104, 105, 112, 90.

- a) ¿Cuál es el rango de los pesos?
- b) Calcule el peso medio aritmético.
- c) Estime la desviación media de los pesos.

En los siguientes ejercicios calcule: a) el rango; b) la media aritmética; c) la desviación media; d) interprete los valores que obtenga.

1. Hubo cinco representantes de servicio al cliente que trabajaron en Electronic Super Store durante la pasada venta de fin de semana. Las cantidades de HDTV que vendieron estos representantes son: 5, 8, 4, 10 y 3.

2. El Departamento de Estadística de la Western State University ofrece ocho secciones de estadística básica. En seguida aparecen los números de estudiantes matriculados en estas secciones: 34, 46, 52, 29, 41, 38, 36 y 28.

3. Dave's Automatic Door instala puertas automáticas para cocheras. La siguiente lista indica el número de minutos que se requieren para instalar una muestra de 10 puertas automáticas: 28, 32, 24, 46, 44, 40, 54, 38, 32 y 42.

5. Una muestra de ocho compañías de la industria aeronáutica participaron en una encuesta sobre la recuperación de la inversión que tuvieron el año pasado. Los resultados (en porcentaje) son los siguientes: 10.6, 12.6, 14.8, 18.2, 12.0, 14.8, 12.2 y 15.6.

Varianza y desviación estándar.

La varianza y la desviación estándar también se fundamentan en las desviaciones de la media. Sin embargo, en lugar de trabajar con el valor absoluto de las desviaciones, la varianza y la desviación estándar lo hacen con el cuadrado de las desviaciones.

VARIANZA

Media aritmética de las desviaciones de la media elevadas al cuadrado.

La varianza es no negativa y es cero sólo si todas las observaciones son las mismas.

DESVIACIÓN ESTÁNDAR.

Raíz cuadrada de la varianza.

Varianza de la población.

Las fórmulas de la varianza poblacional y la varianza de la muestra son ligeramente diferentes. La varianza de la población se estudia primero. (Recuerde que una población es la totalidad de las observaciones estudiadas.) La varianza de la población se determina de la siguiente manera:

$$\sigma^2 = \frac{\sum(x - \mu)^2}{N}$$

En esta fórmula:

σ^2 es la varianza de la población (σ es la letra minúscula griega sigma); se lee *sigma al cuadrado*.

X es el valor de una observación de la población.

μ es la media aritmética de la población.

N es el número de observaciones de la población.

Observe el proceso de cálculo de la varianza:

1. Comience por determinar la media.
2. Calcule la diferencia entre cada observación y la media, y eleve al cuadrado dicha diferencia.
3. Sume todas las diferencias elevadas al cuadrado.
4. Divida la suma de las diferencias elevadas al cuadrado entre el número de elementos de la población.

Desviación estándar de la población.

Tanto el rango como la desviación media resultan fáciles de interpretar. El rango es la diferencia entre los valores alto y bajo de un conjunto de datos, y la desviación media es la media de las desviaciones de la media.

$$\text{Desviación estándar: } \sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

Ejemplo

Los años de servicio de una muestra de siete empleados en la oficina de quejas de State Farm Insurance, de Cleveland, Ohio, son: 4, 2, 5, 4, 5, 2 y 6.

- A) ¿Cuál es la varianza de la muestra?
- B) Calcule la desviación estándar de la muestra.

Observe el proceso de cálculo de la varianza:

Comience por determinar la media.

$$\mu = \frac{4+2+5+4+5+2+6}{7} = \frac{28}{7} = 4$$

$\mu = 4$

2. Calcule la diferencia entre cada observación y la media, y eleve al cuadrado dicha diferencia.

$$\sum(x - \mu)^2 = (4 - 4)^2 + (2 - 4)^2 + (5 - 4)^2 + (4 - 4)^2 + (5 - 4)^2 + (2 - 4)^2 + (6 - 4)^2$$

$$\sum(x - \mu)^2 = (0)^2 + (-2)^2 + (1)^2 + (0)^2 + (1)^2 + (-2)^2 + (2)^2$$

3. Sume todas las diferencias elevadas al cuadrado.

$$\sum(x - \mu)^2 = 4 + 1 + 1 + 4 + 4$$

$$\sum(x - \mu)^2 = 14$$

4. Divida la suma de las diferencias elevadas al cuadrado entre el número de elementos de la población.

$$\frac{\sum(x - \mu)^2}{N} = \frac{14}{7} = 2$$

$$\sigma^2 = \frac{\sum(x - \mu)^2}{N} \rightarrow \sigma^2 = 2$$

La varianza es 2

Desviación estándar: $\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}$

$$\sigma = \sqrt{2}$$

$$\sigma = 1.4142$$

La desviación estándar es **1.4142**

Realiza los ejercicios siguientes.

En los ejercicios 47-52, efectúe lo siguiente:

a) Calcule la varianza de la muestra.

b) Determine la desviación estándar de la muestra.

47. Considere los siguientes valores como una muestra:

7	2	6	2	3
---	---	---	---	---

48. Los siguientes cinco valores son una muestra:

11	6	10	6	7
----	---	----	---	---

49. Dave's Automatic Door, que se mencionó en el ejercicio 37, instala puertas automáticas para cocheras. Sobre la base de una muestra, los siguientes son los tiempos, en minutos, que se requieren para instalar 10 puertas automáticas:

28	32	24	46	44	40	54	38	32	42
----	----	----	----	----	----	----	----	----	----

50. A la muestra de ocho compañías en la industria aeronáutica (ejercicio 38), se le aplicó una encuesta referente a su recuperación de inversión del año pasado. Los resultados son los siguientes

10.6	12.6	14.8	18.2	12.0	14.8	12.2	15.6
------	------	------	------	------	------	------	------

51. La Asociación de Propietarios de Moteles de Houston, Texas, llevó a cabo una encuesta relativa a las tarifas de motel entre semana en el área. En seguida aparece la tarifa por cuarto para huéspedes de negocios en una muestra de 10 moteles.

\$101	\$97	\$103	\$110	\$78	\$87	\$101	\$80	\$106	\$88
-------	------	-------	-------	------	------	-------	------	-------	------

52. Una organización de protección al consumidor se ocupa de las deudas de las tarjetas de crédito. Una encuesta entre 10 adultos jóvenes con una deuda con la tarjeta de crédito de más de \$2 000 mostró que éstos pagan en promedio un poco más de \$100 mensuales como abono a sus saldos. En la siguiente lista aparecen las sumas que cada adulto joven pagó el mes pasado.

\$110	\$126	\$103	\$93	\$99	\$113	\$87	\$101	\$109	\$100
-------	-------	-------	------	------	-------	------	-------	-------	-------