

Medidas estadísticas



Indagación

Recuerda las principales medidas estadísticas que has estudiado desde los cursos pasados y en tu cuaderno realiza una lista que compararás con dos o tres compañeros.



Conceptualización Medidas estadísticas

Haciendo una recopilación de lo estudiado en cursos anteriores, a cerca de las medidas usadas en estadística, tendremos en cuenta lo siguiente:

Las medidas descriptivas se dividen en dos grandes grupos, las medidas de tendencia central y las medidas de dispersión o variación.

Las medidas de tendencia central corresponden a aquellas que nos dan una idea de los valores medios, valores centrales o más frecuentes de una determinada distribución de valores. La media, moda y mediana son ejemplos de ellas.

Recordemos lo que significan:

- La media aritmética o promedio, es la medida de tendencia central más utilizada, un ejemplo de esta utilización, es el sacar el promedio de las notas de una materia. Esta medida de tendencia central es un dato que se ubica en el centro de los datos y representa las características del grupo. Podemos decir que esta medida es el punto de equilibrio del conjunto de datos. Recuerda que se representa con el símbolo \bar{x} .
- La mediana es el dato que divide un conjunto de datos en dos partes proporcionalmente iguales. Se representa por: \tilde{x} .
- La moda es el dato que más se repite. Se puede representar de dos formas Mo y \bar{x} , pero la más utilizada es la primera.

Las medidas de dispersión son aquellas que nos informan sobre el grado de variabilidad o variación presente en un grupo de datos u observaciones y como ejemplo tenemos al rango, varianza, desviación estándar, desviación media y el coeficiente de variación.

Otros autores hacen la siguiente clasificación:

Las medidas de centralización sirven para determinar los valores centrales de la distribución o conjunto de datos. Estas son moda, media y media aritmética o promedio.

Las medidas de dispersión dan una idea sobre la representatividad de las medidas centrales, a mayor dispersión menor representatividad. Ellas son: varianza, desviación estándar y desviación media.

Las medidas de localización son útiles para encontrar determinados valores importantes, para una “clasificación” de los elementos de la muestra o población. Ellas son los cuartiles, deciles y percentiles.

Resumiendo:

Estadística: es la rama de la matemática que nos permite recoger, organizar y analizar datos. Existen dos conceptos importantes dentro de la estadística que nos permiten analizar y estudiar dichos datos, estos son: población y muestra.

Población: es el conjunto de datos que caracteriza el fenómeno que se desea estudiar.

Muestra: es un subconjunto de la población a estudiar, el cual es necesario que sea representativo de toda la población.

Gráfica: es una representación de la relación entre variables, muchos tipos de gráficos aparecen en estadística, según la naturaleza de los datos involucrados y el propósito de la gráfica, es la de representar los valores tabulados obtenidos de los muestreos o los datos del total de la población. Las gráficas más usadas son: pictogramas, barras, histogramas, polígonos de frecuencias y gráficas circulares.

Distribución de frecuencia: Al resumir grandes colecciones de datos, es útil distribuirlos en clases o categorías, y determinar el número de individuos que pertenecen a cada clase llamado frecuencia de clase.

Una disposición tabular de los datos por clases, junto con las frecuencias correspondientes de clase, se llaman distribuidores de frecuencia o tablas de frecuencia.

Analicemos el caso siguiente:

En un taller de automóviles, su dueño quiere realizar un estudio para saber que el taller se puede sostener sin tener pérdidas, para esto, ha establecido que el promedio de carros que debe entrar diario de lunes a domingo, para poder pagar a sus empleados debe ser mínimo de 5 carros o mayor, para esto tomó los siguientes datos en una semana de lunes a domingo:

Dados los datos de carros que entran de lunes a domingo en su orden: 3, 5, 2, 7, 6, 4, 9.

Para poder responder a la pregunta del dueño, es necesario establecer un estudio estadístico que me permita responder a la pregunta. Para esto, es necesario calcular:

- La **moda**, la mediana y la media.
- La **varianza**, la **desviación media** y la **desviación típica**.
- Los **cuartiles** 1° y 3°.
- Los **deciles** 2° y 7°.
- Los **percentiles** 32 y 85.

Solución

- Moda:** no existe **moda** porque todas las puntuaciones tienen la misma frecuencia, es decir no hay números que se repitan.

Mediana 2, 3, 4, 5, 6, 7, 9. Me = 5 porque es el dato que queda en el centro del conjunto ordenado.

Conclusión: el 50% de los días que se abrió, 5 carros o menos fueron al taller.

Media aritmética, media o promedio

$$\bar{X} = \frac{2 + 3 + 4 + 5 + 6 + 7 + 9}{7} = 5.143$$

Suma de todos los valores, dividido entre el número total de dato.

Conclusión: durante los siete días, en promedio fueron 5.143 carros al taller y puede redondearse a 5.

b. **Varianza**

$$\sigma^2 = \frac{2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 9^2}{7} = 31.43$$

La **varianza** es la **media aritmética del cuadrado de las desviaciones respecto a la media aritmética** de una distribución estadística. La varianza se representa por σ^2 o por S^2 .

La **Desviación típica o Estándar** es la raíz cuadrada de la varianza. Luego va esa igualdad.

$$\sigma = \sqrt{4.978} = 2.231$$

Desviación media:

Es la sumatoria de las diferencias entre los datos y la media, dividida entre el número de datos.

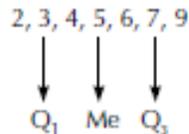
$$d_x = \frac{|2-5.143| + |3-5.143| + |4-5.143| + |5-5.143| + |6-5.143| + |7-5.143| + |9-5.143|}{7} = 1.878$$

Rango:

Rango = Dato mayor - dato menor

$$R = 9 - 2 = 7$$

c. **Cuartiles:**



Q₁ = El 25% de los días en que se tomó la asistencia de carros, 3 carros o menos asistieron al taller, mientras que el 75% de los días más de 3 carros asistieron al taller.

Q₃ = El 75% de los días en que se tomó la asistencia de carros, aproximadamente 7 carros o menos asistieron al taller, mientras que el 25% de los días más de 7 personas asistieron.

d. **Deciles:**

$$7 (2/10) = 1.4 \quad \text{entonces} \quad D_2 = 3$$

$$7 (7/10) = 4.9 \quad \text{entonces} \quad D_7 = 6$$

e. **Percentiles**

$$7 (32/100) = 2.2 \quad \text{entonces} \quad P_{32} = 4$$

$$7 (85/100) = 5.9 \quad \text{entonces} \quad P_{85} = 7$$

Problema 2:

En la empresa de gaseosas el manantial, un trabajador considera que una de las dos máquinas que tiene la empresa, se encuentra dañada pues en las botellas de su producto estrella, no están saliendo los 470cc que se ofrecen. Como el trabajador quiere saber cuál de las máquinas está dañada, escoge una muestra de 10 botellas llenas de cada máquina, mide su contenido y anota en la siguiente tabla.

Máquina 1	470	453	465	460	458	468	465	470	467	455
Máquina 2	460	445	470	455	465	463	470	487	445	470

Lo primero que tenemos que hacer, es hallar la media de la muestra que se tomó de las dos máquinas, para poder establecer un elemento de juicio para saber que máquina tiene el menor promedio de llenado.

Es el promedio de c.c. con el cual la máquina 1 llena las botellas.

$$\bar{X}_{\text{Máquina 1}} = \frac{470 + 453 + 465 + 460 + 458 + 468 + 465 + 470 + 467 + 455}{10}$$

$$\bar{X}_{\text{Máquina 1}} = \frac{4,630}{10} = 463 \quad \text{Es el promedio de c.c. con el cual la máquina 1 llena las botellas.}$$

$$\bar{X}_{\text{Máquina 2}} = \frac{460 + 445 + 470 + 455 + 465 + 463 + 470 + 487 + 445 + 470}{10}$$

$$\bar{X}_{\text{Máquina 2}} = \frac{4,630}{10} = 463 \quad \text{Es el promedio de cc con el cual la máquina 2 llena las botellas.}$$

De esta manera, podemos observar que la media aritmética no es un elemento válido de conocimiento para saber cuál es la máquina que está dañada.

Utilizaremos la desviación estándar que nos permitirá observar cuál es la dispersión de los datos con respecto a la media.

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad \text{desviación estándar}$$

Máquina 1

$$S = \sqrt{\frac{(470-463)^2 + (453-463)^2 + (465-463)^2 + (460-463)^2 + (458-463)^2 + (468-463)^2 + (465-463)^2 + (470-463)^2 + (466-463)^2 + (455-463)^2}{10}}$$

$$S = \sqrt{\frac{(7)^2 + (-10)^2 + (2)^2 + (3)^2 + (-5)^2 + (5)^2 + (2)^2 + (7)^2 + (3)^2 + (8)^2}{10}}$$

$$S = \sqrt{\frac{49 + 100 + 4 + 9 + 25 + 25 + 4 + 49 + 9 + 64}{10}}$$

$$S = \sqrt{\frac{338}{10}}$$

$$S = \sqrt{33.8}$$

$$S = 5.81$$

Máquina 2

$$S = \sqrt{\frac{(460-463)^2 + (445-463)^2 + (470-463)^2 + (455-463)^2 + (465-463)^2 + (463-463)^2 + (470-463)^2 + (487-463)^2 + (445-463)^2 + (470-463)^2}{10}}$$

$$S = \sqrt{\frac{9 + 64 + 49 + 64 + 4 + 0 + 49 + 576 + 64 + 49}{10}}$$

$$S = \sqrt{\frac{880}{10}}$$

$$S = \sqrt{88}$$

$$S = 9.38$$

Como resultado tenemos, que la desviación estándar de las máquinas son: máquina 1: $S = 5.81$ y la máquina 2: $S = 9.38$, es así como podemos concluir que la máquina 2, es la que tiene mayor dispersión. De esta manera, podemos deducir que es la máquina que se debe enviar a reparación.



Aplicación

En tu cuaderno resuelve los siguientes ejercicios y analízalos con tus compañeros.

1. El maestro de música de la escuela de Luna nueva, ha conformado una banda con sus 20 estudiantes quienes tienen edades que oscilan entre los 11 y los 15 años.

Las edades son:

14 15 11 13 14 14 12 15 15 14 13 14 12 11 14 14 13 12 14 15

Sigue cada instrucción:

1. Ordena la distribución de edades, de mayor a menor.
 2. Construye la tabla de frecuencias.
 3. Señala la mediana (Me).
 4. Identifica la modas o modas, si las hay.
 5. Expresa el rango de la distribución.
 6. Encuentra los cuartiles.
 7. Calcula la varianza.
 8. Encuentra la desviación típica.
 9. Calcula la desviación media.
 10. Representa la distribución dada en un histograma.
2. Una empresa petrolera desea contratar a una persona que sea quien dirija las inversiones en la ciudad capital, para esto, ya solo dos aspirantes están en la etapa de las últimas pruebas y los resultados de cada una, se muestran a continuación:

	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5	Prueba 6	Prueba 7	Prueba 8	Prueba 9	Prueba 10
Aspirante 1	9.5	8.3	8.7	9.1	9.3	9.7	8.7	9.2	9.8	10
Aspirante 2	10	10	8.0	9.2	9.8	7.0	9.5	9.8	9.8	9.5

¿Qué aspirante elegiría usted para quedarse con el cargo?