



REPÚBLICA DE COLOMBIA  
SECRETARÍA DE EDUCACIÓN MUNICIPAL DE PALMIRA  
"INSTITUCIÓN EDUCATIVA "DE ROZO"  
Aprobada por Resolución N° 0835 del 20 de FEBRERO de 2.017



## GUÍA DE APRENDIZAJE No. 2

Grado: Noveno	
Área o asignatura: Matemáticas	ESTADÍSTICA
Fecha de recibido:	31 de Agosto de 2020
Fecha de entrega:	30 de septiembre 2020
Docentes:	Diego Bernal, Frederick Ramírez
Nombre del estudiante:	
Objetivo de aprendizaje y/o DBA:	Propone un diseño estadístico adecuado para resolver una pregunta que indaga por la comparación sobre las distribuciones de dos grupos de datos, para lo cual usa comprensivamente diagramas de caja, medidas de tendencia central, de variación y de localización.

## INTRODUCCIÓN

Este tema es muy interesante ya que su empleo está en nuestra vida diaria. Es muy útil para entender y tomar decisiones en muchos trabajos sobre todo lo que compete a la calidad de los procesos. Lo que se busca es que el alumno logre desarrollar como se hallan las medidas de dispersión haciendo uso de las formulas, su uso para plantear justificaciones a procedimientos propuestos y tomar decisiones acertadas para solucionar situaciones en la vida diaria sobre todo de las empresas.

Teniendo en cuenta su practicidad y utilización en la vida cotidiana, se ha diseñado esta guía con el propósito de adquirir las competencias suficientes con relación a este tema.

Por lo anterior, esta guía está diseñada para trabajarla aproximadamente en dos semanas, teniendo en cuenta que usted le va a dedicar aproximadamente dos horas semanales. Cada tema de la guía contiene una explicación, ejemplos y unos ejercicios de práctica al final.

El procedimiento es que ustedes vayan leyendo la explicación que se proporciona en cada tema, pongan atención a los ejemplos y los ejercicios los copien y desarrollen en el cuaderno.

## TEMAS

- ✓ Rango o recorrido
- ✓ Desviación media
- ✓ Varianza
- ✓ Desviación estándar

## MEDIDAS DE DISPERSIÓN

A pesar de la gran importancia de las medidas de tendencia central y de la cantidad de información que aportan individualmente, no hay que dejar de señalar que en muchas ocasiones esa información, no sólo no es completa, sino que puede inducir a errores en su interpretación. Veamos algunos ejemplos.

Consideremos dos grupos de personas extraídos como muestras respectivas de dos poblaciones distintas: el primero está compuesto por 100 personas que asisten a la proyección de una película para niños, y el segundo por 100 personas elegidas entre los asistentes a una discoteca juvenil. Pudiera ocurrir que, aun siendo las distribuciones de las edades de ambos grupos muy distinta, la media y la mediana coincidieran para ambas.

Igualmente ocurre en este otro ejemplo. La caja de un kiosco registra las siguientes entradas en miles de pesos, a lo largo de dos semanas correspondientes a épocas distintas del año

1ª semana	2ª semana
10	30
20	40
30	50
50	50
60	60
80	60
<u>100</u>	<u>60</u>
<b>350</b>	<b>350</b>

La media y la mediana de ambas distribuciones coinciden (el valor de ambas es 50 en los dos casos) y, sin embargo, las consecuencias que se podrían derivar de una y otra tabla son bien distintas.

Comprendemos pues, a la vista de estos ejemplos, la necesidad de conocer otras medidas, aparte de los valores de centralización, que nos indiquen la mayor o menor desviación de cada observación respecto de aquellos valores.

Las medidas de desviación, variación o dispersión que estudiaremos a continuación son: Rango o amplitud, desviación media, varianza y desviación estándar.

### RANGO, AMPLITUD TOTAL O RECORRIDO

El rango se suele definir como la diferencia entre los dos valores extremos que toma la variable.

$$R = V_{\max} - V_{\min}$$

Es la medida de dispersión más sencilla y también, por tanto, la que proporciona menos información. Además, esta información puede ser errónea, pues el hecho de que no influyan más de dos valores del total de la serie puede provocar una deformación de la realidad.

Comparemos, por ejemplo, estas dos series:

Serie 1: 1 5 7 7 8 9 9 10 17

Serie 2: 2 4 6 8 10 12 14 16 18

$$R_1 = 17 - 1 = 16$$

$$R_2 = 18 - 2 = 16$$

Ambas series tienen rango 16, pero están desigualmente agrupadas, pues mientras la primera tiene una mayor concentración en el centro, la segunda se distribuye uniformemente a lo largo de todo el recorrido.

El uso de esta medida de dispersión, será pues, bastante restringido.

### DESVIACIÓN MEDIA

En teoría, la desviación puede referirse a cada una de las medidas de tendencia central: media, mediana o moda; pero el interés se suele centrar en la medida de la desviación con respecto a la media, que llamaremos desviación media.

Puede definirse como la media aritmética de las desviaciones de cada uno de los valores con respecto a la media aritmética de la distribución, y se indica así:

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

Nótese que se toman las desviaciones en valor absoluto, es decir, que la fórmula no distingue si la diferencia de cada valor de la variable con la media es en más o es menos.

Ya se habrá advertido que esta expresión sirve para calcular la desviación media en el caso de datos sin agrupar. Veamos un ejemplo:

Se tiene los valores 2,2,4,4,4,5,6,7,8,8. Averiguar la desviación media de estos valores.

$x$	$x - \bar{x}$	$ x $
2	-3	3
2	-3	3
4	-1	1
4	-1	1
4	-1	1
5	0	0
6	1	1
7	2	2
8	3	3
8	3	3

Hallamos la media  $\bar{x} = \frac{2+2+4+4+4+4+5+6+7+8+8}{10} = \frac{50}{10} = 5$

$$DM = \frac{18}{10} = 1,8$$

Veamos ahora cómo se calcula la desviación media en el caso de datos agrupados en intervalos.

Donde observamos que ahora las desviaciones van multiplicadas por las frecuencias de los intervalos correspondientes.

Además, las desviaciones son de cada centro, o marca de clase, a la media aritmética. Es decir,

$$DM = \frac{\sum |f(x-\bar{x})|}{N}$$

Ejemplo: Para hallar la desviación media de la siguiente tabla referida a las edades de los 100 empleados de una cierta empresa:

Clase	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

Veamos cómo se procede:

Clase	f	X <sub>(clase media)</sub>	f · x	x - $\bar{x}$	f ·  x - $\bar{x}$
16-20	2	18	36	16,72	33,44
20-24	8	22	176	12,72	101,76
24-28	8	26	208	8,72	69,76
28-32	18	30	540	4,72	84,96
32-36	20	34	680	0,72	14,4
36-40	18	38	684	3,28	59,04
40-44	15	42	630	7,28	109,2
44-48	8	46	368	11,28	90,24
48-52	3	50	150	15,28	45,84
	100		3.472		608,64

$$\text{Media } (\bar{X}) = \frac{3472}{100} = 34,72$$

$$\text{DM} = \frac{608,64}{100} = 6,0864$$

La desviación media viene a indicar el grado de concentración o de dispersión de los valores de la variable. Si es muy alta, indica gran dispersión; si es muy baja refleja un buen agrupamiento y que los valores son parecidos entre sí.

La desviación media se puede utilizar como medida de dispersión en todas aquellas distribuciones en las que la medida de tendencia central más significativa haya sido la media. Sin embargo, para las mismas distribuciones es mucho más significativa la desviación estándar, que estudiaremos a continuación, y eso hace que el uso de la desviación media sea cada vez más restringido.

### VARIANZA Y DESVIACIÓN TÍPICA O ESTÁNDAR

La varianza y la desviación estándar son medidas de dispersión o variabilidad, es decir, indican la dispersión o separación de un conjunto de datos. Hay que tener en cuenta que las fórmulas de la varianza y la desviación estándar son diferentes para una muestra que para una población. Se diferencian en que el denominador de la población es (N) y el de la muestra es (n-1).

#### Varianza ( $s^2$ )

La varianza se define como la media aritmética de los cuadrados de las diferencias de los datos con su media aritmética.

$$S^2 = \frac{\Sigma(x-\bar{x})^2}{N} \quad s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$$

#### Desviación estándar (s)

La desviación estándar es la raíz cuadrada positiva de la varianza.

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \quad s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

Se recomienda calcular primero la varianza de la población y luego sacar su raíz cuadrada para obtener la desviación estándar.

Ten en cuenta que, si tienes una serie de valores de una población y necesitas calcular su varianza y su desviación estándar, deberás calcular primero la media

Ejemplo. Hallar la varianza y la desviación estándar de la serie: 5, 8, 10, 12, 16.

x	$(x - \bar{x})$	$(x - \bar{x})^2$
5	-5,2	27,04
8	-2,2	4,84
10	-0,2	0,04
12	1,8	3,24
16	5,8	33,64

		68,8
--	--	------

Primero hallamos la  $\bar{x} = \frac{5+8+10+12+16}{5} = 10,2$

Varianza =  $\frac{68,8}{5} = 13,76$

Desviación estándar =  $\sqrt{13,76} = 3,7094$

### EJEMPLO

El gerente de una empresa de alimentos desea saber que tanto varían los pesos de los empaques (en gramos), de uno de sus productos; por lo que opta seleccionar al azar cinco unidades de ellos para pesarlos. Los productos tienen los siguientes pesos (490, 500, 510, 515 y 520) gramos respectivamente. Hallar la varianza y la desviación estándar. Por lo que primero hallamos su media que es:

$$\frac{490+500+510+515+520}{5} = \frac{2535}{5} = 507$$

#### Varianza

$$S^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

$$S^2 = \frac{(490-507)^2+(500-507)^2+(510-507)^2+(515-507)^2+(520-507)^2}{(5-1)}$$

$$S^2 = \frac{(-17)^2+(-7)^2+(3)^2+(8)^2+(13)^2}{4} \quad S^2 = \frac{289+49+9+64+169}{4} \quad S^2 = \frac{580}{4}$$

$$S^2 = 145$$

#### Desviación estándar

$$S = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

$$S = \sqrt{\frac{(490-507)^2+(500-507)^2+(510-507)^2+(515-507)^2+(520-507)^2}{(5-1)}}$$

$$S = \sqrt{\frac{(-17)^2+(-7)^2+(3)^2+(8)^2+(13)^2}{4}} \quad S = \sqrt{\frac{289+49+9+64+169}{4}} \quad S = \sqrt{\frac{580}{4}}$$

$$S = \sqrt{145} \quad S = 12,04$$



### ACTIVIDAD

1. ¿Cuáles son las medidas de dispersión?
2. ¿Qué nos permite determinar las medidas de dispersión?
3. ¿Qué relación se puede establecer entre las medidas de tendencia central y las medidas de dispersión?
4. ¿Cómo se determinan las medidas de dispersión en las diferentes distribuciones de datos?
5. La siguiente tabla muestra un registro histórico a lo largo de 40 años de la temperatura (T) promedio en Celsius que se alcanzó en el mes de julio en la ciudad de Medellín, junto a la cantidad de veces que se presentó dicha temperatura promedio.

Límite inferior	Límite superior	Marca de clase	Frecuencia
14	15	14,5	4
16	17	16,5	10
18	19	18,5	15
20	21	20,5	11

Con base en los datos que nos suministra la situación, determinemos las medidas de dispersión para este caso: Rango, Varianza y Desviación Típica



## AUTOEVALUACIÓN

Vas a reflexionar respecto a cómo te sentiste y qué tanto aprendiste en el desarrollo de esta guía.

En tu cuaderno registra las conclusiones a las que llegaste *¡Debes de ser muy sincero!*

1. ¿Qué fue lo que más te causó dificultad al resolver las actividades de la guía?
2. ¿Por qué crees que te causó dificultad?
3. ¿Qué fue lo que te gustó del trabajo en casa en la guía?
4. Con tus palabras escribe qué aprendiste
5. ¿Qué crees que puedes hacer en la próxima guía para que entiendas mejor lo que se te propone?

## BIBLIOGRAFÍA

[https://contenidosparaaprender.colombiaaprende.edu.co/G\\_11/M/M\\_G11\\_U05\\_L04/M\\_G11\\_U05\\_L04\\_01\\_01\\_01.html](https://contenidosparaaprender.colombiaaprende.edu.co/G_11/M/M_G11_U05_L04/M_G11_U05_L04_01_01_01.html)