



# Guía para el aprendizaje Nº5

Nombre de alumno/a: \_\_\_\_\_

Curso: \_\_\_\_\_

Asignatura: **Matemáticas**

Nivel: **III medio**

Unidad: **Estadística y probabilidades**

Contenido: **Medidas de tendencia central y dispersión**

Objetivo de aprendizaje: *Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión y probabilidades condicionales.*

## Instrucciones:

*En los próximos días deberás resolver la guía, la cual te ayudará para prepararte para la prueba del mes. Debes hacer tus consultas al profesor, a continuación se presentan el correo electrónico para que puedas ponerte en contacto con el docente.*

## Estadística descriptiva

### Lección I. Medidas de tendencia central

Las medidas de tendencia central por lo general corresponden a los valores que se encuentran en la parte central de un conjunto de datos, nos ayudan a resumir la información recibida en un solo número, los cuales nos ayudarán a comprender de mejor forma los datos.

#### Aprende 1. Media aritmética ( $\bar{x}$ )

Se define como el cociente (división) entre la suma de los valores de los datos, en el total de datos. Por lo general esta definición es para datos NO tabulados.

##### Ejemplo 1.

*Las notas obtenidas por 5 estudiantes son: 5,6 ; 6,8 ; 4,2 ; 5,7 ; 6,5. La media aritmética es:*

$$\bar{x} = \frac{5,6 + 6,8 + 4,2 + 5,7 + 6,5}{5} = \frac{28,8}{5} = 5,76$$

*RESPUESTA: Debido a que trabajamos notas, solo debemos trabajar con 1 decimal, la media aritmética es **5,8***

**Resuelve 1.** *Calcula la media aritmética de cada uno de los grupos presentados, recuerda tener en cuenta la pertinencia del número obtenido como resultado.*

1. *Un grupo comenta la cantidad de hermanos que tienen: [2, 3, 2, 5, 2, 3, 4, 4]*
2. *En un grupo familiar, comentan sus edades: [20, 25, 30, 13, 15, 25, 40, 27, 25]*
3. *Se habla de los kilos de arroz que se consumen en casa al mes [1, 4, 5, 6, 7, 8, 9, 2, 4, 2, 10, 1, 4]*

**Resuelve 2.** *A continuación, se presentan problemas que involucran el cálculo de la media aritmética.*

4. *Daniel quiere obtener un promedio final de 4,5. El promedio entre sus primeras 5 notas es 3,8. ¿Qué nota debiese obtener en la 6ta prueba para obtener un 4,5 de promedio?*

## Aprende 2. Mediana ( $M_e$ )

Se define como el valor central de una distribución, una vez ordenados los datos de manera decreciente o creciente. El dato que representa la mediana divide la distribución en dos grupos uno superior y otro inferior.

Se debe tener en cuenta si la cantidad de datos son par o impar.

### Ejemplo 1.

Si la cantidad de datos es par, existen dos datos centrales, por tanto debes hallar el valor medio entre ambos.

Se presentan los siguientes datos: [2, 5, 2, 3, 4, 4]

Se ordenan de menor a mayor [2, 2, 3, 5, 4, 4] y se identifican los valores centrales [2, 2, 3, 5, 4, 4]

$$M_e = \frac{3 + 5}{2} = 4$$

### Ejemplo 2.

Si la cantidad de datos es impar, existirá solo un dato central, siendo aquel la mediana.

Se presentan los siguientes datos: [1, 7, 5, 7, 8, 3, 4]

Se ordenan de menor a mayor [1, 3, 4, 5, 7, 7, 8] y se identifica el valor central [~~1~~, ~~3~~, ~~4~~, 5, ~~7~~, ~~7~~, ~~8~~]

$$M_e = 5$$

### Resuelve 3. Indica la mediana de los siguientes grupos de números

1. [5, 7, 8, 6, 4, 5, 3, 5, 6, 7, 5, 7, 2, 5]	2. [2, 3, 2, 5, 2, 3, 4, 4]
3. [1, 4, 5, 6, 7, 8, 9, 2, 4, 2, 10, 1, 4]	4. [20, 25, 30, 13, 15, 25, 40, 27, 25]

## Aprende 3. Moda ( $M_o$ )

Se define como el valor de la variable con mayor frecuencia absoluta, o el valor que mas se repite. Puede haber más de una moda.

### Ejemplo 1.

Se presentan los siguientes datos [1, 3, 4, 5, 5, 7, 7, 8]

Se identifican aquellos valores con mayor frecuencia, [5, 5] y [7, 7]

Ya que 5 y 7 se repiten (en la misma cantidad) ambos son moda de los datos presentados.

$$M_o = 5 \text{ y } 7$$

### Resuelve 4. Indica la moda de los siguientes grupos de números

1. [5, 7, 8, 6, 4, 5, 3, 5, 6, 7, 5, 7, 2, 5]	2. [2, 3, 2, 5, 2, 3, 4, 4]
3. [1, 4, 5, 6, 7, 8, 9, 2, 4, 2, 10, 1, 4]	4. [20, 25, 30, 13, 15, 25, 40, 27, 25]



## Lección II. Medidas de dispersión

Las medidas de dispersión indican cuán cercanos o alejados están los datos de un valor central, por lo general la media aritmética. Muestran el grado de variabilidad de los datos (cuán confiable son). Por lo general las medidas de dispersión se trabajan de forma contextualizada, en problemas matemáticos comparando dos grupos de datos.

### Aprende 4. Definiciones medidas de dispersión

FORMULARIO ( I )	
<b>Rango:</b> corresponde a la diferencia entre el mayor y el menor de los datos de la distribución.	$R = x_{máx} - x_{mín}$
<b>Desviación media:</b> corresponde a la media aritmética de los valores absolutos de las desviaciones.	$D_{\bar{x}} = \frac{\sum  x_i - \bar{x} }{n}$
<b>Varianza:</b> corresponde a la media aritmética de los cuadrados de las desviaciones de los n datos. Se expresa en unidades cuadradas.	$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$
<b>Desviación estándar:</b> se obtiene extrayendo la raíz cuadrada de la varianza. Se expresa en la misma unidad que la variable, por lo que nos puede dar una idea más cercana de los dispersos que el conjunto.	$\sigma = \sqrt{\sigma^2}$
<b>Coefficiente de variación:</b> permite realizar comparaciones entre conjuntos con respecto a la dispersión de sus datos, e incluso entre variables que se miden con diferentes unidades de medida.	$CV = \frac{\sigma}{ \bar{x} }$

#### Ejemplo 1.

A continuación se presentan las notas de dos grupos de personas, calculamos sus medidas de dispersión para compararlas.

Las notas son 6,4 5,8 6,8 6,2 6,5

- Primero debemos calcular el promedio de los datos,  $\bar{x} = 6,34$  como es una nota, trabajamos con 6,4
- Luego calculamos el rango. Al dato máximo (6,8) debemos restarle el dato mínimo (5,8)  
 $R = 6,8 - 5,8 = 1,0$
- Para hallar el resto de los datos, trabajaremos en una tabla.

Notas Grupo A	$ x_i - \bar{x} $ valor absoluto, de la diferencia entre el dato y el promedio.	$(x_i - \bar{x})^2$ Se calcula el cuadrado de la diferencia anterior
6,4	$ 6,4 - 6,4  = 0$	$0^2 = 0$
5,8	$ 5,8 - 6,4  = 0,6$	$0,6^2 = 0,36$
6,8	$ 6,8 - 6,4  = 0,4$	$0,4^2 = 0,16$
6,2	$ 6,2 - 6,4  = 0,2$	$0,2^2 = 0,04$
6,5	$ 6,5 - 6,4  = 0,1$	$0,1^2 = 0,01$

- Para hallar la desviación media debemos sumar los valores de la primera columna y dividirlo en la cantidad de datos (5)

$$D_{\bar{x}} = \frac{0 + 0,6 + 0,4 + 0,2 + 0,1}{5} = \frac{1,3}{5} = 0,26$$

- Para hallar la varianza debemos sumar los valores de la segunda columna y dividirlo en la cantidad de datos (5)

$$\sigma^2 = \frac{0 + 0,36 + 0,16 + 0,04 + 0,01}{5} = \frac{0,57}{5} = 11,4$$

- Para hallar la desviación estándar debemos calcular la raíz de la varianza

$$\sigma = \sqrt{11,4} = 3,376388 \dots$$

Para trabajar de una forma acotada, trabajaremos con 2 decimales, (redondear si es necesario), por lo tanto la desviación estándar es: 3,38

- Por último, para el coeficiente de variación debemos dividir la desviación estándar por el promedio.

$$CV = \frac{3,38}{6,4} = 0,528125 \approx 0,53$$

**RESPUESTA:** al obtener una desviación estándar de 3,38 eso nos indica cuán disperso están las notas del promedio. Comprender que mientras más bajo la desviación estándar es aún más confiable será el promedio como indicador estadístico.

**Resuelve 5.** A continuación se presentan 2 grupos de datos, calcula sus medidas de tendencia central e indica “cual es más confiable”

Roberto es entrenador de 2 ciclistas, y te presenta los tiempos (en minutos) de los entrenamientos que demoran en recorrer la misma pista.

Ciclista 1: 12, 9, 8, 9, 10, 11, 9, 7

Ciclista 2: 11, 8, 7, 10, 10, 10, 8, 10

Anota aquí tus respuestas, (trabaja solo con 1 decimal)

	Promedio	Rango	Desviación media	Varianza	Desviación estándar	Coficiente de variación
Ciclista 1						
Ciclista 2						

1. Desarrolla los datos de los ciclistas en las siguientes tablas

CICLISTA 1		
Tiempos (en minutos)	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
12		
9		
8		
9		
10		
11		
9		
7		
$\Sigma =$		

CICLISTA 2		
Tiempos (en minutos)	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
11		
8		
7		
10		
10		
10		
8		
10		
$\Sigma =$		

2. Según tus cálculos, ¿A que ciclista escogerías para una futura carrera, por qué?



### Lección III. Datos agrupados

Entendemos que a medida que la cantidad de datos recogidos para realizar un estudio estadístico se mayor, será más difícil realizar los cálculos. Sacar el promedio de 5 notas será sencillo en comparación si calculamos las notas de 500 personas. Para ello existe la agrupación de datos, y con ello, nuevas formas de calcular la media, mediana y moda para una gran cantidad de datos.

Se estudiarán dos tipos de tablas:

#### Ejemplo 1.

*Tabla por frecuencia: Es una gran cantidad de datos, pero las respuestas son limitadas.*

CANTIDAD DE HERMANOS	FRECUENCIA
0	6
1	5
2	7
<u>TOTAL</u>	<u>18</u>

**De la tabla se entiende:** 18 personas respondieron la encuesta, 6 personas tienen 0 hermanos, 5 tienen solo 1 y 7 personas tienen 2.

#### Ejemplo 2.

*Tabla por intervalos: Es una gran cantidad de datos, y existe un abanico muy amplio de respuestas.*

SUELDOS EN UNA EMPRESA	
SUELDO (MILES DE PESOS)	FRECUENCIA
[300, 450[	35
[450, 600[	20
[600, 750[	7
[750, 900[	3
<u>TOTAL</u>	<u>65</u>

**De la tabla se entiende:** 65 personas de la empresa respondieron la encuesta, hay 35 personas que reciben entre \$300.000 y \$450.000, aun que no se especifica claramente el sueldo de cada uno de ellos. Entonces si una persona gana \$388.540 estará en el grupo nº1.

**Resuelve 6.** Observa la siguiente tabla y responde las preguntas a continuación.

Gasto mensual Gastos básicos	
Gasto (miles de pesos)	Frecuencia
[100, 150[	9
[150, 200[	11
[200, 250[	19
[250, 300[	23
[300, 350[	17
[350, 400]	11

- ¿Cuántas personas han sido consultadas?
- ¿Cuántas personas gastan entre \$200.000 y \$300.000?
- ¿Puedes identificar cuantas personas gastan \$325.000?
- ¿En que grupo se ubica la persona que gasta \$250.000?

## Lección IV. Medidas de tendencia central datos agrupados

Como ya se mencionó para lograr calcular la media mediana y moda al trabajar con tablas, debemos trabajar de una forma distinta, para ello se presentarán las siguientes formulas.

<b>FORMULARIO ( II )</b>	
<p><b>Media:</b> Corresponde a la suma de: la multiplicación de cada dato (o marca de clase, en tablas de intervalo) por su frecuencia. Dividido en el total de datos.</p>	$\bar{x} = \frac{\sum x \cdot f_i}{n}$ <p><i>x: dato o marca de clase</i>  <i>f<sub>i</sub>: frecuencia del subgrupo</i>  <i>n: total de datos</i></p>
<p><b>Mediana:</b> Recordar que la mediana se encuentra en la mitad de los datos, por ello, debemos ubicar la mitad de los datos e indicar en que subgrupo se encuentra. Luego se utiliza la fórmula para clarificar e indicar el valor exacto del valor.</p>	$M_e = Li + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot A$ <p><i>L<sub>i</sub>: Limite inferior del subgrupo</i>  <i>f<sub>i</sub>: frecuencia del subgrupo</i>  <i>F<sub>i-1</sub>: Frecuencia acumulada del subgrupo anterior</i>  <i>A: Amplitud del sugfrupo</i></p>
<p><b>Moda:</b> Recordar que la moda es aquel valor que mas se repite, por ello debemos trabajar con aquel subgrupo que tenga mayor frecuencia.</p>	$M_o = Li + \frac{f_i - F_{i-1}}{(f_i - F_{i-1}) + (f_i - F_{i+1})} \cdot A$ <p><i>F<sub>i-1</sub>: Frecuencia acumulada del subgrupo anterior</i>  <i>F<sub>i+1</sub>: Frecuencia acumulada del subgrupo posterior</i></p>

**Resuelve 7.** A continuación se presentará una tabla extendida, la cual debes llenar con los datos solicitados, de ese modo podremos calcular la media, mediana y moda.

<b>Masa corporal de un grupo curso</b>			
Masa corporal (Kg)	x	Frecuencia (f)	Frecuencia acumulada (F)
Datos en grupos	Media de cada grupo	Cantidad de datos clasificado en cada subgrupo	Acumular la frecuencia de cada subgrupo
[40, 50[		9	
[50, 60[		11	
[60, 70[		19	
[70, 80[		23	
[80, 90[		17	
[90, 100]		11	

<p><b>Media:</b></p>	$\bar{x} = \frac{\sum x \cdot f_i}{n}$ <p><i>x: dato o marca de clase</i>  <i>f<sub>i</sub>: frecuencia del subgrupo</i>  <i>n: total de datos</i></p>
<p><b>Mediana:</b></p>	$M_e = Li + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot A$ <p><i>L<sub>i</sub>: Limite inferior del subgrupo</i>  <i>f<sub>i</sub>: frecuencia del subgrupo</i>  <i>F<sub>i-1</sub>: Frecuencia acumulada del subgrupo anterior</i>  <i>A: Amplitud del sugfrupo</i></p>
<p><b>Moda:</b></p>	$M_o = Li + \frac{f_i - F_{i-1}}{(f_i - F_{i-1}) + (f_i - F_{i+1})} \cdot A$ <p><i>F<sub>i-1</sub>: Frecuencia acumulada del subgrupo anterior</i>  <i>F<sub>i+1</sub>: Frecuencia acumulada del subgrupo posterior</i></p>



## Lección V. Medidas de dispersión datos agrupados

### FORMULARIO ( I )

<b>Rango:</b> corresponde a la diferencia entre el mayor y el menor de los datos de la distribución.	$R = x_{máx} - x_{mín}$
<b>Desviación media:</b> corresponde a la media aritmética de los valores absolutos de las desviaciones. En esta instancia se debe trabajar con la marca de clase.	$D_{\bar{x}} = \frac{\sum  mc_i - \bar{x} }{n}$
<b>Varianza:</b> corresponde a la media aritmética de los cuadrados de las desviaciones de los n datos. Se expresa en unidades cuadradas.	$\sigma^2 = \frac{\sum (mc_i - \bar{x})^2}{n}$
<b>Desviación estándar:</b> se obtiene extrayendo la raíz cuadrada de la varianza. Se expresa en la misma unidad que la variable, por lo que nos puede dar una idea más cercana de los dispersos que el conjunto.	$\sigma = \sqrt{\sigma^2}$
<b>Coefficiente de variación:</b> permite realizar comparaciones entre conjuntos con respecto a la dispersión de sus datos, e incluso entre variables que se miden con diferentes unidades de medida.	$CV = \frac{\sigma}{ \bar{x} }$

**Resuelve 8.** Lee con atención la siguiente situación, y responde a la problemática.

1. La cantidad de medicamentos utilizados diariamente en todos los pabellones de un hospital, durante el mes anterior, se registraron en la siguiente tabla.

Cantidad de medicamentos	Frecuencia
[0, 200[	12
[200, 400[	15
[400, 600[	20
[600, 800[	45
[800, 1000]	21



Es necesario tener una desviación estándar menor a 200, de lo contrario tendremos problemas en la organización y logística

¿Deberá preocuparse el director del hospital por la logística y entrega de medicamentos el próximo mes? Calcula la desviación media, la desviación estándar y aclara la duda del director.

2. Resuelve el trabajo resolviendo las siguientes tablas.

Cantidad de medicamentos	Marca de clase ( $X_i$ )	Frecuencia ( $f_i$ )	$X_i \cdot f_i$	$ X_i - \bar{X} $	$ X_i - \bar{X}  \cdot f_i$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^2 \cdot f_i$
[0, 200[		12					
[200, 400[		15					
[400, 600[		20					
[600, 800[		45					
[800, 1000]		21					
	$\sum =$			$\sum =$		$\sum =$	

Anota aquí tus respuestas, (**trabaja solo con 1 decimal**)

	Promedio	Rango	Desviación media	Varianza	Desviación estándar	Coefficiente de variación
Cantidad de medicamentos						

RESPUESTA:

# EJERCICIOS

**Resuelve 9.** Resuelve cada uno de los ejercicios presentados.

**1. Medidas de Tendencia Central**

Tobías juega fútbol en uno de los equipos de su comuna. Los puntos obtenidos por los equipos en el campeonato de apertura de la comuna han sido los siguientes:

**12, 10, 22, 18, 7, 14, 15, 17, 18, 20, 21, 20, 10, 10, 11, 15**

**Determina:**

a. El promedio de los puntos.

b. La mediana de los puntos

c. Moda de los puntos

**2. Medidas de Dispersión**

En el colegio de Antonia la premiación anual será la segunda semana de diciembre, los profesores junto a la UTP del colegio deben escoger al alumno de excelencia académica, los postulantes son, Marco, Antonia y María José cuyas notas se muestran en la siguiente tabla:

Antonia	5,8	6,5	6,7	6,5	6,4	6,3	6,2	6,8	6,1
Marco	6,4	6,3	6,0	6,6	6,7	6,8	6,4	6,1	6,3
María José	6,1	5,9	6,5	6,7	6,6	6,3	6,5	6,7	6,2

**Determina:**

a. El promedio de cada alumno.

b. La desviación estándar para las notas de cada alumno.

c. El coeficiente de variación (CV) de las notas para cada alumno.

d. En tu opinión, ¿quién tiene mejor rendimiento?