Guía Nº2: Medidas de Dispersión

Lección Nº1

Nombre: _____ Curso_____ Fecha: _____

OA 1. Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión y probabilidades condicionales.

Objetivo de la Guía Nº2

Analizar los datos de situaciones usando medidas de dispersión y tomar decisiones a partir de ellos

Instrucciones:

1. **La siguiente es una guía de Contenidos, debes escribir cada concepto y su definición en tu cuaderno.**
2. **Toda duda o consulta se debe informar al mail** profesora.carolina.salort@gmail.com **la cual será respondida a la brevedad**
3. **Todo avance como evidencia fotográfica debe ser enviado al mail** profesora.carolina.salort@gmail.com**, con el asusto " Avance Guía de aprendizaje Nº1: Estadística"**
4. **Puedes apoyar tu estudios con el link** https://www.youtube.com/watch?v=ISbnLcFFrNY **Unicoos: Estadística.**

Toma de decisiones aplicando medidas de dispersión de datos

Medidas de dispersión

Las medidas de dispersión sirven para determinar si los datos se encuentran en torno a la **media** o si están muy dispersos. Para cuantificar la dispersión, estudiaremos las medidas más conocidas: el **Rango**, la **Desviación Media**, la **Varianza** y la **Desviación Estándar**

Ejemplo

Observa la siguiente situación

El Entrenador de un equipo de natación debe elegir su representante para la próxima competencia de 100 m en estilo libre. Para ello, cuenta con información consistente en el tiempo, segundos, de las dos postulantes en las 5 últimas carreras en este estilo.

Competencia de Daniela	
Nº de carrera	Tiempo (s)
1	64
2	58
3	68
4	62
5	65

Competencia de Bárbara	
Nº de carrera	Tiempo (s)
1	69
2	63
3	65
4	50
5	70

a. ¿Cuál es el tiempo promedio de Daniela en las últimas 5 semanas de 100 m estilo libre?, ¿y el de Bárbara?
b. ¿Cómo son los promedios de Daniela y Bárbara?
c. ¿A quién debería elegir el entrenador para participar en la competencia?, ¿Por qué?

Refuerza la **Media aritmética**

Rango

El ***Rango (R)*** corresponde a la diferencia entre el mayor y el menor de los datos de la distribución. Esta medida indica le alguna manera cuan dispersos están los datos de la distribución

$$R = X_{màx} - X_{mìn}$$

Por ejemplo: en el caso anterior, si se denota R_1 y R_2 los rangos de Daniela y Bárbara respectivamente, se tiene:

$$R_1 = 68 - 58 = 10 \rightarrow R_1 = 10\ s$$

$$R_2 = 70 - 50 = 20 \rightarrow R_1 = 20\ s$$

Esto no da indicios de que los tiempos de Daniel pueden ser menos dispersos que los de Bárbara.

- La ***Desviación de una variable x con respecto a su media aritmética*** *está dada por*
$$D = x_i - \bar{x}$$

Del Ejemplo:

Analiza los pasos que realiza el entrenador para comparar los tiempos de las competencias de Daniela con respecto a su tiempo promedio.

Paso 1:

Calcula las desviaciones de los tiempos de Daniela, tal como se muestra a continuación:

Tiempos de Daniela

Tiempo (s)	$\boldsymbol{x}$	**64**	**58**	**68**	**62**	**65**
Desviación con respecto a la media	$\boldsymbol{x - \bar{x}}$	$\mathbf{0,6}$	$\mathbf{-5,4}$	$\mathbf{4,6}$	$\mathbf{-1,4}$	$\mathbf{1,6}$

La **desviación** puede ser calculada con respecto a cualquier valor, no solo con respecto al promedio.

Desarrollo	
$\bar{x} = \dfrac{x_1 + x_2 \ldots + x_n}{n}$	$64 - 63{,}4 = \mathbf{0,6}$ $58 - 63{,}4 = \mathbf{-5,4}$ $68 - 63{,}4 = \mathbf{4,6}$ $62 - 63{,}4 = \mathbf{-1,4}$ $65 - 63{,}4 = \mathbf{1,6}$
$\bar{x} = \dfrac{64 + 58 + 68 + 62 + 64}{5} = \dfrac{317}{5} = \mathbf{63,4}$	

Paso 2:

Calcula la suma de las desviaciones medias:

$$0{,}6 + (-5{,}4) + 4{,}6 + (-1{,}4) + 1{,}6 = 0$$

Actividad 1:

Calcula las desviaciones de los tiempos de Bárbara, guíate por el ejemplo

Tiempo (s) de Bárbara	$\boldsymbol{x}$	**69**	**63**	**65**	**50**	**70**
Desviación con respecto a la media	$\boldsymbol{x - \bar{x}}$					

Desarrollo	
$\bar{x} = \dfrac{x_1 + x_2 \ldots + x_n}{n}$	

- La ***Desviación media*** $(\boldsymbol{D_{\bar{x}}})$ corresponde a la media aritmética de los valores absolutos de las desviaciones $(x_i - \bar{x})$ de los n datos, esto es:

Para **datos no agrupados** se tiene:

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \ldots + |x_n - \bar{x}|}{n}$$

Para **datos agrupados** se tiene:

$$D_{\bar{x}} = \frac{|x_{mc1} - \bar{x}| \cdot f_1 + |x_{mc2} - \bar{x}| \cdot f_2 + |x_{mc3} - \bar{x}| \cdot f_3 + \ldots + |x_{mcn} - \bar{x}| \cdot f_n}{n}$$

Donde

- x_{mci} es la marca de la clase del intervalo i,
- $\bar{x}$ es la media aritmética de la variable
- f_i es la frecuencia absoluta del intervalo i
- n es el número total de datos

Paso 3:

Calcula la **desviación media** de la siguiente manera

$$= \frac{|64 - 63{,}4| + |58 - 63{,}4| + |68 - 63{,}4| + |62 - 63{,}4| + |65 - 63{,}4|}{5}$$

$$= \frac{0{,}6 + 5{,}4 + 4{,}6 + 1{,}4 + 1{,}6}{5}$$

$$= \frac{13{,}6}{5} = \mathbf{2{,}72\ s}$$

La **Desviación media** permite determinar en cuanto varían, en promedio, los datos de una distribución con respecto a la media aritmética

Actividad 2:

a. Calcula las desviaciones media en los tiempos obtenidos por Bárbara, guíate por el ejemplo
b. ¿Qué resultado se obtiene al sumar las desviaciones de Bárbara?, ¿es el mismo que en el caso de Daniela? ¿Qué puedes concluir al respecto?
c. Según los resultados de las actividades 1 y 2, ¿qué datos son más dispersos: Los de Daniela o los de Bárbara?, ¿por qué?

Varianza y desviación estándar

La ***Varianza*** y la ***Desviación estándar*** permite cuantificar la dispersión dada por la desviación media.

- La **Varianza** ($\boldsymbol{\sigma^2}$) corresponde a la media aritmética de los cuadrados de las desviaciones de los n datos. Se expresa en unidades cuadradas.

Para datos NO agrupados se tiene:

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$$

Para datos agrupados se tiene:

$$\sigma^2 = \frac{(x_{mc1} - \bar{x})^2 \cdot f_1 + (x_{mc2} - \bar{x})^2 \cdot f_2 + (x_{mc3} - \bar{x})^2 \cdot f_3 + \ldots + (x_{mcn} - \bar{x})^2 \cdot f_n}{n}$$

Dónde:

- ✓ x_{mci} marca de clase del intervalo i
- ✓ $\bar{x}$ es la media aritmética de la variable
- ✓ f_i es la frecuencia absoluta del intervalo i
- ✓ n es el número total de datos

- La ***Desviación estándar*** (σ) 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 lo disperso que es el conjunto.

Del Ejemplo:

El entrenador continúa su análisis para tomar una adecuada decisión. Para ello, sigue estos pasos para calcular la ***Varianza y la Desviación Estándar*** de los tiempos de Daniela

Paso 1 Varianza

$$\sigma^2 = \frac{(64 - 63{,}4)^2 + (58 - 63{,}4)^2 + (68 - 63{,}4)^2 + (62 - 63{,}4)^2 + (65 - 63{,}4)^2}{5}$$

$$\sigma^2 = \frac{(0{,}6)^2 + (-5{,}4)^2 + (4{,}6)^2 + (-1{,}4)^2 + (1{,}6)^2}{5}$$

$$\sigma^2 = \frac{0{,}36 + 29{,}16 + 21{,}16 + 1{,}96 + 2{,}56}{5}$$

$$\sigma^2 = \frac{55{,}2}{5} = \mathbf{11{,}04\ s^2}$$

Paso 2 Deviación estándar

$$\sigma = \sqrt{11{,}04} \approx \mathbf{3{,}32\ s}$$

Actividad Nº2

a. Calcula la varianza de los tiempos de Bárbara.
b. Calcula la desviación estándar de los tempos de Bárbara.
c. Compara la dispersión entre los datos de Daniela y los de Bárbara. ¿Dónde es mayor dispersión?

> ***Nota***
> A ***mayor dispersión***, mayor valor de la varianza;
> a ***menor dispersión***, menor valor de la varianza

d. Finalmente, con toda la información obtenida acerca de los tiempos de ambas nadadoras, responde:

> ¿Qué decisión debe tomar el entrenador?,¿Quién debería participar en la próxima competencia: Daniela o Bárbara?

Guía Nº2: Medidas de Dispersión

Actividades

Nombre: _____ Curso_____ Fecha: _____

1. Las temperaturas (en grados Celsius) durante dos semanas en Talca fueron las siguientes:

Temperatura semana 2 (ºC)	30	31	30	25	21	20	22

Temperatura semana 2 (ºC)	30	29	29	27	26	20	27

a. Calcula e interpreta las medias de dispersión
b. ¿Qué ocurriría con la dispersión de los datos si las temperaturas se toman en distintas estaciones del año? Justifica.

Economía

2. La cantidad de Cheques cobrados diariamente en todas las sucursales de un banco el mes anterior se registran en la siguiente tabla:

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

Una desviación estándar superior a 200 cheques diarios ocasionara problemas de organización o logística en las sucursales

¿Deberá preocuparse el jefe de operaciones del banco por la cantidad de empleados que se necesitara el mes siguiente?, ¿Qué decidirá?

3. La chef de un restaurante acaba de recibir un encargo de barras de chocolate de su proveedor, pero aun no los acepta. Los gramos de cada barra se muestran en el recuadro

178,60	204,12	206,95
221,13	192,78	209,79
226,80	209,79	215,46
229,63	215,46	218,30

Aceptare las barras si la masa promedio es 212,62 g y la desviación estándar es menor que 14,18 g.

¿Qué decisión tomara la Chef?, ¿Por qué? Desarrolla y argumenta tu respuesta

4. ***Desarrolla en tus cuadernos las páginas 1 y 2 del Cuaderno de Actividades. Actividades 1, 2 y 3***