COLEGIO SAN CARLOS DE QUILICURA
MATEMÁTICA/TERCERO MEDIO
J. VELÁSQUEZ/L. CONTRERAS/2020
SESIÓN 2 (45 minutos)

COLEGIO
SAN CARLOS
QUILICURA

Guía de Trabajo N° 25 Matemática

(Del 13 al 16 de octubre)

Nombre	Curso	Fecha
	IIIº____	___ / 10/ 2020

OA 2: Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión y probabilidades condicionales.

CONTENIDOS QUE SE TRABAJARÁN EN ESTA GUÍA

Unidad I

- Medidas de dispersión

INSTRUCCIONES

- El tiempo estimado para el desarrollo de la guía será de 90 minutos. Puedes realizarla en dos sesiones de 45 minutos.
- Los materiales que necesitaras para el desarrollo de la guía serán: cuaderno de la asignatura, lápiz mina, lápiz pasta, goma, calculadora, saca puntas y una regla.
- El desarrollo de los ejercicios escríbelo con lápiz mina y la respuesta final escríbela con lápiz pasta.
- En la Guía de Trabajo N° 26 se anexará la retroalimentación de esta guía.

¡Hola! Un gusto saludarte de nuevo, deseando que te encuentres muy bien junto a tus familiares y seres queridos.

En esta guía retomaremos el estudio de las **"Medidas de dispersión",** tema que fue trabajado anteriormente en las **Guías de Trabajo N° 2, N° 3, N° 4, N° 6 y N° 10.** Comenzaremos por estudiar el concepto de medidas de dispersión, el rango, la desviación media, varianza y desviación estándar, algunos ejemplos y 4 ejercicios propuestos para que practiques.

¡ÁNIMO Y MUCHOS ÉXITOS!

Medidas de dispersión

Llegado el verano, Nicolás decide tomarse unas vacaciones. Dos son los destinos que baraja, los cuales denotaremos como A y B. Dado que a Nicolás no le gusta mucho el sol, decide optar por el destino cuyas temperaturas máximas sean en promedio más bajas. Para ello observa el registro de temperaturas máximas del último mes para ambos destinos y calcula su promedio (figura 1). Para su sorpresa, tanto A como B promedian la misma temperatura máxima, lo cual lo deja muy contento pues podrá elegir libremente su destino.

Destino A				$\bar{x}_A = 21°C$		
Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
		12°C	19°C	10°C	36°C	15°C
30°C	27°C	8°C	28°C	24°C	9°C	33°C
23°C	30°C	27°C	10°C	27°C	18°C	24°C
6°C	35°C	17°C	21°C	8°C	29°C	20°C
11°C	22°C	36°C	15°C			

Destino B				$\bar{x}_B = 21°C$		
Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
		22°C	25°C	23°C	18°C	23°C
21°C	24°C	22°C	20°C	17°C	19°C	25°C
24°C	18°C	24°C	16°C	21°C	22°C	22°C
20°C	21°C	20°C	19°C	22°C	21°C	17°C
17°C	23°C	21°C	23°C			

Figura 1. Temperaturas máximas registradas para los destinos A y B.

¿Es correcto el procedimiento que Nicolás adopta para elegir entre ambos destinos? De no serlo, ¿qué otro procedimiento debió considerar en su decisión?

Por un lado, el procedimiento realizado por Nicolás es correcto, ya que sabemos que el promedio (o media aritmética) de un conjunto de datos es aquel valor que los "representa", sin embargo, no debe dejar pasar que el promedio es "sensible" a valores extremos. En el destino A por ejemplo, las temperaturas máximas registradas varían enormemente: desde los 6◦C hasta los 36◦C, contrario a las temperaturas máximas registradas en el destino B, las cuales se concentran próximas a los 21◦C.

Sigue que, si Nicolás elige el destino A, lo más probable es que un día sea muy poco soleado y al siguiente haya un calor insoportable, mientras que si elige el destino B se asegura que las temperaturas máximas durante su estancia serán similares. En este sentido, se espera que las temperaturas máximas sean más homogéneas (similares) para el destino B que para el A.

En consecuencia, para comparar dos o más conjuntos de datos debemos considerar no solo sus medidas de tendencia central (en particular el promedio) sino que también necesitamos de algún instrumento matemático que nos indique acerca de su dispersión, es decir, qué tanto varían los datos del conjunto, tanto entre sí como con respecto a su media. Dichas herramientas corresponden a las medidas de dispersión, entre las que se estudiará el rango, la desviación media, la varianza y la desviación estándar.

El promedio (media aritmética) es un indicador sensible a los valores extremos. En otras palabras, la media puede resultar no representativa en presencia de datos cuyos valores sean mucho mayores o mucho menores que la mayoría de ellos.

1.1. RANGO

El rango se define como la diferencia entre el mayor y el menor de los datos de un conjunto. Se designa con la letra R.

El rango nos informa la "amplitud" que abarca el conjunto en base a sus valores extremos, pero ignora todo aquello que pueda ocurrir dentro de los mismos, razón por la cual no resulta un indicador de dispersión suficiente.

Siguiendo el ejemplo inicial, el rango de las temperaturas máximas para los destinos A y B son:

$$\mathrm{R_A} : 36°\mathrm{C} - 6°\mathrm{C} = 30°\mathrm{C}$$

$$\mathrm{R_B} : 25°\mathrm{C} - 16°\mathrm{C} = 9°\mathrm{C}$$

Lo anterior nos indica que la variabilidad de temperaturas máximas es mayor para el destino A. Sin embargo, dado que solo se consideran dos datos del total (el mayor y el menor), no tenemos certeza de lo que ocurre con el resto de ellos (evidentemente, suponiendo que no tenemos el registro de la figura 1). Para averiguarlo se define la desviación media.

1.2. DESVIACIÓN MEDIA

La desviación media de un conjunto de datos se define como el promedio de las distancias de cada uno de los datos a su media aritmética. Se representa por Dm y matemáticamente se calcula mediante:

$$\mathrm{D_m} = \frac{\sum_{i=1}^{\mathrm{N}} |x_i - \overline{x}|}{\mathrm{N}}$$

Donde $\overline{x}$ corresponde al promedio de los N datos $x_1, x_2, \ldots, x_{\mathrm{N}}$ que conforman el conjunto.

Advierta de la expresión anterior que el valor absoluto corresponde a la distancia entre el dato i − ésimo del conjunto y la media aritmética del mismo. Dicha distancia se calcula para cada uno de los N datos del conjunto y luego se promedian. El resultado final representa la distancia promedio que existe entre cada uno de los datos del conjunto y su media aritmética.

Para dejar aún más claro de qué se trata la desviación media, la figura 2 ilustra las temperaturas del ejemplo inicial para ambos destinos, A y B. Sobre el mismo se destaca una recta horizontal que representa la temperatura máxima promedio respectiva.

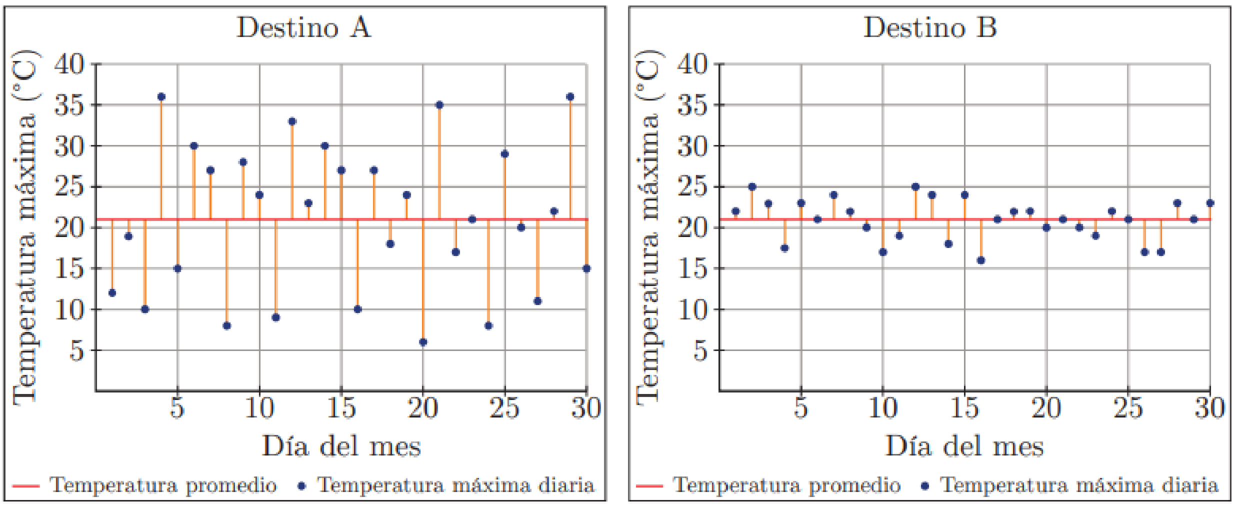

Figura 2. Representación gráfica de las distancias de cada una de las temperaturas máximas registradas y su temperatura promedio para los destinos A y B.

Las rectas verticales corresponden (literalmente) a la distancia que existe entre cada uno de los datos del conjunto y su respectiva media aritmética. Así, lo que la desviación media hace es obtener el promedio de las rectas verticales, vale decir, de las distancias de cada uno de los datos del conjunto a su media aritmética.

Como consecuencia y en términos generales, a menor valor de la desviación media, más próximos se encuentran los datos del conjunto respecto a su media aritmética (son más homogéneos). Al contrario, a mayor valor de la desviación media, más lejanos se encuentran los datos del conjunto respecto a su media aritmética (son más dispersos).

Con lo anterior y volviendo al ejemplo original, la desviación media de las temperaturas para el destino A es aproximadamente 7,7◦C, mientras que para el destino B es 2◦C. El cálculo numérico no se detalla pues es extenso, pero sencillo.

El resultado anterior nos anuncia lo que para nosotros es evidente al observar la figura 2: la fluctuación (dispersión) de las temperaturas con respecto a su media aritmética es mayor para el destino A que para el destino B. En otras palabras, las temperaturas del destino B son más estables (u homogéneas).

1.3. VARIANZA

Otra forma de cuantificar y poder así comparar la dispersión entre dos o más conjuntos de datos es mediante la varianza, la cual se define como el promedio de las distancias al cuadrado entre los datos y su media aritmética respectiva. Se representa por σ^2 o por s^2 y matemáticamente se calcula mediante:

$$\boxed{\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \overline{x})^2}{N}}$$

Alternativamente, se define la expresión equivalente:

$$\boxed{\sigma^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} - \overline{x}^2}$$

Donde $\overline{x}$ corresponde al promedio de los N datos $x_1, x_2, \ldots, x_N$ que conforman el conjunto.

Conceptualmente, la varianza es exactamente igual a la desviación media. La diferencia radica en la precisión de la escala de medición que se ocupa. Por lo tanto, a mayor varianza, mayor dispersión de los datos. Al contrario, a menor varianza, más homogéneos son.

A diferencia del rango o de la desviación media, la varianza se expresa en unidades cuadradas: la varianza del problema inicial se expresa en grados Celsius cuadrados ($°C^2$). Puesto que esta unidad de medida resulta poco usual (y por tanto poco comprensible), se define la desviación estándar (o desviación típica) como la raíz cuadrada de la varianza, de modo que la unidad de medida sea la misma que la de los datos del conjunto.

1.4. DESVIACIÓN ESTÁNDAR (O DESVIACIÓN TÍPICA)

Como se enuncio previamente, se define la desviación estándar (o desviación típica) de un conjunto de datos como la raíz cuadrada de la varianza. Como consecuencia, su unidad de medida es la misma que la de los datos del conjunto. Se representa por σ o s y matemáticamente se calcula mediante:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{\mathrm{N}} (x_i - \overline{x})^2}{\mathrm{N}}}$$

Alternativamente, se define la expresión equivalente:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{\mathrm{N}} x_i^2}{\mathrm{N}} - \overline{x}^2}$$

Donde $\overline{x}$ corresponde al promedio de los N datos $x_1, x_2, \ldots, x_{\mathrm{N}}$ que conforman el conjunto.

Conceptualmente, la varianza y la desviación estándar son equivalentes, con la diferencia de sus unidades de medida. Por lo tanto, a mayor desviación estándar, mayor dispersión de los datos. Al contrario, a menor desviación estándar, más homogéneos son.

Tres propiedades importantes referidas a la varianza y la desviación estándar se desprenden de su definición:

- *La varianza (y por tanto la desviación estándar) toma valores mayores o iguales a cero.*
- *Si a cada uno de los elementos de un conjunto se les suma un número, la varianza no se altera (por lo tanto la desviación estándar tampoco).*
- *Si cada uno de los elementos de un conjunto se multiplica por un número, la varianza queda multiplicada por el cuadrado de dicho número, mientras que la desviación estándar queda multiplicada por él.*

Retomando el ejemplo original, la varianza y la desviación estándar de A y B son los mostrados a continuación. Su desarrollo se propone como ejercicio.

$$\sigma_{\mathrm{A}}^2 \approx 81°\mathrm{C}^2 \qquad \sigma_{\mathrm{A}} \approx 9°\mathrm{C}$$

$$\sigma_{\mathrm{B}}^2 \approx 6°\mathrm{C}^2 \qquad \sigma_{\mathrm{B}} \approx 2,5°\mathrm{C}$$

Nuevamente los cálculos avalan los resultados previos: las temperaturas son más homogéneas (y en este sentido más estables) para el destino B que para el destino A.

Para concluir, Nicolás no debió alegrarse de que el promedio de temperaturas de ambos destinos sean los mismos, pues la media aritmética al ser sensible a valores extremos, por sí sola no es un buen representante del conjunto. Ahora, dado que la media de ambos destinos es la misma, Nicolás debe optar por aquel destino cuyas temperaturas sean lo más estables posible (de lo contrario se esperaría cualquier cosa, incluido días muy calurosos). En este sentido, se debe apoyar en las medidas de dispersión, las cuales nos entregan información valiosa acerca de, como su nombre lo indica, la dispersión de los datos del conjunto.

✎ ***Ejemplo***

1. **Los conjuntos S1 y S2 representan la variación del precio del dólar para dos semanas consecutivas y cuyos promedios son \$539 y \$565 respectivamente. Al respecto, ¿en cuál conjunto el precio del dólar es más estable?**

$$\mathrm{S}_1 = \{\$572,\ \$523,\ \$521,\ \$532,\ \$522,\ \$567,\ \$536\}$$
$$\mathrm{S}_2 = \{\$550,\ \$600,\ \$581,\ \$585,\ \$530,\ \$570,\ \$539\}$$

SOLUCIÓN: La pregunta del enunciado hace referencia a la estabilidad del precio del dólar. Por estabilidad se entiende cuando su precio varía la menor cantidad posible, lo que en otras palabras es equivalente a decir que su dispersión debe ser lo más cercana a cero que se pueda.

Cualquiera de los indicadores descritos previamente (con excepción del rango) podría ser bueno para el cálculo de la dispersión, sin embargo, vamos a resolver el problema utilizando la desviación estándar, tanto porque no presenta dificultades con las unidades de medida, como también por su precisión. Luego, calculemos la desviación estándar para S1:

$$\begin{aligned}
&\sqrt{\frac{(\$572-\$539)^2+(\$523-\$539)^2+(\$521-\$539)^2+(\$532-\$539)^2+(\$522-\$539)^2+(\$567-\$539)^2+(\$536-\$539)^2}{7}}\\
&=\sqrt{\frac{(\$33)^2+(-\$16)^2+(-\$18)^2+(-\$7)^2+(-\$17)^2+(\$28)^2+(-\$3)^2}{7}}\\
&=\sqrt{\frac{\$1.089+\$256+\$324+\$49+\$289+\$784+\$9}{7}}\\
&=\sqrt{\frac{\$2.800}{7}}\\
&=\sqrt{\$400}\\
&=\$20
\end{aligned}$$

Análogamente, la desviación estándar para S2 es:

$$\begin{aligned}
&\sqrt{\frac{(\$550-\$565)^2+(\$600-\$565)^2+(\$581-\$565)^2+(\$585-\$565)^2+(\$530-\$565)^2+(\$570-\$565)^2+(\$539-\$565)^2}{7}}\\
&=\sqrt{\frac{(-\$15)^2+(\$35)^2+(\$16)^2+(\$20)^2+(-\$35)^2+(\$5)^2+(-\$26)^2}{7}}\\
&=\sqrt{\frac{\$225+\$1.225+\$256+\$400+\$1.225+\$25+\$676}{7}}\\
&=\sqrt{\frac{\$4.032}{7}}\\
&=\sqrt{\$576}\\
&=\$24
\end{aligned}$$

Con lo anterior, la desviación estándar es menor para S1, lo que indica que los valores que lo componen se encuentran en promedio más cerca de su media aritmética que en el caso de S2. Como consecuencia, el precio registrado del dólar es más estable para S1.

Cabe destacar también que el rango de S1 y S2 es 51 y 70 respectivamente. Es de esperarse entonces que, dados dos conjuntos con una misma cantidad de elementos, sea más homogéneo aquel cuyo rango sea menor, pues sus datos se concentran dentro de un intervalo más pequeño, confirmando de esta forma el resultado previo. No obstante, esta afirmación no siempre es cierta, pues recuerde que el rango solo considera los valores extremos del conjunto, de modo que no es un buen indicador de dispersión y puede inducir a errores.

2. **La estatura promedio de los estudiantes de dos cuartos medios de un colegio, en conjunto con sus respectivas desviaciones estándar, se muestran en la tabla 1. Al respecto, ¿en qué curso es más acertado afirmar que la estatura promedio de los estudiantes es 1,61 m?**

Tabla 1. Estadística de las estaturas de los estudiantes de los 4° medios A y B.

	Estatura Promedio (m)	Desviación Estándar (m)
4° medio A	1,61	0,11
4° medio B	1,61	0,34

SOLUCIÓN: Si bien en ambos cursos la estatura media es la misma, la desviación estándar del 4◦ medio B es más del triple que la del 4◦ medio A. Por otro lado, recuerde que la desviación estándar nos entrega información acerca de la dispersión de los datos de un conjunto (qué tan alejados se encuentran éstos en relación a su valor central). Luego, se espera que las estaturas de los estudiantes del 4◦ medio A sean más próximas a 1,61 m que los estudiantes del 4◦ medio B (porque $0,11 < 0,34$). Por lo tanto, es más acertado afirmar que los estudiantes del 4◦ medio A tienen una estatura promedio igual a 1,61 m.

✍ *Ejercicios*

1. Calcule el rango, desviación media, varianza y desviación estándar de un conjunto conformado por cinco enteros pares consecutivos mayores que 0.

2. Posterior al noticiero, la periodista que informa el tiempo dijo “hoy tuvimos un agradable día con temperaturas entre los 15 °C y los 25 °C". Al respecto, ¿cuál es el rango de temperaturas registrado en ese día?

3. Determine el rango de un conjunto de datos cuya desviación media es nula.

4. Se aplica una misma prueba de álgebra a dos segundos medios, ambos con la misma cantidad de estudiantes. Los resultados obtenidos para cada curso se ilustran en la figura 3. Al respecto, ¿en cuál curso la dispersión de las calificaciones obtenidas es menor?

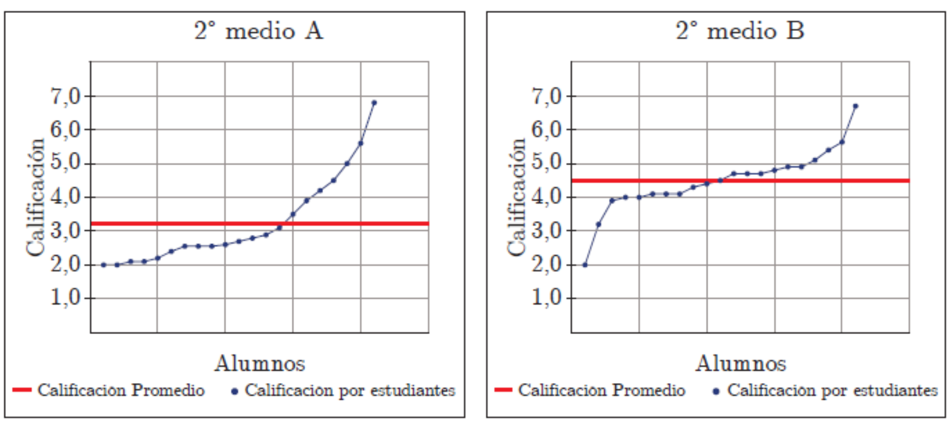

Figura 3. Estadística de las calificaciones obtenidas por los 2° medios A y B.

NUESTRA CLASE ONLINE N° 14 SE EFECTUARÁ EL PRÓXIMO MARTES 13 DE OCTUBRE PARA III°A Y III° B Y EL DÍA JUEVES 15 DE OCTUBRE PARA III° C, A TRAVÉS DE LA PLATAFORMA GOOGLE MEET, ASI QUE DEBES BUSCAR EL LINK PARA UNIRTE A LA CLASE EN TU CALENDARIO.

CURSO: III° A **Nombre del profesor:** Josimar Velásquez **Día:** Martes 13 de octubre **Hora:** 10:00 – 10:45 am	**CURSO: III° B** **Nombre del profesor:** Josimar Velásquez **Día:** Martes 13 de octubre **Hora:**11:00 am – 11:45am	**CURSO: III° C** **Nombre del profesor:** Loreto Contreras **Día:** Jueves 15 de octubre **Hora:** 4:00 pm – 4:45 pm	Meet

¡TE ESPERAMOS!
CUÍDATE MUCHO