Colegio San Carlos de Quilicura
Matemática/Cuarto Medio
LCG/CSV/2020

Guía n°24 de Matemáticas

(Del 5 al 9 de octubre)

Nombre	Curso	Fecha
	IVº	/ 10 / 2020

Los contenidos de esta actividad estarán en la prueba de admisión transitoria:
Eje temático: ESTADISTICA
OA 2 (III°Medio): Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión y probabilidades condicionales.
CONTENIDOS: • Medidas de dispersión : rango, desviación media, desviación estándar, varianza y coeficiente De variación

Estimada(o) estudiante: La guía n°24 consta de dos partes. La primera consiste en que revises los contenidos de MEDIDAS DE DISPERSIÓN , y la segunda parte consiste en revisar ejercicios resueltos.

ESTA ES UNA PARTE DE LA TABLA DE "CONTENIDOS DE LA PRUEBA DE TRANSICIÓN DE MATEMÁTICA (PTU)". AQUÍ PUEDES EVIDENCIAR EL CONTENIDO QUE ESTAMOS REFORZANDO EN ESTA GUÍA:

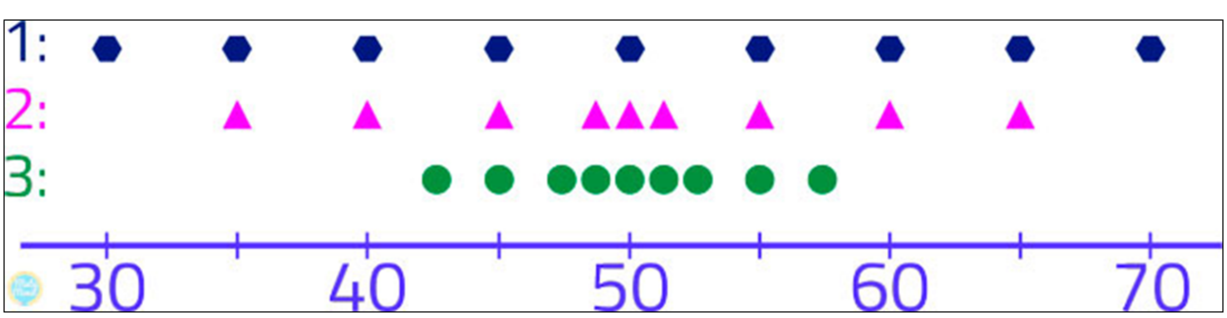
EJE TEMÁTICO	UNIDADES TEMÁTICAS	DESCRIPCIÓN
	Representación de datos a través de tablas y gráficos	○ Tablas de frecuencia absoluta y relativa. ○ Tipos de gráficos que permitan representar datos. ○ Problemas que involucren tablas y gráficos en diversos contextos.
	Medidas de tendencia central y rango	○ Medidas de tendencia central y rango de uno o más grupos de datos. ○ Problemas que involucren medidas de tendencia central y rango en diversos contextos.
PROBABILIDAD Y ESTADÍSTICA	Medidas de posición	○ Cuartiles y percentiles de uno o más grupos de datos. ○ Diagrama de cajón para representar distribución de datos. ○ Problemas que involucren medidas de posición en diversos contextos.

Parte I: Contenido: Medidas de Dispersión

¿QUÉ SON LAS MEDIDAS DE DISPERSIÓN?

Las medidas de dispersión son medidas estadísticas que muestran la variabilidad en la distribución de los datos. Las medidas de tendencia central, como la media, la mediana y la moda, solo describen el centro de los datos, pero no nos dicen nada acerca de la dispersión (separación) de los datos. Y en ocasiones, es muy importante conocer que tan dispersos o separados se encuentran los datos, y esto se consigue con las medidas de dispersión o variabilidad.

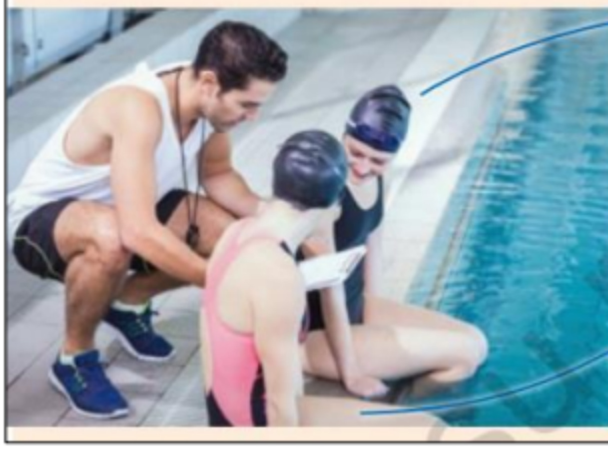
Las medidas de dispersión dan una idea ideal de "alejamiento" de los datos respecto a las medidas de centralización (mediana y media). En otras palabras, sirven para determinar si los datos se encuentran en torno a la media o si están muy dispersos.



Con respecto a la gráfica anterior podemos concluir que los datos de color azul están muy dispersos y los datos de color verde son menos dispersos por que están más cercanos a la media aritmética y a la mediana.

Por ejemplo, observemos 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, en segundos, de las dos postulantes en las 5 últimas carreras en este estilo.



Competencias de Daniela	
N.° de carrera	Tiempo (s)
1	64
2	58
3	68
4	62
5	65

Competencias de Bárbara	
N.° 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 carreras de 100 m estilo libre?, ¿y el de Bárbara?

Recordemos el concepto de promedio o media aritmética que trabajamos en la Guía N°1, donde para calcularlo, primero se suman todos los datos y luego se divide por la cantidad de datos:

$$\bar{x} = \frac{64+58+68+62+65}{5} = \frac{317}{5} = 63{,}4\ segundos\ (Daniela)$$

$$\bar{x} = \frac{69+63+65+50+70}{5} = \frac{317}{5} = 63{,}4\ segundos\ (Bárbara)$$

Por tanto, el tiempo promedio de Daniela en las últimas 5 carreras de 100 m estilo libre es de 63,4 segundos y el de Bárbara también es de 63,4 segundos.

b) ¿Cómo son los promedios de Daniela y Bárbara?

El promedio de los tiempos de Daniela y Bárbara son iguales.

c) ¿A quién debiera elegir el entrenador para participar en la competencia?, ¿por qué?

Respuesta variable. Pero un ejemplo seria, que el entrenador debiera elegir a Daniela ya que sus tiempos son menos dispersos en relación a los de Bárbara.

¿CUÁLES SON LAS MEDIDAS DE DISPERSIÓN MÁS CONOCIDAS?

- El rango
- La desviación media
- La varianza.
- Desviación estándar.

¿QUÉ ES EL RANGO Y CÓMO SE CALCULA?

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

$\boldsymbol{R = Xmax - Xmin}$

Por ejemplo: en el caso anterior, si se denotan por R1 y R2 los rangos de los tiempos de Daniela y Bárbara respectivamente, se tiene:

$$R_1 = X_{máx} - X_{mín} = 68 - 58 = 10 \rightarrow R_1 = 10\text{ s}$$
$$R_2 = X_{máx} - X_{mín} = 70 - 50 = 20 \rightarrow R_2 = 20\text{ s}$$

Esto da indicios de que los tiempos de Daniela pueden ser menos dispersos que los de Bárbara. Sin embargo, no es posible concluir de inmediato: debemos disponer de más información.

¿QUÉ ES LA DESVIACIÓN MEDIA Y CÓMO SE CALCULA?

La desviación media ($\boldsymbol{D\bar{x}}$) es un dato de dispersión que sirve para comparar los datos en relación con el promedio y tomar decisiones. si los datos están más lejos del promedio entonces se puede decir que los datos son más dispersos. La fórmula de la desviación media es la siguiente:

$$D.M. = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \cdots + |x_n - \bar{x}|}{n} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Donde:

- $\bar{x}$: media aritmética de los datos.
- $x_1, x_2, x_3, \ldots, x_n$: datos.
- x_i: cada uno de los datos.
- n: número de datos.

Recuerda calcular la media aritmética $\bar{x}$ antes de aplicar la fórmula de la desviación media.

Por 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 (Recordemos que su promedio era $\bar{x} = 63{,}4$)

Tiempo (s)	x	64	58	68	62	65
Desviación con respecto a la media.	$x - \bar{x}$	$x - \bar{x} =$ $64 - 63{,}4$ $= \mathbf{0{,}6}$	$x - \bar{x} =$ $58 - 63{,}4$ $= \mathbf{-5{,}4}$	$x - \bar{x} =$ $68 - 63{,}4$ $= \mathbf{4{,}6}$	$x - \bar{x} =$ $62 - 63{,}4$ $= \mathbf{-1{,}4}$	$x - \bar{x} =$ $65 - 63{,}4$ $= \mathbf{1{,}6}$

PASO 2: Calcula la suma de las desviaciones medias:

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

PASO 3: Calcula la desviación media de la siguiente manera:

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

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

$$D_{\bar{x}} = \frac{13{,}6}{5} = \mathbf{2{,}72\ segundos}$$

La desviación media permite determinar en cuánto varían, en promedio, los datos de una distribución con respecto a la media aritmética.

Ahora respondemos:

1. **¿Cuáles son las desviaciones con respecto a la media aritmética en los tiempos obtenidos por Bárbara?**

Tiempo (s)	x	69	63	65	50	70
Desviación con respecto a la media.	$x-\bar{x}$	$x-\bar{x}=$ $69-63{,}4$ $=\mathbf{5{,}6}$	$x-\bar{x}=$ $63-63{,}4$ $=\mathbf{-0{,}4}$	$x-\bar{x}=$ $65-63{,}4$ $=\mathbf{1{,}6}$	$x-\bar{x}=$ $50-63{,}4$ $=\mathbf{-13{,}4}$	$x-\bar{x}=$ $70-63{,}4$ $=\mathbf{6{,}6}$

2. **¿Qué resultado se obtiene al sumar las desviaciones de Bárbara?, ¿es el mismo que en el caso de Daniela? ¿Qué puedes concluir al respecto?**

Se obtiene cero. Sí. La suma de las desviaciones respecto a la media es siempre cero.

3. **Calcula la desviación media de los tiempos de Bárbara.**

$$D_{\bar{x}}=\frac{|69-63{,}4|+|63-63{,}4|+|65-63{,}4|+|50-63{,}4|+|70-63{,}4|}{5}=\frac{5{,}6+0{,}4+1{,}6+13{,}4+6{,}6}{5}$$

$$D_{\bar{x}}=\frac{27{,}6}{5}=5{,}52\ segundos$$

4. **Según los resultados de las actividades 2 y 3, ¿qué datos son más dispersos: los de Daniela o los de Bárbara?, ¿por qué?**

Los datos de Bárbara son más dispersos, porque su desviación media es mayor.

¿QUÉ ES LA VARIANZA? ¿Y LA DESVIACIÓN ESTANDAR?

La **varianza** y la **desviación estándar** permiten cuantificar la dispersión dada por la desviación media.

- La varianza (σ^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:

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

Donde x_{mci} es la marca de clase del intervalo i, $\bar{x}$ es la media aritmética de la variable, f_i es la frecuencia absoluta del intervalo i y 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.

Ejemplo:

El entrenador continúa su análisis para tomar una adecuada decisión. Para ello, sigue estos pasos:

PASO 1: Calcula la media de los cuadrados de las diferencias entre cada tiempo de Daniela y el promedio. Obtiene así la varianza (σ^2)

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

$$\sigma^2=\frac{0{,}36+29{,}16+21{,}16+1{,}96+2{,}56}{5}=\frac{55{,}2}{5}=11{,}04\ s^2$$

PASO 2: Calcula la raíz cuadrada del valor anterior y obtiene la desviación estándar (σ):

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

Ahora respondemos,

1. **¿Puede ser negativo el valor de la varianza?, ¿por qué?**

No, ya que es la suma de números positivos.

2. **Calcula la varianza de los tiempos de Bárbara.**

$$\sigma^2 = \frac{(69-63{,}4)^2 + (63-63{,}4)^2 + (65-63{,}4)^2 + (50-63{,}4)^2 + (70-63{,}4)^2}{5}$$

$$\sigma^2 = \frac{31{,}36 + 0{,}16 + 2{,}56 + 179{,}56 + 43{,}56}{5} = \frac{257{,}2}{5} = 51{,}44\ s^2$$

3. **Calcula la desviación estándar de los tiempos de Bárbara.**

$$\sigma = \sqrt{51{,}44} \approx 7{,}17\ segundos$$

4. **Compara la dispersión entre los datos de Daniela y los de Bárbara. ¿Dónde es mayor la dispersión?**

En los tiempos de Bárbara es mayor.

5. **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?**

Daniela debería participar.

Ejercicios resueltos

1. **Matías llega todos los días al paradero fuera de su casa a las 07:30. Decide realizar un reclamo pues la micro demora mucho en pasar, y para llevarlo a cabo con fundamentos, mide los tiempos de espera durante 1 semana. Los tiempos de espera, en minutos, fueron: 5; 12; 27; 1; 6; 32; 15. Debido a la demora, no ha podido asistir a sus clases de estadística, por lo que no sabe determinar algunos estimadores. ¿Cuál es el rango y la desviación estándar de esta muestra, respectivamente?**

A) 31 minutos ; 135.333 minutos
B) 31 minutos ; 11.63 minutos
C) 31 minutos ; 812 minutos
D) 10 minutos ; 135.333 minutos
E) 10 minutos ; 11.63 minutos

SOLUCIÓN

El rango consiste en la diferencia entre el mayor y el menor valor de la muestra. En este caso, corresponde a:

Rango = 29-1 = 31minutos.

Para determinar la desviación estándar, primero se determina la media:

$$\dot{X} = \frac{5+12+27+1+6+32+15}{7} = \frac{98}{7} = 14 minutos$$

Luego, se determina la varianza:

$$Var = \frac{1}{N-1}\sum_i (X_i - \dot{X})^2$$

$$\frac{(5-14)^2 + (12-14)^2 + (27-14)^2 + (1-14)^2 + (6-14)^2 + (32-14)^2 + (15-14)^2}{6}$$

$$Var = \frac{(-9)^2 + (-2)^2 + 13^2 + (-13)^2 + (-8)^2 + 18^2 + 1^2}{6}$$

$$Var = \frac{81 + 4 + 169 + 169 + 64 + 324 + 1}{6} = \frac{812}{6} = 135,33$$

Finalmente, la desviación estándar corresponde a la raíz cuadrada de la varianza:

$$\sigma = \sqrt{Var} = \sqrt{135,333} = 11,63 minutos$$

Por lo tanto, la respuesta correcta es la b.

---

2. **Las notas obtenidas por un alumno son: 4,0 - 5,5 - 3,5 - 6,0 - 5,0 - 6,0**

   El valor de la desviación media es, aproximadamente:

   A) 0,833
   B) 0,9
   C) 0,75
   D) 0,5
   E) 1

SOLUCIÓN

Calculamos primero el promedio:

$$\frac{4{,}0 + 5{,}5 + 3{,}5 + 6{,}0 + 5{,}0 + 6{,}0}{6} = \frac{30}{6} = 5{,}0$$

Luego la desviación media, que es el promedio de las desviaciones, en valor absoluto:

$$\frac{1 + 0{,}5 + 1{,}5 + 1 + 0 + 1}{6} = 0{,}83$$

3. **En la tabla adjunta se muestra, en intervalos, el tiempo que los usuarios utilizaron un computador de una biblioteca durante un fin de semana.**

Tiempo en minutos	Número de usuarios
[0, 5[	45
[5, 10[	38
[10, 15[	30
[15, 20[	45
[20, 25[	36
[25, 30]	15

Según los datos de la tabla, ¿cuál de las siguientes afirmaciones es FALSA?

A) Hubo un total de 209 usuarios ese fin de semana.
B) Los intervalos modales son [0,5[y [15,20[.
C) Hubo 158 usuarios que utilizaron un computador a lo menos 20 minutos.
D) Hubo 96 usuarios que utilizaron un computador 15 o más minutos.
E) La mediana se encuentra en el intervalo [10,15[.

SOLUCIÓN

En esta pregunta se debe determinar cuál de las afirmaciones dadas en las opciones es falsa con respecto a los datos de la tabla adjunta.

Para determinar la veracidad de la afirmación en A) se debe sumar el número de usuarios pertenecientes a cada uno de los intervalos y así obtener el total de usuarios ese fin de semana. Esto es:

$$45 + 38 + 30 + 45 + 36 + 15 = 209 \text{ usuarios}$$

Luego, la opción A) es verdadera.

Para determinar la veracidad de la afirmación en B) recuerde que el intervalo modal es el intervalo de mayor frecuencia.

Se observa de la tabla que los intervalos con mayor frecuencia son los que tienen frecuencia 45, los cuales son [0,5[y [15,20[. **Por lo tanto, la afirmación en B) es verdadera.**

Los usuarios que utilizaron un computador a lo menos 20 minutos, es decir, 20 o más minutos, son los que están en los intervalos [20,25[y [25,30], o sea, 36 + 15 = 51 usuarios. **Luego, la afirmación en C) es falsa.**

Los usuarios que utilizaron un computador 15 o más minutos son los pertenecientes a los intervalos [15,20[, [20,25[y [25,30], es decir, 45+36+15 = 96 usuarios. **Por lo que, la afirmación en D) es verdadera.**

Por último, para determinar el intervalo donde se encuentra la mediana se puede agregar la columna de las frecuencias acumuladas porcentuales a la tabla hasta acumular el 50% de los datos, tal como se muestra a continuación:

Tiempo en minutos	Número de usuarios	Frecuencia acumulada porcentual
$[0, 5[$	45	$\frac{45}{209} \cdot 100 \approx 21{,}5$
$[5, 10[$	38	$\frac{45+38}{209} \cdot 100 \approx 39{,}7$
$\boxed{[10, 15[}$	30	$\frac{45+38+30}{209} \cdot 100 \approx 54{,}1$
$[15, 20[$	45	
$[20, 25[$	36	
$[25, 30]$	15	

Como en este intervalo se acumula el 54,1% de los datos y en el intervalo anterior se acumulaba el 39,7% de los datos, la mediana se encuentra en este intervalo.

Lo que implica que la afirmación en E) es verdadera. Por lo anterior, la clave es C).

5) La formación del equipo chileno en el partido disputado contra España el día 18 de junio del presente año con motivo del mundial de Brasil 2014, es el mostrado en la figura.

Al respecto, ¿cuál es el rango de la numeración de las camisetas de los jugadores?

A) 1
B) 18
C) 19
D) 20
E) 21

Solución

Dado un conjunto estadístico, el rango corresponde a la diferencia entre el máximo y el mínimo de los valores del conjunto. En este caso en particular, el conjunto está conformado por la numeración de las camisetas de los 11 jugadores en cancha. La camiseta con mayor numeración corresponde a Díaz, con la número 21, mientras que la camiseta con menor numeración corresponde a Bravo, con la número 1.
Luego, el rango de la numeración de las camisetas es:

$$21 - 1 = 20$$

Estimados alumnos, les recordamos que nuestra PRÓXIMA CLASE ONLINE SE EFECTUARÁ EL MARTES 6 DE OCTUBRE PARA IV°A Y IV° B Y EL DÍA MIERCOLES 7 DE OCTUBRE PARA IV° C, A TRAVÉS DE LA PLATAFORMA GOOGLE MEET.

CURSO: IV° A **Nombre de profesora:** Loreto Contreras **Día:** martes 6 de octubre. **Hora:** 10:00 – 10:45 am	**CURSO: IV° B** **Nombre de profesora:** Loreto Contreras **Día:** martes 6 de octubre **Hora:** 11:00 am – 11:45 am	**CURSO: IV° C** **Nombre de profesora:** Carol Soto **Día:** miércoles 7 de octubre **Hora:** 11:30 am- 12:30 pm	Meet

¡Cuídate mucho!