INSTITUCIÓN EDUCATIVA REPÚBLICA DE HONDURAS

Aprobada mediante Resolución No 033 del 21 de abril de 2003

SECUENCIA DIDÁCTICA No_3_ 2021

Generado por la contingencia del COVID 19

Título de la secuencia didáctica:	ESTADISTICA DESCRIPTIVA		
Elaborado por:	DANIEL URAZAN		
Nombre del Estudiante:			**Grado:11**
Área/Asignatura	MATEMATICAS	**Duración:** 18	

MOMENTOS Y ACTIVIDADES

EXPLORACIÓN

En la introducción al cálculo debemos tener claros los conceptos básicos como las operaciones con polimonios la factorización. Así que antes de empezar esta guía es bueno tener claridad en estos temas.

Un polinomio es una expresión algebraica. En ella intervienen varios números y letras, relacionados mediante sumas, multiplicaciones y/o potencias. Las variables se escriben con letras (como "x" o "y") porque pueden asumir distintos valores, en tanto que a los números se les llama coeficientes.

Los polinomios se puede clasificar de diferentes formas dependiendo del número de términos que lo conforman o el grado de sus variables… pero esto ya queda como repaso para cada uno de ustedes.

ESTRUCTURACIÓN

RECOLECCION DE DATOS.
Población, muestra, individuo y carácter. Las primeras definiciones necesarias para el inicio de cualquier estudio estadístico son:

- **Población**: Conjunto de todos los elementos que verifican una característica que será objeto de estudio.
- **Individuo**: Cada uno de los elementos de la población.
- **Muestra**: Cualquier subconjunto de la población. Este subconjunto es muy importante que sea *representativo* de la población.
- **Carácter o atributo**: Cada una de las propiedades que poseen los individuos de la población y que pueden ser objeto de estudio.

ORGANIZACION DE DATOS. Esto se puede realizar mediante la utilización la tabulación que no es más que la elaboración de tablas de frecuencias simples o con agrupación de datos.
Ejemplo: En un curso de 40 alumnos, se desea estudiar el comportamiento de la variable estatura, registrándose los siguientes valores:

1,52	1,64	1,54	1,64	1,73	1,55	1,56	1,57	1,58	1,58
1,59	1,53	1,60	1,60	1,61	1,61	1,65	1,63	1,79	1,63
1,62	1,60	1,64	1,54	1,65	1,62	1,66	1,76	1,70	1,69
1,71	1,72	1,72	1,55	1,73	1,73	1,75	1,67	1,78	1,63

Para realizar una tabla de frecuencias simple primero debemos realizar un conteo de todos los datos, para esto podemos hacer uso de las siguientes tablas.

Alumno	Talla	Alumno	Talla	Alumno	Talla	Alumno	Talla
1	1,52	11		21		31	
2	1,53	12		22		32	
3	1,54	13		23		33	
4	1,54	14		24		34	
5	1,55	15		25		35	
6	1,55	16		26		36	
7	1,56	17		27		37	
8	1,57	18		28		38	
9	1,58	19		29		39	
10	1,58	20		30		40	

Luego realizamos la tabla de frecuencias y se registra la frecuencia de cada valor de la variable. La frecuencia puede ser absoluta (**f**), número que indica la cantidad de veces que la variable toma un cierto valor, relativa (**fr**), cociente entre la frecuencia absoluta de cada valor de la variable y el número total de observaciones; relativa porcentual que es el porcentaje de la fr; frecuencia Acumulada la suma de la fi y la acumulada porcentual, que la suma de fr%

x (tallas)	Absoluta fi	Relativa fr = f/n	R. Porcentual (100.fr) %	Acumulada Fa	Ac. Porcentual Fa %
1,52	1	1/40 = 0,025	2,5 %	1	2,5%
1,53	1	1/40 = 0,025	2,5%	2	5%
1,54	2	2/40 = 0,05	5%	4	10%
1,55					
1,56					
1,57					
1,58					
1,59					
1,60					
1,61					
1,62					
1,63					
1,64					
1,65					
1,66					
1,67					
1,68					
1,69					
1,70					
1,71					
1,72					
1,73					
1,74					
1,75					
1,76					
1,77					
1,78					
1,79					

Agrupación de datos por intervalos de clase: intervalos iguales en los que se divide el número total de observaciones. Es conveniente utilizar los intervalos de clase cuando se tiene un gran número de datos de una variable continua**.**

¿Cómo saber cuántos intervalos considerar? ¿Cómo determinar su amplitud?

Primero debemos determinar el rango de los datos, que es la diferencia entre el mayor y el menor de los valores obtenidos.

$$\textbf{Rango} = x_{máx} - x_{mín}$$

- Calcula el rango de los datos de nuestro ejemplo.

Luego debemos establecer el número de intervalos (N) y determinar la amplitud (A) de los mismos.

A = rango / N (N tu lo eliges, pero es conveniente que no sea muy pequeño)

- Si queremos trabajar con 10 intervalos, ¿cuál es, para nuestro caso, la amplitud de cada uno de ellos? De ser necesario, podemos aproximar el valor hallado
- Siendo el primer intervalo [1,52 ; 1.55) completa la tabla con todos los restantes. Observa que el extremo izquierdo del intervalo se usa un corchete " [", lo que indica que tomamos este valor, en cambio en el derecho usamos ") " que nos indica que el intervalo es abierto, o sea, no se toma este valor. La **Marca de clase** es el promedio aritmético de los extremos del intervalo.

Tallas	Marca de clase (MC)	fi	fr	fr%	Fa	Fa%
[1,52 ; 1.55)	1,535					
[1,55 ; 1,58)	1,565					
[1,58 ; 1,61)	1,595					
Totales						

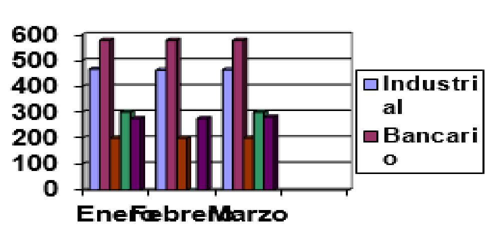
Gráficos: la recopilación de datos y la tabulación pueden traducirse gráficamente mediante representaciones convenientemente elegidas: barras, sectores circulares, mapas curvas, etc.

Los gráficos permiten visualizar e interpretar el fenómeno que se estudia, en forma más clara.

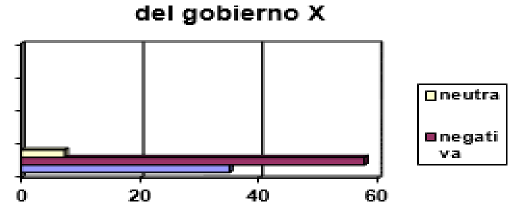
Las **barras** se utilizan generalmente para representar atributos cualitativos o cuantitativos discreto. La longitud es igual a la frecuencia de cada observación. Pueden ser barras simples o múltiples, según se trate de representar uno o más atributos.

Las barras pueden ser horizontales o verticales.

Gráfico de barras compuesto:
Remuneraciones medias (año Z)

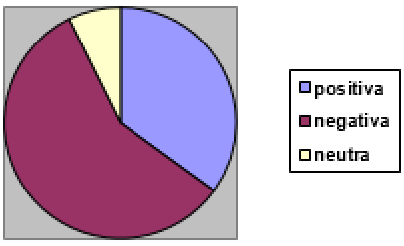


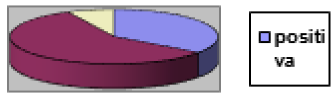
Gráf. de barras: Evaluación del gobierno X



Los **gráficos circulares** o **gráficos de torta** son útiles para comparar datos pues, en general, trabajan con porcentuales. El área de cada sector representa el porcentaje que corresponde a la frecuencia de un cierto valor de la variable. Esta representación es conveniente cuando el número de sectores es pequeño y sus áreas están bien diferenciadas.

Evaluación del gobierno X





El **histograma** se utiliza para representar una tabla de frecuencias de intervalos de clase.
Sobre el eje horizontal se representan los intervalos de clase y sobre el eje vertical, las frecuencias de los intervalos.
El gráfico consiste en un conjunto de rectángulos adyacentes cuya base representa un intervalo de clase y cuya altura representa la frecuencia del intervalo.

El **polígono de frecuencias** se construye uniendo los puntos medios de los lados opuestos de las bases de cada rectángulo. Si se quiere cerrar el rectángulo, se agregan dos intervalos: uno anterior y otro posterior al último y se prolonga el polígono hasta los puntos medios de estos intervalos.

Las **curvas** se utilizan generalmente para representar la variación de una variable a través del tiempo (años, meses, horas, etc.). Sobre el eje horizontal figuran los períodos de tiempo.

<u>Análisis y medición de datos</u>

Para describir un conjunto de datos, se calculan algunas medidas que resumen la información y que permiten realizar comparaciones.

Medidas de posición: se utilizan para encontrar un valor que represente a todos los datos. Las más importantes son: la **media aritmética**, la **moda** y la **mediana**.

- La **media aritmética** o **promedio ($\bar{x}$)** de varios números se calcula como el *cociente entre la suma de todos esos números y la cantidad de números que sumamos*.
- La **moda (Mo)** es el *valor que más se repite.* Puede suceder que haya más de una moda o ninguna (si todos los valores tienen igual frecuencia).
- La **mediana (Me)** es el *valor que ocupa el lugar central al ordenar los datos de menor a mayor*. Si la cantidad de datos es par, la mediana es el promedio entre los dos valores centrales.

- Los sueldos de cinco empleados de una empresa son: $ 400000, $500000, $450000, $600000 y $3500000. Calcula el sueldo medio, la moda, si es que existe, y la mediana e indica cuál representa mejor a los datos.

- El entrenador de un equipo de natación debe elegir a uno de sus integrantes para la próxima competencia de estilo libre. Según los tiempos en segundos que obtuvieron los postulantes de las cinco últimas carreras de 100 m de estilo libre, ¿qué nadador le conviene elegir?

Diego	61,7	61,7	62,3	62,9	63,1
Tomás	61,5	62,9	62,9	63,7	63,7
Sergio	60,7	62,4	62,7	62,7	63,2

Para poder decidir, calcula las medidas de posición de cada uno.

	promedio	moda	mediana
Diego	62,34	61,7	62,3
Tomás			
Sergio			

En promedio, los nadadores más rápidos son .... y ...., pero esto no significa que hayan tenido el mismo rendimiento; por eso necesitamos las otras medidas de posición: de ellos dos, tanto la moda como la mediana indican que .... fue más veloz. Sin embargo, para elegir el nadador adecuado, no basta con considerar las medidas de posición, ya que también es

$$\sigma_{Diego} = \sqrt{\frac{\quad}{5}} \cong \ldots\ldots\ldots\ldots$$

necesario que su rendimiento *sea parejo*, es decir, que los tiempos de sus 100 m libres *no tengan mucha dispersión.*

Medidas de dispersión: nos informan cómo están distribuidos los datos. La más importante es el **desviación estándar (σ),** que mide *la dispersión de los datos con respecto al promedio*. Cuanto menor es el desvío estándar, menos dispersos están los datos con respecto al promedio.
Para calcular el desvío estándar, seguimos los siguientes pasos:

- Calculamos la diferencia entre cada uno y el promedio.
- Elevamos al cuadrado cada una de las diferencias anteriores.
- Sumamos todos los valores hallados en el paso anterior y dividimos el resultado por la cantidad de datos. Así obtenemos la **varianza**.
- Calculamos el **desviación estándar (σ)** como la **raíz cuadrada de la varianza**.

$$\boxed{\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n}}}$$

n: número de datos

- Diego y Sergio, dos de los nadadores del ejercicio anterior, obtuvieron el mismo promedio y si embargo sus tiempos están distribuidos de manera diferente.
  Calcula los desvíos estándares de los tiempos de los nadadores:

Tiempos de Diego

x_i	$(x_i - x)$	$(x_i - x)^2$
61,7	-0,64	
61,7	-0,64	
62,3	-0,04	
62,9	0,56	
63,1	0,76	
total		

Tiempos de Sergio

x_i	$(x_i - x)$	$(x_i - x)^2$
total		

$$\sigma_{Diego} = \sqrt{\frac{\quad}{5}} \cong \ldots\ldots\ldots\ldots \qquad \sigma_{Sergio} = \sqrt{\quad} \cong \ldots\ldots\ldots\ldots$$

Entonces:

Podemos ver que el desvío estándar de .... es menor que el de ...., lo cual indica que el promedio representa mejor los datos de ...., porque sus tiempos fueron menos dispersos.
Entonces, aunque cinco datos son muy pocos para hacer estadística, si con esa información hay que elegir un nadador de ese equipo para la próxima competencia, conviene que sea ....

Si los datos están agrupados ya sea en tablas de frecuencias simples o en intervalos de clase, debemos utilizar un criterio diferente para calcular los distintos estadígrafos. Analicemos el siguiente ejemplo:

Consideremos la siguiente distribución de frecuencias que corresponden a los puntajes de 50 alumnos en una prueba.

Intervalos	M.C. (x)	fi	f·x	Fa		
[60 – 65)	62,5	5	312.5	5		
[65 – 70)	67,5	5	337.5	10		
[70 – 75)	72,5	8	580	18		
[75 – 80)	77,5	12	930	30	←	Intervalo mediano
[80 – 85)	82,5	16	1320	46	←	Intervalo modal
[85 – 90)	87,5	4	350	50		
TOTALES		**50**	**3830**			

La **Media Aritmética:** $\bar{x} = \dfrac{\sum f \cdot x}{\sum f}$ → $\bar{x} = \dfrac{3830}{50} = 76.6$ ptos. ≈ 77 ptos.

Para calcular **La Mediana** necesitamos la siguiente fórmula:

$$Me = L + \frac{\left(\frac{n}{2} - F_a\right) \cdot A}{f_i}$$

Dónde: L es el límite inferior del intervalo mediano.
F_a es la frecuencia acumulada hasta antes del Intervalo mediano.
f_i es la frecuencia absoluta del intervalo mediano.
A es la Amplitud del intervalo.

En el ejemplo, la cantidad de datos es 50, luego 50: 2 = 25, y la Fa 25 se encuentra en el intervalo [75 – 80) ya que el 25 está aquí, en cambio en la anterior (18) no está. Luego el intervalo mediano es [75 – 80)

Entonces: L = 75 (límite inferior)
$f_i = 8$
A = 5 (80 – 75 = 5)
$F_a = 18$ (frecuencia acumulada del intervalo anterior)

$$Me = 75 + \frac{\left(\frac{50}{2} - 18\right) \cdot 5}{8} = 75 + \frac{7 \cdot 5}{8} = 75 + 4.375 = 79.375 \approx 79 \text{ ptos.}$$

y finalmente, para calcular la **Moda** en datos agrupados, utilizamos la siguiente fórmula, teniendo presente que la **clase modal** es la que tiene mayor frecuencia, y esta es la Frecuencia Modal.

$$\boxed{Mo = L + \frac{d_1}{d_1 + d_2} \cdot A}$$

L: Límite real inferior de la clase modal.
d_1: es la diferencia entre la frecuencia modal y la frecuencia anterior.

L = 80 (intervalo modal [80 – 85), ya que la frecuencia es 16, que es la mayor)
d_1= 16 – 12 = 4 (diferencia con la frecuencia anterior)
d_2= 16 – 4 = 12 (diferencia con la frecuencia siguiente)
A = 5

Luego, $Mo = 80 + \dfrac{4}{4+12} \cdot 5 = 80 + \dfrac{20}{16} = 81{,}25$ puntos. ≈ 81 puntos.

Se estima que el valor más repetido de los puntajes de esta prueba fue el 81.

TRANSFERENCIA

❖ Estas son las notas obtenidas por los 100 candidatos que se presentaron a un concurso:

38	51	32	65	25	28	34	12	29	43
71	62	50	37	8	24	19	47	81	53
16	62	50	37	4	17	75	94	6	25
55	38	46	16	72	64	61	33	59	21
13	92	37	43	58	52	88	27	74	66
63	28	36	19	56	84	38	6	42	50
98	51	62	3	17	43	47	54	58	26
12	42	34	68	77	45	60	31	72	23
18	22	70	34	5	59	20	68	55	49
33	52	14	40	38	54	50	11	41	76

Presenta dichos datos en una tabla de intervalos de clase.

❖ En una cierta ciudad de la provincia de Valdivia, se registra el número de nacimientos ocurridos por semana durante las 52 semanas del año, siendo los siguientes los datos obtenidos:

6	4	2	8	18	16	10	6	7	5	12	8	9
12	17	11	9	16	19	18	18	16	14	12	7	10
3	11	7	12	5	9	11	15	9	4	1	6	11
7	8	10	15	3	2	13	9	11	17	13	12	8

Realiza una tabla de intervalos de clase.

- Las edades de veinte chicos son 12, 13, 14, 10, 11, 12, 11, 13, 14, 12, 10, 12, 11, 13, 12, 11, 13, 12, 10 y15. Organiza los datos en una tabla de frecuencias.
- ¿Qué porcentaje de chicos tienen 12 años?
- ¿Cuántos chicos tienen menos de 14 años?
- En cada día del mes de enero, en el camping Iglú hubo la siguiente cantidad de turistas: 12, 14, 17, 16, 19, 15, 15, 21, 24, 26, 28, 24, 25, 26, 20, 21, 34, 35, 33, 32, 34, 38, 40, 43, 41, 45, 50, 53, 58. Construye una tabla de frecuencias para estos datos.
- Construye el histograma y el polígono de frecuencias para la tabla del ejercicio de intervalos de clase (el de las tallas...)
- Realiza para cada punto un gráfico de barras y uno circular y realiza una explicación de cual puede ser más útil.
- Explique que son las medidas de tendencia central y de variación.
- Los siguientes datos numéricos corresponden a la cantidad de veces que cada alumno de un grupo ha ido a un recital o concierto.

2 – 4 – 3 – 2 – 1 – 1 – 6 – 3 – 0 – 3 – 2 – 4 – 6 – 9 – 3 – 2 – 1 – 6

Calcula, sin tabular, Media, moda, mediana, desviación, n, rango.

- En un diagnóstico de educación física se pidió a los alumnos de los cuartos medios que hicieran abdominales durante 3 minutos. Se obtuvieron los siguientes resultados:

4º A: 45 38 43 29 34 60 54 27 32 33 23 34 34 28 56 62 56 57 45 47 48 54
33 45 44 41 34 36 34 54
4º B: 43 45 44 38 34 46 43 42 43 45 57 44 38 38 37 43 61 38 37 45 28 42
41 49 40 37 34 44 41 43

¿Cuál de los dos cursos tiene el rendimiento más parejo? ¿qué distribución estadístiao permite comparar la distribución de este tipo de datos?

AUTOEVALUACIÓN

1. ¿Qué aprendizajes construiste?
2. Lo qué aprendiste, ¿te sirve para la vida? ¿Si/no; por qué?
3. ¿Qué dificultades tuviste? ¿Por qué?
4. ¿Cómo resolviste las dificultades?
5. Si no las resolviste ¿Por qué no lo hiciste?
6. ¿Cómo te sentiste en el desarrollo de las actividades? ¿Por qué?

RECURSOS	COLOMBIAPRENDE ALGEBRA BALDOR CLASSROOM VIDEOS DE YOUTUBE correo electrónico : daniel.urazan@ierepublicadehonduras.edu.co código classroom : yd5abao http://aprende.colombiaaprende.edu.co/cainicio Whatsapp: 3158963635
FECHA Y HORA DE DEVOLUCIÓN	De acuerdo a la programación institucional.