



**Nombre de alumno: Monica Isabel
Morales Toledo**

**Nombre del profesor: Jorge Sebastián
Domínguez Torres**

**Nombre del trabajo: Estadística
descriptiva**

Materia: estadística descriptiva

Grado: 3

Grupo:

Comitán de Domínguez Chiapas a 11 de Junio de 2021.

Instrucciones: Si bien la práctica hace al maestro, pero enfrascarse en la práctica sin una teoría nos volveríamos empíricos y nosotros seremos expertos en la estadística descriptiva. Realiza cada uno de los enunciados que se presentan a continuación.

- I. Investiga una situación o ejemplo ya sea de algún libro, artículo, revista o tesis que abarque los siguientes conceptos.
 - Frecuencia absoluta
 - Frecuencia relativa
 - Frecuencia acumulada
 - Frecuencia relativa acumulada
 - Distribución de frecuencia
 - Marca de clase

- II. ¿En qué momento se utiliza cada una de las diferentes gráficas? Presenta un ejemplo de gráfica de cada una de ellas.
 - Barras
 - Histograma
 - Lineal
 - Circular
 - Cartograma
 - Pentagonal
 - Dispersión

- III. Realiza un ensayo de no más de 5 cuartillas acerca de los siguientes temas:
 - Propiedades de las distribuciones de frecuencia
 - Estadísticos de posición grupal

Ojo

La principal actividad de un estadístico es la investigación, por ende, en esta unidad no les proporciono links de apoyo porque serán ustedes los investigadores que recaudarán la información, estructurarla y presentarla

Correo para consultas personales al Maestro.
Sebastian_dominguez97@hotmail.com

Investiga una situación o ejemplo ya sea de algún libro, artículo, revista o tesis que abarque los siguientes conceptos

FRECUENCIA ABSOLUTA

La frecuencia absoluta es el número de veces que un dato se repite dentro de un conjunto de datos. Se representa como f_i , donde la i corresponde al número de dato.

La forma de obtener la frecuencia absoluta no es otra que contando las veces que aparece el dato en el conjunto de datos.

La suma de las frecuencias absolutas corresponde al número total de datos, representado por la letra N :

$$\sum_{i=1}^n f_i = f_1 + f_2 + f_3 + \dots + f_n = N$$

FRECUENCIA RELATIVA

La frecuencia relativa de un dato es el número que se repite ese dato en relación al número total de datos, o en otras palabras, es la proporción de veces que aparece ese dato con respecto al total.

Se representa como n_i siendo i el número de dato y se calcula dividiendo la frecuencia absoluta de cada dato entre el número total de datos:

$$n_i = \frac{f_i}{N}$$

El valor de la frecuencia relativa siempre va a estar entre 0 y 1 . El valor obtenido está en tanto por uno, pero se puede expresar en tanto por ciento si se multiplica por 100 .

La suma de todas las frecuencias relativas de todos los datos de la muestra es igual a 1 (cuando se expresa en tanto por 1 , que es lo más común):

$$\sum_{i=1}^n n_i = n_1 + n_2 + n_3 + \dots + n_n = 1$$

FRECUENCIA ABSOLUTA ACUMULADA

La frecuencia absoluta acumulada es la suma de las frecuencias absolutas que se va acumulando hasta ese dato, es decir, la frecuencia absoluta acumulada de un dato en concreto se obtiene sumando su frecuencia absoluta a las frecuencias absolutas de los datos que son menores que él.

Se representa como F_i , donde i es el número de dato.

Se calcula sumando la frecuencia absoluta de un dato más la frecuencia absoluta del dato anterior. Por tanto, la frecuencia absoluta acumulada del primer dato coincide con su frecuencia absoluta y la frecuencia absoluta acumulada del último dato coincide con el número total de datos.

FRECUENCIA RELATIVA ACUMULADA

La frecuencia relativa acumulada es el mismo concepto que para la frecuencia absoluta acumulada.

Se representa como N_i , donde la i es el número de dato y se puede obtener como el cociente entre la frecuencia absoluta acumulada para cada dato entre el número de datos totales:

$$N_i = \frac{F_i}{N}$$

O también, como la suma de la frecuencia relativa de un dato más la frecuencia relativa del dato anterior. Así que, la frecuencia relativa acumulada del primer dato coincide con su frecuencia relativa y la frecuencia relativa acumulada es igual a 1 .

CLASES

Si se tiene un número muy grande de datos, éstos se agrupan en intervalos, para no tener que realizar tablas muy largas con muchos datos diferentes. También se agrupan en intervalos cuando las variables son continuas.

En estos casos se realiza una tabla de frecuencias con datos agrupados.

Los datos se agrupan en intervalos, llamados *clases* y es a estos intervalos los que se asignan sus frecuencias correspondientes.

Sobre las clases, resaltan los siguientes conceptos:

Límites de clase: Cada intervalo tiene un límite inferior, que pertenece a ese intervalo (cerrado por la izquierda con un corchete) y un límite superior que no pertenece (abierto por la derecha).

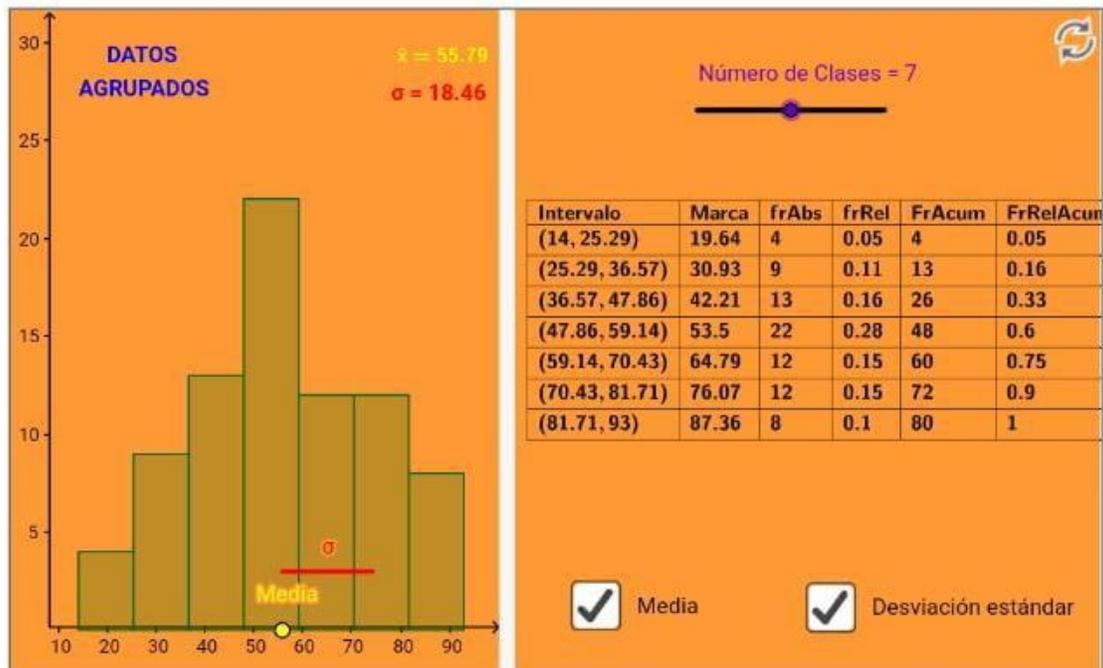
Amplitud de clase: La amplitud es la diferencia entre el límite superior e inferior y debe ser la misma para cada intervalo.

Marca de clase: Es el punto medio de cada intervalo y es el valor que se utiliza para calcular otras medidas.

PROPUESTA DE TRABAJO

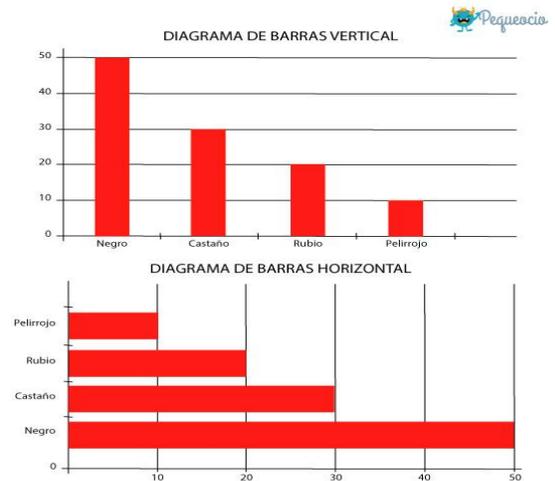
1. Analizar la tabla de frecuencias e identificar los intervalos, las marcas de clase, las frecuencias absolutas, las frecuencias relativas, las frecuencias relativas acumuladas y las frecuencias relativas acumuladas. Notar que es una población de 80 datos.
2. Mover el deslizador para establecer el número de clases.
3. Observar el comportamiento del histograma.
4. Activar las casillas para ver la media y la desviación estándar de la población con datos agrupados. Interpretar el significado de estas medidas.
5. Concluir que la tabla de frecuencias es un instrumento muy útil para el análisis de datos estadísticos.
6. Pulsar el icono que se sitúa arriba a la derecha para regresar a la construcción inicial

Datos Agrupados

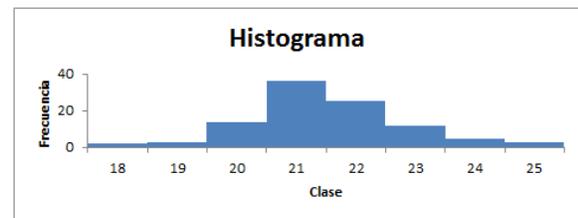


¿En qué momento se utiliza cada una de las diferentes gráficas? Presenta un ejemplo de gráfica de cada una de ellas.

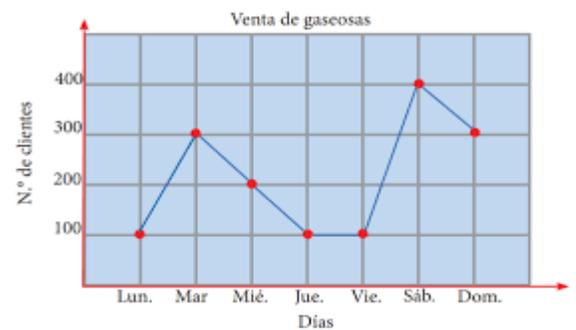
Diagrama de barras: En el gráfico de barras, los datos se representan por medio de rectángulos de igual base sobre el eje de conceptos; en tanto que la longitud del otro lado corresponde al valor del dato, según la escala utilizada en el eje de valores. Cuando se grafica más de una categoría existen diferentes modalidades de presentación.



Histograma: Un histograma representa la distribución de frecuencias (relativas o absolutas) de una variable continua cuya escala es dividida en clases. Para construirlo se dibuja un eje horizontal con la escala de la variable en cuestión y se marcan los límites de clase. Luego, tomando como base el segmento entre los límites de cada clase, se dibujan rectángulos de altura proporcional a la frecuencia de la clase correspondiente.



Lineal: En este tipo de gráfica los valores del indicador se representan con un punto, los cuales se unen mediante líneas para facilitar la visualización del comportamiento del indicador. Estos gráficos se emplean cuando es necesario representar las tendencias de una serie de datos, y éstos son numerosos o continuos; los gráficos de línea pueden cubrir períodos de minutos, horas, días, semanas, meses o años.



Circular: Estos gráficos nos permiten ver la distribución interna de los datos que representan un hecho, en forma de porcentajes sobre un total. Los gráficos circulares son adecuados para recalcar la magnitud relativa de los componentes del total. Consiste en dividir un círculo en sectores cuyas superficies sean proporcionales a las cantidades correspondientes a cada categoría.



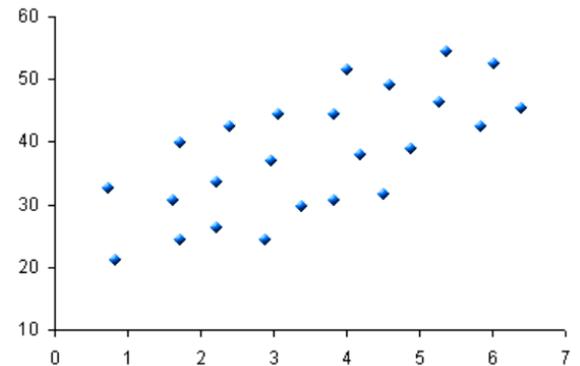
Cartogramas: Son unos gráficos que se realizan de forma de mapa, indicando según sus características las distintas zonas y las cantidades según su representación, es decir que representan las unidades geométricas y sus atributos como son expuestos en el área, solo mostrando los cambios que dependen del incremento, como por ejemplo al no representar el espacio geográfico sino la cantidad de habitantes que tiene si se trata de un cartograma de un país.



Pentagonales: Los gráficos radiales son una forma de comparar múltiples variables cuantitativas. Esto los hace útiles para ver qué variables tienen valores similares o si hay valores atípicos entre cada variable. Los gráficos radiales también son útiles para ver qué variables son altas o bajas dentro de un conjunto de datos, haciéndolos ideales para mostrar el rendimiento.



Los gráficos de dispersión se usan para trazar puntos de datos en un eje vertical y uno horizontal, mediante lo que se trata de mostrar cuánto afecta una variable a otra. Cada fila de la tabla de datos la representa un indicador cuya posición depende de sus valores en las columnas que se establecen en los ejes X e Y. Se pueden usar varias escalas en el eje Y cuando se desea comparar varios indicadores con rangos de valor significativamente distintos. Se puede establecer una tercera variable para que se corresponda con el color o el tamaño (por ejemplo, un gráfico de burbujas) de los indicadores, lo que agregaría otra dimensión más al gráfico



ESTADISTICA DESCRIPTIVA

En estos temas hablaremos sobre la distribución de frecuencias y los estadísticos de posición grupal en el cual en el primero se basan en el datos de diferentes categorías ésta se divide en cuatro etapas de distribución Qué son los que se verán a continuación y en la estadística de posición grupal se basa en revisar la serie de datos que nos permite resumir toda la información contenida en la distribución de frecuencias esos nos ayudarán a proporcionarnos los valores numéricos que reflejaran diferentes características tanto como su posición su variabilidad y su simetría.

- PROPIEDADES DE LA DISTRIBUCION DE FRECUENCIAS

-Tendencia central:

En valor numérico de una variable está posicionado justo en el centro de la distribución de frecuencias llamando a los y como promedio y los ayuda a sintetizar todos los valores de la distribución.

-Viabilidad:

Es una medida que nos permite resumir el grado de concentración de todos los valores en una distribución promedio de acuerdo que si los valores están muy cercanos al promedio los indica una distribución homogénea, pero por lo contrario si los valores están muy lejanos se habla de una distribución heterogénea.

Sesgo o asimetría:

Es el grado de simetría de una distribución de frecuencias. Este tiene dos números uno debajo y uno por encima del promedio que nos indica si la distribución es simétrica o asimétrica si hay una mayor frecuencia de valores bajos que de valores altos se dice que es una distribución asimétrica positiva, pero si hay una mayor frecuencia valores altos que bajos es una distribución asimétrica negativa ya que la distribución asimétrica positiva se basa en las propias tareas o test difíciles y la distribución asimétrica negativa suele ser de tareas difíciles.

Curtosis o apuntamiento:

Es el grado de apuntamiento de distribución de frecuencias cuando es muy apuntada se dice que es leptocurtica y si es muy aplastada se dice que es platicurtica cuando hay un grado intermedio entre estos dos se le denomina como distribución mesocurtica que nos indica que es la distribución normal o de referencia para establecer un grado de apuntamiento.

- ESTADISTICOS DE POSICION GRUPAL

-La moda:

La moda de una variable x es un estadístico de tendencia central que se obtiene como un valor que tiene muchas repeticiones dentro de todos los datos correspondientes a una variable, este es el que tiene la frecuencia absoluta más alta en la distribución de frecuencias.

Puede ser en variables categóricas o también se puede obtener a través de las variables ordinales y cuantitativas donde nos indica que tenemos otro índice estadístico para representar variables categóricas y tenemos que encontrar una manera de descartar los puntos que no captan bien la información contenida de las variables ordinarias y las cuantitativas.

-Variables originales: la mediana.

Para las variables categóricas la mediana es aplicada también en variables cuantitativas pues no es el índice estadístico que mejor resume la tendencia central de estos tipos de variables.

La mediana de una variable x es el valor de un dato obtenido a través de ordenar todos los datos de la variable haciendo que el número quede debajo de estos números en el caso de un número par de todos los datos se obtiene el promedio de los valores de los datos centrales.

-La media.

La media o promedio de una variable x es un índice estadístico que consiste en sumar los valores de los datos de una variable y dividir por el número de ellos mismos.

Si los datos los obtenemos en forma de distribución de frecuencias el cálculo deberá ser como una suma donde el producto de los valores corresponde a la frecuencia absoluta y se divide el resultado por el número de casos.

La media también se aplica como variables ordinales por otro lado se puede obtener a través de una variable ordinal, pero en otras ocasiones hay valores que no se pueden encontrar dentro del Rango de los valores de la escala de la medida de una variable.

Uno de los problemas dentro de la media y sus variables ordinales es que en ocasiones no se pueden encontrar los valores posibles y puede darse como una variable cuantitativa discreta.

El mínimo y el máximo:

El mínimo es una modalidad que adopta una variable donde su valor más bajo en la escala de la medida de las variables y por otro lado el máximo el valor que sea más alto ambos nos permiten identificar de una manera más fácil donde está localizado cada dato de las variables.

-Cuantiles:

Cuantil es el valor de la variable $K\%$ tú ándale sujetos tienen un valor inferior o igual.

El cálculo de un cuantil K resulta sencillo de obtener pues se obtiene a partir de la distribución de frecuencias de la variable dependiendo de la columna de los porcentajes acumulados

La interpretación de las cuales se hace a través del valor del cuántica nos dice el porcentaje de casos que son iguales inferiores hace valor.

Todos y cada uno de estos datos son fáciles de utilizar con las fórmulas correcta haciendo que sea un poco más fácil de obtener la información requerida y que encontremos cada valor de diferente manera nos permite conocer los diferentes cálculos, diferentes fórmulas y casi tengamos un mejor aprendizaje de los temas y que a su vez podamos hacerlo de una manera más eficiente.