MEDIDAS Y DATOS.

ESTADISTICA I

MTRO. ROSARIO GOMEZ LUJANO

PRESENTA EL ALUMNO:

Flor de Liz García Mendoza

GRUPO, SEMESTRE y MODALIDAD:

1er. cuatrimestre “A” Psicología Escolarizado

Pichucalco, Chiapas

13 De octubre del 2020.

INTRODUCCIÓN

En una investigación de cualquier tema de interés con una metodología cuantitativa, es necesario obtener datos para la información que deseamos, los datos En una investigación de cualquier índole o son de dos grupos: Agrupados y No Agrupados, dependiendo de la cantidad de datos que uno obtiene.

De igual forma, los datos agrupados y no agrupados necesitan de las medidas de tendencia central (moda, mediana, media aritmética) y de variabilidad (rango, varianza, desviación estándar, etc.) y otras.

En este trabajo comprenderemos mejor de estos términos estadísticos que usamos en nuestra vida cotidiana aunque no nos demos cuenta, otros si serán extraños y otros conocidos.

Medidas y datos.

Nos hemos preguntado una vez ¿De dónde pudieron sacar ese pequeño porcentaje o dato de una cantidad de población?, la respuesta está en las medidas de tendencia central y medidas de variabilidad, ¿para que se usan esos datos? Una vez recaudado la información o los datos de una muestra o población pasa a ser analizados por esas medidas para determinar la cantidad más clara.

La manera en que los datos pasan de muchos datos obtenidos a un número. Analizaremos más detalle sobre estas medidas.

Media aritmética.

Es conocida como promedio y es de las primeras medidas que nos enseñan en el nivel primario, consiste en que una vez obtenido el conjunto de datos este obtiene el valor representativo de todo ese conjunto en un solo dato o porcentaje. Es un objeto de estudio que parte del principio de la esperanza matemática o valor esperado.

Se calcula de esta forma, recordamos que *n* puede ser cualquier valor numérico: *“es la suma de n valores de la variable y luego dividido entre n, donde n es el número de sumandos, o en el caso de estadística el número de datos que da el resultado*”.

La mediana.

Esta es la más simple que no necesita de alguna fórmula específica para sacar el resultado.

Consiste en el valor que está en medio de un motón de datos de manera ordenada del mayor a menor o viceversa.

Existen dos métodos para el cálculo de la mediana:

1. Considerando los datos en forma individual, sin agruparlos.
2. Utilizando los datos agrupados con intervalos de clase.

Moda.

Se trata de que en un conjunto de datos, el valor moda sea el más repetitivo o el de mayor frecuencia en las distribuciones de datos, el valor que más destaca, se puede distinguir a través de una tabla gráfica.

Eso si hay muchos términos adentro de la moda, si los datos no tienen numero repetido o solo tienen un valor por número se considera amodal (sin moda), si hay una sola moda es unimodal (una moda) si tiene dos modas es bimodal y si hay más de dos modas se considera polimodal (más de dos modas).

Cuantiles.

Los cuantiles son medidas de posición que se determinan mediante un método que determina la ubicación de los valores en partes iguales. Es una de las extensiones de la mediana cuyo datos se mide en cuatro partes y se usan para evaluar rápidamente la dispersión y la tendencia central.

Las cuartiles detonantes son:

Q1- tiene el valor de 25% de los datos, es menor o igual al valor.

Q2- la mediana, es decir el 50% de los datos es menor o igual al valor.

Q3- tiene el valor de 75% de los datos es menor que o igual a este valor.

Nota: los cuartiles son valores calculados, no observables en los datos, a menudo es necesario interpolar entre dos observaciones para calcular un cuartil con exactitud.

Deciles.

Son utilizados para el aprovechamiento académico como los cuartiles, los deciles se tratan de números que dividen la sucesión de datos a diez partes iguales, de manera que cada parte 1/10 de la muestra o población.

Un decil es una de las posibles formas de un cuantil, otras incluyen un cuartil y el porcentil.

Se calculan de esta manera:

- el primer decil separe el juego de datos entre el 10% de los valores inferiores.

- el noveno decil separe los datos entre el 90% de los valores inferiores y el 10% de los valores superiores.

Porcentiles.

Estas medidas son usadas para datos de ubicación o para clasificar personas cuyas variables deben obtener porcentaje o en ese caso lo conocidos valores cuantitativos continuas. Es una medición en el cual ese porcentaje de los valores totales es el mismo o menor que esa medición. Para datos no agrupados de usa una formula:

$$x = \frac{\mathrm{n.i}}{100}$$

Rango.

Permite obtener una idea de la dispersión de los datos, cuanto mayor es el rango, aun mas dispersos los datos. Se calcula restando el valor mayor y el menor de un conjunto de datos con esta manera:

$$R = vM - vm$$

Dónde:

R= rango

VM= valor mayor en la muestra.

Vm= valor menor en la muestra.

Varianza.

La variable aleatoria es una medida de dispersión definida como la esperanza del cuadrado de la desviación de dicha variable. Su unidad de medida corresponde el cuadrado de la unidad.

Este valor se saca de la diferencia de los datos elevados al cuadrado y sumándolos para dividirlo entre valor n-1.

Hay que tener en cuenta que la varianza puede verse muy influida por los valores antipáticos y no se aconseja su uso cuando las distribuciones de los variables aleatorias.

Desviación típica (Estándar)

Es como la varianza con la diferencia que no calculas al cuadrado los datos y sacas la raíz cuadrada del valor obtenido en la suma y la división. Se usa para cuantificar la variación de la dispersión de un conjunto de datos numéricos.

El símbolo σ se utiliza para representar la desviación estándar de una población, mientras que se utiliza para representar la desviación estándar de una muestra.

Se puede utilizar también para un valor de referencia donde se estima la variación general de un proceso.

Coeficiente de variación.

En esta media está relacionado con la desviación típica de una muestra y la media de los datos que obtienen.

Permite comparar las dispersiones de dos distribuciones distintas, siempre que sus medidas sean positivas. Se calcula para cada una de las distribuciones y los valores que se obtienen y comparación entre sí.

Coefciente de Pearson.

Es una prueba que mide la relación estadística entre dos variables continuas, si la asociación entre los elementos no lineales, entonces el coeficiente no se encuentra representado adecuadamente.

Puede tomar un rango de valores de 11 a -1, un valor de 0 indica que no hay que no hay asociación entre los dos variables, un valor mayor que 0 indica una asociación positiva.

Para llevar a cabo la correlación de Pearson.

- la escala de medida debe ser una escala de intervalo o relaciones.
- los variables deben estar distribuida de forma aproximada.

- asociación lineal.

CONCLUSION.

Concluimos que las diferentes medidas que se usan en la estadística son de importancia para los resúmenes de diversos conjuntos de valores y datos de lo que deseemos valor y calcular.

Considerando que esto se usa para diferentes ámbitos, ya sea para determinar porcentajes, lo mas en tendencia entre otras en investigaciones científicas, censos, tasas de consumo de algún producto o calificaciones escolares.

EJEMPLOS.

Medidas de tendencia central para datos no agrupados.

Media 1

Medidas de tendencia central en datos agrupados

1.- Media Aritmética ($\overline{X}$):

Para calcular la media aritmética debemos realizar los siguientes paso:

•Realizar la sumatoria del producto de la frecuencias por las marcas de clases correspondientes.

•Dividir el resultado por el total de datos.

Ejemplo: Calcular la media aritmética de los siguientes datos:

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)
[55,59[	2	2	57
[59,63[	5	7	61
[63,67[	3	10	65
[67,71[	7	17	69
[71,75]	3	20	73

Media 2

Medidas de tendencia central en datos agrupados

Completamos la última columna:

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)	fi • x
[55,59[	2	2	57	104
[59,63[	5	7	61	305
[63,67[	3	10	65	195
[67,71[	7	17	69	483
[71,75]	3	20	73	219

Sumamos la columna: 104 + 305 + 195 + 483 + 219 = 1306

Dividimos el total por la cantidad de datos: 1306/20 =65,3

Por lo tanto $\overline{X}$ = 65,3

Mediana 1

Medidas de tendencia central en datos agrupados

2.- Mediana (Me):

Para calcular la mediana debemos aplicar la siguiente fórmula:

Me = Li + A[(n/2 - Fi-1)/fi]
Donde:
Li= Límite inferior A= Amplitud
n = Cantidad de datos **Fi-1**= Frecuencia acumulada anterior
fi = Frecuencia absoluta

Ejemplo: Calcular la mediana de los siguientes datos:

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)
[55,59[	2	2	57
[59,63[	5	7	61
[63,67[	3	10	65
[67,71[	7	17	69
[71,75]	3	20	73

Mediana 2

Medidas de tendencia central en datos agrupados

- Calculamos n/2, es decir, 20/2 = 10
- Ubicamos el valor 10 en las frecuencias acumuladas y trabajamos con los valores del intervalo

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)
[63,67[	3	10	65

Donde:
Li= 63 A= 4 n= 20 Fi-1= 7 fi = 3

Reemplazamos en la fórmula: Me = Li + A[(n/2 - Fi-1)/fi]

Me = 63 + 4 [(10 – 7) / 3]
Me = 63 + 4
Me = 67

Observación: Si el valor n/2 se encuentra en la lista de las frecuencias acumuladas la mediana corresponderá al límite superior del intervalo correspondiente.

Moda 1

Medidas de tendencia central en datos agrupados

3.- Moda (Mo):

Para calcular la moda debemos aplicar la siguiente fórmula:

Mo = Li + A[(fi - fi-1)/(fi – fi-1) + (fi – fi+1)]
Donde:
Li= Límite inferior A= Amplitud
fi= Frecuencia absoluta fi-1 = Frecuencia absoluta anterior
Fi+1 = Frecuencia absoluta superior

Ejemplo: Calcular la moda de los siguientes datos:

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)
[55,59[	2	2	57
[59,63[	5	7	61
[63,67[	3	10	65
[67,71[	7	17	69
[71,75]	4	21	73

Moda 2

Medidas de tendencia central en datos agrupados

- Busco en las frecuencias absolutas cual es la que mas se repite.

Peso (Kg.)	Frecuencia (fi)	Frecuencia Acum. (Fi)	Marca de clase (x)
[67,71[	7	17	69

Li= 67 A= 4 fi= 7 fi-1= 3 fi+1= 4

Reemplazamos en la fórmula:
Mo = Li + A[(fi - fi-1)/(fi – fi-1) + (fi – fi+1)]

Mo = 67 + 4 [(7 – 3) / (7 – 3) + (7 - 4)]
Mo = 67 + 4 [(4 / 4 + 3]
Mo = 67 + 4 [(4 / 7)]
Mo = 67 + (16 / 7)
Mo = 67 + 2,28
Mo = 69,28

Medidas de variabilidad para datos agrupados.

Varianza

$$\sigma^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+...+(x_n-\bar{x})^2}{N} \quad ó \quad \sigma^2 = \frac{x_1^2+x_2^2+...+x_n^2}{N}-\bar{x}^2$$

$$\sigma^2 = \frac{12^2+6^2+7^2+3^2+15^2+10^2+18^2+5^2}{8}-9.5^2 = 23.75$$

Varianza

$$\sigma^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+...+(x_n-\bar{x})^2}{N} \quad ó \quad \sigma^2 = \frac{x_1^2+x_2^2+...+x_n^2}{N}-\bar{x}^2$$

$$\sigma^2 = \frac{(2-6)^2+(3-6)^2+(6-6)^2+(8-6)^2+(11-6)^2}{5} = \frac{54}{5} = 10.8$$

Desviación estándar.

$$\bar{x} = \frac{3+2+3+3+2+6+3}{7} = \frac{22}{7} = 3.14$$

$$\sigma^2 = \frac{10.86}{7} = 1.55$$

$$\sigma = \sqrt{\frac{10.86}{7}} = 1.25$$

$$\bar{x} = \frac{5+5+12+13+15+15+15+20+20+23}{10} =$$

$$= \frac{143}{10} = 14.3$$

$$\sigma^2 = \frac{322.1}{10} = 32.21$$

$$\sigma = \sqrt{\frac{322.1}{10}} = 5.68$$

Ejercicios.

1. Calcula la media aritmética, mediana, moda para los siguientes siguiente datos no agrupados: 6, 7 ,8 ,9 ,10 ,7 ,8 ,9 ,7.

Orden: 6, 7, 7, 7, 8, 8, 9, 9, 10.

Mediana: 8. Media aritmética: 71/9= 7.8 moda: 7.

2. Calcula el rango de varianza y desviación estándar para los siguientes datos: 45, 55, 50, 47, 52.

Orden: 46, 47, 50, 52, 55. Media aritmética: 250/5= 50

Mediana: 50 Moda: no.

$$\sigma = \frac{(55-50)^2 + (46-50)^2 \ + (50-55)^2 + (47-50)^2 + (52-50)^2}{\mathrm{n}-1}$$

$$\sigma = \frac{(5)^2 + (-4)^2 + (0)^2 + (-3)^2 + (2)^2}{\mathrm{n}-1}$$

$$\sigma = \frac{25+16+0+9+4}{n-1} = 54 \div 5 - 1 = 13.5$$

Desviación estándar: $\sqrt{13.5} = 3.7$

3. La siguiente tabla representa la calificación de matemáticas 1 de un grupo de alumnos de la secundaria.

Calificación.	Número de alumnos. (F)
4	3
5	6
6	9
7	16
8	12
9	5
10	3

Calcula la calificación promedio= 7

Total de alumnos= 54.

Total de puntos por grupo= 379.

$$379 \div 54 = 7$$

Calcula el porcentaje de alumnos que aprobaron la materia.= 83.33%

54 = 100%

$$45 \times 100 = 4500 \div 54 = 83.33$$