



Mi Universidad

Súper nota

Nombre del Alumno: **Leticia Mayo López**

Nombre del tema: **Medidas de tendencia central para datos agrupados.**

Parcial: **primero**

Nombre de la Materia: **Estadística Descriptiva**

Nombre del profesor: **Andrés Alejandro Reyes Molina**

Nombre de la Licenciatura: **Administración y Estrategias de Negocios**

Cuatrimestre: **Tercer Cuatrimestre**

Fecha de entrega: **11/07/2023**

Medidas de tendencia central para datos agrupados.

Las medidas de tendencia central son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Representan un centro en torno al cual se encuentra ubicado el conjunto de los datos. Las medidas de tendencia central más utilizados son: media, mediana y moda.

Las medidas de tendencia central tienen distintos usos, entre ellos: resumir la información. Conocer el elemento promedio o típico de un grupo. Comparar e interpretar los resultados obtenidos al analizar una colección de valores observados.

Media

La media es una operación muy utilizada en matemáticas, economía y estadística para el cálculo del valor medio dentro de un conjunto de valores. Para ello se realizan un conjunto de operaciones a esos valores para conseguir el resultado deseado.

La media es el valor promedio de un conjunto de datos numéricos, calculada como la suma del conjunto de valores dividida entre el número total de valores. La media a diferencia de la esperanza matemática, es un término matemático.

Media aritmética para datos agrupados

La media aritmética también se conoce como promedio, y básicamente se calcula como la suma de todos los datos dividida entre el número total de datos. Pero esto aplica para datos sueltos, es decir, no agrupados.

Para los datos agrupados debemos considerar con un valor representativo de cada intervalo que se denomina marca de clase y asumir que todas las cantidades de la frecuencia absoluta se ven representados por ese valor.

La media se calcula sumando todos los datos y dividiendo entre el total de ellos. Pero para datos agrupados asumimos que por ejemplo en el primer intervalo esas 5 personas todas tienen 14.5 años, entonces queda más práctico multiplicar 5×14.5 o lo que es lo mismo $14.5 + 14.5 + 14.5 + 14.5 + 14.5$.

Mediana

Si se ordenan todos los datos, de menor a mayor, la mediana es el valor que ocupa la posición central. Si el número de datos es par, la mediana es la media aritmética de los dos centrales. La moda es el valor que más se repite o, lo que es lo mismo, el que tiene la mayor frecuencia.

La mediana se utiliza generalmente para devolver la tendencia central en el caso de distribuciones numéricas sesgadas. La media se calcula sumando todos los valores y dividiendo la suma entre el número total de valores.

La mediana es un conjunto es un valor que se encuentra a la mitad de los otros valores, es decir, que al ordenar los números de menor a mayor, este se encuentra justamente en medio entre los que están por arriba.

Moda

La moda se simboliza como M y nuestro paso será identificar el intervalo modal. Es muy sencillo, el intervalo modal corresponde a aquel que posee la frecuencia absoluta más alta.

Si fuesen datos no agrupados, fácilmente diríamos que la moda es el dato que más se repite sin realizar ningún cálculo ni operación matemática. Pero como nuestro interés es calcular la moda para datos agrupados debemos utilizar la siguiente formula:

$$M_o = L_i + \left(\frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \right) \cdot A$$

Varianza y desviación estándar

Como la varianza es el promedio de las distancias al cuadrado que van desde las observaciones a la media, la desviación estándar es la raíz cuadrada del promedio de las distancias al cuadrado que van desde las observaciones a la media.

La varianza y la desviación estándar indican si los alores se encuentran más o menos próximos a las medidas de posición. La desviación estándar es simplemente la raíz cuadrada positiva de la varianza.

La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media. Formalmente se calcula como la suma de os residuos al cuadrado divididos entre en total de observación, también se puede calcular como la desviación típica al cuadrado. Para realizar de manera correcta la operación primero calcular la varianza de la muestra y luego sacra su raíz cuadrada para obtener la desviación estándar.

Varianza y desviación estándar para datos agrupados				
	Varianza	Desviación estándar	Media	Número de elementos
Población	$\sigma^2 = \frac{\sum_{i=1}^k f_i (x_i - \mu)^2}{N}$	$\sigma = \sqrt{\sigma^2}$	$\mu = \frac{\sum_{i=1}^k x_i \cdot f_i}{N}$	$N = \sum_{i=1}^k f_i$
Muestra	$s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n - 1}$	$s = \sqrt{s^2}$	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n}$	$n = \sum_{i=1}^k f_i$

Gráficas para representar datos agrupados

La representación gráfica más usada para datos agrupados es el histograma de frecuencia absoluta o relativa. Un histograma es un conjunto de rectángulos adyacentes, cada uno de los cuales representa un intervalo de clase. La base de cada rectángulo es proporcional a la amplitud del intervalo.

Histogramas: se agrupan los datos en clases, y se cuenta cuantas observaciones (frecuencia absoluta) hay en cada una de ellas. En algunas variables (variables cualitativas) las clases están definidas de modo natural, p.e sexo con dos clases: mujer, varón o grupo sanguíneo con cuatro: A, B, AB, O. en las variables cuantitativas, las clases hay que definir las explícitamente (intervalo de clase). Lo cual existe diferentes histogramas así como se muestra en los siguientes enlistados:

- Histograma simple
- Histograma acumulado
- Histograma por grupos
- Histograma dirigido
- Histograma estratificado
- Grafica de pastel
- Grafica ojiva.