

ESTADÍGRAFOS DE ATRACCIÓN

Los estadígrafos de atracción son medidas estadísticas utilizadas para describir el comportamiento de los datos en función de su centralización, dispersión y forma. Estas medidas nos permiten resumir y analizar conjuntos de datos de manera eficiente, ayudándonos a extraer información clave, ya sea para datos agrupados o no agrupados.

Es decir, después de haber ordenado y descrito un conjunto de datos, aún el análisis resulta todavía un tanto incompleto; es necesario entonces resumir la información y facilitar así su análisis e interpretación utilizando ciertos indicadores.

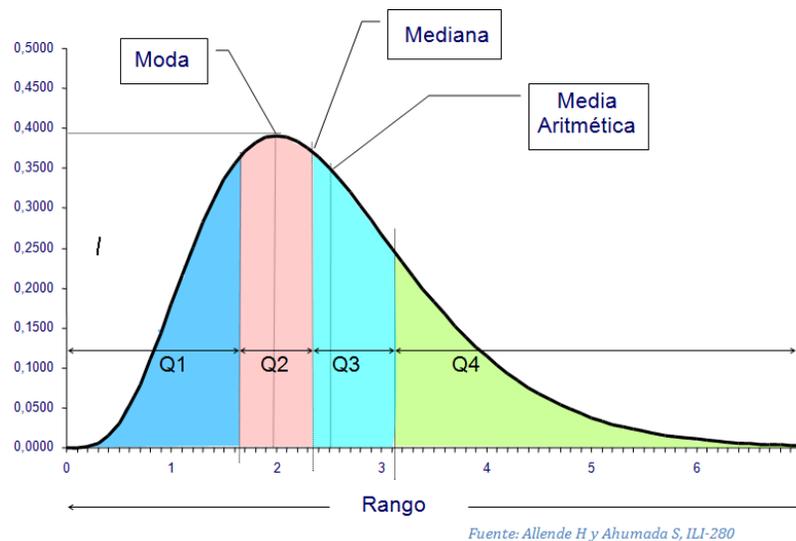
A estos indicadores se les denomina también estadígrafos o **medidas de resumen**, las cuales nos permiten hallar un valor numérico, el mismo que representa a toda la población o muestra en estudio.

Los estadígrafos se clasifican en tres grupos:

- Medidas de tendencia central: Media, mediana, moda.
- Medidas de dispersión: Desviación standard, varianza.
- Medidas de posición: Deciles, cuartiles, percentiles.

ESTADÍGRAFOS DE ATRACCIÓN PARA DATOS NO AGRUPADOS

Los **datos no agrupados** son aquellos que se presentan de manera individual, sin ser clasificados en intervalos o grupos. Las observaciones están listadas tal cual como se recopilaron, lo que significa que cada valor es único y no se encuentra en una categoría.



Las medidas de **tendencia central** se utilizan para identificar el centro de un conjunto de datos, o el valor alrededor del cual se agrupan la mayoría de las observaciones. Las más comunes son:

- **Media.**

La media aritmética es el promedio de todos los valores de los datos.

$$Media = \frac{\sum_{i=1}^n X_i}{n}$$

donde X_i son los valores de las observaciones y n es el número total de observaciones.

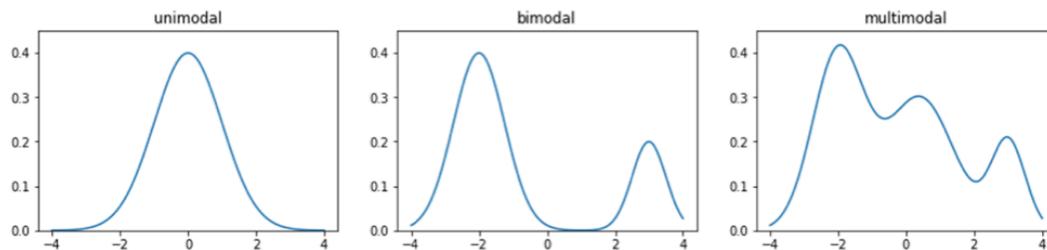
NOTA: En la media aritmética el valor de cada uno de los datos es tomado en cuenta, siendo esta su principal ventaja. Sin embargo, su desventaja es que es sensible a los valores extremos.

- **Mediana.**

Es el valor que se encuentra en el centro de los datos cuando están ordenados de menor a mayor. Si el número de observaciones es impar, la mediana es el valor del medio; si es par, es el promedio de los dos valores centrales.

- **Moda.**

Es el valor que más veces se repite en el conjunto de datos. La moda puede no existir e incluso puede no ser única. Cuando hay dos modas en un conjunto de datos, se le llama conjunto bimodal; cuando hay más, multimodal.



http://materias.df.uba.ar/mytda2020c2/files/2013/03/Apunte_estadistica_v2.pdf

Ejemplo:

Supongamos que tenemos los siguientes datos sobre las edades de un grupo de 7 personas:

Datos: 21, 23, 25, 21, 26, 21, 22.

- **Media:**

$$Media = \frac{(21 + 23 + 25 + 21 + 26 + 21 + 22)}{7} = 22.71$$

Entonces: La **media** de las edades es aproximadamente **22.71 años**.

- **Mediana:**

Primero, ordenamos los datos: 21, 21, 21, 22, 23, 25, 26. El valor central es **22**, por lo que la **mediana** es **22**.

- **Moda:**

El número que más se repite es **21**, por lo que la **moda** es **21**.

ESTADÍSTGRAFOS DE ATRACCIÓN PARA DATOS AGRUPADOS

Cuando los datos están **agrupados** significa que las observaciones se han organizado en intervalos o clases. Este tipo de organización es útil para manejar grandes cantidades de datos, especialmente cuando las observaciones son continuas. En este caso, es necesario utilizar **frecuencias** y **clases** para describir los datos.

Las principales medidas de **tendencia central** para datos agrupados son:

- **MEDIA:**

La media para datos agrupados se calcula utilizando las **frecuencias** de cada clase. La fórmula general para la media en datos agrupados es:

$$Media = \frac{\sum_{i=1}^k f_i * x_i}{n}$$

Donde:

f_i es la frecuencia de la i -ésima clase; x_i es el punto medio de la i -ésima clase; n es el número total de observaciones (que se obtiene sumando todas las frecuencias $f_1+f_2+\dots+f_k$).

Pasos para calcular la media en datos agrupados:

- **Paso 1:** Calcular el punto medio de cada clase. El punto medio de una clase se calcula como el promedio de los límites superior e inferior de la clase:

$$Punto\ medio = \frac{Límite\ inferior + Límite\ superior}{2}$$

- **Paso 2:** Multiplicar el punto medio de cada clase por la frecuencia de esa clase para obtener el producto de la frecuencia y el punto medio.
- **Paso 3:** Sumar todos los productos obtenidos en el paso anterior.
- **Paso 4:** Dividir entre el número total de observaciones (n) para obtener la media.

- **MEDIANA:**

Para encontrar la **mediana** de los datos agrupados, es necesario encontrar el intervalo de la mediana y luego aplicar una fórmula específica para calcularla. Esto se hace a partir de la frecuencia acumulada. La fórmula general para calcular la mediana en datos agrupados es:

$$Mediana = L + \left(\frac{\frac{n}{2} - F}{f} \right) \times h$$

Donde:

L es el límite inferior de la clase mediana (el límite inferior de la clase que contiene la mediana); **n** es el número total de observaciones; **F** es la frecuencia acumulada de la clase anterior a la clase mediana. **f** es la frecuencia de la clase mediana y **h** es el ancho de la clase (la diferencia entre el límite superior e inferior de la clase).

Pasos para calcular la mediana en datos agrupados:

- **Paso 1:** Ordenar las clases de menor a mayor.
- **Paso 2:** Calcular la frecuencia acumulada. Se calcula sumando las frecuencias de las clases de manera acumulada. La **frecuencia acumulada** es la suma de las frecuencias de todas las clases anteriores a la clase que estamos considerando.

Por ejemplo:

Intervalo de clase	Frecuencia (f)	Frecuencia acumulada (F)
0 - 10	5	5
10 - 20	8	5 + 8 = 13
20 - 30	12	13 + 12 = 25
30 - 40	10	25 + 10 = 35
40 - 50	5	35 + 5 = 40

- **Paso 3.** Encontrar la clase mediana. Identificar el intervalo de clase que contiene la posición de la mediana. La mediana es el valor que divide el conjunto de datos en dos partes iguales, por lo que se debe buscar el intervalo cuya frecuencia acumulada sea igual o mayor a la mitad del número total de observaciones ($n/2n$).

○

Siguiendo el ejemplo:

El número total de observaciones es $n=40$; La mediana se encuentra en la posición $n=40/2=20$

Observamos que la frecuencia acumulada de la clase **20 - 30** es 25 y la frecuencia acumulada de la clase anterior (10 - 20) es 13. Como el valor 20 cae dentro de la clase 20 - 30 (porque $13 < 20 \leq 25$), la clase mediana es **20 - 30**.

- **Paso 4.** Aplicar la fórmula de la mediana.

Siguiendo el ejemplo:

$L=20$ (el límite inferior de la clase mediana).

$F=13$ (la frecuencia acumulada de la clase anterior).

$f= 12$ (la frecuencia de la clase mediana).

$h=10$ (el ancho de la clase, ya que $30-20=10$).

Aplicando la fórmula:

$$Mediana = 20 + \left(\frac{\frac{40}{2} - 13}{12} \right) \times 10 = 25.83$$

Por lo tanto, la **mediana** de este conjunto de datos agrupados es **25.83**.

- **MODA:**

La **moda** para datos agrupados se encuentra en el intervalo de clase con la **mayor frecuencia**. Utilizamos la fórmula de la moda para calcularla dentro de ese intervalo:

$$Moda = L_{mod} + \left(\frac{f_{mod} - f_{mod-1}}{(f_{mod} - f_{mod-1}) + (f_{mod} - f_{mod+1})} \right) \times h$$

Donde:

L_{mod} es el límite inferior de la clase modal (el límite inferior de la clase con la mayor frecuencia).

f_{mod} es la frecuencia de la clase modal.

$f_{\text{mod}-1}$ es la frecuencia de la clase anterior a la clase modal.

$f_{\text{mod}+1}$ es la frecuencia de la clase posterior a la clase modal.

h es el ancho de las clases (la diferencia entre el límite superior e inferior de las clases).

Pasos para calcular la moda en datos agrupados:

- **Paso 1:** Ordenar las clases de menor a mayor.

Ejemplo. Si tenemos los siguientes datos:

Intervalo de Clase	Frecuencia (f_i)
0 - 10	5
10 - 20	8
20 - 30	12
30 - 40	18
40 - 50	7

- **Paso 2.** Identificar la clase modal. La clase modal es el intervalo con la mayor frecuencia f_{mod} . En caso de que haya varias clases con la misma frecuencia máxima, se toma la clase que aparece primero en el orden.

Siguiendo los datos del ejemplo, observamos que la mayor frecuencia es **18**, que corresponde a la clase **30 - 40**.

Por lo tanto:

$L_{\text{mod}}=30$ (el límite inferior de la clase modal).

$f_{\text{mod}}=18$ (la frecuencia de la clase modal).

$f_{\text{mod}-1}=12$ (la frecuencia de la clase anterior, 20 - 30).

$f_{\text{mod}+1}=7$ (la frecuencia de la clase posterior, 40 - 50).

$h=10$ (el ancho de las clases, ya que $40-30=10$).

- **Paso 3.** Aplicar la fórmula de la moda.

Calculando la moda:

$$\text{Moda} = 30 + \left(\frac{18 - 12}{(18 - 12) + (18 - 7)} \right) \times 10 = 33.53$$

Por lo tanto, la **moda** de este conjunto de datos agrupados es **33.53**.



TOMA NOTA Conocer el valor de la moda en datos agrupados permite conocer cuál es la clase o intervalo más común, así como: identificar patrones, tomar decisiones informadas sobre distribución de recursos, mejorar productos o servicios y detectar tendencias en los datos. Sin embargo, es importante complementar la moda con otras medidas de tendencia central, como la media y la mediana para obtener una visión más completa de la distribución de los datos.

Referencias:

Walpone, Ronald. (2012) Probabilidad y estadística para ingenierías y ciencias. México. Pearson.

Redacción. (s.f.) Estadígrafos, Medidas descriptivas o de resumen. Angel fire. Recuperado en:

<https://www.angelfire.com/sc/matasc/EyD/bioesta/medidas.htm>

Redacción. (2025) Repaso de media, moda y mediana. Khan Academy. Recuperado de:

<https://es.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/mean-median-basics/a/mean-median-and-mode-review>

Quevedo, Ricardi. Fernando. (2011) Medidas de tendencia central y dispersión. Medwave. Recuperado de:

<https://dsp.facmed.unam.mx/wp-content/uploads/2013/12/Quevedo-F.-Medidas-de-tendencia-central-y-dispersion.-Medwave-2011-Ma-113..pdf>

Cavada, Gabriel. (2008) Introducción a la estadística. Revista Chilena de Endocrinología y Diabetes.

Recuperado de: <https://revistasoched.cl/3-2008/10.html>