## 1- ¿Qué es estadística?

La **estadística** es una ciencia referente a la recolección, análisis e interpretación de datos, ya sea para ayudar en la resolución de la toma de decisiones o para explicar condiciones regulares o irregulares de algún fenómeno o estudio aplicado, de ocurrencia en forma aleatoria o condicional.

Un estudio estadístico consta de las siguientes fases:

- Recogida de datos.
- Organización y representación de datos.
- Análisis de datos.
- Obtención de conclusiones.

## 2- ¿Cómo se divide la estadística?

La estadística se divide en dos elementos:
**2.1- Estadística descriptiva**
Se dedica a los métodos de recolección, descripción, visualización y resumen de datos originados a partir de los fenómenos en estudio.
**2.2- Estadística inferencial**
Se dedica a la generación de los modelos, inferencias y predicciones asociadas a los fenómenos en cuestión teniendo en cuenta la aleatoriedad de las observaciones. Se usa para modelar patrones en los datos y extraer inferencias acerca de la población bajo estudio.

## 3- Conceptos de estadística

**- Población:**
Una población es el conjunto de todos los elementos a los que se somete a un estudio estadístico.

**- Individuo:**
Un individuo o unidad estadística es cada uno de los elementos que componen la población.

**- Muestra:**
Una muestra es un conjunto representativo de la población de referencia, el número de individuos de una muestra es menor que el de la población.

**- Muestreo:**
El muestreo es la reunión de datos que se desea estudiar, obtenidos de una proporción reducida y representativa de la población.

**- Valor:**
Un valor es cada uno de los distintos resultados que se pueden obtener en un estudio estadístico. Si lanzamos una moneda al aire 5 veces obtenemos dos valores: cara y cruz.

**- Dato:**

Un dato es cada uno de los valores que se ha obtenido al realizar un estudio estadístico. Si lanzamos una moneda al aire 5 veces obtenemos 5 datos: cara, cara, cruz, cara, cruz.

## 4- Variable estadística

**- Definición de variable**
Una variable estadística es cada una de las características o cualidades que poseen los individuos de una población.

### 4.1- Tipos de variable estadísticas

**a) Variable cualitativa**
Las variables cualitativas se refieren a características o cualidades que no pueden ser medidas con números. Podemos distinguir dos tipos:

**- Variable cualitativa nominal:**
Una variable cualitativa nominal presenta modalidades no numéricas que no admiten un criterio de orden.

***Por ejemplo:***

- El estado civil, con las siguientes modalidades: soltero, casado, separado, divorciado y viudo.

**- Variable cualitativa ordinal o variable cuasicuantitativa:**
Una variable cualitativa ordinal presenta modalidades no numéricas, en las que existe un orden.

***Por ejemplo:***
- La nota en un examen: suspenso, aprobado, notable, sobresaliente.
- Puesto conseguido en una prueba deportiva: 1º, 2º, 3º, ...
- Medallas de una prueba deportiva: oro, plata, bronce.

**b) Variable cuantitativa**
Una variable cuantitativa es la que se expresa mediante un número, por tanto se pueden realizar operaciones aritméticas con ella. Podemos distinguir dos tipos:

**- Variable discreta**
Una variable discreta es aquella que toma valores aislados, es decir no admite valores intermedios entre dos valores específicos.

***Por ejemplo:***
El número de hermanos de 5 amigos: 2, 1, 0, 1, 3.

**- Variable continua**
Una variable continua es aquella que puede tomar valores comprendidos entre dos números.

***Por ejemplo:***

- La altura de los 5 amigos: 1.73, 1.82, 1.77, 1.69, 1.75.

En la práctica medimos la altura con dos decimales, pero también se podría dar con tres decimales.

**MEDIA ARITMETICA .**

La **media aritmética** es la suma de todos los datos dividida entre el número total de datos. Se calculan dependiendo de cómo vengan ordenados los datos.

Ejemplo:

¿Cuál es la media de las edades de Andrea y sus primos?

## MEDIA ARITMETICA

**Sean n datos de una variable...**

$$x_1 + x_2 + \ldots + x_n$$

**El promedio de la muestra se denomina** $\overline{x}$

**Dicho promedio es un estimador de la media de la población. La media de la población es un parámetro y se denomina** $\mu$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

| L | M | M | J | V | S | D |
|---|---|---|---|---|---|---|
| 6 | 3 | 4 | 7 | 5 | 4 | 8 |

$$\bar{x} = \frac{6+3+4+7+5+4+8}{7} = 5.28$$

| N° de hermanos | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| N° de veces | 4 | 3 | 2 | 1 |

1° ) $1 \cdot 4 + 2 \cdot 3 + 3 \cdot 2 + 4 \cdot 1 = 20$

2° ) N° de datos: $4 + 3 + 2 + 1 = 10 \Rightarrow 20 \div 10 = 2$

La media de los datos es 2

| | Nota N°1 | Nota N°2 | Nota N°3 | Nota N°4 | Nota N°5 | Nota N°6 |
|---|---|---|---|---|---|---|
| Notas | 6,5 | 7 | 5,5 | 8 | 8,5 | 9 |

$$\frac{6{,}5 + 7 + 5{,}5 + 8 + 8{,}5 + 9}{6} = \frac{44{,}5}{6} = 7{,}416.. = \overline{X}$$

Así se denota la media aritmética

## Moda

La moda de un conjunto de datos es el dato que **más veces se repite**, es decir, aquel que tiene **mayor frecuencia absoluta**. Se denota por **Mo**. En caso de existir dos valores de la variable que tengan la mayor frecuencia absoluta, habría dos modas. Si no se repite ningún valor, no existe moda.

**- Ejemplo1:**

¿Cuál es el dato que más se repite en el ejemplo anterior?
El dato que más se repite es el **1,** es el que tiene mayor frecuencia absoluta (4 veces).

**La moda del número de hermanos es 1**

**- Ejemplo 2:**

2, 3, 4, 5 , 6 , 9

En este conjunto de datos **no** existe ningún valor que se repita, por lo tanto, este conjunto de valores **no tiene** moda.

**- Ejemplo 3:**

1, 1, 1, 4, 4, 5, 5, 5, 7, 8, 9, 9, 9 **$M_o$= 1, 5, 9**

Si en un grupo hay dos o varias puntuaciones con la misma frecuencia y esa frecuencia es la máxima, la distribución es bimodal o multimodal, es decir, **tiene varias modas.**

**- Ejemplo 4:**

0, 1, **3, 3, 5, 5**, 7, 8 **$M_o$ = 4**

**Si dos puntuaciones adyacentes tienen la frecuencia máxima**, la moda es el promedio de las dos puntuaciones adyacentes.

## La mediana

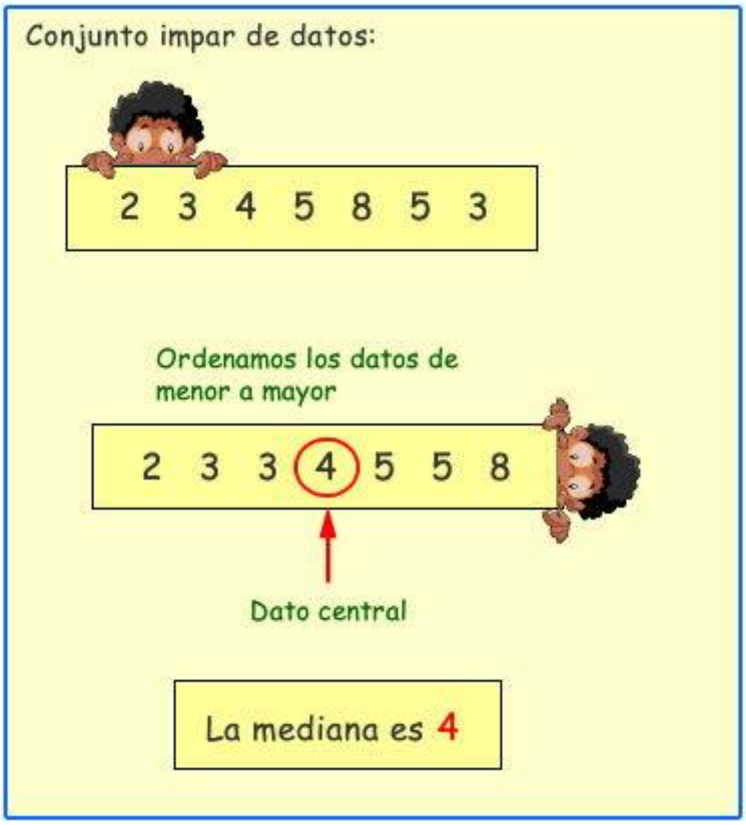
La mediana es el valor que ocupa el lugar central entre todos los valores del conjunto de datos, cuando estos están ordenados en forma creciente o decreciente.
La mediana se representa por **Me**.

**Calculo de la mediana:**

1° Ordenamos los datos de menor a mayor.

- La mediana de un conjunto **con un número impar de datos** es, una vez ordenados los datos, el dato que ocupa el lugar central.

**Ejemplo:**
Calcular la mediana del conjunto de datos:

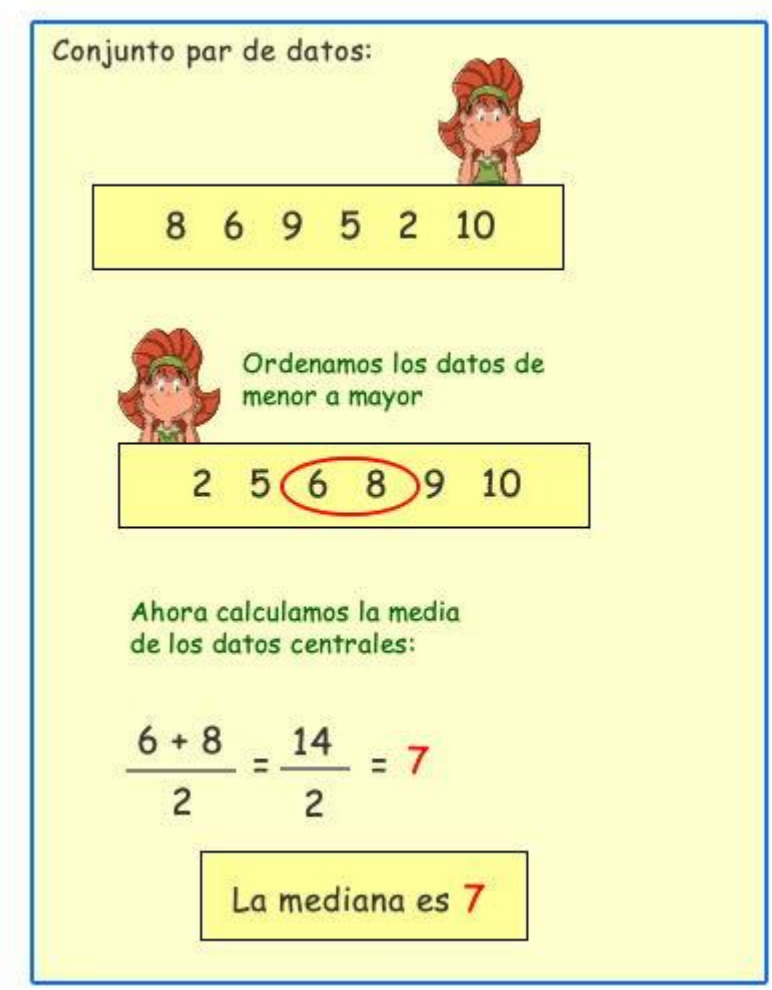



- También podemos usar la siguiente fórmula para determinar la posición del dato central:

**(n + 1) /2 = mediana datos impares.**

- La mediana de un conjunto **con un número par de datos es**, una vez ordenados, la media de los dos datos centrales.

**Ejemplo:**
Calcular la mediana del conjunto de datos:




## Rango

**El rango** da la idea de proximidad de los datos a la media. Se calcula restando el dato menor al dato mayor.
Este dato permite obtener una idea de la **dispersión de los datos**, cuanto mayor es el rango, más dispersos están los datos de un conjunto.

**Ejemplo:**

Se preguntó a 9 familias cuántas bicicletas tenían en total, dieron las respuestas ordenadas en la siguiente tabla:

| Nº de bicicletas | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Frecuencia absoluta | 1 | 5 | 2 | 1 |

**- ¿Cómo hallarías el rango?**

Se resta el dato mayor al dato menor: 3 - 0 = 3; Por lo tanto el rango sería 3 en este caso.

Si el conjunto de datos que se recolecta es muy numeroso, o bien, si el rango es muy amplio, es conveniente agruparlos y ordenarlos en **intervalos** o clases.

**La amplitud** o tamaño de cada intervalo se puede calcular dividiendo el valor del rango por la cantidad de intervalos que se desean obtener.

**4- Ejercicios:**

1- Se le pregunta a un grupo de personas acerca de la cantidad de libros que leyó durante el año 2015, y las respuestas son: 4; 3; 2; 7; 10; 8; 2; 9; 3; 6; 8; 1; 1; 9; 2. La moda de la muestra es:

a) 2 b) 3 c) 4 d) 5 e) 9

2- Halla la mediana de las siguientes series estadísticas.

a) 1, 7, 3, 2, 4, 6, 2, 5, 6

b) 4, 2, 1, 3, 8, 5, 3, 1, 6, 7

3- Se tienen dos distribuciones cuyos datos son los siguientes:

Distribución A: 9, 5, 3, 2, 1, 2, 6, 4, 9, 8, 1, 3, 5, 4, 2, 6, 3, 2, 5, 6, 7

Distribución B: 1, 1, 3, 2, 5, 6, 7, 2, 5, 4, 3, 1, 2, 1, 5, 7, 8, 9, 9, 2, 1

a) Halla el rango de ambas distribuciones.

4- Se tiene el siguiente conjunto de datos:
10, 13, 4, 7, 8, 11, 10, 16, 18, 12, 3, 6, 9, 9, 4, 13, 20, 7, 5, 10, 17, 10, 16, 14, 8, 18

a) Obtén la mediana

***Respuestas:***
1- a

2- a) 1, 2, 2, 3, 4, 5, 6, 6, 7 M = 4
b) 1, 1, 2, 3, 3, 4, 5, 6, 7, 8; La mediana es la media aritmética de los dos valores centrales, M = 3,5.

3- Rango de A: 9 - 1 = 8
Rango de B: 9 - 1 = 8

4- a) Ordenamos los datos de menor a mayor:

3, 4, 4, 5, 6, 7, 7, 8, 8, 9, 9, 10, 10, 10, 10, 11, 12, 13, 13, 14, 16, 16, 17, 18, 18, 20

Como hay 26 valores, la mediana es la media de los dos valores centrales: M= 10 + 10 / 2 = 10

## Tablas de frecuencias con datos agrupados

Cuando los valores de la variable son muchos, conviene agrupar los datos en **intervalos o** clases para así realizar un mejor análisis e interpretación de ellos.

• Para construir una tabla de frecuencias con datos agrupados, **conociendo los intervalos,** se debe determinar la frecuencia absoluta (**fi)** correspondiente a cada intervalo, contando la cantidad de datos cuyo valor está entre los extremos del intervalo. Luego se calculan las frecuencias relativas y acumuladas, si es pertinente.

• **Si no se conocen los intervalos,** se pueden determinar de la siguiente manera: (recuerda que los intervalos de clase se emplean si las variables toman un número grande de valores o la variable es continua).

- Se busca el valor máximo de la variable y el valor mínimo. Con estos datos se determina el **rango**.

- Se divide el rango en la cantidad de intervalos que se desea tener,(por lo general se determinan 5 intervalos de lo contrario es ideal que sea un numero impar por ejemplo 5, 7, 9) obteniéndose así la **amplitud** o tamaño de cada intervalo.

- Comenzando por el mínimo valor de la variable, que será el extremo inferior del primer intervalo, se suma a este valor la amplitud para obtener el extremo superior y así sucesivamente.

• Otra forma de calcular la cantidad de intervalos es aplicando los siguientes métodos:

**Método Sturges**: $k = 1 + 3{,}332 \log n$

donde:
k= número de clases
n= tamaño muestral

Debemos tener en cuenta 2 cosas. Primero que el número de intervalos me tiene que dar **impar,** segundo que el resultado se redondea generalmente a la baja. Si al redondear a la baja nos da como resultado un número par debemos redondear al alza. Este es el método que tiene mayor precisión.

**Método Empírico:** este método depende del criterio del evaluador de los datos, por lo tanto es arbitrario. Dice lo siguiente.

$$5 \geq k \geq 20$$

**Veamos cómo se resuelve el siguiente ejercicio**
En un centro comercial, se consultó la edad a todas las personas que entraban entre las 12:00 h y 12:30 h. Los resultados obtenidos fueron los siguientes:

| | | | | | | |
|---|---|---|---|---|---|---|
| 15 | 73 | 1 | 65 | 16 | 3 | 42 |
| 36 | 42 | 3 | 61 | 19 | 36 | 47 |
| 30 | 45 | 29 | 73 | 69 | 34 | 23 |
| 22 | 21 | 33 | 27 | 55 | 58 | 17 |
| 4 | 17 | 48 | 25 | 36 | 11 | 4 |
| 54 | 70 | 51 | 3 | 34 | 26 | 10 |

- **Construye una tabla de frecuencias cuyos datos estén agrupados en ocho intervalos**.

1°Para poder construir la tabla de frecuencias lo primero que debemos hacer es calcular el **rango**.

**El rango** da la idea de proximidad de los datos a la media. Se calcula restando el **dato menor al dato mayor**.

El dato mayor y el menor lo hemos destacado con color rojo:

Dato mayor - dato menor = **73 - 1** = 72

Por lo tanto; **Rango = 72**

2° En el problema nos dicen que debemos agruparlo en 8 **intervalos o clases,** con este dato podemos calcular la amplitud o tamaño de cada intervalo, dividiendo el valor del rango por la cantidad de intervalos que se desean obtener (en este caso son 8).

**Amplitud:** La amplitud de un intervalo es la diferencia entre el límite superior y el límite inferior. La amplitud(A) de los intervalos puede calcularse mediante la expresión:

$$\text{Amplitud} = \frac{\text{Rango}}{\text{N° de intervalos}}$$

**72 / 8 = 9**

**Por lo tanto la amplitud de cada intervalo será de 9**

- El valor de la amplitud se redondea al número inmediato superior de acuerdo a la cantidad de decimales que tienen los datos o según la precisión con que se desea trabajar.

- Puede haber intervalos con distinta amplitud.
- Puede haber intervalos con amplitud indefinida (intervalos abiertos)

***3° Ahora podemos comenzar a construir la tabla de frecuencias:***

Hay distintas formas de construir los intervalos dependiendo del tipo de variable que estemos trabajando.

**a) Variables cuantitativas discretas**: solo pueden tomar un número finito de valores. Siendo por lo general estos valores los números naturales 1, 2, 3...Un ejemplo son el número de hijos, el número de habitaciones de una vivienda, el número de matrimonios de una persona. Cuando categorizamos variables discretas los límites de clase son idénticos a los límites reales. Por ejemplo, el número de personas que viven en una familia podemos agruparlo, De 1 hasta 2 (0 es imposible no hay ninguna familia sin ningún miembro) De 3 hasta 4, De 5 hasta 7.

**b) Variables cuantitativas continuas**: Las variables continuas, por el contrario, pueden, tomar un número infinito de valores en cualquier intervalo dado. En este caso los valores se agrupan en intervalos cuyos límites inferior y superior serían los siguientes:

Inferior : $Li_i$
Superior: $Ls_{i-1}$

Habitualmente, los intervalos se consideran cerrados a la izquierda y abiertos a la derecha, es decir que el extremo inferior está incluido en el intervalo, pero el extremo superior no.

Es importante mencionar que las clases o intervalos para las variables continuas pueden ser de tres tipos:

**abiertas**: clases abiertas tienen límites determinados (a,b), pero los valores que la contienen comprenden valores muy cercanos a estos límites sin comprenderlos a ellos mismos, esto se representa con un intervalo definido entre paréntesis (). Esto quiere decir que esta clase contiene valores desde a hasta b pero no contiene exactamente a ni b solo valores muy cercanos.

**cerradas**: las clases cerradas, además de los valores que están entre a y b, los contiene a ellos, y se representa con corchetes [a,b].

**semiabiertas:** pueden contener a o b más los valores que están entre ellos, y se puede representar con un corchete y un paréntesis, por ejemplo, (a,b], en este caso no contiene el valor **a** y si los valores de **b, además de los valores que están entre estos.**

**C) Registro discreto de variables continuas:** Cuando la variable considerada es continua pero ocurre que la precisión del instrumento de medida se limita a un número finito de datos, existe la opción de construir los intervalos de tal forma que ambos extremos estén incluidos en él.

Ej 50 a 52, 53 a 55, 56 a 58, 59 al 61 y 62 al 64

Estos serían los límites aparentes de los intervalos.

→ Con esta información construiremos la tabla en esta ocasión con el último método explicado.

| Intervalo (i) | Edades | Frecuencia Absoluta (fi) | Frecuencia Acumulada (Fi) | Frecuencia Relativa (hi) | Frecuencia Relativa Acumulada (Hi) |
|---|---|---|---|---|---|
| 1 | 1 - 10 | 7 | 7 | 7 / 42: 0,17 | 7 / 42: 0,17 |
| 2 | 11 - 20 | 6 | 7+6: 13 | 6 / 42: 0,14 | 13 / 42: 0,31 |
| 3 | 21 - 30 | 8 | 13+8: 21 | 8 / 42: 0,19 | 21 / 42: 0,5 |
| 4 | 31 - 40 | 6 | 21+6: 27 | 6 / 42: 0,14 | 27 / 42: 0,64 |
| 5 | 41 - 50 | 5 | 27+5: 32 | 5 / 42: 0,12 | 32 / 42: 0,76 |
| 6 | 51 - 60 | 4 | 32+4: 36 | 4 / 42: 0,1 | 36 / 42: 0,86 |
| 7 | 61 - 70 | 4 | 36+4: 40 | 4 / 42: 0,1 | 40 / 42: 0,95 |
| 8 | 71 - 80 | 2 | 40+2: 42 | 2 / 42: 0,05 | 42 / 42: 1 |
| | Amplitud 9 | N: 42 | | hi: fi /N | Hi: Fi /N |

**- Marca clase o centro de la clase:** es la semisuma de los límites de cada clase. Representa a todos los datos que están contenidos en una clase.

Responder las siguientes preguntas:

a) Del total de personas encuestadas, ¿cuántas personas tienen entre 31 y 40 años?

**Respuesta:** Observamos los datos obtenidos en la tabla y tenemos que:

| | | | | | |
|---|---|---|---|---|---|
| 4 | 31 - 40 | 6 | 21+6: 27 | 6 / 42: 0,14 | 27 / 42: 0,64 |

El dato lo obtenemos de la columna de la frecuencia absoluta.

Recuerda que:

**Frecuencia absoluta** Corresponde a la cantidad de veces que se repite un dato. Denotamos este valor por $f_i$.

Por lo tanto la respuesta es **6** personas.

b) Del total de personas encuestadas, ¿cuántas personas tienen 60 o menos años?

**Respuesta:** Observamos los datos obtenidos en la tabla y tenemos que:

| | | | | | |
|---|---|---|---|---|---|
| 6 | 51 - 60 | 4 | 32+4= 36 | 4 / 42= 0,1 | 36 / 42= 0,86 |

El dato lo obtenemos de la columna de frecuencia absoluta acumulada.

Recuerda que:

**Frecuencia absoluta acumulada** es la suma de las frecuencias absolutas observadas hasta el intervalo **i**.

En este caso es el intervalo 6. Por lo tanto la respuesta es 36 personas tienen 60 o menos años.

c) ¿Cuál es la probabilidad de, que al elegir al azar a un persona consultada, esta tenga entre 11 y 20 años?

**Respuesta:** Observamos los datos obtenidos en la tabla y tenemos que:

| | | | | | |
|---|---|---|---|---|---|
| 2 | 11 - 20 | 6 | 7+6= 13 | 6 / 42= 0,14 | 14% |

El dato lo obtenemos de la columna de frecuencia relativa.

Recuerda que:

**Frecuencia relativa** Corresponde a la probabilidad de pertenecer a cierta categoría. Se puede expresar en tantos por ciento.

En este caso es el intervalo 2, ya que es ahí donde se encuentran las edades entre 11 y 20 años.

**Entonces la respuesta es: La probabilidad es 14%.**

Por último vamos a repasar el concepto de:

**Frecuencia relativa acumulada (Hi),** Es la probabilidad de observar un valor menor o igual al valor que toma la variable en estudio en ese intervalo.

Se calcula dividiendo Fi por el número total de datos. También puedes calcularlo Sumando la frecuencia relativa de cada grupo con la frecuencia relativa acumulada del grupo anterior.

Si haces correctamente estos cálculos, el último grupo tendrá una frecuencia acumulada de 1, o muy cerca de 1, permitiendo redondear el error.

Recuerda que este valor se puede expresar como porcentaje, para esto solo debes multiplicar el valor obtenido por 100 y listo!!!

Este cálculo te sirve en el caso de que te pregunten:

d) Si le preguntas a una persona cualquiera ¿Cuál es la probabilidad de que tenga 50 años o menos?

| | | | | | | |
|---|---|---|---|---|---|---|
| 5 | 41 - 50 | 5 | 27+5= 32 | 5 / 42= 0,12 | 32 / 42= 0,76 | 76% |

**Respuesta:** La probabilidad es de un 76%

## 1- Media aritmética para datos agrupados

Se calcula sumando todos los productos de **marca clase** con la frecuencia absoluta respectiva y su resultado dividirlo por el número total de datos:

$$\bar{x} = \frac{\text{Suma (marca clase} \times \text{frecuencia absoluta)}}{\text{Total de datos}}$$

La **marca clase** de una tabla para datos agrupados en intervalos corresponde al promedio de los extremos de cada intervalo.

El intervalo $\boxed{26 - 30}$

26: corresponde al extremo inferior del intervalo.
30: corresponde al extremo superior del intervalo

En el intervalo anterior la **marca de clase** es 28 es decir $\bar{x} = \frac{26+30}{2} = 28$

## 2- Moda

Es el valor que representa la **mayor frecuencia absoluta**. En tablas de frecuencias con datos agrupados, hablaremos de intervalo modal.

La moda se representa por **Mo**.

**2.1- Todos los intervalos tienen la misma amplitud.**

$$Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot t_i$$

$L_i$ Extremo inferior del intervalo modal (intervalo que tiene mayor frecuencia absoluta).

$f_i$ Frecuencia absoluta del intervalo modal.

$f_{i-1}$ Frecuencia absoluta del intervalo anterior al modal.

$f_{i+1}$ Frecuencia absoluta del intervalo posterior al modal.

$t_i$ Amplitud de los intervalos.

**2.2 Si los intervalos tienen amplitudes distintas.**

En primer lugar tenemos que hallar las alturas.

$$h_i = f_i / t_i$$

**Donde:**

**hi:** altura correspondiente a cada intervalo.

**fi:** Frecuencia absoluta del intervalo (también se puede utilizar la frecuencia acumulada o relativa)

**ti:** Amplitud de los intervalos

Luego la clase modal es la que tiene mayor altura.

## 3- Mediana

Es el valor que ocupa el lugar central de todos los datos cuando éstos están ordenados de menor a mayor. La mediana se representa por **Me**. La mediana se puede hallar sólo para **variables cuantitativas.**

**Cálculo de la mediana para datos agrupados**

La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas.

Es decir tenemos que buscar el intervalo en el que se encuentre. **N / 2**

**Luego calculamos según la siguiente fórmula:**

$$Me = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot t_i$$

$L_{i-1}$ es el límite inferior de la clase donde se encuentra la mediana.

**N / 2** es la semisuma de las frecuencias absolutas.

$F_{i-1}$ es la frecuencia acumulada anterior a la clase mediana.

**fi** es la frecuencia absoluta del intervalo mediano.

$t_i$ es la amplitud de los intervalos.

**Ahora veamos un ejemplo:**

- En la siguiente tabla se muestran las edades de un grupo de personas.

**1° Calculemos la media aritmética:**

| Edad | Marca clase ($X_i$) | Frecuencia absoluta ($f_i$) | Frecuencia acumulada ($F_i$) |
|---|---|---|---|
| [0 - 10) | 5 | 3 | 3 |
| [10 - 20) | 15 | 6 | 9 |
| [20 - 30) | 25 | 7 | 16 |
| [30 - 40) | 35 | 12 | 28 |
| [40 - 50) | 45 | 3 | 31 |

N = 31

$$\bar{X} = \frac{5 \times 3 + 15\times 6 + 25\times 7 + 35\times 12 + 45\times 3}{31} =$$

$$\bar{X} = \frac{15 + 90 + 175 + 420 + 135}{31} = \frac{835}{31} = 26{,}94$$

$$\boxed{\bar{X} = 26{,}94}$$

**2° Ahora calculemos la mediana (Me)** según las fórmulas explicadas más arriba:

Lo primero que debemos hacer para poder calcular la mediana es identificar la **clase mediana**. Para esto tenemos que buscar el intervalo en el que se encuentre. **N / 2**

**en este caso N / 2 = 31 / 2 ⇒ 15,5**

Ahora debemos buscar el intervalo donde la frecuencia acumulada ($F_i$) contenga el valor obtenido (15,5).

Veamos:

| Edad | Marca clase ($X_i$) | Frecuencia absoluta ($f_i$) | Frecuencia acumulada ($F_i$) |
|---|---|---|---|
| [0 - 10) | 5 | 3 | 3 |
| [10 - 20) | 15 | 6 | 9 |
| [20 - 30) | 25 | 7 | 16 |
| [30 - 40) | 35 | 12 | 28 |
| [40 - 50) | 45 | 3 | 31 |

$\frac{N}{2} = 15,5$

N = 31

Ahora reemplazamos los datos en la fórmula:

$$Me = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot t_i$$

$$Me = 20 + \frac{15,5 - 9}{7} \cdot 10$$

$$Me = 20 + 9,29$$

$$\boxed{Me = 29,285}$$

**Recuerda:**

**$L_{i-1}$** :es el límite inferior de la clase donde se encuentra la mediana, en este caso el límite inferior es 20.

**N / 2** :es la semisuma de las frecuencias absolutas, en este caso es 15,5.

**$F_{i-1}$** :es la frecuencia acumulada anterior a la clase mediana, en este caso es 9.

**fi** : es la frecuencia absoluta del intervalo mediano, en este caso es 7

**$t_i$** :es la amplitud de los intervalos. Se calcula restando el extremo superior menos el inferior del intervalo, en este caso es:

30 - 20 = **10**

**3° Calculemos la moda $M_o$ :**

Lo primero que debemos hacer es identificar el intervalo modal:

| Edad | Marca clase ($X_i$) | Frecuencia absoluta ($f_i$) | Frecuencia acumulada ($F_i$) |
|---|---|---|---|
| [0 - 10) | 5 | 3 | 3 |
| [10 - 20) | 15 | 6 | 9 |
| [20 - 30) | 25 | 7 | 16 |
| [30 - 40) | 35 | 12 | 28 |
| [40 - 50) | 45 | 3 | 31 |

N = 31

Intervalo modal: mayor frecuencia absoluta

Ahora podemos reemplazar los datos en la fórmula:

$$Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot t_i$$

$$Mo = 30 + \frac{12 - 7}{(12 - 7) + (12 - 3)} \cdot 10$$

$$Mo = 30 + 3,57$$

$$\boxed{Mo = 33,6}$$

- Si la moda está en el primer intervalo, entonces **$f_{i-1}$= 0.** Si la moda está en el último intervalo, entonces **$f_{i+1}$= 0.**

- Puede haber más de una moda en el caso en que dos o más valores de la variable presenten la misma frecuencia (distribuciones bimodales o multimodales).

## 1- Interpretación de tablas de frecuencias

**Una tabla de frecuencias** resume la información acerca de la cantidad de veces que una variable toma un valor determinado. Además permite Organizar e interpretar de manera más rápida y eficiente.

**1.1- La frecuencia absoluta**
Corresponde a la cantidad de veces que se repite un dato. Denotamos este valor por $f_i$.

La suma de las frecuencias absolutas es igual al número total de datos, que se representa por **N**.

Por Ejemplo:

Si hacemos una encuesta a 20 personas para saber cuál es su color favorito obtenemos lo siguiente:

[Tabla 1]

| Color | Frecuencia Absoluta |
|---|---|
| Azul | 5 |
| Rojo | 4 |
| Verde | 6 |
| Rosa | 3 |
| Violeta | 2 |
| Total | 20 |

*El color Azul le gustaba a 5 personas*

*El color Rojo le gustaba a 4 personas*

**1.2- La Frecuencia Absoluta Acumulada**
Se obtiene sumando sucesivamente las frecuencias absolutas. Denotamos este valor por $F_i$.

[Tabla 2]

| Color | Frecuencia Absoluta | Frecuencia Acumulada |
|---|---|---|
| Azul | 5 | 5 |
| Rojo | 4 | 5 + 4 = 9 |
| Verde | 6 | 9 + 6 = 15 |
| Rosa | 3 | 15 + 3 = 18 |
| Violeta | 2 | 18 + 2 = 20 |
| Total | 20 | |

## 1.3- La Frecuencia Relativa

Es la probabilidad de obtener cierto dato, se obtiene calculando la razón entre la frecuencia absoluta de un dato con el total. Se puede expresar como fracción, decimal o porcentaje. Denotamos este valor por $h_i$.

[Tabla3]

| Color | Frecuencia Absoluta | Frecuencia Acumulada | Frecuencia Relaiva |
|---|---|---|---|
| Azul | 5 | 5 | $\frac{5}{20}$ - 0,25 - 25% |
| Rojo | 4 | 9 | $\frac{4}{20}$ - 0,2 - 20% |
| Verde | 6 | 15 | $\frac{6}{20}$ - 0,3 - 30% |
| Rosa | 3 | 18 | $\frac{3}{20}$ - 0,15 - 15% |
| Violeta | 2 | 20 | $\frac{2}{20}$ - 0,1 - 10% |
| Total | 20 | | |

→ *5 de cada 20 personas les gusta el color azul*

0,3x100 =30%

3 : 20 = 0,15

Para obtener el numero en decimal se divide la frecuencia absoluta por el total y para obtener el porcentaje se multiplica este decimal por 100.

Los ejemplos representan una tabla de frecuencias de datos **No agrupados,** en el caso de las tablas de datos Agrupados representan las frecuencias en rangos de datos, como en el siguiente caso.

Se entrevistan a 28 personas que realizan un taller preguntándoles la edad que tengan:

[tabla 4]

*10 personas tienen entre 1 y 8 años*

| Edad | Frecuencia Absoluta | Frecuencia Acumulada | Frecuencia relativa |
|---|---|---|---|
| 1-4 | 4 | 4 | $\frac{4}{28}$ - 0,14 - 14% |
| 5-8 | 6 | 10 | $\frac{6}{28}$ - 0,21 - 21% |
| 9-12 | 15 | 25 | $\frac{15}{28}$ - 0,53 - 53% |
| 13-16 | 3 | 28 | $\frac{3}{28}$ - 0,1 - 10% |

*6 de cada 28 personas que hacen el taller tienen entre 5 y 8 años (5, 6,7 y 8)*

*4 personas tienen entre 1 y 4 años (tienen 1,2,3 y 4)*

### 1.4- Frecuencia relativa acumulada

La frecuencia relativa acumulada es el cociente entre la frecuencia acumulada de un determinado valor y el número total de datos. Se puede expresar en tantos por ciento.