



Probabilidad y Estadística

Profesor: Sergio Díaz

Tablas de frecuencia: Conceptos básicos

Dato o intervalo: Información (variable) que se estudia en estadística

Marca de clase: Promedio entre los extremos de un intervalo

Amplitud de un intervalo: Es la diferencia entre el límite superior y el límite inferior

Frecuencia absoluta (f): Número de veces que se repite un dato

Frecuencia acumulada (F): Suma de las frecuencias absolutas de todos los valores menores o iguales al valor considerado

Tablas de frecuencia: Conceptos básicos

Frecuencia relativa (f_r): Es el cociente entre la frecuencia absoluta de uno de los valores de la variable y el total de datos. Se puede representar como fracción, número decimal o porcentaje y se denota por $f_r(\%)$, donde $f_r(\%) = 100f_r$ (probabilidad frecuentista o experimental)

Frecuencia acumulada relativa (F_r): Es la que se obtiene sumando ordenadamente las frecuencias relativas hasta la que ocupa la última posición.

Ejemplos

x	f	F	fr(%)	Fr(%)
1	5	5	20%	20%
2	8	13	B%	52%
3	A	17	16%	68%
4	8	25	32%	C%

Automóviles según su color

Color	f	f_r	f_r(%)
verde	5		
gris	8		
rojo	2		
blanco	5		

La tabla adjunta muestra la distribución de frecuencias de los ingresos familiares en miles de pesos, (agrupados por intervalos) que tienen 200 familias

Tramo	Ingreso familiar en miles de pesos	Marca de clase	Frecuencia (n° de familias)
1	[0 – 100 [67
2	[100 – 200[66
3	[200 – 300[39
4	[300 – 400[28

Medidas de tendencia Central

Son **indicadores** **que** **representan** **valores** **cuantitativos o cualitativos en** **torno a los cuales tienden a** **agruparse los valores** de una variable estadística. Las medidas de tendencia central mas utilizadas son: media aritmética, mediana y moda



Media Aritmética (\bar{x})

Es el cociente entre la suma de todos los datos y el número total de datos. Si se tienen n datos x_1, x_2, \dots, x_n :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Obs: \bar{x} se aplica sólo a variables cuantitativas. Más adelante veremos que corresponde a la esperanza experimental

Media Aritmética (\bar{x}): En tablas de frecuencia

Dato	Frecuencia
x_1	f_1
x_2	f_2
x_3	f_3
\vdots	\vdots
x_n	f_n

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \cdots + x_n \cdot f_n}{f_1 + f_2 + \cdots + f_n}$$

Estimación de la media aritmética en datos agrupados en intervalos

Marca de clase	Frecuencia
c_1	f_1
c_2	f_2
c_3	f_3
\vdots	\vdots
c_n	f_n

$$\bar{x} = \frac{c_1 \cdot f_1 + c_2 \cdot f_2 + \cdots + c_n \cdot f_n}{f_1 + f_2 + \cdots + f_n}$$

Media Aritmética

Interpretación:

- Repartición equitativa

https://geogebra.es/~ms/cvg_primaria/02/html/media.html

- "Centro de gravedad"



Media Aritmética: Propiedades

Si se tiene la variable X de n datos x_1, x_2, \dots, x_n con media aritmética \bar{x} , entonces:

- Si cada dato aumenta en a su valor, entonces la nueva media es $\bar{x} + a$
- Si cada dato se incrementa en k veces su valor, entonces la nueva media es $k\bar{x}$

En consecuencia si una variable está dada por $Y = kX + a$, entonces $\bar{y} = k\bar{x} + a$

- Si se agregan datos iguales a la media, la media permanece invariante.

La siguiente tabla de frecuencia corresponde a la estatura de 10 personas. ¿Cuál es la media aritmética de las estaturas?

Altura (m)	f
1,50	3
1,60	2
1,70	5

La siguiente tabla de frecuencia corresponde a las edades de diez personas. ¿Cuál es la media aritmética de las edades?

Edad en años	Marca de clase	f
[10, 20[2
[20, 30[1
[30, 40[
[40, 50[1
[50, 60[2

El número de computadores que hay en los hogares de un grupo de personas, viene dado en la tabla adjunta:

N° de computadores	f
0	1
1	5
2	7
3	5
4	1

- a. ¿Cuántos hogares registra la muestra?
- b. ¿Cuántos computadores en total tiene la muestra?
- c. ¿Cuál es el promedio del n° de computadores por hogar que tiene la muestra?
- d. Si en cada hogar se agregan dos nuevos computadores, ¿cuál es el nuevo promedio?
- e. Si la cantidad de computadores en cada hogar se triplica, ¿cuál es el nuevo promedio?

651) Las fichas del peso de 10 niños, marcan en promedio 20 kg. En la oficina de control se pierde una ficha y se sabe que el promedio del resto es 19 kg. ¿Cuál es el peso del niño al que le perdieron la ficha?

652) De 50 controles acumulativos, Juan lleva promedio 6,3. Si le dan la posibilidad de borrar las tres peores pruebas, que son: 3,1; 2,7 y 3,7; entonces, su nuevo promedio será:

653) En un curso de 50 personas, 25 alumnos obtuvieron promedio 5,2; 20 alumnos obtuvieron promedio 5,7 y los demás promedio 6,4. El promedio del curso fue:

657) La distribución de notas de un curso de 100 estudiantes es la indicada en la tabla. Entonces, con la información disponible, es posible estimar que el promedio aritmético de las notas es:

- A) 3,73
- B) 4,23
- C) 4,53
- D) 5,03
- E) 5,53

Intervalo	Frecuencia Absoluta
$1,5 \leq N < 2,5$	5
$2,5 \leq N < 3,5$	22
$3,5 \leq N < 4,5$	30
$4,5 \leq N < 5,5$	31
$5,5 \leq N < 6,5$	12
Total	100

Mediana (M_e)

Es el dato que **ocupa la posición central** de la muestra cuando estos se encuentran ordenados en forma creciente o decreciente. **Si la muestra tiene un número par de datos, la mediana es la media aritmética de los dos términos centrales.**

Se aplica a variables cuantitativas

Mediana: Tablas de Frecuencia

Sea n el número de datos, entonces:

- Si n es impar: $M_e = x_{\frac{n+1}{2}}$
- Si n es par: $M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$

Calcular la mediana para las siguientes tablas de frecuencia que en a) muestra las edades de un grupo de personas y en b) el número de tarjetas comerciales de un grupo de personas

a)

Edad	f
17	5
18	10
19	7
20	8
Total	

b)

Número de tarjetas comerciales	f
5	5
6	10
7	7
8	8
9	23
Total	

De los siguientes datos: $p + q$, $8p + 16q$, $10p + 20q$, $6p + 12q$, $2p + 4q$ y $4p + 8q$ enteros positivos, con $p < q$, la mediana es

Cálculo de la mediana para datos agrupados en intervalos

Para n datos agrupados en intervalos, la mediana se encuentra en el primer intervalo en que la frecuencia acumulada es mayor o igual a $\frac{n}{2}$. En intervalos del tipo $[, [$

Estimación de la mediana para datos agrupados

Donde:

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

N : Es el número total de datos.

L_i : Es el límite inferior de la clase donde se encuentra la mediana.

$\frac{N}{2}$: Es la semisuma de las frecuencias absolutas.

F_{i-1} : Es la frecuencia acumulada anterior a la clase mediana.

f_i : Es la frecuencia absoluta del intervalo mediano.

t_i : Es la amplitud de los intervalos. Se calcula restando el extremo superior menos el extremo inferior del intervalo.

Ejemplo:

N : Es el número total de datos.

L_i : El es límite inferior de la clase donde se encuentra la mediana.

$\frac{N}{2}$: Es la semisuma de las frecuencias absolutas.

F_{i-1} : Es la frecuencia acumulada anterior a la clase mediana.

f_i : Es la frecuencia absoluta del intervalo mediano.

t_i : Es la amplitud de los intervalos. Se calcula restando el extremo superior menos el extremo inferior del intervalo.

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

DATOS	f	F
[10, 15[20	20
[15, 20[25	45
[20, 25[45	90
[25, 30[10	100
Total	100	

Ejemplo:

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

DATOS	f	F
[10, 15[20	20
[15, 20[25	45
[20, 25[45	90
[25, 30[10	100
Total	100	

Intervalo de la mediana : [20 , 25[

$$\frac{N}{2} = \frac{100}{2} = 50$$

Desarrollo

$$Me = 20 + \frac{50-45}{45} \cdot 5$$

$$Me = 20 + \frac{5}{45} \cdot 5$$

$$Me = 20 + \frac{25}{45} = 20 + 0,5\bar{5} = 20,5\bar{5}$$

Moda

Es el dato o los datos que presentan la mayor frecuencia absoluta. La muestra puede ser:

- Amodal
- Unimodal
- Bimodal
- Polimodal

Está definida tanto para variables cuantitativas como cualitativas

Intervalo Modal

Es el intervalo que presenta la mayor frecuencia absoluta.

Estimación de la moda para datos agrupados en un intervalo modal

Todos los intervalos son de igual amplitud, donde:

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.

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

Ejemplo

Todos los intervalos son de igual amplitud, donde:

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.

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

Datos	f
[10, 15[20
[15, 20[25
[20, 25[45
[25, 30[10
Total	100

Ejemplo

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

Datos	f
[10, 15[20
[15, 20[25
[20, 25[45
[25, 30[10
Total	100

Desarrollo

$$Mo = 20 + \frac{(45-25)}{(45-25) + (45-10)} \cdot 5$$

$$Mo = 20 + \frac{20}{20 + 35} \cdot 5$$

$$Mo = 20 + \frac{100}{55} = 20 + 1,81 = 21,81$$

Observación

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

Ejercicios

654) Los datos corresponden al número de alfajores que se venden diariamente en un quiosco durante 18 días. De las siguientes afirmaciones ¿Cual(es) es(son) verdadera(s)?

31	22	13	19	6	31	9	19	16
7	22	25	11	28	18	30	15	31

- I) La moda es menor que la mediana y que la media
- II) La media es menor que la moda y la mediana
- III) La media es mayor que la mediana.

- A) Solo I
- B) Solo II
- C) Solo III
- D) Solo I y III
- E) Ninguna

655) La tabla de distribución de frecuencias de la figura corresponde a las estaturas de un grupo de 100 personas. $C = \text{clase}$, $f = \text{frecuencia}$. La moda, media y mediana se encuentran, respectivamente, en las clases:

- A) b, b, b
- B) c, b, c
- C) b, d, c
- D) b, c, c
- E) c, c, b

C	Estatura (cm)	f
a	$[1,2 - 1,4[$	10
b	$[1,4 - 1,6[$	34
c	$[1,6 - 1,8[$	28
d	$[1,8 - 2,0[$	24
e	$[2,0 - 2,2]$	4

650) La tabla adjunta muestra las edades de 220 alumnos de un colegio. ¿Cuál(es) de las siguientes afirmaciones es(son) verdadera(s)?

Edad (años)	15	16	17	18	19
Alumnos	50	40	60	50	20

- I) La moda es 17 años
- II) La mediana es mayor que la media
- III) La mitad de los alumnos del colegio tiene 17 o 18 años.

- A) Solo I
- B) Solo II
- C) Solo I y II
- D) Solo II y III
- E) I, II y III

656) Si las notas de Esteban en una asignatura son: 3, 4, 6, 3, 5, 5, 6, 3, 4 y de estas notas se cambia un 6 por un 7. ¿Cuál(es) de las siguientes medidas de tendencia central cambia(n)?

- I) La moda
- II) La mediana
- III) La media aritmética

- A) Solo II
- B) Solo III
- C) Solo I y III
- D) Solo II y III
- E) Ninguna

658) La siguiente tabla muestra los valores de una variable X y sus respectivas frecuencias. ¿Cuál es el valor de la mediana?

- A) 5,5
- B) 6
- C) 6,5
- D) 7
- E) 7,5

X	Frecuencia
4	4
5	8
6	10
7	20
8	8

659) De acuerdo a la siguiente muestra: $a + 2$, $a + 4$, $a + 6$, $a + 6$, $a + 6$, $a + 4$, $a + 2$, la suma de la mediana y la moda es:

- A) $2(a + 6)$
- B) $2a + 10$
- C) $a + 12$
- D) $2a$
- E) $a + 2$

660) Los datos de una muestra son todos números naturales consecutivos, si no hay ningún dato repetido y la mediana de la muestra es 11,5, entonces ¿Qué cantidad de datos no puede ser?

- A) 2
- B) 4
- C) 6
- D) 7
- E) 8