

Aritmética

ESTADISTICA

DEFINICIÓN

Es la ciencia que nos proporciona un conjunto de métodos y procedimientos para la recolección, clasificación e interpretación de datos, lo cual sirve para sacar conclusiones que permitan tomar decisiones y aplicar los correctivos en caso fuera necesario.

POBLACIÓN

Es un conjunto de elementos con una característica común. Por ejemplo: todos los alumnos matriculados en los colegios MENTOR.

MUESTRA

Es una parte o subconjunto de la población. Generalmente se elige en forma aleatoria (al azar). Por ejemplo: una muestra de 40 alumnos del colegio MENTOR de Las Flores elegidos al azar.

VARIABLE ESTADÍSTICA

Es una característica de la población y puede tomar diferentes valores. Se clasifican en:

A. CUALITATIVA

Son variables cuyos valores son *cualidades* que representa la población. Por ejemplo: la variable "profesión" puede adoptar las modalidades: Ingeniero, Médico, Profesor, etc.

B. CUANTITATIVA

Son variables que pueden ser expresadas mediante números. Por ejemplo: número de alumnos matriculados, estatura, peso, edad, etc. Las variables cuantitativas pueden ser a su vez:

1. Discretas

Cuando toma valores enteros. Por ejemplo: número de alumnos, número de colegios en el distrito de Miraflores, número de hijos, etc.

2. Continuas

Cuando puede tomar cualquier valor numérico enteros o decimales. Por ejemplo: el peso, la talla, el tiempo, el sueldo, etc.

DISTRIBUCIÓN DE FRECUENCIAS

Consideramos una muestra de tamaño "n" (número de elementos de la muestra) y la variable estadística "x" que puede tomar "k" valores diferentes: $x_1, x_2, x_3, \dots, x_k$.

FRECUENCIA ABSOLUTA SIMPLE (f_i)

También llamada simplemente frecuencias es el número de veces que aparece repetido el valor " x_i ".

Se cumple:

$$f_1 + f_2 + f_3 + \dots + f_k = n$$

en notación sigma: $\sum_{i=1}^k f_i = n$

FRECUENCIA ACUMULADA (F_i)

Es la que resulta de acumular sucesivamente las frecuencias absolutas simples.

Así tenemos:

$$F_1 = f_1$$

$$F_2 = f_1 + f_2$$

$$F_3 = f_1 + f_2 + f_3$$

.

.

.

$$F_i = f_1 + f_2 + f_3 + \dots + f_i$$

FRECUENCIA RELATIVA SIMPLE (h_i)

Es el cociente de la frecuencia absoluta simple y el número total de datos. Sus valores son números reales que oscilan entre 0 y 1. La suma de todas las frecuencias relativas es igual a 1.

$$h_i = \frac{f_i}{n} \quad 0 \leq h_i \leq 1$$

FRECUENCIA RELATIVA ACUMULADA (H_i)

Es la que resulta de acumular sucesivamente las frecuencias relativas simples.

Así tenemos:

$$H_1 = h_1$$

$$H_2 = h_1 + h_2$$

$$H_3 = h_1 + h_2 + h_3$$

·

·

·

$$H_i = h_1 + h_2 + h_3 + \dots + h_i$$

Nota

Las frecuencias relativas también se pueden expresar en tanto por ciento (%), bastará con multiplicar por 100 la frecuencia relativa.

Ejemplos :

$$\begin{array}{lcl} 0,32 & \xrightarrow{\times 100} & 32\% \\ 0,07 & \xrightarrow{\times 100} & 7\% \end{array}$$

REPRESENTACIÓN DE DATOS

Los datos pueden ser representados por:

TABLAS ESTADÍSTICAS

Es un arreglo de filas y columnas en los cuales se encuentran distribuidos los datos.

Ejemplo 1:

De un grupo de 200 alumnos se obtuvo la siguiente información, respecto a sus edades.

x_i [edades]	x_i [frecuencia]
14	25
15	45
16	65
17	38
18	27
	200

x_i = Variable estadística

f_i = Frecuencia absoluta simple

GRÁFICOS ESTADÍSTICOS

Se pueden representar mediante barras o sectores circulares.

Ejemplo 2:

Con los datos del ejemplo 1, construimos los siguientes diagramas

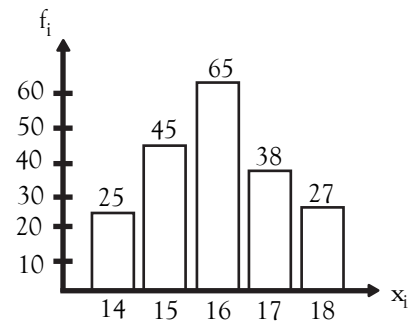
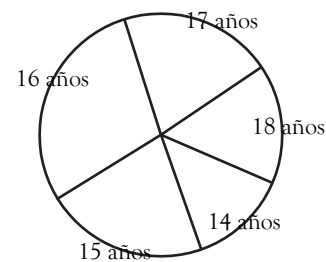


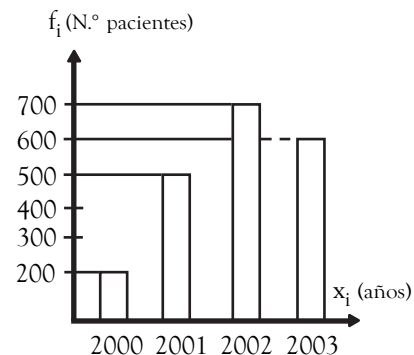
Diagrama de Barras



Sector Circular

Ejemplos prácticos:

- 1) El siguiente gráfico nos muestra el número de pacientes atendidos en un centro de salud, en los años 2000; 2001; 2002 y 2003.



Ahora construimos la tabla de datos estadísticos.

x_i	f_i	F_i	h_i	H_i
2000	200	200	$0,1 = 10\%$	$0,10 = 10\%$
2001	500	700	$0,25 = 25\%$	$0,35 = 35\%$
2002	700			
2003	600			
	$n=2000$		$1=100\%$	

Cálculo de:

$$F_1 = f_1 = 200$$

$$F_2 = f_1 + f_2 = 200 + 500 = 700$$

$$F_3 = \dots\dots\dots$$

$$F_4 = \dots\dots\dots$$

$$h_1 = \frac{f_1}{n} = \frac{200}{2000} = 0,10$$

$$h_2 = \frac{f_2}{n} = \frac{500}{2000} = 0,25$$

$$h_3 = \dots\dots\dots$$

$$h_4 = \dots\dots\dots$$

$$H_1 = h_1 = 0,10$$

$$H_2 = h_1 + h_2 = 0,10 + 0,25 = 0,35$$

$$H_3 = \dots\dots\dots$$

$$H_4 = \dots\dots\dots$$

- 2) A un seminario empresarial asistieron 80 personas y se registró las edades de los participantes en los siguientes intervalos:

Intervalos	Conteo	f_i	F_i	h_i	H_i
[20; 25>	IIII III	8	8		
[25; 30>	IIII IIII	10	18		H_2
[30; 35>	IIII IIII IIII	20		h_3	
[35; 40>	IIII IIII IIII IIII	24	F_4		$0,775$
[40; 45>	IIII IIII II	12		$0,15$	
[45; 50]	IIII I	6			
		$n=80$			

Luego de completar el cuadro interpretar los siguientes datos:

$f_4 = 24$; hay 24 personas cuyas edades varían entre 35 y 40 años.

$F_4 = 62$; hay 62 personas cuyas edades varían entre 20 y 40 años.

$h_3 = 0,25 = 25\%$; el 25% de los asistentes tienen entre 30 y 35 años.

$H_2 = 0,225 = 22,5\%$; el 22,5% de los asistentes tienen entre 20 y 30 años.

Nota

Cuando la variable toma muchos valores, como el caso anterior, imagínese hacer una tabla con cada una de las edades desde los 20 años hasta los 50 años, entonces la variable se agrupa adecuadamente en intervalos.

Al punto medio de cada intervalo se denomina **Marca de clase**, que es un valor representativo para el intervalo.

$$[20; 25>: x_1 = \frac{20+25}{2} = 22,5$$

[1.ª Marca de clase]

$$[25; 30>: x_2 = \frac{25+30}{2} = 27,5$$

[2.ª Marca de clase]

MÉTODO PARA DETERMINAR EL NÚMERO DE INTERVALOS PARA UNA VARIABLE CONTINUA

A continuación se muestra las notas obtenidas por 40 alumnos de un aula en el último Examen Bimestral de Aritmética.

10 15 11 08 12 10 13 10 12 10
12 17 10 12 11 14 15 20 10 12
10 20 14 13 06 16 06 06 14 18
07 05 12 11 02 04 14 18 16 17

1. DETERMINACIÓN DEL RANGO (R)

Es la diferencia entre el mayor valor y el menor valor de los datos:

$$R = \text{Máximo} - \text{Mínimo}$$

Del ejemplo: $R = 20 - 02$

$$R = 18$$

2. DETERMINACIÓN DEL NÚMERO DE INTERVALOS (k)

Consiste en dividir el rango en un número conveniente de intervalos, llamados también "**Intervalos de clase**". Estos intervalos son generalmente del mismo tamaño. Podemos aplicar las siguientes alternativas:

- a) Si "n" es el número de datos, entonces $k = \sqrt{n}$; en el ejemplo:

$$n = 40 \rightarrow k = \sqrt{n} = \sqrt{40} = 6,3$$

Puede considerarse 6; 7 y 8 intervalos.

b) Si "n" es el número de datos, entonces:

$$k = 1 + 3,3 \log n$$

en el ejemplo: $n = 40$, entonces

$$k = 1 + 3,3 \log 40 = 6,28; \text{ puede considerarse } 6; 7 \text{ y } 8 \text{ intervalos.}$$

Los dos métodos no dan el posible número de intervalos, la elección es arbitraria. Tomaremos en este caso: $k = 6$ intervalos, porque el rango es $R = 18$ y nos daría una cantidad exacta.

3. DETERMINACIÓN DEL TAMAÑO DE LOS INTERVALOS (C)

Dividimos el rango (R) entre el número de intervalos (k). También se le denomina **amplitud de clase**.

$$C = \frac{R}{k}$$

Del ejemplo:

$$C = \frac{R}{k} = \frac{18}{6} = 3$$

4. DETERMINACIÓN DE LOS LÍMITES DE LOS INTERVALOS

Generalmente el límite inferior del primer intervalo es el menor de los datos, luego se agrega la amplitud de clase (C) para obtener el límite superior del intervalo.

En el ejemplo:

Mínimo = 02 [límite inferior]

$$C = 3$$

$$02 + 03 = 05 [\text{límite superior}]$$

1.º intervalo: [02; 05>

2.º intervalo: [05; 08>

Finalmente tendremos:

(Realiza el conteo y completa el siguiente cuadro)

Intervalos	Conteo	f_i	F_i	h_i	H_i
[02; 05>					
[05; 08>					
[08; 11>					
[11; 14>					
[14; 17>					
[17; 20>					

$n=40$

Ejercicios Resueltos

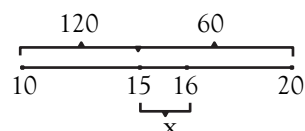
1) Las notas de un grupo de 400 alumnos fueron:

Notas	Cantidad
< 0 - 5]	100
< 5 - 10]	120
< 10 - 15]	120
< 15 - 20]	6

Indica qué porcentaje de alumnos obtuvo una nota comprendida en el intervalo < 10; 16]

- a) 32% d) 35%
b) 33% e) 36%
c) 34%

Resolución:



Por proporciones:

$$\frac{x}{1} = \frac{60}{5} \rightarrow x = 12$$

Me piden: $120 + 12 = 132$

Total de alumnos:

$$100 + 120 + 120 + 60 = 400$$

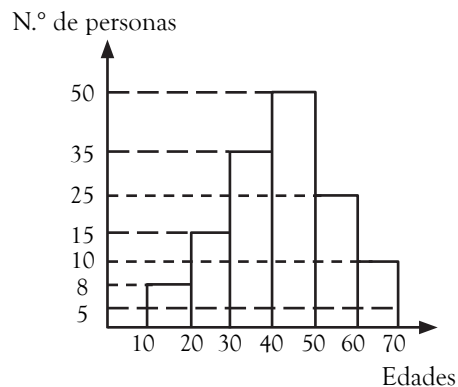
$$400 \longrightarrow 100\%$$

$$132 \longrightarrow x$$

$$x = 33\%$$

Rpta.: b

2) En el siguiente histograma



Determina el tamaño de la muestra.

- a) 140 b) 143 c) 141
d) 145 e) 142

Resolución:

Llevamos el histograma a un cuadro estadístico.

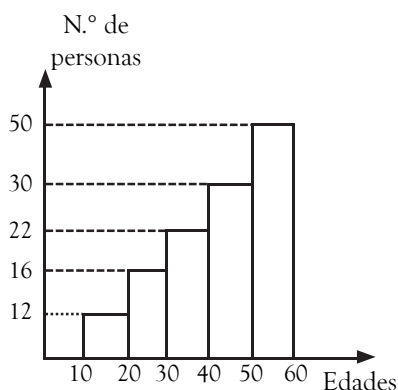
Intervalos	Frecuencias
[10 - 20>	8
[20 - 30>	15
[30 - 40>	35
[40 - 50>	50
[50 - 60>	25
[60 - 70>	10

Tamaño de la muestra

$$8 + 15 + 35 + 50 + 25 + 10 = 143$$

Rpta.: b

- 3) En el siguiente diagrama escalonado referente a las edades de un grupo de personas.



¿Cuántas personas son mayores de 29 años?

- a) 32 d) 36
b) 33 e) 38
c) 34

Resolución:

Llevando el diagrama escalonado a un cuadro estadístico.

[$L_i - L_s$ >	f_i	F_i
[10 - 20>	12	12
[20 - 30>	4	16
[30 - 40>	6	22
[40 - 50>	8	30
[50 - 60>	20	50

Me piden:

$$6 + 8 + 20 = 34$$

Rpta.: c

- 4) Dado el siguiente cuadro estadístico sobre los ingresos anuales de cierto número de personas.

[$L_i - L_s$ >	x_i	f_i
[20 - 30>		10
[30 - 40>		
[40 - 50>		
[50 - 60>		10

Además:

$$\sum_{i=1}^4 x_i f_i = 42 \quad ; \quad \frac{f_2}{f_3} = \frac{1}{3}$$

Calcula el número de familias con ingresos no menos de 30 soles.

- a) 50 d) 80
b) 60 e) 90
c) 70

Resolución:

Hallando las marcas de clase y completando las frecuencias.

[$L_i - L_s$ >	x_i	F_i	$x_i f_i$
[20 - 30>	25	10	250
[30 - 40>	35	k	35k
[40 - 50>	45	13	135k
[50 - 60>	55	10	550

Por dato: $\frac{800 + 170k}{20 + 4k} = 42$

$$k = 20$$

Me piden: $k + 3k + 10 = 90$

Rpta.: e

Resolviendo en clase

1 ¿Cuántas de las siguientes variables estadísticas son cualitativas?

- Edad
- Nacionalidad
- Profesión
- Años de servicio
- Horas trabajadas

Resolución:

Rpta:

2 ENUNCIADO 1:

En una encuesta a 30 alumnos se obtuvo los siguientes datos, respecto a sus pesos en kilogramos:

54	42	58	64	70	46
46	52	62	66	58	47
45	40	56	55	64	66
54	52	48	61	63	60
47	58	52	54	57	56

Determina el rango (R).

Resolución:

Rpta:

3 Del ENUNCIADO 1:

El posible número de intervalos (k) es:

Resolución:

Rpta:

4 Del ENUNCIADO 1:

Si consideramos el número de intervalos $k = 6$, ¿cuál será los límites del primer intervalo?

Resolución:

Rpta:

5 ENUNCIADO 2:

Considera la tabla siguiente:

[10; 20>	a	b	0,1	c
[20; 30>	d	e	f	g
[30; 40>	h	i	0,3	j
[40; 50>	24	k	m	0,85
[50; 60>	30	p	q	r

Determina $a+b$

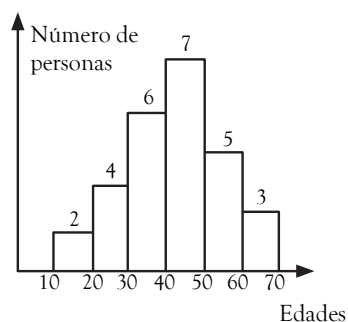
Resolución:

Rpta:

Ahora en tu cuaderno

7. ENUNCIADO 3:

Se hizo una encuesta sobre el número de personas aficionadas a la lectura y se las clasificó por edades y se obtuvo el siguiente histograma.



Determina el tamaño de la muestra.

8. Del ENUNCIADO 3:

Determina la cantidad de personas menores de 58 años aficionados a la lectura.

6 Del ENUNCIADO 2:

Determina " $d - h + g$ "

Resolución:

Rpta:

9. Del ENUNCIADO 3:

¿Cuántas personas que no sean mayores de 60 años ni menores que 30 años son aficionados a la lectura?

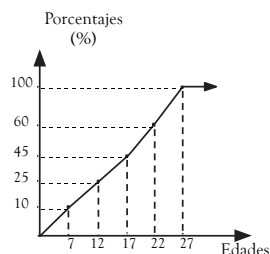
10. Se tiene las temperaturas observadas en el hemisferio norte durante 24 días.

°Centígrados	f_i	h_i
[-19; - 17>		
[-17; - 15>	2	
[-15; - 13>	8	
[-13; - 11>		0,125
[-11; - 9>	4	
[-9; - 7>		0,2083

¿Durante cuántos días se obtuvo una temperatura de -16°C a -10°C ?

11. ENUNCIADO 4:

En el siguiente cuadro muestra la ojiva de la frecuencia relativa acumulada de las edades de cierto número de alumnos.



¿Qué porcentaje de alumnos es mayor de edad y menor de 23 años?

12. Del ENUNCIADO 4:

¿Qué porcentaje de alumnos tienen edades comprendidas entre 14 y 25 años inclusive?

Para reforzar

1. ¿Cuántas de las siguientes variables estadísticas son cuantitativas continuas?

- Estatura
- Peso
- Número de cursos
- Número de hijos
- Sueldo

- a) 1
- b) 4
- c) 2
- d) 5
- e) 3

* ENUNCIADO 1:

Se tomó una evaluación a un grupo de alumnos de cuarto año y los resultados obtenidos fueron:

97 80 75 120 92 78 105 82 79 87
82 92 105 81 76 70 84 87 91 84

2. Del ENUNCIADO 1:

Determina el rango (R).

- a) 40
- b) 55
- c) 35
- d) 60
- e) 50

3. Del ENUNCIADO 1:

El posible número de intervalos es:

- a) 6; 7
- b) 2; 3
- c) 7; 8
- d) 8; 9
- e) 4; 5

4. Del ENUNCIADO 1:

Si consideramos como número de intervalos $k = 5$, ¿cuál será los límites del último intervalo?

- a) [115; 120>
- b) [112; 120>
- c) [116; 120>
- d) [105; 120>
- e) [110; 120>

* ENUNCIADO 2:

La distribución de frecuencias mostradas corresponde a los pesos de 60 paquetes registrados en una empresa de encomiendas.

Pesos	f_i	F_i	h_i	H_i
[250; 400>	6			
[400; 550>	12			
[550; 700>				
[700; 850>	18			
[850; 1000>	6			
[1000; 1050>	3			

5. Del ENUNCIADO 2:

¿Cuál es el intervalo con la mayor frecuencia absoluta?

- a) 1.° b) 4.° c) 2.°
d) 5.° e) 3.°

6. Del ENUNCIADO 2:

Halla " $F_3 + f_4$ ".

- a) 51 b) 42 c) 33
d) 55 e) 48

7. Del ENUNCIADO 2:

Halla " $h_1 + h_4 + H_2$ ".

- a) 0,5 b) 0,7 c) 0,6
d) 1,2 e) 0,3

* ENUNCIADO 3:

Se muestra la siguiente tabla de distribución del número de trabajadores de un Ministerio, de acuerdo a su ocupación.

x_i ocupación	f_i N.º de personas	F_i	h_i
Administradores	120		
Ingenieros	50		
Abogados	80		
Obreros	90		
Secretarias	60		
	$n = 400$		

8. Del ENUNCIADO 3:

¿Cuál es la frecuencia relativa de los abogados?

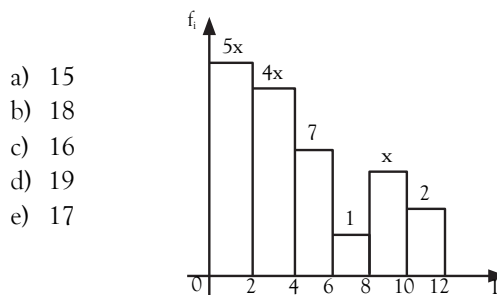
- a) 0,25 b) 0,70 c) 0,20
d) 0,80 e) 0,40

9. Del ENUNCIADO 3:

Halla " F_3 ".

- a) 200 b) 400 c) 220
d) 180 e) 250

10. Dado el siguiente histograma de frecuencias absolutas, tomados de una muestra de tamaño 40, halla: $f_1 + f_4 + f_5$.



- a) 15
b) 18
c) 16
d) 19
e) 17

11. La tabla muestra la distribución del ingreso familiar correspondiente a 60 familias.

Intervalo de Ingreso S/.	f_i	F_i	h_i
250 - 260	14		
260 - 270	16	30	
270 - 280	12		0,2
280 - 290	10		0,16
290 - 300	8		

Determina el número de familias que ganan menos de 280 soles.

- a) 22 b) 42 c) 32
d) 52 e) 34

12. Dada la siguiente distribución de frecuencias:

Puntaje	f_i
[10 - 20>	
[20 - 40>	
[40 - 50>	50
[50 - 70>	
[70 - 80>	
Total	100

determina la suma: " $h_5 + h_2$ ".

Se sabe además que: $h_1 = h_5$ y $h_2 = h_4$.

- a) 1/2 b) 1/5 c) 1/3
d) 3/4 e) 1/4