

ESTADÍSTICA I

DEFINICIÓN

Es una ciencia que nos proporciona un conjunto de métodos y procedimientos para la recolección, clasificación, organización, presentación, análisis e interpretación de datos en forma adecuada con el fin de realizar una teoría de decisiones más efectivas.

CLASES DE ESTADÍSTICA

ESTADÍSTICA DESCRIPTIVA

Es la parte de la estadística que trata de recopilar, clasificar, presentar y describir datos estadísticos.

ESTADÍSTICA INFERENCIAL

Es la parte de la estadística cuyo objeto de estudio es investigar cómo deben ser utilizados los datos para producir resultados o probar alguna hipótesis.

Observación:

La diferencia entre la estadística descriptiva y la inferencial es que la segunda usa el cálculo de la probabilidad.

Población:

Es un conjunto de datos referentes a determinadas características de un grupo de individuos o elementos.

Ejemplo:

Las edades de los alumnos de la UNI.

Muestra

Es el subconjunto tomado al azar de los elementos de una determinada población.

Ejemplo:

Las edades de los alumnos de la facultad de Mecánica.

DISTRIBUCIÓN DE FRECUENCIAS

Consiste en distribuir los datos de la muestra de clase o categorías e ir colocando el número de datos que caen en cada intervalo.

DEFINICIONES PREVIAS

ALCANCE (A):

Es el intervalo definido por los datos extremos (mayor y menor valor).

RANGO O RECORRIDO (R):

Es la longitud de alcance que resulta de la diferencia entre el mayor y menor valor.

INTERVALO DE CLASE O CATEGORÍA (L_i):

Son grupos que resultan de dividir el alcance o recorrido; el número de grupos (K) se determina por la regla propuesta por Sturges.

$$K = 1 + 3,32 \log n$$

(redondeando al entero superior e inferior según convenga).

Donde:

n: número total de datos disponibles.

ANCHO DE CLASE (W):

Es la diferencia que hay entre los extremos de cada intervalo de clase.

Ejemplo:

Sea el intervalo

$$[L_i ; L_{i+1}^* \\ W = L_{i+1} - L_i$$

también :

$$W = \frac{R}{K}$$

MARCA DE CLASE (x):

Son los puntos medios de los intervalos de clase.

Ejemplo:

Sea el intervalo $[L_i; L_{i+1}^*]$

$$x_i = \frac{L_i + L_{i+1}^*}{2}$$

Problema aplicativo:

Se tiene los pesos de 50 estudiantes de la UNI con una aproximación de 1 kg.

73	67	67	60	61
67	57	59	57	77
69	76	52	69	72
76	77	94	77	93
79	70	68	72	63
47	82	70	67	80
70	85	70	73	58
58	67	68	66	86
79	88	67	54	56
64	46	63	84	74

Calculando el alcance:

Dato mayor : 94

Dato menor : 46

$$A = [46; 94]$$

Calculando el Rango:

$$R = 94 - 46 = 48$$

Calculando el número de intervalos:

Si $n = 50$; ($n =$ número de datos)

$$K = 1 + 3,32 \log(50) = 6,61 \text{ E } 7$$

Calculando el ancho de clase:

$$W = \frac{R}{K} = \frac{48}{7} = 6,86 \text{ E } 7$$

Con los siguientes datos encontrados haremos una tabla de distribución de frecuencia.

$L_i; L_{i+1}$	Tabulación	f_i	F_i	h_i	H_i	x_i
[46 ; 53 *]		3	3	$\frac{3}{50}$	$\frac{3}{50}$	49,5
[53 ; 60 *]		7	10	$\frac{7}{50}$	$\frac{10}{50}$	56,5
[60 ; 67 *]		6	16	$\frac{6}{50}$	$\frac{16}{50}$	63,5
[67 ; 74 *]		19	35	$\frac{19}{50}$	$\frac{35}{50}$	70,5
[74 ; 81 *]		8	43	$\frac{8}{50}$	$\frac{43}{50}$	77,5
[81 ; 88 *]		4	47	$\frac{4}{50}$	$\frac{47}{50}$	84,5
[88 ; 95 *]		3	50	$\frac{3}{50}$	$\frac{50}{50}$	91,5

FRECUENCIA ABSOLUTA (f_i):

Es el número de datos que caen dentro de cada intervalo de clase.

FRECUENCIA RELATIVA (h_i):

Viene a ser el cociente entre la frecuencia absoluta y el número total de datos.

$$h_i = \frac{f_i}{n}$$

FRECUENCIA ABSOLUTA ACUMULADA (F_i):

Es aquella que resulta de sumar sucesivamente las frecuencias absolutas.

Ejemplo:

Suponiendo «k» intervalos.

$$F_1 = f_1$$

$$F_2 = f_1 + f_2$$

$$F_3 = f_1 + f_2 + f_3$$

⋮

$$F_{(k)} = f_1 + f_2 + f_3 + \dots + f_k = n$$

Donde n : número de intervalos.

$$F_i = \sum_{j=1}^i f_j$$

Donde : $i = 1, 2, 3, \dots, k$

FRECUENCIA RELATIVA ACUMULADA (H)

Es aquella que resulta de sumar sucesivamente las frecuencias relativas.

$$H_1 = h_1$$

$$H_2 = h_1 + h_2$$

$$H_3 = h_1 + h_2 + h_3$$

⋮

$$H_{(k)} = h_1 + h_2 + h_3 + \dots + h_k = 1$$

$$H_i = \sum_{j=1}^i h_j$$

Donde : $i = 1, 2, 3, \dots, k$

Calculando las frecuencias absolutas

Del conjunto de datos se puede observar cuantos de éstos caen en cada intervalo de la distribución de frecuencias; este número de datos se irá colocando en sus respectivos casilleros hasta llenar toda la columna.

Calculando las frecuencias absolutas acumuladas

$$\begin{aligned} F_1 &= 3 \\ F_2 &= 3 + 7 \\ F_3 &= 3 + 7 + 6 \\ F_4 &= 3 + 7 + 6 + 19 \\ F_5 &= 3 + 7 + 6 + 19 + 8 \\ F_6 &= 3 + 7 + 6 + 19 + 8 + 4 \\ F_7 &= 3 + 7 + 6 + 19 + 8 + 4 + 3 \end{aligned}$$

Calculando las frecuencias relativas

$$h_1 = \frac{3}{50} \quad h_2 = \frac{7}{50}$$

$$h_3 = \frac{6}{50} \quad h_4 = \frac{19}{50}$$

$$h_5 = \frac{8}{50} \quad h_6 = \frac{4}{50}$$

$$h_7 = \frac{3}{50}$$

Calculando las frecuencias relativas acumuladas

$$H_1 = \frac{3}{50}$$

$$H_2 = \frac{3}{50} + \frac{7}{50}$$

$$H_3 = \frac{3}{50} + \frac{7}{50} + \frac{6}{50}$$

$$H_4 = \frac{3}{50} + \frac{7}{50} + \frac{6}{50} + \frac{19}{50}$$

$$H_5 = \frac{3}{50} + \frac{7}{50} + \frac{6}{50} + \frac{19}{50} + \frac{8}{50}$$

$$H_6 = \frac{3}{50} + \frac{7}{50} + \frac{6}{50} + \frac{19}{50} + \frac{8}{50} + \frac{4}{50}$$

$$H_7 = \frac{3}{50} + \frac{7}{50} + \frac{6}{50} + \frac{19}{50} + \frac{8}{50} + \frac{4}{50} + \frac{3}{50}$$

GRÁFICOS

Histograma: Son diagramas de barras o rectángulos cuyas bases son los intervalos de clase y las alturas las frecuencias absolutas o relativas.

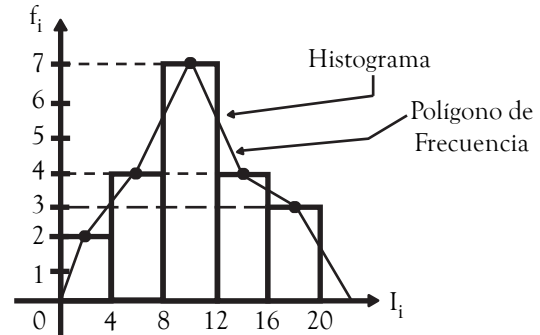


GRÁFICO DE SECTORES

Preferencia	N.º de Personas	Expresado en grados	Expresado en porcentajes
Aritmética (A)	60	144°	40 %
Álgebra (X)	20	48°	13,3 %
Geometría (G)	30	72°	20 %
Trigonometría (T)	40	96°	26,7 %
Total	150		

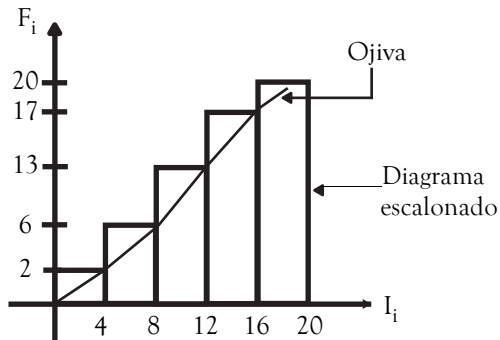
En grados:

$$\begin{array}{l} 150 \quad 360^\circ \\ \quad \quad \quad x \\ 60 \end{array} \Rightarrow x = \frac{60 \cdot 360^\circ}{150} = 144^\circ$$

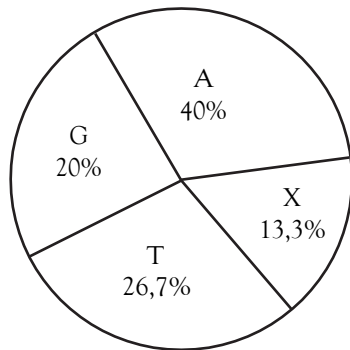
En porcentaje:

$$\begin{array}{l} 360^\circ \quad 100\% \\ \quad \quad \quad y \\ 144^\circ \end{array} \Rightarrow y = \frac{144^\circ \cdot 100\%}{360^\circ} = 40\%$$

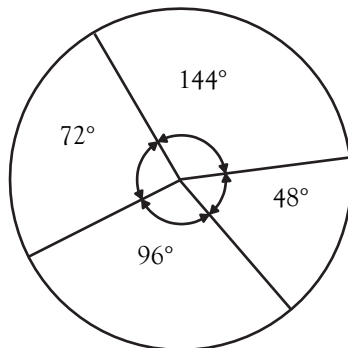
Diagrama Escalonado: Son diagramas similares al histograma, con la diferencia de que las alturas son frecuencias absolutas o relativas acumuladas.



Porcentajes



Grados



Ejercicios Resueltos

1) En una empresa se realizó una encuesta sobre las edades de los empleados, obteniéndose:

Edades	N° de empleados
[25; 30>	60
[30; 35>	75
[35; 40>	120
[40; 45>	85
[45; 50>	60

Donde A es el porcentaje de empleados con 30 años o más y B es el porcentaje de empleados con menos de 45 años. Halla: A+B.

Resolución:

Empleados con más de 30 años:
 $f_2 + f_3 + f_4 + f_5 = 75 + 120 + 85 + 60 = 340$

⊗ $A = \frac{340}{400} \times 100\% = 85\%$

Empleados con menos de 45 años:
 $f_1 + f_2 + f_3 + f_4 = 60 + 75 + 120 + 85 = 340$

⊗ $B = \frac{340}{400} \times 100\% = 85\%$

⊗ $A+B = 85\% + 85\% = 170\%$

Rpta.: 170%

2) Dada la siguiente distribución de frecuencias en base al ingreso familiar de 450 familias:

Ingreso	f	F
[; >		a
[240; >	80	
[; >	2a+40	5a
[; >		
[; 400>	a-20	

Si el ancho de clase es constante, ¿cuántas familias tienen un ingreso comprendido entre 300 y 380 soles?

Resolución:

Se cumple: $240 + 4w = 400$
 $w = 40$

También: $f_1 = F_1 = a$
 $F_3 = f_1 + f_2 + f_3$

⊗ $5a = a + 80 + 2a + 40$
 $a = 40$

También: $f_1 = 450$
 $i=1$

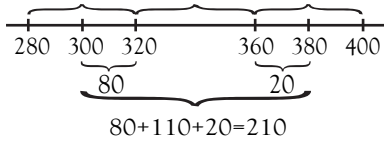
⊗ $60 + 80 + 160 + f_4 + 40 = 450$
 $f_4 = 110$

En la tabla:

Ingreso	f_i
[200; 240>	60
[240; 280>	80
[280; 320>	160
[320; 360>	110
[360; 400>	40

} Entre 300 y 380 soles.

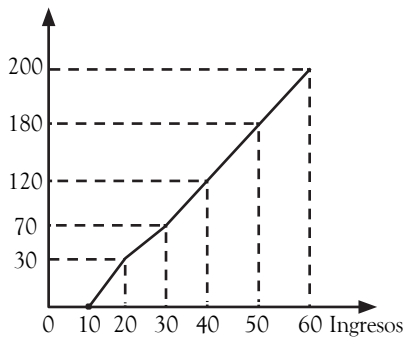
$$f_3 = 160 \quad f_4 = 110 \quad f_5 = 40$$



Rpta.: 210

3) La siguiente ojiva muestra las frecuencias absolutas acumuladas correspondientes al ingreso diario (en soles) de un cierto número de empleados.

N.º de empleados



¿Cuántos empleados ganan entre 20 y 45 soles?

Resolución:

De la ojiva se forma la tabla:

Sueldos	F_i	f_i
[10; 20>	30	30
[20; 30>	70	40
[30; 40>	120	50
[40; 50>	180	60
[50; 60>	200	20

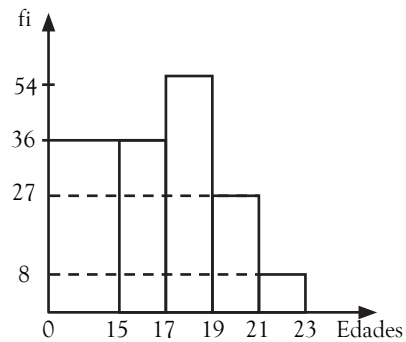
En [20;40> hay $40+50=90$, lo que corresponde a [40;45> lo averiguamos mediante una interpolación.

$$\left. \begin{array}{l} [40;45> \square x \\ [40;50> \square 60 \end{array} \right\} x = \frac{60(45-40)}{50-40} = 30$$

Entonces entre 20 y 45 hay $90+30=120$

Rpta.: 120

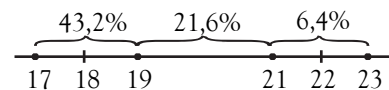
de las edades de los ingresantes a cierta facultad. ¿Qué porcentaje de ingresantes tienen entre 18 y 22 años?



Resolución:

I	f_i	h_i (%)
[15; 17>	36	28,8%
[17; 19>	54	43,2%
[19; 21>	27	21,6%
[21; 23>	8	6,4%

Total 125



De 18 a 22 se tendrá:

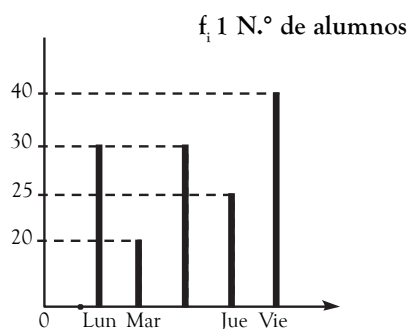
$$\frac{43,2\%}{2} + 21,6\% + \frac{6,4\%}{2} = 46,4\%$$

Rpta.: 46,4%

Resolviendo en clase

1 ENUNCIADO 1:

El siguiente diagrama muestra el número de alumnos que llegaron tarde la semana pasada al colegio.



¿Cuántas tardanzas se registraron en toda la semana?

Resolución:

Rpta:

2 Del ENUNCIADO 1:

¿Qué porcentaje del total de tardanzas se registró el día viernes?

Resolución:

Rpta:

3 Del ENUNCIADO 1:

El número de tardanzas del viernes, con respecto al día jueves, ¿en qué porcentaje es mayor?

Resolución:

Rpta:

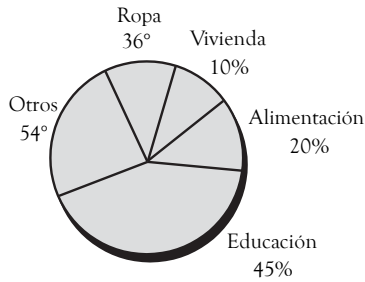
4 Al averiguar el número de hijos de 20 familias se obtuvo los siguientes datos:

3, 2, 5, 0, 1, 3, 2, 1, 0, 1, 3, 4, 2, 4, 4, 3, 4, 3, 2, 3
Determina qué porcentaje de familias tiene dos hijos más, pero menos de cuatro.

Resolución:

Rpta:

5 Para el siguiente gráfico:



¿Qué se puede afirmar si la persona tiene un ingreso de S/. 300?

- I. La persona gasta S/. 135 en educación.
- II. Gasta igual en vivienda y en ropa.
- III. En alimentación gasta S/.50.

Resolución:

Rpta:

Ahora en tu cuaderno

7. La tabla muestra la distribución de frecuencias (f_i) de un conjunto de niños, sobre sus edades:

Edades (años)	f_i
6	4
8	6
10	2
12	7

- ¿Cuántos niños tienen 8 años?
- ¿Cuántos niños tienen más de 8 años?
- ¿Cuántos niños tienen por lo menos 10 años?
- ¿Cuál es la frecuencia absoluta de los niños que tienen 12 años?

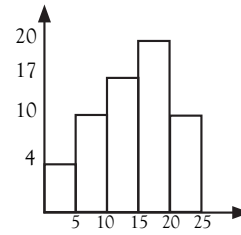
8. La inversión anual, en miles de dólares, de una muestra de 40 pequeñas empresas fueron:

31 17 27 20 28 10 34 25 4 24
 15 39 18 30 41 26 12 46 18 23
 36 19 29 37 33 27 24 26 27 31
 25 28 33 28 22 23 31 29 35 21

Si se construye una distribución de frecuencias de 7 intervalos de clase.

- I. Determina la mayor frecuencia.
- II. Calcula qué porcentaje de empresas invirtieron entre 14 mil y 20 mil dólares.

6 Observa el siguiente histograma de frecuencias y completa la tabla de frecuencias que está debajo.



Clases	x_i	f_i	F_i
0 - 5			
5 - 10			
10 - 15			
15 - 20			
20 - 25			

Calcula $x_3 + f_5 + F_4 + n$.

Resolución:

Rpta:

9. El siguiente cuadro corresponde al ingreso semanal (en soles) de cierto número de obreros. Calcula cuántos empleados se estima que ganan entre S/.125 y S/.260.

I	f_i	h_i	F_i	H_i
[100, 150>	50			
[150, 200>		0,30		
[200, 250>				0,95
[250, 300>	5			

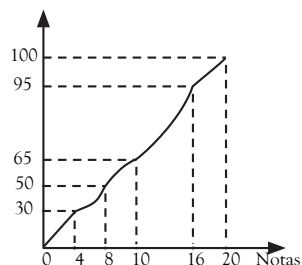
10. Dada la siguiente distribución de frecuencia halla: $f_1 + f_3 + F_4$

I_i	f_i	F_i	h_i	H_i
[10, 20>			0,1	
[20, 30>				
[30, 40>	24		0,3	
[40, 50>	30			0,85
[50, 60>				

11. En la siguiente tabla de distribución simétrica de frecuencias, de ancho de clase constante, calcula:
 $a+d+e+q+k+m+f_3+x_1$

I_i	x_i	f_i	F_i
$[a ; b >$			
$[c ; d >$		10	
$[80 ; 100 >$		m	
$[e ; f >$		2k	
$[p ; q >$		k	
$n = 60$			

12. El siguiente cuadro muestra la ojiva de las frecuencias relativas acumuladas de las notas de un examen de ingreso a la U.N.M.S.M. Determina qué tanto por ciento de alumnos tuvieron una nota entre 9 y 15.



Para reforzar

- * ENUNCIADO 1:
 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		
$n=400$			

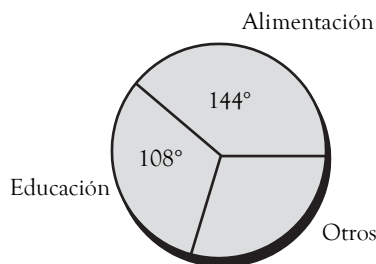
Completa la tabla.

- Del ENUNCIADO 1:
 ¿Cuál es la frecuencia relativa de los abogados?
 a) 0,25 b) 0,20 c) 0,40
 d) 0,70 e) 0,80
- Del ENUNCIADO 1:
 Halla el porcentaje de administradores.
 a) 30% b) 40% c) 25%
 d) 50% e) 20%
- Del ENUNCIADO 1: Halla F_3 .
 a) 200 b) 220 c) 250
 d) 400 e) 180

4. Dados los datos: 10; 4; 5; 7; 7; 4; 6; 10; 5; 10; 6
 Calcula:

- La frecuencia del 4.
 Rpta.: _____
- La frecuencia del 6.
 Rpta.: _____
- La frecuencia absoluta del 10.
 Rpta.: _____

5. Dado el gráfico:
 (presupuesto de un estudiante)



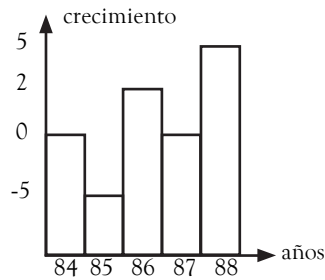
Se afirma:

- Destina el 40% a la educación.
- Destina un 30% a la alimentación.
- El 30% lo dedica a otras actividades.

Son ciertas:

- Sólo I b) I y II c) Sólo II
 d) II y III e) Sólo III

6. Del siguiente gráfico:



Se afirma:

- I. El mayor crecimiento ocurre durante el año 88.
- II. Tiene dos años con iguales crecimientos.
- III. Su mínimo crecimiento se da en el año 86.

Son ciertas:

- a) Sólo I
- b) Sólo II
- c) Sólo III
- d) I y II
- e) II y III

7. La tabla muestra la distribución de frecuencias de 50 familias según el número de hijos que poseen.

N.º de hijos	f_i
0	8
1	12
2	m
3	n
4	20

¿Cuántas familias tiene un solo hijo?

- a) 20
- b) 12
- c) 15
- d) 10
- e) 9

8. Del siguiente cuadro acerca del número de entenados que tienen los casados con viudas.

N.º Entenados	f_i	h_i
2	20	0,16
3	x	
4	15	
5	30	

Halla x.

- a) 30
- b) 60
- c) 20
- d) 50
- e) 70

9. Completa la siguiente tabla de distribución de frecuencias e indica qué tanto por ciento del total tienen edades desde 20 hasta 33 años.

Edades	f_i	h_i	F_i
[12, 18>			0,10
[, 24>		0,30	
[, 30>	40		
[, 36>	20		

- a) 40%
- b) 50%
- c) 60%
- d) 70%
- e) 80%

10. De la siguiente tabla de distribución de frecuencias, calcula: $f_2 \cdot f_1 + n$

Clases	f_i	h_i	F_i	H_i
[10, 20>		0,1		
[20, 30>				
[30, 40>		0,3		
[40, 50>	25			0,8
[50, 60>	20			

- a) 102
- b) 103
- c) 104
- d) 105
- e) 106

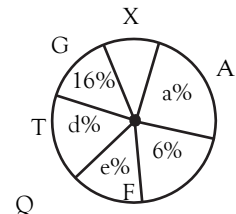
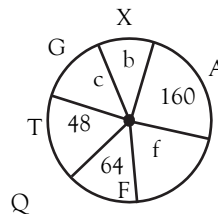
11. La distribución de edades de 100 trabajadores es:

l_i	f_i	F_i	h_i	H_i
[21 ; 27 >		8		
[27 ; 33 >			0,1	
[33 ; 39 >				0,3
[39 ; 45 >				
[45 ; 51 >			0,25	

¿Cuántas personas tienen menos de 33 años?

- a) 21
- b) 20
- c) 18
- d) 16
- e) 25

12. En el siguiente pictograma se muestra las preferencias de la cantidad de alumnos referentes a los cursos A; X; G; T; F y Q; así mismo se indica los porcentajes: Calcula $a+b+c+d+e+f$.



- a) 135
- b) 136
- c) 137
- d) 138
- e) 139