



DEFINICIONES DE LA ESTADÍSTICA

DEFINICIÓN

Es la ciencia que nos proporciona un conjunto de métodos y procedimientos para la recolección, clasificación análisis e interpretación de datos.

Definiciones previas

1. Población

Es la unidad de elementos del conjunto estudiado, en la cual se presentarán determinadas características, (edades, pesos, estatura de los habitantes de una ciudad, el número de artículos defectuosos producidos por una industria, etc.). Dependiendo del número de elementos que forman una población, esta puede ser finita o infinita.

Ejemplos:

- ❖ El número de habitantes que viven en una ciudad (finito).
- ❖ Cantidad de número en el conjunto de los números reales (infinito).

2. Muestra

Es una parte de la población que será sometida a un estudio. Se suelen tomar muestras cuando es difícil o costosa la observación de todos los elementos de la población estadística.

3. Dato estadístico

Son números o medidas que han sido recopilados como resultado de observación.

4. Variables estadísticas

Es la característica de la población que interesa al investigador y que puede tomar distintos valores. Las variables estadísticas se denotan en las letras x , y , z , etc.

Ejemplos: peso, estatura, sexo, demanda de un producto, sintonía de un programa de TV o radio, etc.

- ❖ La variable estadística puede ser cualitativa o cuantativa.

a) Variable cualitativa

Consiste en la clasificación por categorías, no lleva clasificación numérica.

Ejemplo: La variable estado civil, puede adoptar las modalidades: soltero, casado, divorciado y viudo.

b) Variable cuantitativa

Son variables que se obtienen como resultado de mediciones o conteos.

- ❖ La variable cuantitativa puede ser directa o continua.

a) Variable discreta:

Esta variable no puede tomar cualquier valor comprendido entre otros dos enteros.

Ejemplo:

El número de miembros de una familia puede ser 4 o 5, pero no un valor entre ellos.

b) Variable continúa:

Esta variable puede tomar cualquier valor comprendido entre dos números enteros.

Ejemplo: Una persona puede pesar 70 kg o 71 kg o cualquier valor comprendido entre ellos.

Tabla de frecuencias de una variable discreta

Consiste en presentar los datos en cuadros o tablas.

A continuación, se detallarán los conceptos previos para la construcción de cuadros o tablas.

Ejemplo: Se ha encuestado a 20 familias con respecto al número de miembros que lo conforman (padres e hijos incluidos)

4	2	3	5	6
2	3	3	4	5
3	7	2	2	3
6	3	3	4	2

a) Rango o amplitud de los datos

Llamado también recorrido de los datos, el rango es la diferencia entre el mayor y el menor de los valores que forman la variable.

Del ejemplo, el rango de los datos es: $R = 7 - 2 = 5$

b) **Frecuencia absoluta de un dato (f_j)**

Es el número de veces que aparece un valor de la variable estadística.

Del ejemplo, la frecuencia absoluta de 3 es 7.

c) **Frecuencia absoluta acumulada de un dato (F_j)**

La frecuencia absoluta acumulada de un dato de la variable es la cantidad de datos hasta un determinado valor de la variable.

Del ejemplo, la frecuencia absoluta acumulada de 3 es: $5 + 7 = 12$

d) **Frecuencia relativa de un dato (h_j)**

Es el cociente entre la frecuencia absoluta del dato y el total de datos.

$$h_j = \frac{f_j}{n}$$

Del ejemplo, la frecuencia relativa de 3 es:

$$\frac{7}{20} = 0,35 \text{ o } 35\%$$

e) **Frecuencia relativa acumulada de un dato (H_j)**

Es el cociente entre la frecuencia absoluta acumulada y el total de datos.

$$H_j = \frac{F_j}{n}$$

Del ejemplo, la frecuencia acumulada de 3 es:

$$\frac{12}{20} = 0,60 \text{ o } 60\%$$

Nota:

No confundir frecuencia absoluta con frecuencia relativa. ($f_j \neq h_j$)

Resumiendo los datos de una tabla:

Número de miembros de una familia (X_j)	Número de familias f_j	F_j	h_j	H_j
2	5	5	0,25	0,25
3	7	12	0,35	0,60
4	3	15	0,15	0,75
5	2	17	0,10	0,85
6	2	19	0,10	0,95
7	1	20	0,05	1
Total	20		1	

Trabajando en clase

Integral

Enunciado:

En una encuesta realizada a 800 personas sobre la simpatía por un color (suponiendo que a ninguno le gusta 2 colores a la vez). Se obtienen los siguientes datos.

Color	f_j	h_j
Amarillo	120	c
Rojo	A	0,25
Azul	B	0,30
Negro	160	D
Verde	F	E
	$n = 800$	

1. Calcular el valor de A.
2. ¿A cuántas personas les gusta el color verde?
3. ¿Qué porcentaje del total de encuestados no simpatizan con el color amarillo?

PUCP

En una encuesta realizada a 400 alumnos de la universidad sobre qué profesiones están estudiando (suponiendo que ninguno puede estudiar 2 carreras a la vez) se obtienen los siguientes datos.

Carreras	f_j	h_j
Administración	60	
Contabilidad		0,25
Economía		0,30
Ingeniería	80	
Medicina		
	$n = 400$	

4. Completa el cuadro y determina el porcentaje de alumnos que estudian Administración.

Resolución:

Carreras	f_j	h_j
Administración	60	h_1
Contabilidad	f_2	0,25
Economía	f_3	0,30
Ingeniería	80	h_4
Medicina	f_4	h_5
	400	1

- ▶ $f_2 = 0,25 \times 400 = 100$
- ▶ $f_3 = 0,30 \times 400 = 120$
- ▶ $f_5 \rightarrow 60 + f_2 + f_3 + 80 + f_5 = 400$
 $60 + 100 + 120 + 80 + f_5 = 400$
 $\rightarrow f_5 = 40$
- ▶ $h_1 = \frac{60}{400} \Rightarrow h_1 = 0,15$
- ▶ $h_4 = \frac{80}{400} \Rightarrow h_4 = 0,20$
- ▶ $h_5 = \frac{40}{400} \Rightarrow h_5 = 0,10$

→ Luego, nos piden el porcentaje, y lo obtenemos de la frecuencia relativa con respecto a la carrera de Administración.

→ $h_1 = 0,15$

→ El porcentaje que estudian Administración es el 15%.

5. En una encuesta realizada a los alumnos del colegio pamer sobre a qué universidad preferirían postular (suponiendo que ninguno postula a 2 universidades a la vez), se obtuvo la siguiente información.

Universidad	f_j	h_j
UNMSM	430	
UNI		0,23
PUCP		0,27
AGRARIA	10	
UPC		
	1000	

Completa el cuadro y determina el porcentaje de los alumnos que quieren postular a la Universidad Agraria.

Enunciado

Respecto al siguiente cuadro sobre hinchas de equipo de fútbol.

Equipos	f_j	h_j
Universitario	40	
Alianza Lima		0,30
Cristal		
	100	

- 6. ¿Cuál es el porcentaje de hinchas a los que les gusta Cristal?
- 7. ¿Cuántos hinchas simpatizan con Alianza Lima?

UNMSM

8. Se ha encuestado a 20 familias con respecto al número de miembros que lo conforman (padres e hijos incluidos).

2	2	5	5	2
3	5	2	3	6
2	6	3	6	3
5	2	3	5	5

Resume los datos en una tabla y calcula el porcentaje que representan las familias compuestas por 5 miembros.

Resolución

Resumiendo los datos en una tabla, tenemos:

Número de miembros de una familia (X_j)	Número de familias (f_j)	F_j	h_j	H_j
2	6	6	0,30	0,30
3	5	11	0,25	0,55
5	6	17	0,30	0,25
6	3	20	0,15	1
Total	20		1	

Luego, de acuerdo con la encuesta, el número de familias compuesta por 5 miembros es 6. Calcula el porcentaje es determinar su frecuencia relativa mediante la siguiente formula:

$$h_j = \frac{f_j}{n} \Rightarrow h_{\text{miembros}} = \frac{6}{20} = 0,30$$

∴ las familias compuestas por 5 miembros representan el 30%.

9. Se ha encuestado a 30 familias con respecto al número de miembros que lo conforman (padres e hijos incluidos).

3 2 3 4 3 2
 2 2 5 5 2 3
 3 3 6 6 6 2
 2 2 3 4 4 2
 4 3 3 5 6 2

Resume los datos en una tabla y calcula el porcentaje que representan las familias conformadas por 4 miembros.

Enunciado

En una encuesta realizada a 20 personas sobre su estado civil se obtuvo la siguiente información.

C S C C S
 S C C S C
 C S C C S
 S C C S C

Donde S(soltero) y C(casado)

10. ¿Qué porcentaje representan las personas casadas?
 11. ¿Qué porcentaje de los casados representan los solteros?

UNI

12. Relaciona mediante una flecha según corresponda:

- ❖ La estatura de una persona
- ❖ Los miembros de una familia
- ❖ h_j
- ❖ F_j
- ❖ El sexo de una persona
- ❖ Variable cuantitativa discreta
- ❖ Frecuencia absoluta acumulada
- ❖ Variable cuantitativa continua
- ❖ Variable cualitativa
- ❖ Frecuencia relativa de un dato

Resolución:

- ❖ La estatura es una variable cuantitativa continua por que además de tomar valores enteros, también toma un valor comprendido entre ellas.
- ❖ La cantidad de los miembros de una familia es variable cuantitativa discreta, porque solo toma valores enteros positivos.
- ❖ h_j es la representación de la frecuencia relativa de un dato.
- ❖ F_j es la representación de la frecuencia absoluta acumulada.
- ❖ El sexo es una variable cualitativa porque señala categoría.

13. Relaciona mediante una flecha según corresponda:

- ❖ Variable cualitativa
- ❖ Variable discreta
- ❖ Rango o amplitud
- ❖ Frecuencia absoluta
- ❖ Variable continua
- ❖ Estado civil de las personas
- ❖ El peso
- ❖ Es el número de veces aparece un valor de la variable estadística.
- ❖ Números de empleados de una empresa.
- ❖ Es la diferencia entre el mayor y el menor de los valores que forman la variable.

14. Se muestra la siguiente tabla de distribución de los trabajadores en una empresa de acuerdo con su ocupación:

Ocupación	# de personas
Abogados	18
Ingenieros	32
Obreros	40
Secretarias	10

- Si se despide a 8 abogados y 12 ingenieros. ¿Cuál será la frecuencia relativa correspondiente a los obreros?

Nota: Tanto la frecuencia absoluta acumulada como la relativa acumulada se obtienen sumando los valores de las frecuencias absolutas y relativas anteriores.