

PRESENTACIÓN DE INFORMACIÓN EN ESTADÍSTICA.

- Una vez el plan de muestreo es aplicado y los datos son recolectados, es necesario que los datos sean:
 - Organizados,
 - Presentados y
 - Resumidosadecuadamente con el objetivo de obtener información, la que sirve como apoyo a la toma de decisiones.
- Existen tres formas de resumir datos mediante:
 - **Tablas,**
 - **Gráficos y**
 - **Estadísticas** (como por ejemplo el promedio).

- La **organización de datos** consiste en

determinar qué unidades de análisis (UAs) pertenecen a qué atributos de la variable bajo estudio,

estableciendo para ello las frecuencias con las que estas UAs pertenecen a esos atributos.
- Una vez que los datos están **organizados**, se procede a su presentación a través de:
 - Tablas,
 - Gráficos y
 - Estadísticas.

TABLAS ESTADÍSTICAS

La **organización de datos** depende del tipo de variable estadística a ser analizada. Considere las siguientes definiciones.

Clase

Una “clase” o categoría es uno de los posibles atributos de una variable cualitativa a la que una UA puede pertenecer. Las clases deben ser excluyentes o disjuntas, ya que de esta forma no existe ambigüedad en la clasificación de las UAs.

Marca e intervalo de clase

Una “clase” se denomina “marca de clase” cuando la variable es discreta. En el caso de variables continuas, la “clase” se denomina “intervalo de clase” y el punto medio del intervalo se llama “marca de clase”.

TABLAS DE FRECUENCIA

Se desea estudiar una variable agrupada en k clases excluyentes, c_1, c_2, \dots, c_k , para variables cualitativas, o k marcas de clases excluyentes, x_1, x_2, \dots, x_k , para variables cuantitativas.

Frecuencia absoluta

Corresponde al número de UAs que pertenecen a la clase c_i y se denota por n_i , para $i = 1, \dots, k$, donde

$$\sum_{i=1}^k n_i = n.$$

Frecuencia acumulada

Corresponde al número acumulado de UAs que pertenecen a las clases c_1, c_2, \dots, c_i y está dada por $N_i = \sum_{j=1}^i n_j$, para $i = 1, \dots, k$. Así,

$$N_1 = n_1 \text{ y } N_k = \sum_{i=1}^k n_i = n.$$

Frecuencia relativa

Corresponde a la proporción de UAs que pertenecen a la clase c_i y está dada por $f_i = n_i/n$, para $i = 1, \dots, k$, donde

$$\sum_{i=1}^k f_i = \sum_{i=1}^k \frac{n_i}{n} = \frac{1}{n} \sum_{i=1}^k n_i = 1.$$

Si multiplicamos la frecuencia relativa por 100, obtenemos la frecuencia porcentual.

Frecuencia relativa acumulada

Corresponde a la proporción acumulada de UAs que pertenecen a las clases c_1, c_2, \dots, c_i y está dada por $F_i = \sum_{j=1}^i f_j$, para $i = 1, \dots, k$. Así,

$$F_1 = f_1 \text{ y } F_k = \sum_{i=1}^k f_i = \sum_{i=1}^k \frac{n_i}{n} = \frac{1}{n} \sum_{i=1}^k n_i = 1.$$

TABLAS PARA VARIABLES CUALITATIVAS

Una tabla estadística distribuye las UAs en cada clase y en el caso de una variable cualitativa toma la siguiente forma.

Clase (c_i)	Frecuencia (n_i)	Frecuencia relativa (f_i)	Frecuencia relativa acumulada (F_i)
c_1	n_1	f_1	F_1
c_2	n_2	f_2	F_2
\vdots	\vdots	\vdots	\vdots
c_k	n_k	f_k	$F_k = 1$
Total	n	1	

- Esta tabla se puede usar para variables discretas con poca variabilidad reemplazando c_i por x_i .
- Para facilitar la interpretación se suelen usar las frecuencias relativas porcentuales.
- F_i no se calcula si la variable se mide en escala nominal.

EJEMPLO: RENDIMIENTO ACADÉMICO

Ejemplo

Se realizó un estudio con 15 estudiantes de la UCM para analizar su rendimiento académico medido en una escala de: 1 = bajo, 2 = medio y 3 = alto. Los 15 datos de rendimiento son los siguientes.

Estudiante	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Rendimiento	3	2	1	1	2	1	3	3	2	3	3	2	2	2	2

- Identifique la UA, la variable y su recorrido, su tipo y escala de medición.
- Construya una tabla estadística que incluya frecuencias absoluta (n_i), relativa (f_i), relativa porcentual ($h_i = f_i \times 100$), relativa acumulada (F_i) y relativa porcentual acumulada (H_i).
- Interprete los valores n_2 , f_3 , h_1 , H_2 de la tabla.

SOLUCIÓN DEL EJEMPLO ANTERIOR

La variable en estudio es X : rendimiento académico de los 15 estudiantes de la UCM, $R_X : \{bajo(1), medio(2), alto(3)\}$, variable cualitativa medida en escala ordinal. La tabla es la siguiente.

Tabla: distribución del rendimiento académico para los 15 estudiantes

Rendimiento laboral		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Bajo	3	20,0	20,0	20,0
	Medio	7	46,7	46,7	66,7
	Alto	5	33,3	33,3	100,0
	Total	15	100,0	100,0	

Observación: en la tabla anterior, la columna “Porcentaje” representa la frecuencia relativa porcentual (h_i), para $i = 1, 2, 3$. Además, esta tabla entrega el “Porcentaje válido”, que considera en el conteo los casos con respuesta (en este caso, los 15 estudiantes que fueron medidos). La columna “Porcentaje acumulado” representa la frecuencia relativa acumulada porcentual (H_i).

¿Cómo se interpretan los elementos de esta tabla?