

# TEMA 14: ESTADÍSTICA Y PROBABILIDAD

## 1.- INTRODUCCIÓN Y DEFINICIONES (p 278, 279)

Si queremos saber, por ejemplo, la talla de zapato que tienen las niñas de 13 años de la Comunidad Valenciana, necesitaríamos consultarles a todas. Ponerse en contacto con todas y solicitarles ese dato, puede ser muy difícil y costoso. Para solucionar este problema, utilizamos la Estadística.

La **Estadística** se basa en recoger datos de una muestra, de forma que estos datos nos permitan representar la totalidad de la población objeto de estudio.

Vamos a definir unos cuantos conceptos básicos:

- **Población:** es el conjunto formado por todos los elementos del estudio. En nuestro ejemplo, la población serían todas las niñas de 13 años de la Comunidad Valenciana.
- **Muestra:** es la parte de la población que estudiamos y que nos sirve para deducir características de toda la población. Por ejemplo, como no podemos preguntarles a todas las niñas, vamos a estudiar los datos de las niñas de 13 años de nuestra clase. Esa sería la muestra: las niñas de la clase que tienen 13 años.
- **Individuo:** cada elemento de la población o de la muestra es un individuo. Cada una de las niñas de 13 años a las que consultemos, será un individuo.
- **Tamaño de la muestra:** el número de individuos que componen una muestra. Por ejemplo, si hay 16 niñas de 13 años en clase, el tamaño de la muestra es 20.
- **Variable estadística:** cualquier cualidad de los individuos de la muestra o de la población que queremos estudiar. Por ejemplo, el número de talla de zapato.

Las variables estadísticas pueden ser:

- **Cualitativas:** los valores que toma no son números, sino cualidades. Por ejemplo, si lo que quisiéramos estudiar fuera el color de ojos de una población. Los valores que puede tomar no serían números, sino azules, verdes, marrones, negros,...
- **Cuantitativas:** los valores que toma la variable son números. A su vez, se diferencia entre:
  - \* **Discretas:** sólo puede tomar unos valores determinados. Por ejemplo, el número de hermanos sólo pueden ser 0, 1, 2, 3, etc., sabemos que no podrán ser 2,3 ó 1,5. La talla de zapato también sería una variable discreta.
  - \* **Continuas:** pueden tomar infinitos valores, o al menos, tantos, que es necesario agruparlos por tramos para poder estudiarlos.. Por ejemplo, la estatura sería continua, pues podríamos obtener todos los valores entre 140 (cm) y 200 por ejemplo. No podríamos estudiar tantos valores y tendríamos que hacer un estudio por tramos, por ejemplo, los que miden de 140 a 150 cm formarían parte del mismo grupo, los que están entre 150 y 160, serían otro grupo,...

## 2.- CÓMO TRABAJAR CON LOS DATOS ESTADÍSTICOS

### 2.1.- Recuento de datos (p 280, 281).

Ya hemos preguntado a todas las niñas de la clase su talla de zapato. Hemos obtenido estos resultados: { 38, 36, 39, 40, 36, 37, 38, 35, 41, 37, 37, 38, 39, 36, 37, 40 }

Lo siguiente que tenemos que hacer es construir una tabla de frecuencias.

- En la primera columna pondremos todos los valores posibles que hemos obtenido, sin repetirlos, la nombraremos  $x_i$ , y en ella escribiremos, ordenados, los valores de la variable.
- En la segunda columna pondremos cuántas veces se repite cada valor, la nombraremos  $f_i$  y la llamaremos frecuencia absoluta.
- Y una tercera columna con nombre  $h_i$ , llamada frecuencia relativa, donde calcularemos el cociente entre la frecuencia absoluta y el número total de muestras:

$x_i$	$f_i$	$h_i = \frac{f_i}{N}$
35	1	$1/16 = 0,0625$
36	3	$3/16 = 0,1875$
37	4	$4/16 = 0,25$
38	3	$3/16 = 0,1875$
39	2	$2/16 = 0,125$
40	2	$2/16 = 0,125$
41	1	$1/16 = 0,0625$
TOTAL	$N = 16$	1

La suma de las frecuencias absolutas (N) es el tamaño de la muestra.

La suma de las frecuencias relativas es 1.

### 2.2.- Medidas estadísticas (p 286)

- La **media (mitjana)** de un conjunto de datos, se obtiene sumando todos los valores y dividiéndolos por el número total de datos. Se nombra mediante la letra x y una línea encima:  $\bar{x}$ .  
Según la definición, lo primero que se nos ocurriría sería calcularla así:

$$\bar{x} = \frac{38+36+39+40+36+37+38+35+41+37+37+38+39+36+37+40}{16} = 37,75$$

PERO, hemos hecho la tabla anterior para que sea más sencillo. En la tabla vemos que por ejemplo el 36 se repite 3 veces, el 37, 4 veces, etc. Luego si multiplicamos cada valor por su frecuencia absoluta y sumamos todos estos productos, tenemos lo mismo que en la fórmula de arriba:

$$\bar{x} = \frac{35*1+36*3+37*4+38*3+39*2+40*2+41*1}{N} = \frac{604}{16} = 37,75$$

La media me indica el valor promedio. Evidentemente, no ha salido un número entero y nadie tiene ese número, pero los valores bajos se van compensando con los altos de forma que ese valor es el promedio de todos.

Es como cuando hacemos la media de las notas de los exámenes. Las notas altas van compensando a las bajas y la media es la nota que habría sacado si en todos los exámenes tuviera la misma nota.

- La **moda**, llamada **Mo**, es el valor (o valores) con mayor frecuencia absoluta., o sea, el que más se repite.

En nuestro ejemplo, la moda sería  $Mo = 37$

- La **mediana**, llamada **Me**, es el valor que ocupa la posición central de los valores.

Para calcularla tenemos que ordenar todos los valores, con todas sus repeticiones, y buscar el valor que ocupa el puesto central. Si el número de valores es par, habrá 2 valores centrales, así que tendremos que hacer la media de esos dos valores:

35 36 36 36 37 37 37 **37** **38** 38 38 39 39 40 40 41



$$Mo = \frac{37+38}{2} = 37,5$$

- El **rango**, llamado **R**, se calcula restando al mayor valor, el menor. En nuestro ejemplo:

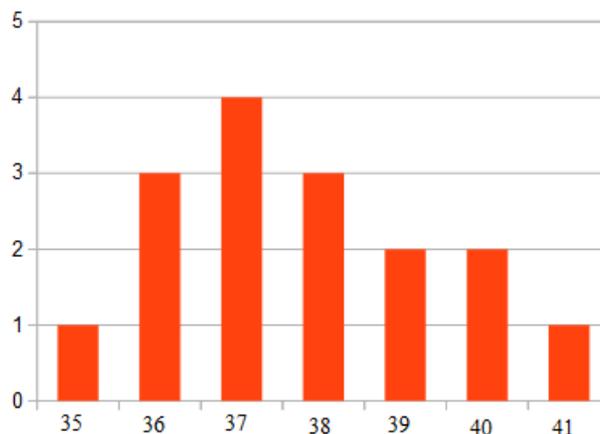
$$R = 41 - 35 = 6$$

### 2.3.- Gráficos estadísticos

#### DIAGRAMA DE BARRAS ( p 282, 283 )

Se compone de dos ejes cartesianos: en el eje horizontal se marcan los valores de la variable y en el eje vertical se marcan las frecuencias absolutas.

En cada valor se dibuja una barra con la altura correspondiente a su frecuencia absoluta.



## DIAGRAMA DE SECTORES (p 284, 285)

Se trata de un círculo dividido en sectores, que representan cada uno de los valores de la variable. La amplitud de cada sector (el ángulo asignado) es proporcional a la frecuencia del valor que representa.

En nuestro ejemplo, si hay un total de 16 datos, a esos 16 datos le corresponderían los 360°. Pues a cada frecuencia, le corresponderá el ángulo proporcional:

x=35:

$$\begin{array}{r} 16 \text{ ----- } 360^\circ \\ 1 \text{ ----- } x \\ x = \frac{1 \cdot 360}{16} = 22,5^\circ \end{array}$$

x= 36:

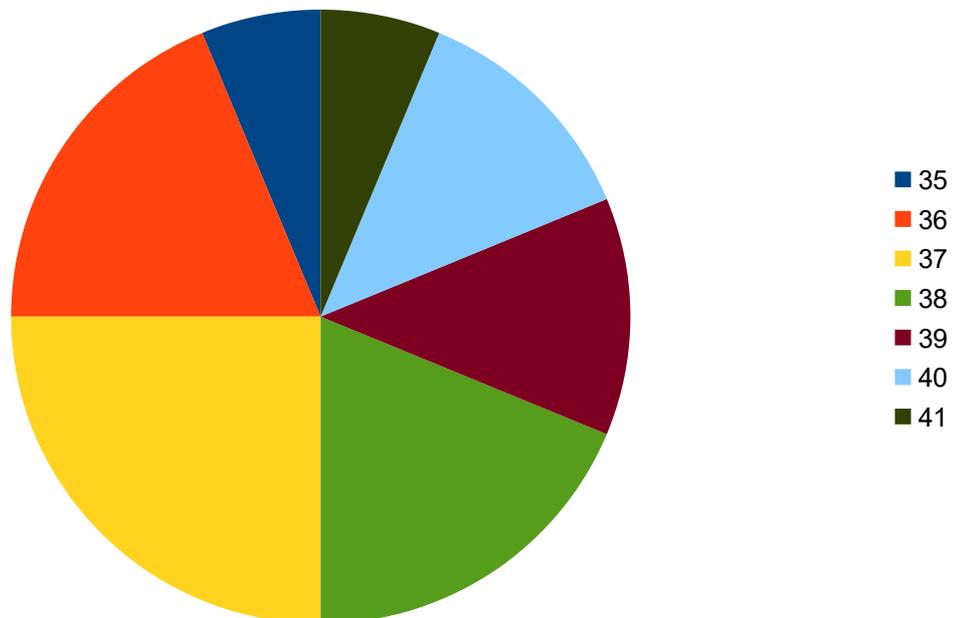
$$\begin{array}{r} 16 \text{ ----- } 360^\circ \\ 3 \text{ ----- } x \\ x = \frac{3 \cdot 360}{16} = 67,5^\circ \end{array}$$

x=37:

$$\begin{array}{r} 16 \text{ ----- } 360^\circ \\ 4 \text{ ----- } x \\ x = \frac{4 \cdot 360}{16} = 90^\circ \end{array}$$

Y así para cada valor.

Estos gráficos son muy sencillos hacerlos con el ordenador:



### **EJERCICIO:**

Queremos estudiar la edad que tienen los padres de los niños de 1º de ESO de la Comunidad Valenciana. Vamos a tomar como muestra nuestra clase. Estos son las contestaciones que obtenemos: { 43, 49, 50, 48, 46, 45, 43, 44, 42, 50, 42, 42, 47, 48, 47, 46, 49, 45, 46, 45 }.

Tienes que:

- Completar la tabla de valores y frecuencias.
- Calcular la media, la moda, la mediana y el rango.
- Realiza un gráfico de barras.
- Realiza un gráfico de sectores.

NOTAS: Puedes usar la calculadora.

Tienes que hacerlo en limpio para entregarlo.