

## Índice

|  |    |
|--|----|
| Ocho apelli2 gallegos.....   | 2  |
| Estadística.....   | 2  |
| 1. Terminología.....   | 2  |
| 1.1. ¿Por qué hablan las encuestas de adolescentes si a mí no me preguntaron?..... | 2  |
| 1.2. Variables estadísticas.....   | 2  |
| 1.3. Tipos de variables estadísticas.....  | 3  |
| 2. Tablas de frecuencias.....  | 4  |
| 2.1. Recuento y agrupamiento de datos.....   | 4  |
| 2.2. Tablas de frecuencias.....  | 5  |
| 2.3. Ejemplos.....   | 6  |
| 3. Gráficos estadísticos.....  | 8  |
| 3.1. Definiciones. Gráficos estadísticos.....                                      | 8  |
| 3.2. Tablas de frecuencias.....  | 8  |
| 3.3. Diagrama de barras.....   | 9  |
| 3.4. Diagrama de sectores.....   | 9  |
| 3.5. Otros gráficos estadísticos.....  | 10 |
| 4. Medidas de centralización.....  | 11 |
| 4.1. Parámetros estadísticos.....  | 11 |

# Ocho apelli2 gallegos

## Estadística

La Estadística es la rama de las Matemáticas que se encarga del estudio y análisis de datos. La palabra estadística significa "ciencia del Estado" y este significado tiene que ver con sus orígenes. Ya las primeras civilizaciones sentían la necesidad de conocer las características de la población, datos como ¿cuántos nacimientos había? o ¿cuántos hombres en edad de ser reclutados? eran relevantes, así que los gobiernos buscaron la forma de recopilar, organizar y procesar la información.

A día de hoy esa labor de recopilar datos sigue siendo importante para los estados, es por ello que existen organismos públicos que se encargar de esta labor. En Galicia tenemos el IGE y en este vídeo puedes ver su presentación. <https://youtu.be/QZYpqBBXKdQ>

### 1. Terminología

#### 1.1. ¿Por qué hablan las encuestas de adolescentes si a mí no me preguntaron?

Seguro que alguna vez te has preguntado por qué en noticias o estudios se habla de lo que opinan los adolescentes pero a ti nunca te han preguntado.

La estadística es capaz de generalizar resultados de individuos de una población con los datos obtenidos a partir de una muestra. ¿A que no has entendido nada? Eso es porque todavía no sabemos lo que significan las palabras clave del texto. Veremos con un ejemplo qué son.

**Ejemplo 1:** Imagina que te encuentras con este titular en un periódico: "Según un estudio estadístico del INE realizado en 2022 a 25.000 personas de hogares españoles, el 70,6% de los jóvenes entre 10 y 15 años tienen dispositivo móvil".

- **Característica:** La característica se refiere a aquella cualidad que se quiere estudiar. En nuestro caso sería "tener teléfono móvil".
- **Población:** La población es el conjunto de personas, animales o cosas sobre las que se quiere conocer determinada característica. En nuestro ejemplo serían todas las personas jóvenes entre 10 y 15 años de España.
- **Muestra:** La muestra es el conjunto de personas, animales o cosas elegidas para realizar el estudio. En nuestro ejemplo sería el conjunto de personas jóvenes a las que sí se les ha preguntado (aunque no fueses tú, seguro que le preguntaron a otros/as).
- **Individuo:** Los individuos son cada uno de los elementos de la población. En nuestro ejemplo sería cada uno de los jóvenes españoles entre 10 y 15 años (les hayan preguntado o no).
- **Tamaño de la muestra:** El tamaño de la muestra es el número de personas elegidas para realizar el estudio. En nuestro ejemplo serían las 25.000 personas preguntadas.

¿Y si leemos otra vez el titular teniendo en cuenta estas definiciones?

- "Según un estudio estadístico del INE": Es decir, que el estudio lo realizó el Instituto Nacional de Estadística.
- "realizado en 2022": Es decir que ya tiene unos años. Como muchas veces hay muchos datos que procesar y lleva su tiempo, la información no siempre está actualizadísima al año en que vivimos.
- "a 25.000 personas de hogares españoles": El tamaño de la muestra es de 25.000 personas y la muestra serían esas 25.000 personas a las que les preguntaron, ya que llevaría mucho tiempo y muchos recursos preguntar a todas las personas que viven en España, además a algunas podríamos no localizarlas.
- "el 70,6% de las personas jóvenes entre 10 y 15 años tienen dispositivo móvil": Este es el resultado final que es extensible a todas las personas españolas, es decir, a la población española.

#### 1.2. Variables estadísticas

La **variable estadística** es aquel aspecto de una población que se quiere estudiar. En el ejercicio anterior la variable estadística sería "Uso diario de internet".

Los posibles valores que puede tomar la variable estadística reciben el nombre de **modalidades de la variable**. En el ejercicio anterior, serían las respuestas a la pregunta ¿Usas diariamente internet?

Las variables estadísticas se suelen representar matemáticamente con el carácter  $x$ . Para los distintos valores  $x_i$  tendremos las modalidades de la variable:  $x_1, x_2, \dots$

### 1.3. Tipos de variables estadísticas

Según las respuestas que se puedan dar al carácter estadístico, tendremos dos tipos de variables:

- **Variables cualitativas:** La respuesta es siempre una cualidad, es decir una o varias palabras que no son números.
- **Variables cuantitativas:** La respuesta es siempre una cantidad, es decir un número.

### 1.4. Ejemplos:

**Ejemplo 1: Carácter y variable:** Indica los posibles valores para los siguientes caracteres estadísticos ¿Es la variable cuantitativa o cualitativa?

| Encuesta/<br>Cuestionario     | Variable<br>estadística ( $x$ ) | Modalidades de la variable                         | Tipo de variable<br>¿Cuantitativa o cualitativa? |
|-------------------------------|---------------------------------|--|--|
| ¿Tienes móvil?                | Tener móvil                     | Sí, No   | Cualitativa                                      |
| ¿Cuántos hermanos tienes?     | Número de hermanos              | 0, 1, 2, 3   | Cuantitativa                                     |
| ¿De qué color tienes el pelo? | Color de pelo                   | Negro, Castaño, Pelirrojo, Rubio, Blanco, Sin pelo | Cualitativa                                      |
| ¿Cuánto mides?                | Altura                          | 1,50, 1,48, 1,62, 1,55                             | Cuantitativa                                     |

**Ejemplo 2: Tipos de variables:** Indica el tipo de las siguientes variables o caracteres estadísticos:

- |  |  |
|--|--|
| • Trabajo de tu madre: Cualitativa.    | • Número de abejas de las colmenas: Cuantitativa         |
| • Nivel de estudios: Cualitativa       | • Comida favorita: Cualitativa                           |
| • Tamaño de las hormigas: Cuantitativa | • Días que se tardan en construir un barco: Cuantitativa |

### 1.5. Curiosidades curiosas

**¿Puedo saber el tamaño de una muestra?** Pues deberías, ya que no es lo mismo una muestra de los 3 cruceiros de un tramo del camino que nos queda cerca de nuestra aldea, que si recorremos Galicia analizando los materiales, formas, altura,... de muchos cruceiros.

Los sitios que publican los estudios deberían tener en algún sitio accesible una ficha técnica del estudio, en el que indique varias cosas sobre cómo se elaboró el estudio, entre ellas la muestra. Por ejemplo, si visitas la web del CIS y eliges cualquier estudio, puedes ver que en la parte de la derecha, donde hay varios documentos, uno de ellos pone "Ficha técnica", ahí puedes ver el tamaño de la muestra.

**¿Puedo saber quién contestó a las encuestas?** Pues... no, eso sí que no. Los datos de las personas que contestan los cuestionarios o encuestas son privados y sujetos a muchos mecanismos de seguridad para que nunca se sepa quién contestó a qué. Así que tranquilo/a, si alguna vez un organismo oficial como el INE, IGE o CIS te quieren hacer una encuesta, contesta ya que nadie va a saber qué contestaste.

**¿Puede una mujer tener 1,16 hijos?** Lo sabemos, nadie tiene un hijo y 0,16 de otro. Las variables cuantitativas nos van a permitir calcular una serie de medidas que las resumen y que por tanto nos ayudarán a interpretarlas.

Aún así no debemos confundir estas medidas que son un apoyo para nuestro análisis, con las posibles respuestas a una pregunta y que son las que determinan qué tipo de variable es.

Esto es lo que ocurre por ejemplo cuando se nos dice que las mujeres tienen 1,16 hijos cada una.

## 2. Tablas de frecuencias

### 2.1. Recuento y agrupamiento de datos

Una vez recogidos los datos de la encuesta o cuestionario, procedemos a realizar el recuento. Para ello debemos diferenciar si la variable es cualitativa o cuantitativa:

- **Variable cualitativa:** Observamos cuáles son los valores que toma la variable,  $x_i$ , y contamos el número de veces que aparece cada uno de los datos,  $f_i$ .
- **Variable cuantitativa:** En primer lugar, ordenamos los valores de la variable,  $x_i$ , de menor a mayor. Una vez que estén ordenados, contaremos el número de veces que aparece cada uno de los valores, y lo indicaremos en  $f_i$ .

**Ejemplo 1: Práctica de deportes.** Preguntamos en clase cuál es el deporte que practica cada uno de nuestros compañeros y obtenemos los siguientes resultados: natación, judo, fútbol, baloncesto, voley, baloncesto, voley, atletismo, natación, baloncesto, fútbol, atletismo, baloncesto, voley, baloncesto

Recuento: contar cuántas personas realizan cada uno de los deportes

- natación: 2
- judo: 1
- fútbol: 2
- baloncesto: 5
- voley: 3
- atletismo: 2

Revisamos que el número total de datos se corresponda con los que recogimos:

$$2+1+2+5+3+2=15$$

Tabla de frecuencias:

| Deportes, $x_i$ | Nº de personas que lo practican, $f_i$ |
|-----------------|--|
| natación        | 2                                      |
| judo            | 1                                      |
| fútbol          | 2                                      |
| baloncesto      | 5                                      |
| voley           | 3                                      |
| atletismo       | 2                                      |
|                 | N=15                                   |

**Ejemplo 2: Horas de práctica:** En la misma encuesta preguntamos el número de días a la semana que practican deporte: 2, 3, 1, 5, 1, 3, 4, 6, 4, 5, 2, 3, 4, 4, 3

Lo primero que tenemos que hacer es ordenar los valores obtenidos de menor a mayor. A continuación, vamos a contar cuántas personas realizan deporte tantos días a la semana, revisamos que el número total de datos se corresponda con los que tenemos recogidos:  $2+2+4+4+2+1=15$

Y por último, trasladamos estos datos a una tabla:

| Días en los que se practica deporte, $x_i$ | Nº de personas que lo practican esos días, $f_i$ |
|--|--|
| 1  | 2  |
| 2  | 2  |
| 3  | 4  |
| 4  | 4  |
| 5  | 2  |
| 6  | 1  |
|  | N=15   |

## 2.2. Tablas de frecuencias

Una vez realizado el recuento de datos, veamos cómo elaborar las tablas de frecuencias. Antes de comenzar, es necesario saber qué representa cada una de las columnas:

- **Variable estadística,  $x$ :** Es la característica de la población que estamos estudiando.
- **Modalidades o valores de la variable estadística,  $x_i$ :** Cada uno de los valores que puede tomar la variable.
- **Frecuencia absoluta,  $f_i$ :** Número de veces que se repite el dato  $x_i$ . La suma de las frecuencias absolutas es el número total de datos:  $N$
- **Frecuencia relativa,  $h_i$ :** Es el cociente entre la frecuencia absoluta de  $x_i$  y el total de datos,  $N$ :

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

La suma de las frecuencias relativas vale 1.

- **Porcentaje:  $p_i$ :** Para el dato  $x_i$  se calcula multiplicando su frecuencia relativa  $h_i$  por 100, es decir,
- 

$$p_i = h_i \cdot 100$$

La suma de los porcentajes vale 100

- **Frecuencia absoluta acumulada,  $F_i$ :** Es la suma de las frecuencias absolutas de los valores que son menores o iguales que el dato  $x_i$ .
- 

$$F_1 = f_1 ; F_2 = f_1 + f_2 ; \dots ; F_i = f_1 + f_2 + \dots + f_i$$

- **Frecuencia relativa acumulada,  $H_i$ :** Para el dato  $x_i$  es la suma de las frecuencias relativas de los valores que son menores o iguales que el dato  $x_i$
- 

$$H_1 = h_1 ; H_2 = h_1 + h_2 ; \dots ; H_i = h_1 + h_2 + \dots + h_i$$

Si la variable es **cualitativa**, no tiene sentido calcular las frecuencias acumuladas. En caso de que la variable sea **cuantitativa** lo primero que se tiene que hacer es ordenar los datos de la variable de menor a mayor. La tabla que obtendremos será como esta:

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia relativa, $h_i$ | Porcentaje, $p_i$     | Frecuencia absoluta acumulada, $F_i$ | Frecuencia relativa acumulada, $H_i$ |
|-----------------------------|----------------------------|----------------------------|-----------------------|--------------------------------------|--------------------------------------|
| $x_1$                       | $f_1$                      | $h_1$                      | $p_1 = h_1 \cdot 100$ | $F_1$                                | $H_1$                                |
| $x_2$                       | $f_2$                      | $h_2$                      | $p_2 = h_2 \cdot 100$ | $F_2$                                | $H_2$                                |
| ...                         | ...                        | ...                        | ...                   | ...                                  | ...                                  |
| $x_n$                       | $f_n$                      | $h_n$                      | $p_n = h_n \cdot 100$ | $F_n$                                | $H_n$                                |
|                             | $N$                        | 1                          | 100                   |                                      |                                      |

### 2.3. Ejemplos

**Ejemplo 1: Ejemplo de variable cualitativa:** Los 25 alumnos de 1º de ESO provienen de cuatro parroquias diferentes, según se muestra en esta lista: Oza, Lérez, Oza, Oza, Oza, Lérez, Castro, Oza, Oza, Oza, Castro, Vilar, Castro, Oza, Lérez, Castro, Lérez, Castro, Oza, Lérez, Lérez, Oza, Oza, Castro, Lérez.

#### 1. Identificamos la variable estadística y hacemos el recuento:

Variable: Parroquias

Recuento:

- Oza: 11
- Lérez: 7
- Castro: 6
- Vilar: 1

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ |
|-----------------------------|----------------------------|
| Oza                         | 11                         |
| Lérez                       | 7                          |
| Castro                      | 6                          |
| Vilar                       | 1                          |
|                             | N=25                       |

#### 2. Calculamos la frecuencia relativa, $h_i = \frac{f_i}{N}$ y los porcentajes $p_i = h_i \cdot 100$

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia relativa, $h_i$ | Porcentajes, $p_i$ |
|-----------------------------|----------------------------|----------------------------|--------------------|
| Oza                         | 11                         | 11/25                      | 44%                |
| Lérez                       | 7                          | 7/25                       | 28%                |
| Castro                      | 6                          | 6/25                       | 24%                |
| Vilar                       | 1                          | 1/25                       | 4%                 |
|                             | N=25                       | 1                          | 100%               |

**Ejemplo 2: Ejemplo de variable cuantitativa.** Realiza la tabla de datos de las calificaciones en matemáticas de 50 estudiantes han sido: 5, 2, 4, 9, 7, 4, 5, 6, 5, 7, 7, 5, 5, 2, 10, 5, 6, 5, 4, 5, 8, 8, 4, 0, 7, 8, 4, 8, 6, 6, 3, 6, 7, 6, 6, 7, 6, 7, 3, 5, 6, 9, 6, 1, 4, 6, 3, 5, 5, 6

#### 1. Identificamos la variable estadística y hacemos el recuento:

La variable estadística es la nota en matemáticas del alumnado.

Hacemos el recuento:

- Nota de examen 0: 1
- Nota de examen 1: 1
- Nota de examen 2: 2
- Nota de examen 3: 3
- Nota de examen 4: 6
- Nota de examen 5: 11
- Nota de examen 6: 12
- Nota de examen 7: 7
- Nota de examen 8: 4
- Nota de examen 9: 2
- Nota de examen 10: 1

Traslademos estos datos a una tabla de frecuencias:

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ |
|-----------------------------|----------------------------|
| 0                           | 1                          |
| 1                           | 1                          |
| 2                           | 2                          |
| 3                           | 3                          |
| 4                           | 6                          |
| 5                           | 11                         |
| 6                           | 12                         |
| 7                           | 7                          |
| 8                           | 4                          |
| 9                           | 2                          |
| 10                          | 1                          |
|                             | N=50                       |

**2. Calculamos la frecuencia absoluta acumulada,  $F_i$  y la relativa  $h_i = \frac{f_i}{N}$**

Recuerda, que para calcular la frecuencia absoluta acumulada debemos hacer:  $F_k = \sum_{i=1}^k f_i = f_1 + f_2 + \dots + f_k$

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia absoluta acumulada, $F_i$ | Frecuencia relativa, $h_i$ |
|-----------------------------|----------------------------|--------------------------------------|----------------------------|
| 0                           | 1                          | 1                                    | 1/50                       |
| 1                           | 1                          | 2                                    | 1/50                       |
| 2                           | 2                          | 4                                    | 2/50                       |
| 3                           | 3                          | 7                                    | 3/50                       |
| 4                           | 6                          | 13                                   | 6/50                       |
| 5                           | 11                         | 24                                   | 11/50                      |
| 6                           | 12                         | 36                                   | 12/50                      |
| 7                           | 7                          | 43                                   | 12/50                      |
| 8                           | 4                          | 47                                   | 4/50                       |
| 9                           | 2                          | 49                                   | 2/50                       |
| 10                          | 1                          | 50                                   | 1/50                       |
| N=50                        |                            |                                      |                            |

**3. Obtenemos la frecuencia relativa acumulada,  $H_i = \frac{F_k}{N}$  y los porcentajes  $p_i = h_i \cdot 100$**

| Variable estadística, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia absoluta acumulada, $F_i$ | Frecuencia relativa, $h_i$ | Frecuencia relativa acumulada, $H_i$ | Porcentajes, $p_i$ |
|-----------------------------|----------------------------|--------------------------------------|----------------------------|--------------------------------------|--------------------|
| 0                           | 1                          | 1                                    | 1/50                       | 1/50                                 | 2%                 |
| 1                           | 1                          | 2                                    | 1/50                       | 2/50                                 | 2%                 |
| 2                           | 2                          | 4                                    | 2/50                       | 4/50                                 | 4%                 |
| 3                           | 3                          | 7                                    | 3/50                       | 7/50                                 | 6%                 |
| 4                           | 6                          | 13                                   | 6/50                       | 13/50                                | 12%                |
| 5                           | 11                         | 24                                   | 11/50                      | 24/50                                | 22%                |
| 6                           | 12                         | 36                                   | 12/50                      | 36/50                                | 24%                |
| 7                           | 7                          | 43                                   | 12/50                      | 43/50                                | 14%                |
| 8                           | 4                          | 47                                   | 4/50                       | 47/50                                | 8%                 |
| 9                           | 2                          | 49                                   | 2/50                       | 49/50                                | 4%                 |
| 10                          | 1                          | 50                                   | 1/50                       | 1                                    | 2%                 |
| N=50                        |                            |                                      |                            |                                      | 100%               |

### 3. Gráficos estadísticos

Una vez que se tienen los datos organizados en tablas, se suele hacer una representación gráfica de ellos. Estos gráficos deben conseguir que un simple análisis visual ofrezca la mayor información posible. Según el tipo del carácter que estemos estudiando, usaremos una representación gráfica u otra. En los siguientes apartados verás cómo construir algunas de ellas.

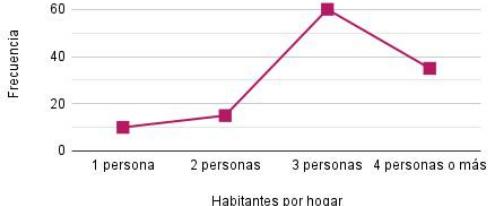
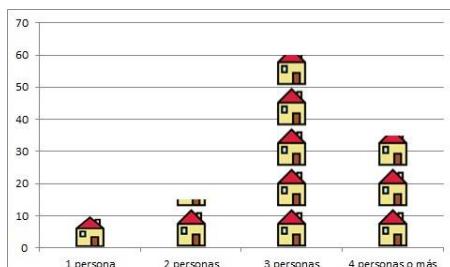
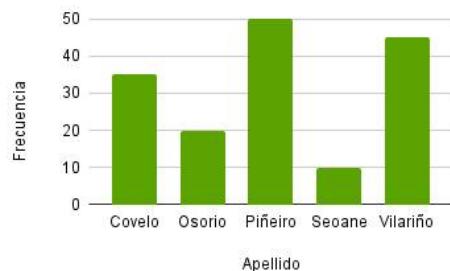
#### 3.1. Definiciones. Gráficos estadísticos

##### Diagrama de barras:

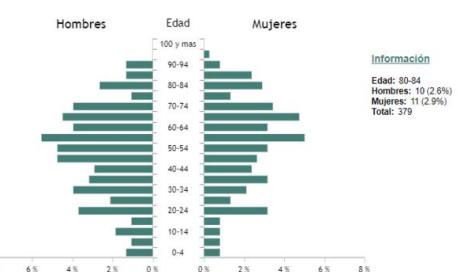
En los diagramas de barras vamos a representar las modalidades de la variable estadística en el eje horizontal (el eje de abscisas). Estos valores pueden ser cualitativos o cuantitativos. Además, distribuimos los valores en el eje de manera que estén equidistantes.

El eje vertical (eje de ordenadas) será numérico siempre y representará las frecuencias absolutas de cada una de las modalidades de la variable.

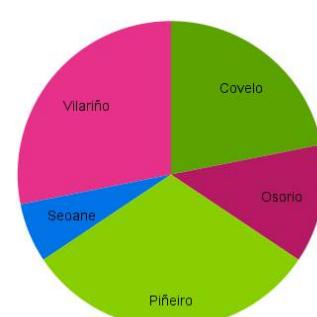
**Pictograma:** Un pictograma es un gráfico que representa mediante figuras o símbolos las frecuencias de una variable. Las figuras son alusivas a la variable objeto de estudio.



**Polígono de frecuencias:** Son gráficos que se utilizan en el caso de una variable cuantitativa y en los que se van trazando segmentos indicando cómo varían las frecuencias entre los valores de la variable. Se construyen uniendo los extremos medios superiores del diagrama de barras.



**Pirámide poblacional:** Una pirámide de población es un diagrama de barras pero con orientación horizontal y con los datos de la variable para hombres y mujeres uno para cada lado, de manera que se puede hacer una comparativa entre ambos. En uno de esos diagramas se representaría a la población de los hombres y en el otro a la población de las mujeres. En el eje vertical se representaría la edad.



**Diagrama de sectores:** Son gráficos circulares en los que cada valor de la variable se representa mediante un sector circular proporcional a la frecuencia que representan.

#### 3.2. Tablas de frecuencias

Ya incluídas en el apartado 2.2

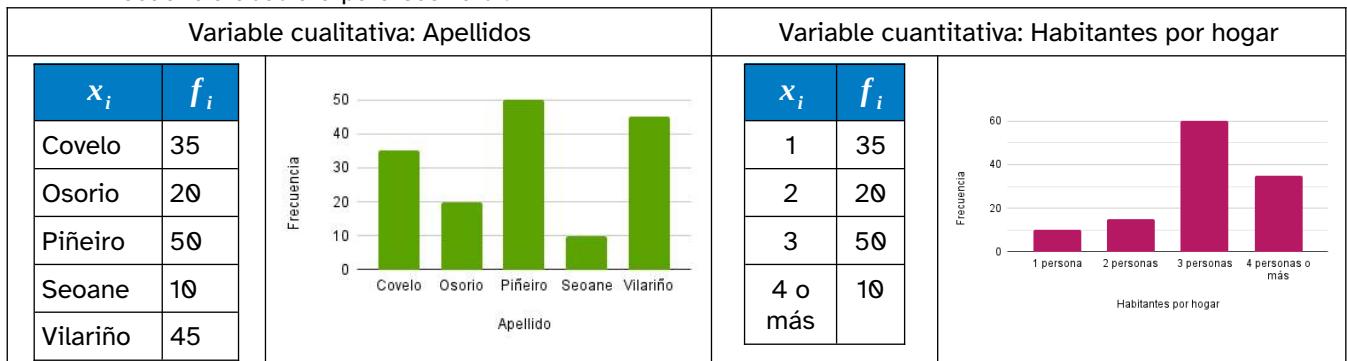
### 3.3. Diagrama de barras

#### Hacer un diagrama de barras verticales

- Trazar los ejes: En el eje horizontal se marcan los posibles valores de la variable estadística, en este caso, los diferentes apellidos. En el eje vertical van las frecuencias. Se observa entre qué valores está la frecuencia y se elige una escala adecuada para poder representarlas todas.

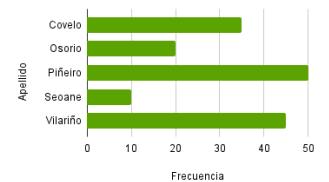


- Dibujar las barras: Sobre cada variable se levanta una barra cuya altura se corresponde con la frecuencia absoluta para ese valor.



#### Hacer un diagrama de barras horizontales

Los diagramas de barras también puede aparecer representados con las barras en horizontal. En este caso el procedimiento para hacerlos sería similar intercambiando los ejes: el eje horizontal se correspondería con las frecuencias y en el eje vertical estarían los valores de la variable estadística.



### 3.4. Diagrama de sectores

#### Hacer un diagrama de sectores

El diagrama de los sectores sirve para representar tanto variables cualitativas como cuantitativas. En este caso se representará la frecuencia de cada uno de los valores con un sector circular.

Para dibujar un diagrama de sectores se establece una relación de proporcionalidad entre la frecuencia y la amplitud del sector circular que le corresponde.

- Calcular la amplitud de los sectores: Hay que repartir los grados de una circunferencia,  $360^\circ$ , entre los diferentes valores de la variable. Para ello, se reparten los grados de circunferencia de manera proporcional a las diferentes frecuencias.

| Apellido | Frecuencia |
|----------|------------|
| Covelo   | 35         |
| Osorio   | 20         |
| Piñeiro  | 50         |
| Seoane   | 10         |
| Vilariño | 45         |

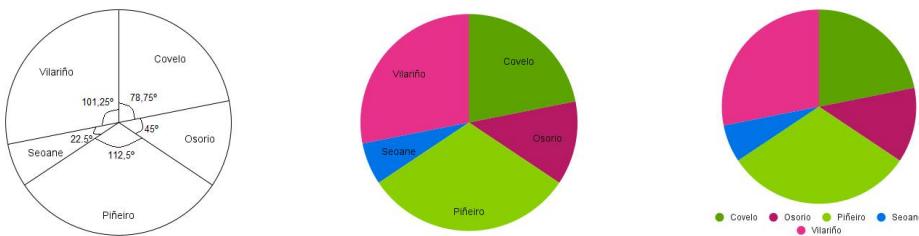
Cálculo de la amplitud del sector:

$$\text{Covelo: } \frac{35}{160} = \frac{A}{360} \Rightarrow A = 78,75^\circ \quad \text{Piñeiro: } \frac{20}{160} = \frac{C}{360} \Rightarrow C = 45^\circ$$

$$\text{Osorio: } \frac{20}{160} = \frac{B}{360} \Rightarrow B = 45^\circ \quad \text{Seoane: } \frac{10}{160} = \frac{D}{360} \Rightarrow D = 22,5^\circ$$

$$\text{Vilariño: } \frac{45}{160} = \frac{E}{360} \Rightarrow E = 101,25^\circ$$

2. Dibujar el diagrama de sectores: Se dibuja un círculo. Se traza un radio y a partir de él se va dibujando cada uno de los sectores con la amplitud que se ha calculado. Después se puede identificar cada sector con un color. También indicar en el sector qué valor de la variable le corresponde o poner una leyenda al lado del gráfico con dicho valor de la variable.



### 3.5. Otros gráficos estadísticos

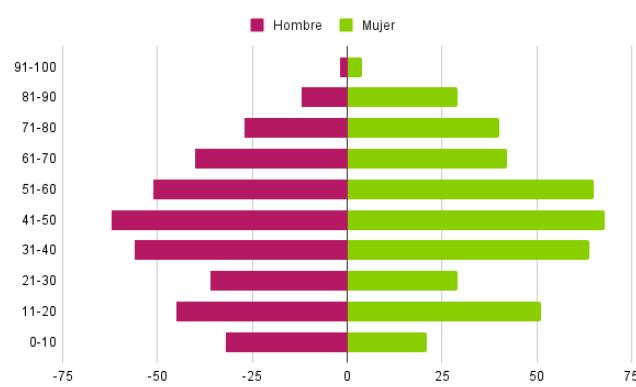
**Polígono de frecuencias:** Los polígonos de frecuencias se trazan uniendo los puntos medios de la base superior de las barras. También se pueden marcar en el eje de coordenadas los puntos que forman el par (valor de la variable, frecuencia) y unirlos.



**Polígono de frecuencias y tiempo:** Es habitual usar un polígono de frecuencias para hacer representaciones en las que la variable es una variable relacionada con el tiempo (meses, años,...) y así poder ver cómo con el tiempo cambia un determinado valor. Estos gráficos son muy útiles para ver las tendencias en el tiempo y comparar diferentes años.

**Pirámide poblacional:** En este caso se necesita tener una tabla de frecuencias en la que se indique cuántas personas de cada edad hay diferenciando según su género. Después se construirían dos diagramas de barras en horizontal, uno para las edades de cada sexo. En cada uno de ellos las barras crecen en un sentido (izquierdo y derecho). Las edades van en orden creciente de abajo hacia arriba.

| Edades   | Hombres | Mujeres |
|----------|---------|---------|
| 0 - 10   | 32      | 21      |
| 11 - 20  | 45      | 51      |
| 21 - 30  | 36      | 29      |
| 31 - 40  | 56      | 64      |
| 41 - 50  | 62      | 68      |
| 51 - 60  | 51      | 65      |
| 61 - 70  | 40      | 42      |
| 71 - 80  | 27      | 40      |
| 81 - 90  | 12      | 29      |
| 91 - 100 | 2       | 4       |



## 4. Medidas de centralización

### 4.1. Parámetros estadísticos

A continuación estudiaremos los parámetros estadísticos, estos nos permitirán analizar y comparar los resultados de esas variables para extraer conclusiones de la población.

### 4.2. Definiciones. Medidas de centralización

Las medidas de centralización son los valores alrededor de los cuales se distribuyen los datos de la variable objeto de estudio. Estas medidas tienen por objetivo resumir todos los valores de la variable en uno solo que los represente.

Las medidas de centralización más usadas son la media, la mediana y la moda.

**Media aritmética:** Media aritmética,  $\bar{x}$ , de un conjunto de datos es el valor medio que los representa, donde  $x$  es la variable objeto de estudio. Sólo es posible calcularla con datos cuantitativos.

**Mediana:** La mediana,  $M_e$ , es el valor que se sitúa en la posición central una vez ordenados todos los datos de menor a mayor. La mediana sólo se puede obtener en variables cuantitativas.

**Moda:** La moda,  $M_o$ , es el valor que más se repite, es decir, con mayor frecuencia absoluta. Podremos obtener la moda tanto en variables cuantitativas como cualitativas. Puede haber más de una moda. En caso de haber dos tendremos una distribución bimodal; si hay tres será trimodal, y en caso de haber más de tres, será multimodal.

### 4.4. Media aritmética

#### Cálculo de la media

Se calcula sumando todos los valores obtenidos de las variables y dividiéndolos por el número de datos que haya.

**Ejemplo:** Sean las respuestas de cinco estudiantes a la pregunta ¿Cuántas personas vivís en vuestra casa?: 2, 2, 3, 3, 5.

$$\text{Su media es: } \bar{x} = \frac{2+2+3+3+5}{5} = 3$$

#### Cálculo de la media para tablas de frecuencias

**Ejemplo:** Si los datos recogidos de la misma pregunta ¿Cuántas personas viven en tu hogar? para toda la clase son: 2, 2, 2, 3, 3, 4, 2, 2, 4, 3, 5, 2, 2, 11, 2, 3, 2, 5, 2, 2, 5, 3, 5, 2, 4, 4, 3, 3, 5, 2, 6, 3, 5, 2, 3, 4, 3, 3, 5, 2, 2, 2, 3, 5, 9

#### 1º Se ordenan los datos en una tabla de frecuencias

El número total de datos es 50. A continuación los ordenamos en una tabla de frecuencias y los vamos contando para calcular las frecuencias absolutas de cada valor:

| N.º personas, $x_i$ | Frecuencia absoluta, $f_i$ |
|---------------------|----------------------------|
| 2                   | 20                         |
| 3                   | 12                         |
| 4                   | 4                          |
| 5                   | 11                         |
| 6                   | 1                          |
| 9                   | 1                          |
| 11                  | 1                          |
|                     | 50                         |

#### 2º Multiplicamos cada valor de la variable por su frecuencia absoluta y lo sumamos

$$\sum_{i=1}^7 x_i \cdot f_i$$

$$2 \cdot 20 + 3 \cdot 12 + 4 \cdot 4 + 5 \cdot 11 + 6 \cdot 1 + 9 \cdot 1 + 11 \cdot 1 = 173$$

#### 3º Dividimos la suma entre el número total de datos, N

$$\bar{x} = \sum_{i=1}^7 \frac{x_i \cdot f_i}{N}$$

$$\bar{x} = \frac{173}{50} = 3,46$$

#### 4.5. Mediana

**Cálculo de la mediana:** Se calcula empezando por ordenar los datos. Si el número de datos es impar, sólo habrá uno en el medio y ese valor de la variable será la mediana. En caso de que el número de datos sea par, hallamos la media aritmética de los dos datos centrales de la variable.

**Ejemplo 1: N impar:** Calcula la mediana de los datos de las notas de 5 alumnos/as en el examen de estadística: 6, 5, 8, 4, 5

Ordenamos los datos para localizarlo: 4, 5, 5, 6, 8. El dato que ocupa la posición central es el 5 y por tanto será la mediana.  $Mediana(x)=5$

Es decir, el 50 % del alumnado ha obtenido un 5 o menos y el otro 50 % ha obtenido un 5 o más.

**Ejemplo 2: N par:** Calcula la mediana de los datos correspondientes a las pagas semanales de 6 alumnos/as de la clase de primero de la eso: 10, 6, 4, 5, 5, 6

Hallaremos la media aritmética de los datos que ocupen las posiciones centrales una vez ordenados: 4, 5, 5, 6, 6, 10. Los datos centrales son 5 y 6, por tanto la mediana será la media de ambos.

$$Mediana = \frac{5+6}{2} = 5,5$$

Es decir, la mitad del alumnado tiene de paga 5,5 € o menos y la otra mitad tiene 5,5 € o más.

#### Cálculo de la mediana para tablas de frecuencias

La mediana es el primer valor cuya frecuencia acumulada sea mayor que  $N/2$  si N es impar. En caso de ser N par será la media aritmética de los valores cuya frecuencia absoluta acumulada contenga a  $N/2$  y  $\frac{N}{2} + 1$ .

**Ejemplo:** Volvamos a trabajar sobre los datos de la pregunta ¿Cuántas personas viven en tu hogar? de toda la clase:

2, 2, 2, 3, 3, 4, 2, 2, 4, 3, 5, 2, 2, 11, 2, 3, 2, 5, 2, 2, 5, 3, 5, 2, 4, 4, 3, 3, 5, 2, 6, 3, 5, 2, 3, 4, 3, 3, 5, 2, 2, 2, 3, 5, 9

#### 1º Se hace la tabla de frecuencias

| Nº personas, $x_i$ | Frecuencia absoluta, $f_i$ |
|--------------------|----------------------------|
| 2                  | 20                         |
| 3                  | 12                         |
| 4                  | 4                          |
| 5                  | 11                         |
| 6                  | 1                          |
| 9                  | 1                          |
| 11                 | 1                          |
|                    | N=50                       |

#### 2º Calculamos las frecuencias acumuladas

| Nº personas, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia absoluta relativa, $F_i$ |
|--------------------|----------------------------|-------------------------------------|
| 2                  | 20                         | 20                                  |
| 3                  | 12                         | 32                                  |
| 4                  | 4                          | 36                                  |
| 5                  | 11                         | 47                                  |
| 6                  | 1                          | 48                                  |
| 9                  | 1                          | 49                                  |
| 11                 | 1                          | 50                                  |
|                    | N=50                       |                                     |

#### 3º Calculamos N y nos fijamos si es par o impar y calculamos también N/2

En este caso,  $N=50$ , que es un número par. Por tanto,  $N/2=25$

#### 4º Calculamos la mediana

| Nº personas, $x_i$ | Frecuencia absoluta, $f_i$ | Frecuencia absoluta relativa, $F_i$ |
|--------------------|----------------------------|-------------------------------------|
| 2                  | 20                         | 20                                  |

El valor de la variable cuya frecuencia acumulada contiene a  $\frac{N}{2}$  es 3 y el valor de la variable cuya frecuencia acumulada contiene a  $\frac{N}{2} + 1$  es también 3.



|    |      |    |
|----|------|----|
| 3  | 12   | 32 |
| 4  | 4    | 36 |
| 5  | 11   | 47 |
| 6  | 1    | 48 |
| 9  | 1    | 49 |
| 11 | 1    | 50 |
|    | N=50 |    |

### Material descargable Ocho apelli2 gallegos

Así que la mediana será la media aritmética de 3 y 3, que es 3.

$$\text{Mediana}(x) = \frac{3+3}{2} = 3$$

La interpretación en ese ejemplo de la mediana es que el 50 % de los hogares de este grupo cuenta con 3 habitantes o menos y el otro 50 % con 3 habitantes o más.

## 4.5. Moda

### Cálculo de la moda

Para el cálculo de la moda o modas tan sólo habrá que fijarse en el valor o valores de la variable objeto de estudio que tenga la frecuencia absoluta máxima (dicho de otra manera el valor o valores que más se repiten). La moda es una medida poco representativa para variables cuantitativas y suele emplearse para las cualitativas, en las que ni la media ni la mediana se pueden calcular.

**Ejemplo 1: Moda de una variable cualitativa:** Si las provincias de residencia de diez alumnos/as de Galicia elegidos al azar son las siguientes: A Coruña, Pontevedra, A Coruña, Lugo, A Coruña, Pontevedra, Ourense, Ourense, Pontevedra, Lugo. Podríamos calcular la moda de la variable provincia de residencia para ese grupo de alumnos.

| Provincia, $x_i$ | Frecuencia absoluta, $f_i$ |
|------------------|----------------------------|
| A Coruña         | 3                          |
| Lugo             | 3                          |
| Ourense          | 2                          |
| Pontevedra       | 2                          |
|                  | N=10                       |

Si hacemos un recuento, podemos observar que tres residen en A Coruña, otros tres en Pontevedra, dos en Lugo y los dos restantes en Ourense.

Por tanto en este caso habrá dos modas, que serán A Coruña y Pontevedra. La moda en ese caso nos aporta la información de las provincias en las que más alumnado de ese grupo reside.

**Ejemplo 2: Moda de una variable cuantitativa** Como ya adelantamos, la moda no es una medida de centralización muy representativa de las variables cuantitativas, aunque se puede calcular. La moda nos dará entre todos los valores de la variable aquel que más repite, pero eso no siempre quiere decir que sea el más representativo.

Supongamos que tenemos los datos del número de hermanos/as de un grupo de 12 alumnos/as del instituto: 0,0,1,1,1,1,2,4,3,5,6,6.

Si calculamos la moda, esta será 1, ya que es el valor que más se repite, sin embargo no es muy representativo ya que el número medio de hermanos/as de ese grupo es de 2,5.



“Material descargable Ocho apelli2 gallegos”, del proyecto *cREAgal*, se publica con [Licencia Creative Commons Reconocimiento No-comercial Compartir igual 4.0](#)