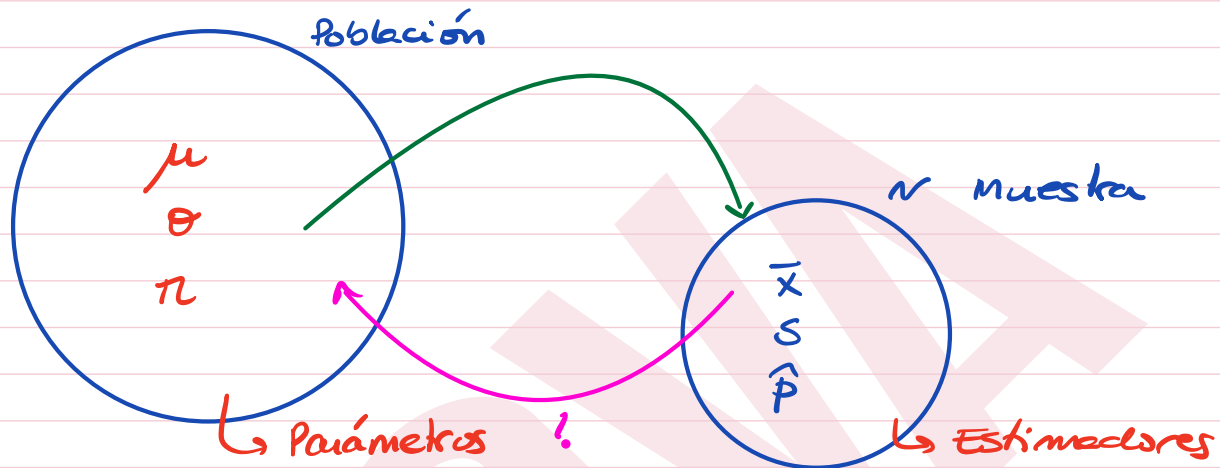


## ESTADÍSTICA I

- ESTADÍSTICA DESCRIPTIVA } Unidimensional  
Bidimensional
- PROBABILIDAD } AXIOMAS  
TIPOS SUCEOS  
TEOREMAS
- VARIABLE ALEATORIA } V. A. DISCRETAS }  $T_S$  } Binomial  
Poisson  
V. A. CONTINUAS }  $T_C$  } Uniforme  
Normal

## T1. CONCEPTO Y CONTENIDO DE LA ESTADÍSTICA

- ESTADÍSTICA DESCRIPTIVA
- ESTADÍSTICA INFERENCIAL } ESTAD II  
ECONOMETRÍA



Población : conjunto de individuos

Muestra : subconjunto de la población

## TIPOS DE DATOS

• DATOS CUANTITATIVOS : se representan con valores numéricos  
no tiene sentido

Ej: edad, salario, ventas,...

• DATOS CUALITATIVOS : expresan una cualidad

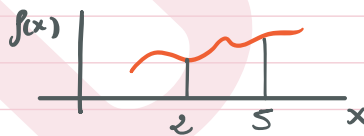
Ej: DNI, nº telefónico, rango ejecutivo,

• DATOS ORDINALES : permiten orden

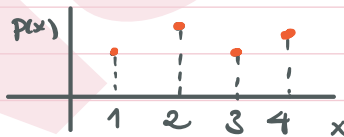
• DATOS NOMINALES : no permiten orden

Ej: colores

• DATOS CONTINUOS



• DATOS DISCRETOS



## T2. DISTRIBUCIONES DE FRECUENCIAS / GRÁFICAS

### \* TABULACIÓN DE DATOS CUANTITATIVOS

$X_i \sim$  edad

$$N = 100$$

$X_i$	$m_i$	$N_i$	$f_i$	$F_i$	$p_i$	$P_i$
17	20	20	0'2	0'2	20%	20%
18	30	50	0'3	0'5	30%	50%
19	29	79	0'29	0'79	29%	79%
21	20	99	0'2	0'99	20%	99%
42	1	100=N	0'01	1	1%	100%
	$\sum m_i = N$ $\sum m_i = 100$		$\sum f_i = 1$			

- $m_i \rightarrow$  frecuencia absoluta  $\rightarrow$  nº repeticiones de cada valor de  $X_i$   
 $\sum m_i = N$
- $N_i \rightarrow$  frecuencia absoluta acumulada

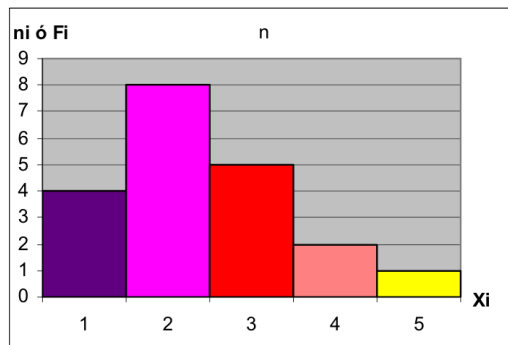
- $f_i \rightarrow$  frecuencia relativa  $\rightarrow f_i = \frac{m_i}{N}$   $0 \leq f_i \leq 1$   $\sum f_i = 1$
- $F_i \rightarrow$  frecuencia relativa acumulada

## Métodos gráficos para datos cualitativos

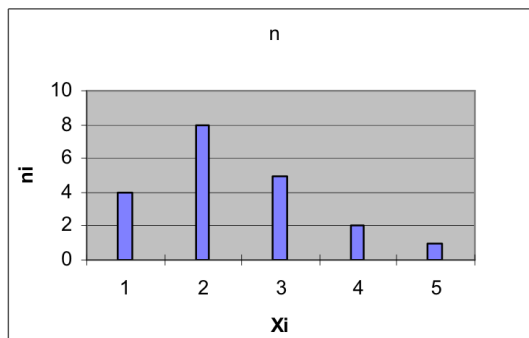
- **Diagramas de barras** (no histogramas)

Se construye colocando las distintas modalidades de la variable cualitativa sobre el eje de abscisas y sobre cada una de ellas se levanta un rectángulo de igual base y altura igual a su frecuencia (absoluta o relativa).

- **Caso vertical:**

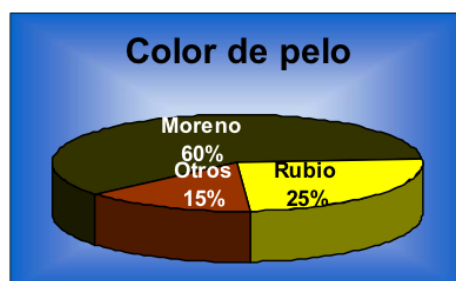


- **Caso horizontal:**



- **Diagramas de sectores**

Sirve para variables cualitativas no agrupadas. Se construye repartiendo el área del círculo en sectores de tamaño proporcional a la frecuencia de cada modalidad. Hay tantos sectores como valores de la variable y los ángulos se calculan de forma proporcional a las frecuencias relativas de cada sector:  $\alpha = fi \cdot 360^\circ$



## ▪ Pictogramas

Consiste en que se emplean figuras relacionadas con el fenómeno que se está estudiando de forma que su tamaño (tipo 1) o número (tipo 2) nos indique la frecuencia asociada a cada modalidad. Las figuras tienen un tamaño proporcional a la frecuencia en el tipo 1.

Tipo de viviendas	ni
Casas	200
Apartamentos	400
Pisos	600



El segundo tipo de pictograma sería el siguiente:



## ▪ Cartogramas

Cada variante identifica una variable de la modalidad que queremos representar.

## Métodos gráficos para datos cuantitativos

### Datos cuantitativos sin agrupar

- **Diagrama de dispersión o nube de puntos**

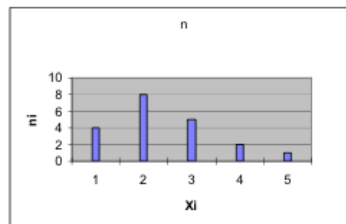
Se colocan los valores de la variable en el eje de las ordenadas y en el eje de las abcisas, la frecuencia (absoluta o relativa) asociada a cada valor.

- **Diagrama de barras** (para tablas de frecuencias no agrupadas)

Se construye igual que el anterior pero en lugar de puntos dibujamos barras. Se anotan sobre el eje de abcisas los distintos valores del carácter y sobre cada uno de ellos se levanta una barra con altura igual a su frecuencia absoluta o relativa.

Ejemplo:

$x_i$	$n_i$	$f_i$	$N_i$	$F_i$
1	4	0'20	4	0'2
2	8	0'40	12	0'6
3	5	0'25	17	0'85
4	2	0'10	19	0'95
5	1	0'05	20	1



Observaciones: la altura es proporcional a las frecuencias, y la base es constante (ya que corresponde a un valor puntual). Al ser la base constante, el área será proporcional a la frecuencia.

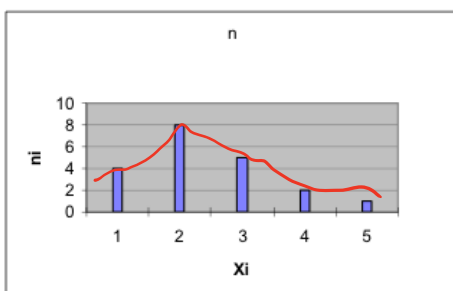
Si unimos los puntos medios de cada barra obtenemos el **polígono de frecuencias**

- **Polígono de frecuencias:**

- **Caso no agrupado en intervalos y frecuencias no acumuladas**

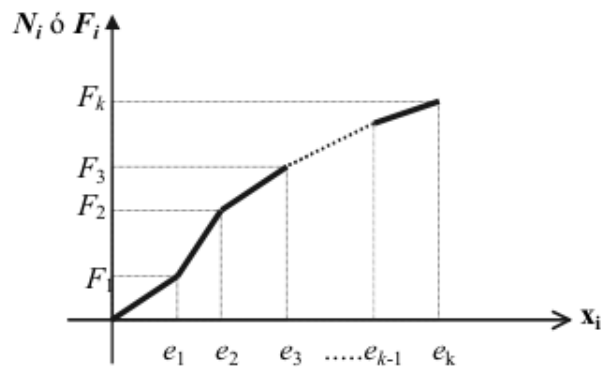
En este caso se construye sobre el diagrama de barras uniando los extremos superiores de barras consecutivas mediante una línea.

Para variables discretas



- **Caso agrupado en intervalos y frecuencias acumuladas**

Se construye marcando los extremos de los intervalos en el eje de abscisas y las frecuencias acumuladas (relativas o absolutas) con la marca en el eje y.





## ▪ Box-Plot (Diagrama de caja y bigotes)

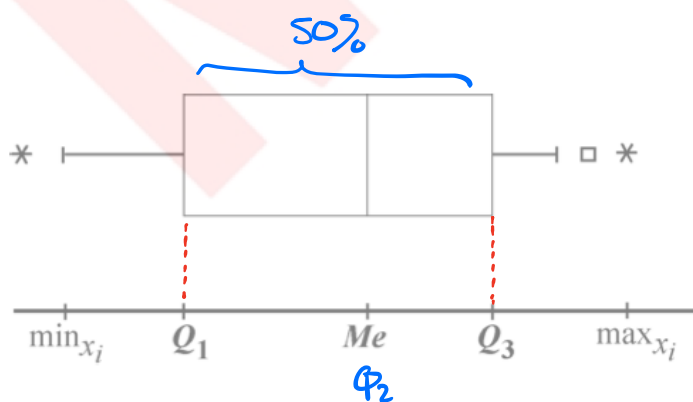
Nos permite hacer representaciones semigráficas de una distribución que nos dejan ver sus características principales y señalar los posibles datos atípicos ("outliers"). El gráfico no se ve afectado por esos datos atípicos.

Proceso de construcción del boxplot:

- se ordenan los datos de menor a mayor y se calculan los cuartiles y la mediana. La caja central queda limitada por  $Q_1$  y  $Q_3$
- se calculan los límites inferior y superior (LI y LS)
- se consideran atípicos los datos fuera del intervalo  $[LI, LS]$
- se dibuja una línea desde cada extremo de la caja central hasta el valor más lejano no atípico
- identificamos los datos fuera de  $[LI, LS]$ , señalándolos como atípicos.

Límite Superior:  $LS = Q_3 + 1,5 \cdot IQR$

Límite Inferior:  $LI = Q_1 - 1,5 \cdot IQR$



Los boxplots son útiles para comparar la distribución de una variable entre diferentes poblaciones.

