

Medidas de Tendencia Central

Santiago Pérez Moncada

20/6/2020

Las **Medidas de tendencia central** son las que dan un valor representativo a todas las observaciones. Algunas de las mas importantes son:

- La **Media aritmética** o **valor medio**

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{j=1}^k n_j X_j}{n} = \sum_{j=1}^k f_j X_j$$

- La **mediana**, que representa el valor central en la lista ordenada de observaciones.
- La **moda** es el valor (o valores) de máxima frecuencia (absoluta o relativa, el resultado será el mismo).

La mediana

La definición formal de la mediana es la siguiente. Denotando por

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

Los datos de la variable cuantitativa ordenados de menor a mayor, la mediana es

- Si n es par, la mediana es

$$\frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

* Si n es impar, el dato central es $x_{\frac{n+1}{2}}$

Ejemplo 1

```
edades = c(13,40,12,27,18,11,40,23,19,18,30,32,40,26,17,12,22,17,40,12)
sort(edades)
```

```
## [1] 11 12 12 12 13 17 17 18 18 19 22 23 26 27 30 32 40 40 40 40
```

```
table(edades)
```

```
## edades
## 11 12 13 17 18 19 22 23 26 27 30 32 40
##  1  3  1  2  2  1  1  1  1  1  1  1  4
```

En este caso la moda es 40, la mediana es $\frac{19+22}{2} = 20.5$ y la media aritmetica es:

$$\frac{13 + 40 + 12 + 27 + 18 + 11 + 40 + 23 + 19 + 18 + 30 + 32 + 40 + 26 + 17 + 12 + 22 + 17 + 40 + 12}{20} = 23.45$$

Ejercicio 2

```
set.seed(2020)
datos = sample(1:6,size = 25,replace = TRUE)
sort(datos)
```

```
## [1] 1 1 1 2 2 2 2 2 3 4 4 4 4 4 4 5 5 5 6 6 6 6 6 6
```

```
table(datos) #Frecuencia absoluta
```

```
## datos
## 1 2 3 4 5 6
## 3 6 1 6 3 6
```

```
round(prop.table(table(datos)),2) #Frecuencia Relativa global
```

```
## datos
##   1   2   3   4   5   6
## 0.12 0.24 0.04 0.24 0.12 0.24
```

```
cumsum(table(datos)) #Frecuencia absoluta acumulada
```

```
## 1 2 3 4 5 6
## 3 9 10 16 19 25
```

```
round(cumsum(prop.table(table(datos))),2) #Frecuencia relativa acumulada
```

```
##   1   2   3   4   5   6
## 0.12 0.36 0.40 0.64 0.76 1.00
```

```
datos.df = data.frame(Puntuacion = 1:6,
                      Fr.abs = as.vector(table(datos)),
                      Fr.rel = as.vector(round(prop.table(table(datos)),2)),
                      Fr.acu = as.vector(cumsum(table(datos))),
                      Fr.rel.acu = as.vector(round(cumsum(prop.table(table(datos))),2)))

datos.df
```

##	Puntuacion	Fr.abs	Fr.rel	Fr.acu	Fr.rel.acu
## 1	1	3	0.12	3	0.12
## 2	2	6	0.24	9	0.36
## 3	3	1	0.04	10	0.40
## 4	4	6	0.24	16	0.64
## 5	5	3	0.12	19	0.76
## 6	6	6	0.24	25	1.00

```
set.seed(NULL)
```

En este caso la **Moda** son tres valores 2,4,6 la **mediana** es $x_{(13)} = 4$ y la **media aritmetica** es 3.72
Vamos a calcular la media aritmetica, mediana, moda de los ejemplos anteriores con instrucciones de R.

```
mean(edades) # media aritmetica
```

```
## [1] 23.45
```

```
mean(dados)
```

```
## [1] 3.72
```

```
median(edades) #mediana
```

```
## [1] 20.5
```

```
median(dados)
```

```
## [1] 4
```

```
as.numeric(names(which(table(edades) == max(table(edades))))) #moda
```

```
## [1] 40
```

```
as.numeric(names(which(table(dados) == max(table(dados)))))
```

```
## [1] 2 4 6
```

Cuando trabajamos con datos cuantitativos, es conveniente que el resultado lo demos con un número. De ahí que hayamos aplicado la función `as.numeric`