

Tema 1

Estadística descriptiva:

Medidas de posición

Curso 2017/18
Grados en biología sanitaria
Departamento de Física y Matemáticas
Marcos Marv Ruiz



MEDIDAS DE POSICIÓN

Mediana, cuartiles,...., y cuantiles variables cuantitativas y cualitativas ordenadas

Mediana: valor de la variable que, una vez ordenados de menor a mayor, deja la mitad de los datos por debajo de sí:

- * Si hay una cantidad **impar** de datos, se toma el valor del centro
- * Si hay una cantidad **par** de datos, se toma la media entre los dos centrales.

Es *robusta* frente a (unos pocos) valores extremos.

MEDIDAS DE POSICIÓN

Mediana, cuartiles,..., y cuantiles variables cuantitativas y cualitativas ordenadas

Mediana: valor de la variable que, una vez ordenados de menor a mayor, deja la mitad de los datos por debajo de sí:

- * Si hay una cantidad **impar** de datos, se toma el valor del centro
- * Si hay una cantidad **par** de datos, se toma la media entre los dos centrales.

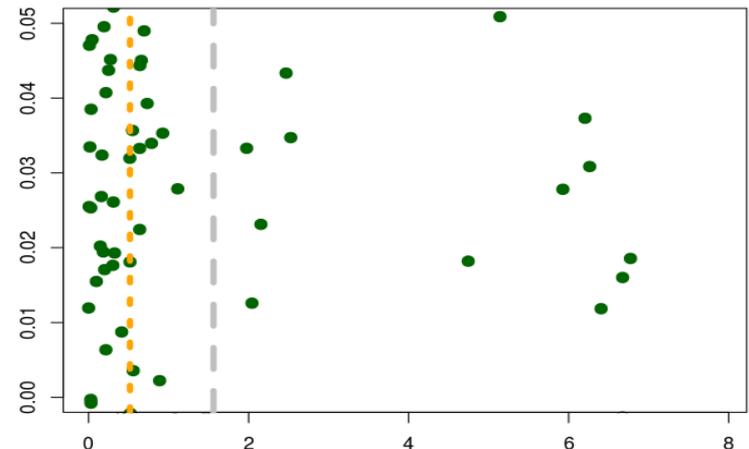
Es *robusta* frente a (unos pocos) valores extremos.

Ejemplos:

* Para $\{1, 4, 5, 8, 27\}$ la mediana vale 5

* En los dos casos

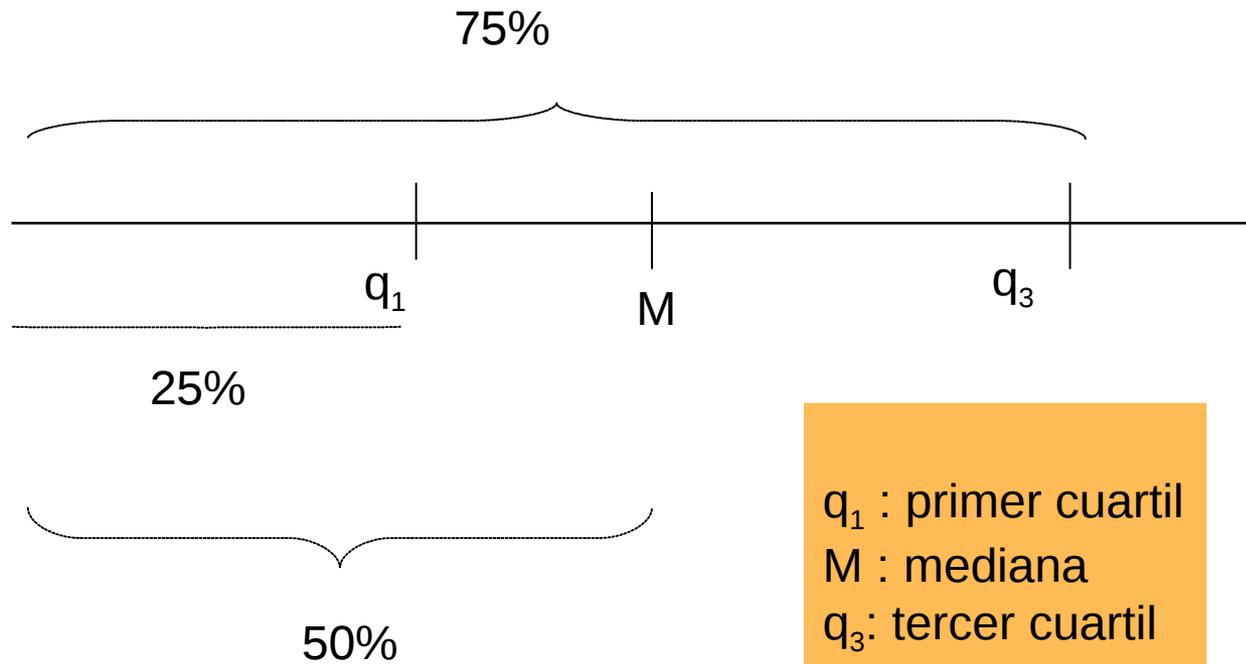
$\{8, 9, 9, 9, 10, 10\}$ $\{1, 1, 9, 9, 10, 10\}$
la mediana vale 9



Linea a puntos (izquierda) la mediana
Linea a guiones (derecha) la media

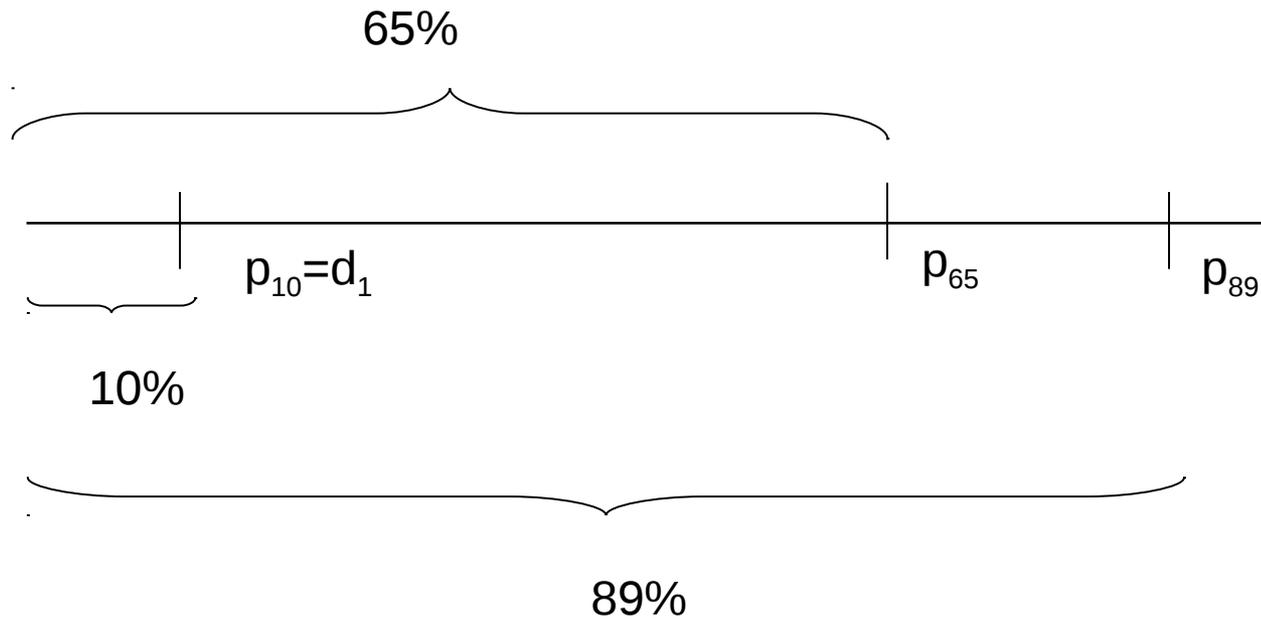
MEDIDAS DE POSICIÓN

Mediana, cuartiles, percentiles



MEDIDAS DE POSICIÓN

Deciles, percentiles



MEDIDAS DE POSICIÓN

Mediana, cuartiles,...., y cuantiles

Cuartiles: valores de la variable que, una vez ordenados de menor a mayor, dividen la muestra en cuatro partes con el mismo número de elementos.

Ejemplo: considera los datos

- 20, 21, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 71, 89, 108.
- La mediana es $Md = Q2 = (40+40)/2 = 40$
- El primer cuartil es $Q1 = (32+35)/2 = 33.5$
- El tercer cuartil es $Q3 = (47+50)/2 = 48.5$

MEDIDAS DE POSICIÓN

Cuartiles: diagrama de caja y bigotes (boxplot) Idea: representar los cuartiles y señalar los puntos muy alejados de los que hay entre los cuartiles 1º y 3º

MEDIDAS DE POSICIÓN

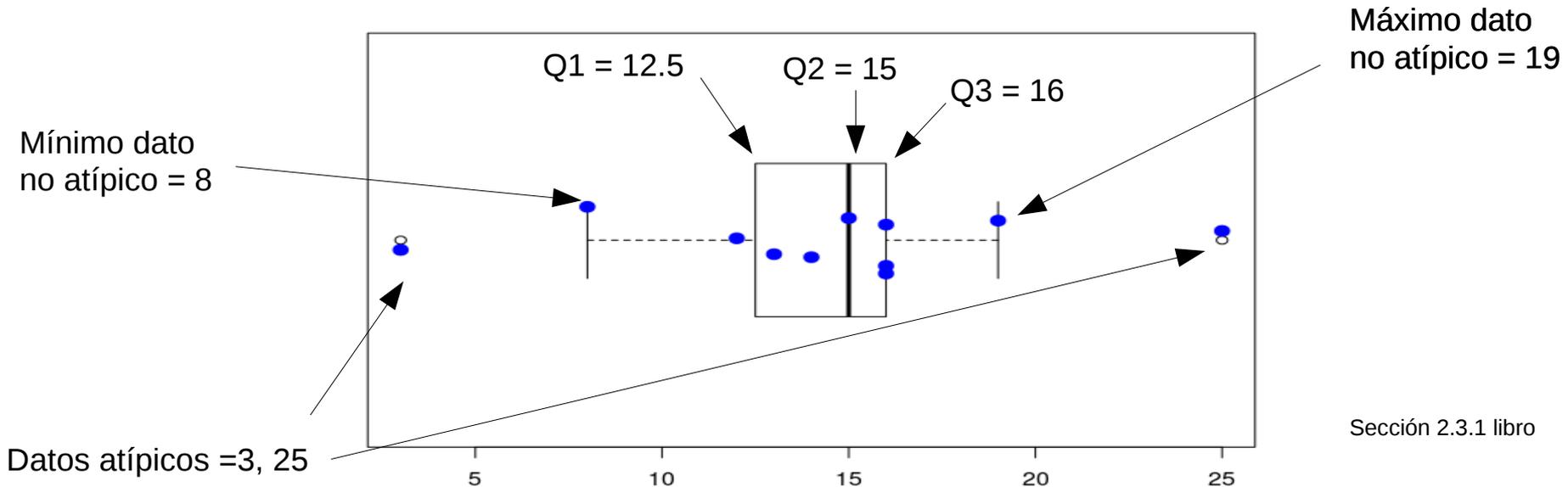
Cuartiles: diagrama de caja y bigotes (boxplot) Idea: representar los cuartiles y señalar los puntos muy alejados de los que hay entre los cuartiles 1º y 3º

Ejemplo: 3, 8, 12, 13, 14, 15, 16, 16, 16, 19, 25

Recorrido intercuartilico: $IQR = Q3 - Q1 = 16 - 12.5 = 3.5$

Es el rango del 50% "central" de la muestra; ¿hay datos muy grandes/pequeños?

- Los que sean menores que $Q1 - 1.5 * IQR = 7.25$
- Los que sean mayores que $Q3 + 1.5 * IQR = 21.25$



MEDIDAS DE POSICIÓN

Deciles, centiles y cuantiles

Ejercicio: Se ha contado el número de especies diferentes que hay en 10 parcelas
{13, 3, 11, 9, 4, 5, 8, 17, 15, 2}

¿Cuántas especies hay, como mucho, en el 30% de parcelas con menor diversidad?

MEDIDAS DE POSICIÓN

Deciles, centiles y cuantiles

Ejercicio: Se ha contado el número de especies diferentes que hay en 10 parcelas
{13, 3, 11, 9, 4, 5, 8, 17, 15, 2}

¿Cuántas especies hay, como mucho, en el 30% de parcelas con menor diversidad?

El **cuantil q** (x_q) es el valor de la variable tal que, una vez ordenados de menor a mayor el $q\%$ de los datos es menor o igual que x_q

Ejemplo: dados los valores

23, 24, 44, 54, 58, 64, 69, 73, 85, 90

el 6º decil es $d_6 = 64$

¿Cuánto vale el percentil 82?

MEDIDAS DE POSICIÓN

Cuantiles y tablas de frecuencias acumuladas

Frecuencia absoluta acumulada de un elemento o clase (si agrupados):

Nº de elementos menor o igual que él.

$$F_i = f_1 + f_2 + \dots + f_i = \sum_{i \leq j} f_j$$

Frecuencia relativa acumulada de un elemento o clase (si agrupados):

es el tanto por uno de elementos menor o igual que él

$$F'_i = f'_1 + f'_2 + \dots + f'_i = \sum_{i \leq j} f'_j$$

Ejemplo:

¿Cuánto vale la mediana?

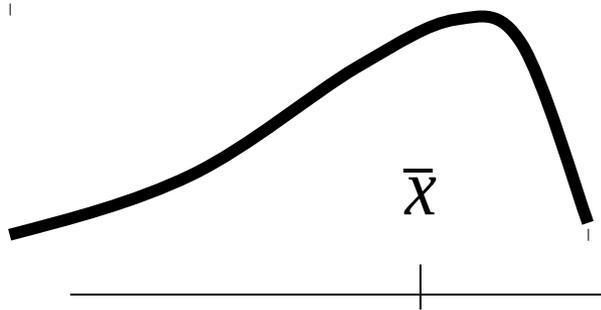
¿Cuánto vale el 3^{er} decil? ¿y el 6^o?

| Nº analgésicos | f _i | F _i | f' _i | F' _i |
|----------------|----------------|----------------|-----------------|-----------------|
| 1 | 1 | 1 | 0,04 | 0,04 |
| 2 | 2 | 3 | 0,08 | 0,12 |
| 3 | 3 | 6 | 0,12 | 0,24 |
| 4 | 4 | 10 | 0,16 | 0,4 |
| 5 | 5 | 15 | 0,2 | 0,6 |
| 6 | 4 | 19 | 0,16 | 0,76 |
| 7 | 3 | 22 | 0,12 | 0,88 |
| 8 | 2 | 24 | 0,08 | 0,96 |
| 9 | 1 | 25 | 0,04 | 1 |

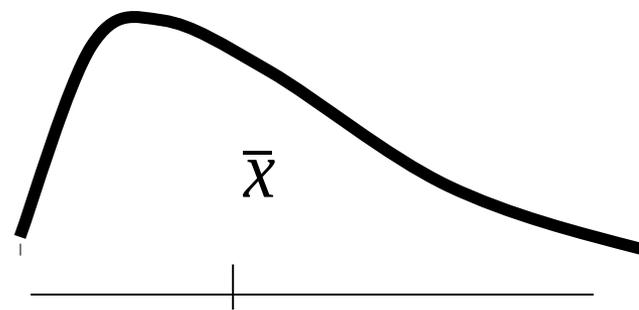
¿Cuánto vale el primer cuartil?

¿Cuánto vale el percentil 23?

Diagrama de dispersión datos univariados y simetrías



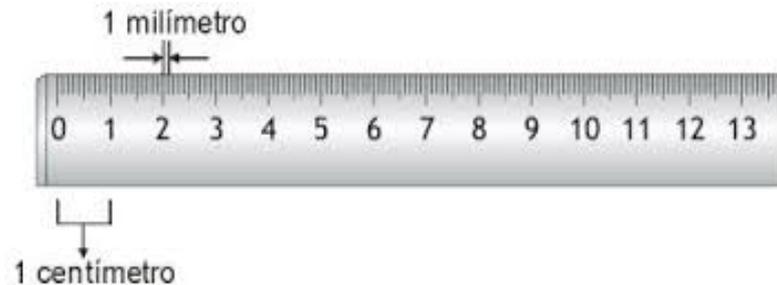
Asimetría negativa (a la izqda.)



Asimetría positiva (a la dcha.)

Redondeo a cifras significativas (sección 1.3 libro)

La variable "altura de un individuo" ¿es continua, o discreta?



Algo que mide 5cm = 50mm sabemos que está entre 49mm y 51mm. Error: 1mm

Si dividimos esa longitud entre 3, la calculadora dice que cada parte mide

16.66666667mm

¿Tiene sentido (significado) usar tantos decimales?

$$456 / 45.4 = \mathbf{10.0440528634}$$

$$456 / 45.5 = \mathbf{10.021978022}$$

$$456 * 45.4 = \mathbf{20702,4}$$

$$456 * 45.5 = \mathbf{20748}$$

Redondeo a cifras significativas (sección 1.3 libro)

Ejemplo 1: redondear 2.712798 a 4 cifras significativas.

1.- Buscar primera cifra (más a la izquierda) no nula

2.712798

2.- contar 4 cifras (hacia la derecha)

2.712798

3.- Qué hacer con el resto de cifras: redondear la 5^a (el resto se obvian):

* Si es menor o igual que 4, se elimina sin más.

* Si es mayor o igual que 5, se elimina y se suma 1 a la 4^a cifra.

2.713

Posible acarreo

Redondeo a cifras significativas (sección 1.3 libro)

Ejemplo 2: redondear 0.023899275 a 5 cifras significativas.

1.- Buscar primera cifra (más a la izquierda) no nula

0.0**2**3899725

2.- contar 5 cifras (hacia la derecha)

0.0**23899**725

3.- Qué hacer con el resto de cifras: redondear la 6^a (el resto se obvian):

0.0**23899****7**25

* Si es menor o igual que 4, se elimina sin más.

* Si es mayor o igual que 5, se elimina y se suma 1 a la 5^a cifra. Si la 5^a cifra es un 9, anota un 0 y suma 1 a la 4^a cifra, y así sucesivamente (*acarreo*)

0.0**23900**