

MEDIDAS DE DISPERSIÓN

→ 12, 13, 14, 15, 16, 17, 18

Grupo 1: 15, 16, 14, 17, 13, 12, 18

$$\rightarrow \bar{X}_1 = 15$$

$$me_1 = 15$$

Grupo 2: 10, 16, 7, 16, 18, 19, 19

$$\rightarrow \bar{X}_2 = 15$$

$$me_2 = 16$$

+ medidas de dispersión o variabilidad

↘ 7, 10, 16, 16, 18, 19, 19

Grupo 3: 7, 7, 7, 16, 19, 19, 19

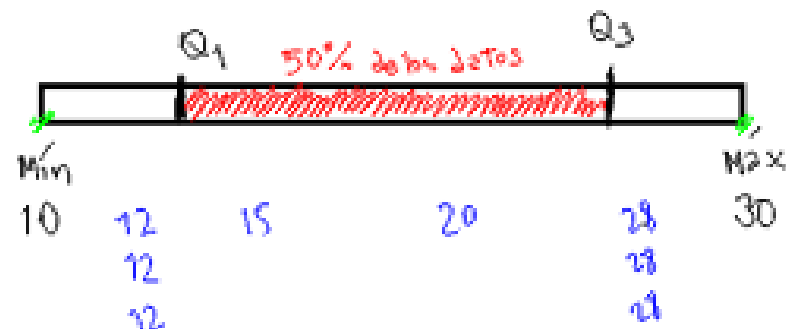
1) $\text{Rango} = \max - \min$
r

Grupo 1 $18 - 12 = 6$

Grupo 2 $19 - 7 = 12$

Grupo 3 $19 - 7 = 12$

2) $\text{RIC} = Q_3 - Q_1 = P_{75} - P_{25}$



Q_1 — Q_2
 Q_2 — Q_3

Temp (°C): -2, -5, -4, 3

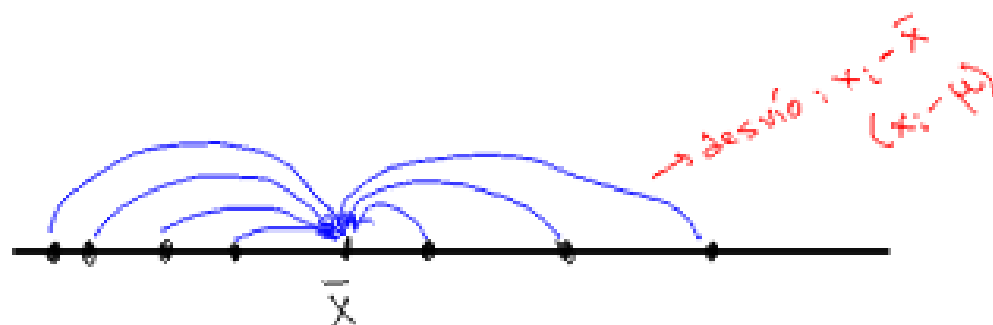
$$r = 3 - (-5) = 8$$

-2, -5, -4, -7

$$r = -2 - (-7) = 5$$

10 12 13 15 17 20 25 28 29 30

3) Varianza



$$\sum \text{desvíos} = 0$$

$$\frac{\sum \text{desvíos}^2}{N} \rightarrow \checkmark$$

$$IMC \rightarrow \text{kg/m}^2$$

$$s_{mc}^2 = 4(\text{kg/m}^2)^2 = 4 \text{ kg}^2/\text{m}^4$$

$$\text{Ingresos} \rightarrow \$$$

$$s_{ingresos}^2 = ? \text{ 200 \$}^2?$$

Ejmi:

Grupo 2: 10, 16, 7, 16, 18, 19, 19 $\rightarrow \bar{x} = 15$

desvíos: -5, 1, -8, 1, 3, 4, 4, $\sum \text{desvíos} = 0$

desvíos²: 25, 1, 64, 1, 9, 16, 16

\downarrow
puntos²

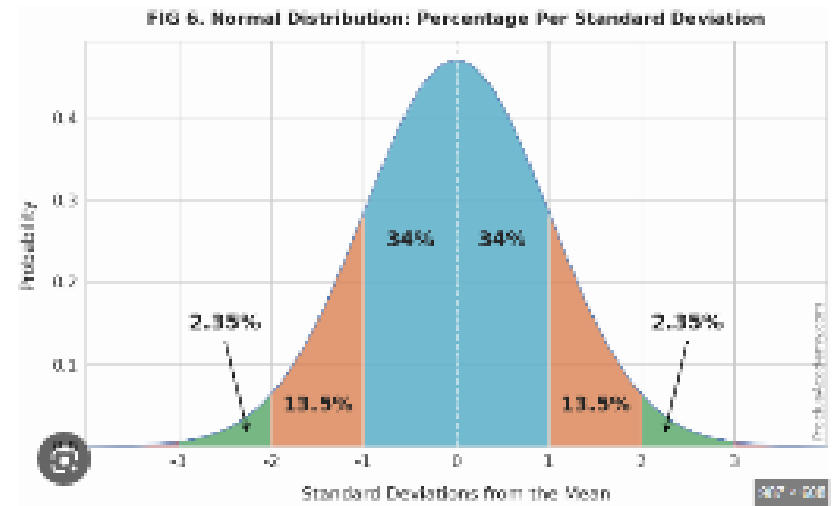
$$* \sum \frac{(e_i - \bar{e})^2}{n-1} : \text{varianza}$$

$$* \sum \frac{|e_i - m_e|}{n-1} \rightarrow \text{Análisis más complejo}$$

desviación absoluta

Ejm: $\bar{x} = 15$, $S = 2$ $(13, 19) \rightarrow$

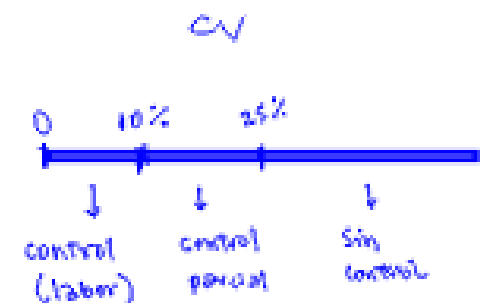
$\bar{x} = 30$, $S^2 = 9$ $(27, 33)$
 $S = 3$

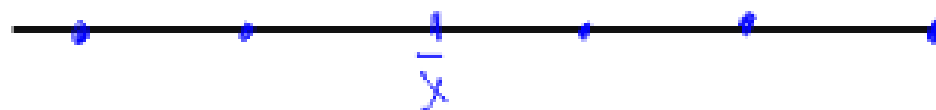


		S	\bar{x}	$CV = \frac{S}{\bar{x}} \cdot 100\%$
Estatura: \nearrow cm	170, 158, 160, 165, 176	7.36 cm	165.8 cm	4.44 %
Peso: \searrow kg	72, 55, 68, 60, 70	7.21 kg	65 kg	11.09% \rightarrow mayor dispersión
Peso insecto: \nearrow g	12, 15, 18, 12, 15, 16	234 g	14.67 g	15.94%
Peso mamífero: \searrow g	1500, 1250, 1600, 985, 1540, 1320	229.66 g	1365.83 g	16.81% \rightarrow mayor dispersión

MAGNITUDES:

Rango \rightarrow	$0 \leq \infty$ (sin límite)	$(\mathbb{R}^+ \cup \{0\})$
RIC \rightarrow	$0 \leq \infty$ (sin límite)	$(\mathbb{R}^+ \cup \{0\})$
$s^2 \rightarrow$	$0 \leq \infty$ (sin límite)	$(\mathbb{R}^+ \cup \{0\})$
$s \rightarrow$	$0 \leq \infty$ (sin límite)	$(\mathbb{R}^+ \cup \{0\})$
CV \rightarrow	$0 \leq \infty$ (sin límite)	$(\mathbb{R}^+ \cup \{0\})$





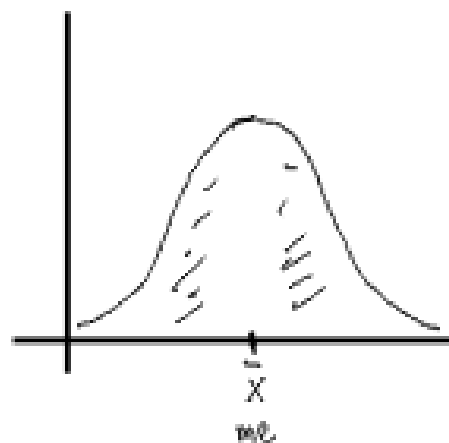
$$S = \sqrt{\frac{\quad}{4}}$$



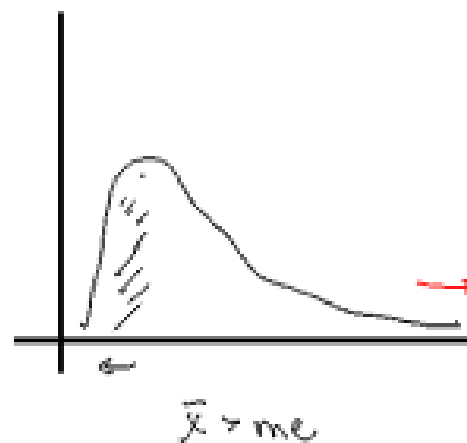
$$S = \sqrt{\frac{\quad}{8}} \rightarrow CV \downarrow$$

MEDIDAS DE ASIMETRÍA

- ▶ Distribución simétrica: La curva es equilibrada respecto a la media
- ▶ Distribución asimétrica positiva. La distribución de los datos presenta una curva con cola a la derecha. La mayor cantidad de los datos son de menores valores.
- ▶ Distribución asimétrica negativa. La distribución de los datos presenta una curva con cola a la izquierda. La mayor cantidad de los datos son de mayores valores.

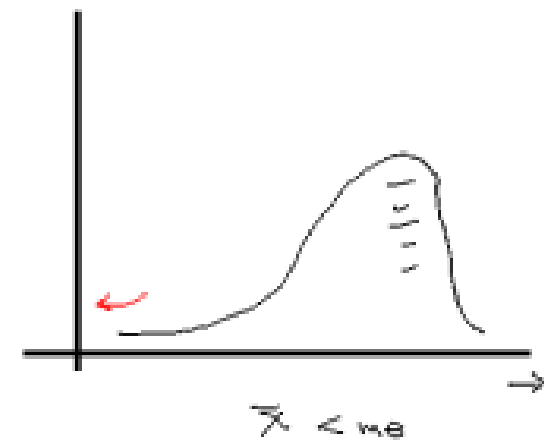


Simetría



Asimetría positiva

$$as = 0.96$$



Asimetría negativa

$$as = -1.7$$

$$as_{FP} = \frac{1}{n} \times \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{s^3}$$

$as_{FP} \in \mathbb{R}$

$-0.5 \text{ a } 0.5 \rightarrow \text{Simetría}$

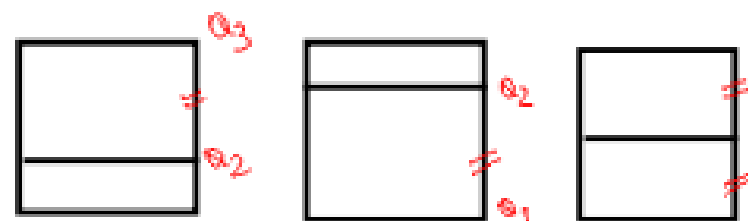
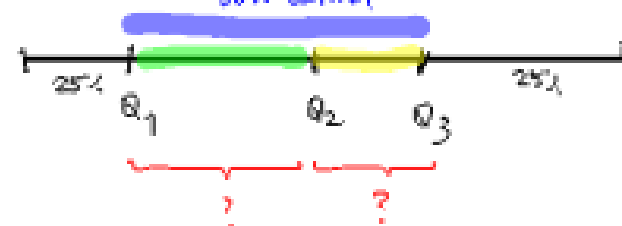
Ejm: $as_1 = 0.19 \rightarrow \text{Más simétrico}$

$as_2 = -1.70 \rightarrow \text{Más asimétrico}$

$as_3 = 0.96$

$$as_B = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$$

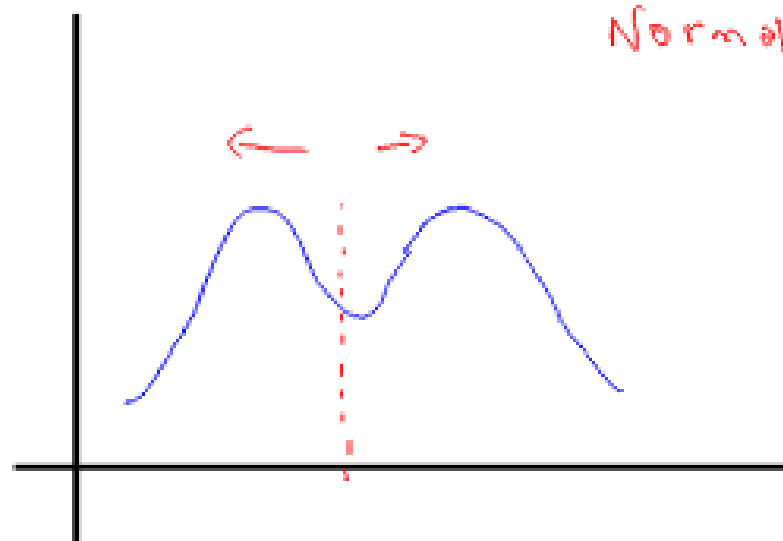
> 0 : Asimetría positiva en el 50% central de los datos
 < 0 : Asimetría negativa en el 50% central de los datos
 $= 0$: Simetría en el 50% central de los datos



As. positiva
50% central

As. negat.
50% central

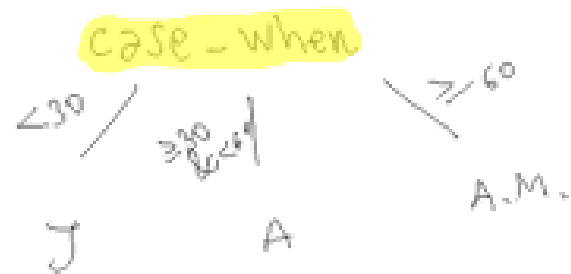
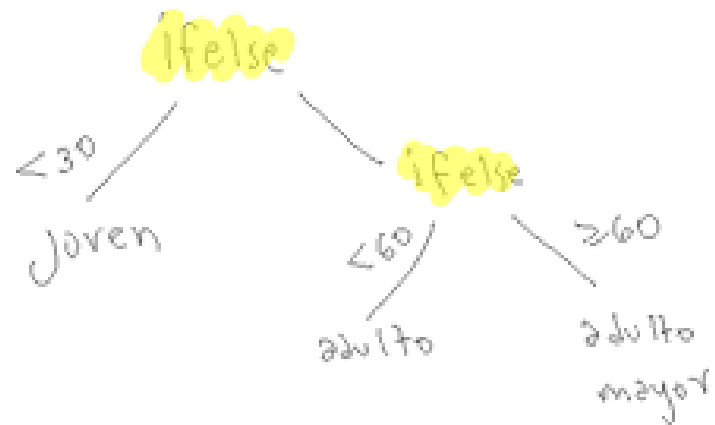
Simetría
50% central



Normal $\xrightarrow{\text{SÍ}}$ Simétrica
 $\xleftarrow{\text{NO}}$



Si

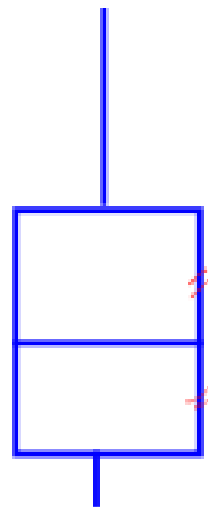


✓✓✓
Las cuatro variables presentan distribuciones que tienden a la simetría, siendo la edad la que más se acerca a la simetría y el IMC la que más se aleja.

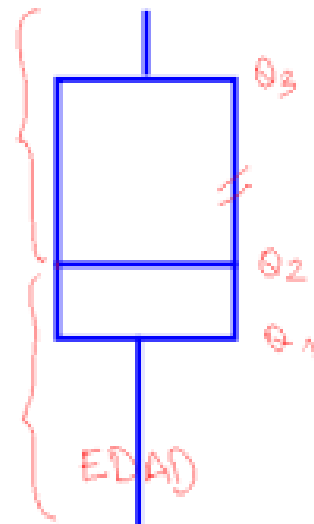
100% datos

Las cuatro variables presentan distribuciones que tienden a la simetría en el 50% central de los datos, siendo el IMC la que más se acerca a la simetría y la edad la que más se aleja. ¿Es una contradicción respecto a la interpretación con el coeficiente de asimetría de Fisher Pearson?

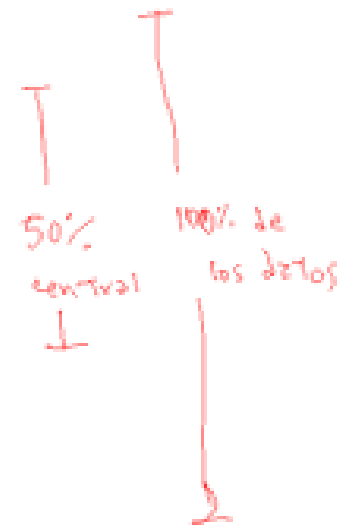
50% central

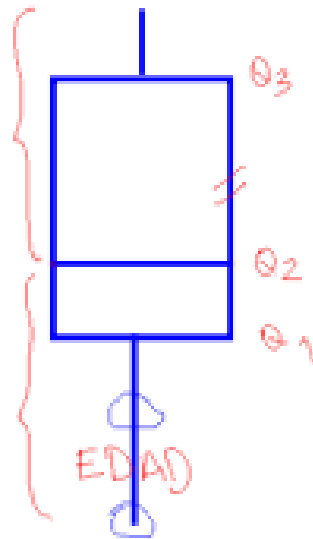
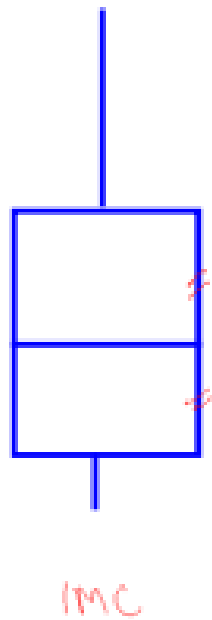


IMC



EDAD





En el caso del IMC, los valores cercanos al centro están distribuidos de manera simétrica o uniforme (50% central), pero a medida que se alejan, hay más datos altos de IMC que bajos (100% central). [¿datos altos = outliers?]

En el caso de la edad, los valores cercanos al centro no están distribuidos de manera simétrica, sino que hay más valores por encima. Sin embargo, al analizar la distribución completa de los datos, la distribución por debajo y por encima de la media sí es similar.