

Para hallar la media y la desviación estándar con la calculadora (Para fx-82MS o similares):

**1° paso:** Poner en modo SD, haciendo: MODE -> 2 (SD)

**2° paso:** Borrar datos viejos: SHIFT -> MODE -> 1 -> = -> AC (SIEMPRE)

**3° paso:** Cargar los datos: 3 -> M+ -> 5 -> M+ ->....-> AC

**4° paso:** Calcular la media: SHIFT -> 2 -> 1 -> =

Calcular la desviación estándar: SHIFT -> 2 -> 3 -> =

#### PARA TABLA DATO A DATO:

\*fa: frecuencia absoluta correspondiente a cada fila en la tabla.

#### Ejemplo:

Datos	fa
3	3
5	2
8	1
9	7

Cargamos los datos:

3 -> SHIFT -> ; -> 3 (fa) -> M+ -> 5 -> SHIFT -> ; -> 2 (fa)-> M+ -> 8 -> SHIFT -> ; -> 1 (fa) -> M+ -> 9 -> SHIFT -> ; -> 7 (fa) -> M+ -> AC

Calcular la media: SHIFT -> 2 -> 1 -> =. Media del ejemplo: **6,92**.

Calcular la desviación estándar: SHIFT -> 2 -> 3 -> =. Desviación estándar del ejemplo: **2,66**.

#### PARA TABLA CON DATOS AGRUPADOS:

\*PM1: Punto medio 1, PM2: Punto medio 2, etc.

#### Ejemplo:

Datos	PM	fa
[25-33)	29	8
[33-41)	37	9
[41-49)	45	5

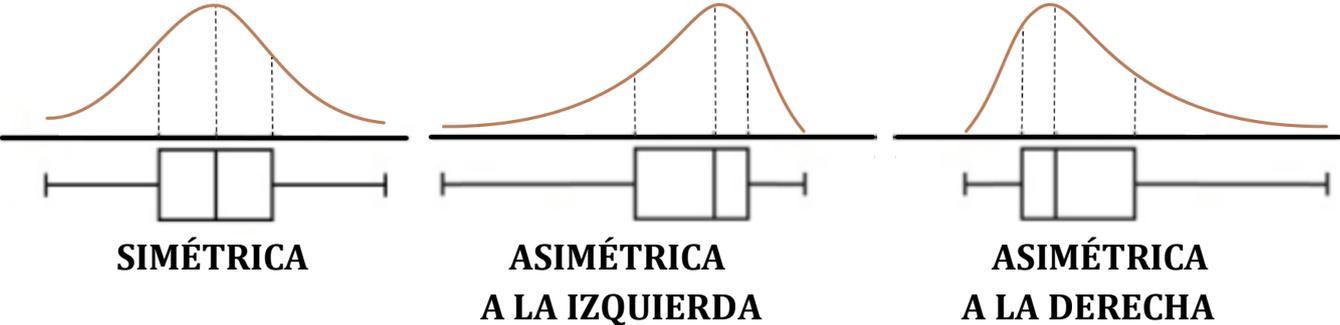
Cargamos los datos:

29 (PM1) -> SHIFT -> ; -> 8 (fa) -> M+ -> 37 (PM2) -> SHIFT -> ; -> 9 (fa) -> M+ -> 45 (PM3) -> SHIFT -> ; -> 5 (fa) -> M+ -> AC

Calcular la media: SHIFT -> 2 -> 1 -> =. Media del ejemplo: **35,9**.

Calcular la desviación estándar: SHIFT -> 2 -> 3 -> =. Desviación estándar del ejemplo: **6,19**.

**SIMETRÍA O ASIMETRÍA EN UN DIAGRAMA DE CAJA**



## Ejercicios resueltos TP2 Parte 3

**Actividad 12:** Los datos que se presentan a continuación detallan el porcentaje de familias que vive por debajo del umbral de pobreza en 20 países del Tercer Mundo:

20	46	65	38	52	67	29	36	45	51
53	64	61	60	55	45	39	45	66	23

a) Realiza el diagrama de tallo y hojas y el diagrama de caja correspondiente.

A continuación, realizamos el diagrama de tallos y hojas:

<b>2</b>		0	3	9					
<b>3</b>		6	8	9					
<b>4</b>		5	5	5	6				
<b>5</b>		1	2	3	5				
<b>6</b>		0	1	4	5	6	7		

Unidad: 1

Por último, hacemos el diagrama de caja correspondiente pero primero calculamos los elementos necesarios para hacerlo:

$$\text{Posición } Q_1: \frac{21}{4} = 5,25$$

$$\text{Posición } M_e: 2 \cdot \frac{21}{4} = 10,5$$

$$\text{Posición } Q_3: 3 \cdot \frac{21}{4} = 15,75$$

$$Q_1 = 38 + 0,25 \cdot (39 - 38) = \mathbf{38,25}$$

$$M_e = 46 + 0,5 \cdot (51 - 46) = \mathbf{48,5}$$

$$Q_3 = 60 + 0,75 \cdot (61 - 60) = \mathbf{60,75}$$

Veamos si existen valores atípicos:

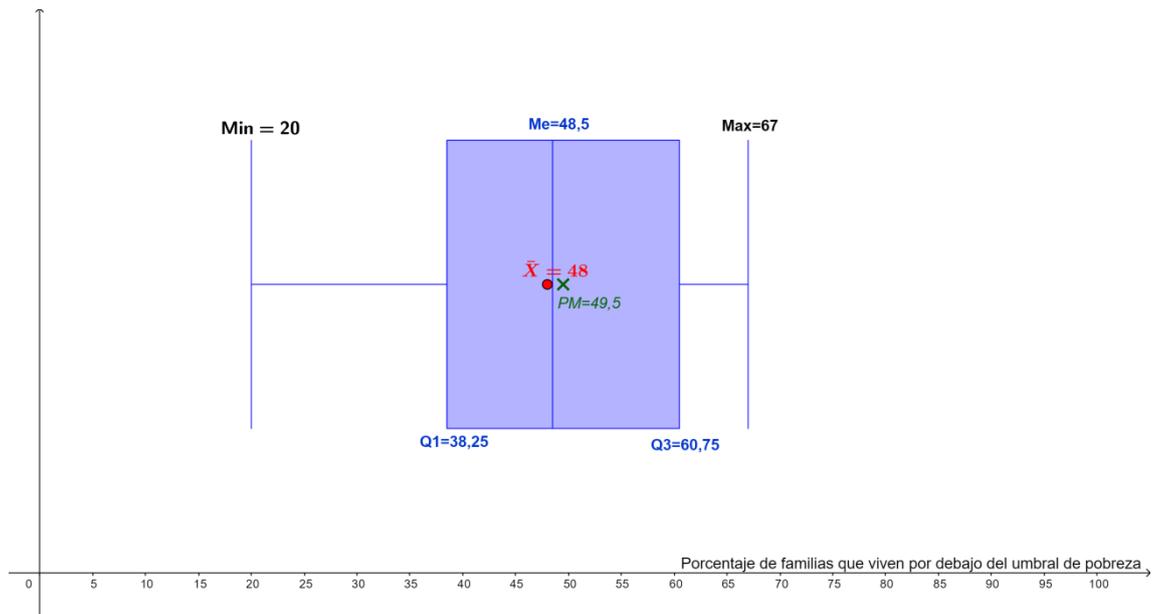
$60,75 + 1,5 \cdot (60,75 - 38,25) = \mathbf{94,5}$ . Este valor es el límite para detectar valores atípicos por exceso, el valor más alto de la muestra es 67, que no supera dicho límite.

Veamos qué sucede por defecto:

$38,25 - 1,5 \cdot (60,75 - 38,25) = \mathbf{4,5}$ . Este valor es el límite para detectar valores atípicos por defecto, el valor más bajo de la muestra es 20, que no está por debajo de dicho límite.

Por lo tanto, **no existen valores atípicos**.

Ahora, graficamos el diagrama de caja:



b) ¿Hay valores alejados en esta distribución? ¿Cuáles son?

Como se explicó en el ítem anterior, **no hay valores alejados en esta distribución.**

c) Basándote en los diagramas, ¿se podría decir que la distribución es simétrica?

Como se puede observar en los diagramas, la distribución **NO ES SIMÉTRICA**. Se puede ver que en la caja es aproximadamente simétrica pero podemos notar que en los bigotes hay una asimetría a la izquierda. Y en cuanto al diagrama de tallo y hojas, se puede apreciar una asimetría a la izquierda.

d) Calcular las medidas de tendencia central y de dispersión adecuadas.

Como hemos mostrado en el ítem a), no hay valores atípicos, por lo tanto, lo mejor es calcular la **media** como medida de tendencia central y la **desviación estándar** como medida de dispersión:

$$\bar{X} = 48 \text{ and } S=14,08$$

Además, en el ítem a), ya hemos calculado la mediana  $M_e=48,5$  y el RIC=22,5.

e) Calcular el coeficiente de asimetría e interpretarlo.

Como no dice cuál de los coeficientes de asimetría usar, podemos calcular los dos:

COEFICIENTE DE ASIMETRÍA DE PEARSON:

$$C_{AP} = \frac{3(\bar{X} - Me)}{S} = \frac{3.(48 - 48,5)}{14,08} = -0.10$$

COEFICIENTE DE ASIMETRÍA DE BOWLEY:

$$AsB = \frac{Q_3 + Q_1 - 2.Me}{Q_3 - Q_1} = \frac{60,75 + 38,25 - 2.48,5}{60,75 - 38,25} = 0,08$$

Para que una distribución sea asimétrica, el resultado de los coeficientes tienen que ser o muy mayor a 0 o muy menor a 0. Si los coeficientes se encuentran en el intervalo  $[-2; 2]$ , la distribución se considera aproximadamente simétrica.

Por tanto, podríamos decir que la distribución es aproximadamente simétrica pero ¡¡OJO!! estos coeficientes proporcionan información sobre la simetría o asimetría de la parte central de la distribución. Es apropiado acompañar estas medidas con el gráfico para corroborar si la información dada por los coeficientes se verifica en las partes extremas de la distribución.

Así observando también los diagramas de tallo y hoja y los diagramas de caja, podemos decir, que en la parte central de la distribución es aproximadamente simétrica mientras que en las partes extremas es asimétrica a la izquierda.

**Actividad 13:** Los siguientes datos representan los tiempos de reacción frente a un determinado estímulo que se aplicó en un experimento a una muestra aleatoria de 14 ratas de laboratorio:

5.4 - 1.3 - 4.1 - 2.9 - 0.7 - 2.6 - 8.5 - 0.6 - 7.1 - 1.9 - 8.2 - 3.9 - 14.8 - 9.8

a) Traza un diagrama de tallo y hojas.

Realizamos el diagrama de tallo y hojas:

0	6	7
1	3	9
2	6	9
3	9	
4	1	
5	4	
6		
7	1	
8	2	5
9	8	
10		
11		
12		
13		
14	8	

Unidad: 0,1

b) Realiza el resumen de los cinco números y el diagrama de caja.

RESUMEN DE LOS CINCO NÚMEROS:

Min: 0,6  
Max: 14,8  
Q<sub>1</sub>: 1,75  
Me: 4  
Q<sub>3</sub>: 8,28

Veamos si existen valores atípicos:

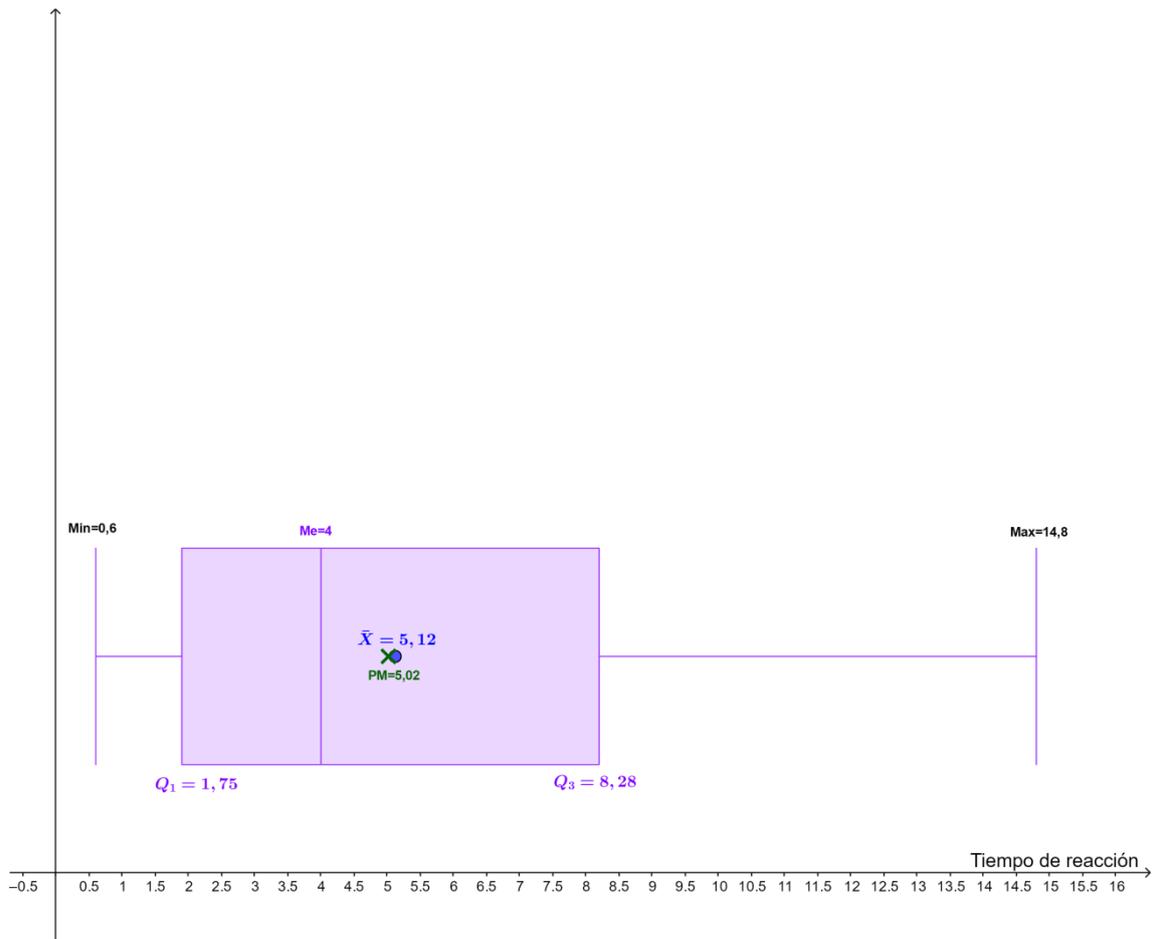
$8,28 + 1,5 \cdot (8,28 - 1,75) = 18,08$ . Este valor es el límite para detectar valores atípicos por exceso, el valor más alto de la muestra es 14,8, que no supera dicho límite.

Veamos qué sucede por defecto:

$1,75 - 1,5 \cdot (8,28 - 1,75) = -8,04$ . Este valor es el límite para detectar valores atípicos por defecto, el valor más bajo de la muestra es 0,6, que no está por debajo de dicho límite.

Por lo tanto, **no existen valores atípicos.**

Ahora, graficamos el diagrama de caja:



c) ¿Es asimétrica esta distribución? Justificar.

Como podemos observar en el diagrama de caja, la distribución ES ASIMÉTRICA. La distribución es ASIMÉTRICA A LA DERECHA ya que como se puede observar, la mediana se posiciona a la izquierda del punto medio de la caja. Además, el bigote a la izquierda es mucho más corto que el bigote a la derecha.

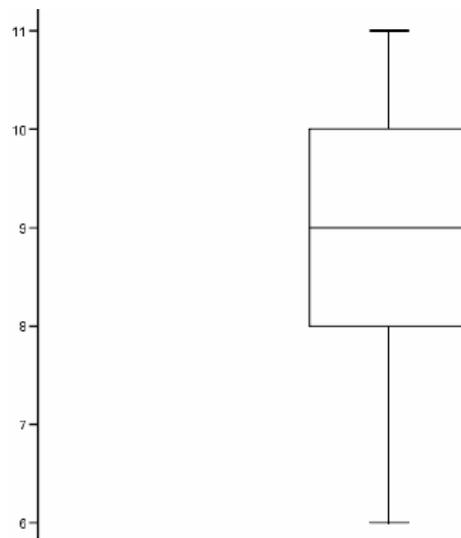
d) ¿Cuál de los dos gráficos representa mejor la información en este caso? ¿Por qué?

De los dos gráficos hechos, representa mejor la información en este caso, el diagrama de caja ya que los datos están muy dispersos y no se puede observar de forma óptima las características de la muestra con el diagrama de tallo y hojas.

e) ¿Sería adecuada una tabla de frecuencia con intervalos? ¿Por qué?

Si, sería adecuada una tabla de frecuencia con intervalos porque los datos están muy dispersos.

**Actividad 14:** El siguiente diagrama de caja corresponde a una muestra aleatoria de 20 niños con desórdenes de conducta anotando el tiempo necesario en horas que requirió para lograr un plan integral de tratamiento con cada uno de ellos:



a) Realiza un breve informe a partir de lo que observas en el diagrama.

Se tomó una muestra aleatoria de 20 niños con desórdenes de conducta y la variable en estudio es el tiempo necesario (en horas) que requirió para lograr un plan integral de tratamiento con cada uno de ellos.

Observemos que hay una simetría en el 50% central de los datos (es decir, en la caja) pero el 25% inferior y el 25% superior de los datos no hay simetría. Cuando ocurren estas cosas se tiene que particionar el diagrama y analizar cada parte.

El 25% inferior y el 25% superior de los datos no hay simetría, hay una asimetría a la derecha.

El valor mínimo es 6, es decir, lo mínimo que se requirió para lograr un plan integral de tratamiento para cada niño con desórdenes de conducta es de 6 horas.

El valor máximo es 11, esto es, lo máximo que se requirió para lograr un plan integral de tratamiento para cada niño con desórdenes de conducta es de 11 horas.

No hay valores atípicos.

El 50% central de los datos se ubica entre 8 horas y 10 horas.

Se puede decir que en el 25% de los niños con desórdenes de conducta se requirió hasta 8 horas para lograr un plan integral de tratamiento para cada uno de ellos. ( $Q_1$ )

En el 50% de los niños con desórdenes de conducta se requirió hasta 9 horas para lograr un plan integral de tratamiento para cada uno de ellos. ( $M_e$ )

En el 75% de los niños con desórdenes de conducta se requirió hasta 10 horas para lograr un plan integral de tratamiento para cada uno de ellos. ( $Q_3$ )

**b)** ¿Hay valores atípicos en esta distribución?

Como podemos observar en la gráfica, no hay valores atípicos en esta distribución.

**c)** Interpreta los valores de los cuartiles y la mediana en el contexto del problema.

Esto ya se interpretó en el ítem a).

**d)** ¿Qué se puede decir respecto de la simetría de la distribución?

Lo que se puede decir respecto de la simetría de la distribución es que hay una simetría en la caja y una asimetría a la derecha en los bigotes.