

Estadística aplicada al Turismo

Ing. Henry Villa Yáñez, MsC.

Gráficos y su interpretación

- ◈ En este apartado se explicarán los tipos de diagramas más empleados cuando se trata de variables cuantitativas.

Histogramas

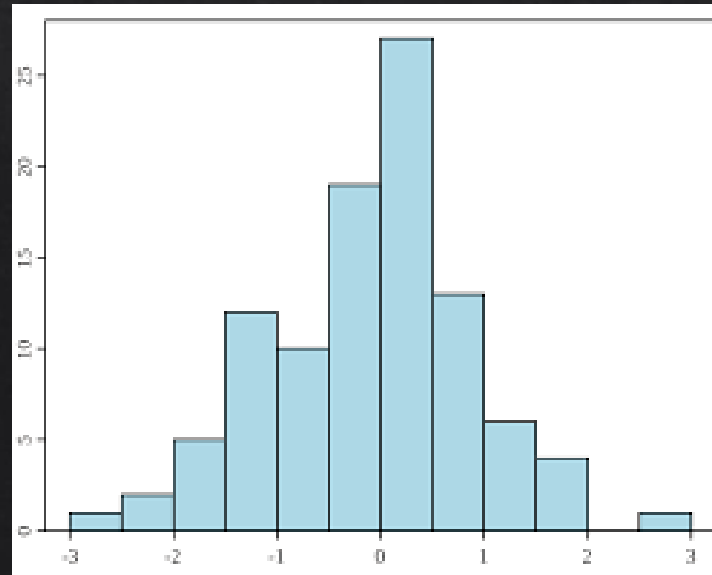
Ojivas

Polígonos de
Frecuencia

Diagramas
de Cajas

Histogramas

- ◊ Los histogramas son un tipo de diagramas de barras para variables cuantitativas.
- ◊ El término histograma se emplea cuando se trata de datos agrupados en intervalos.
- ◊ En el caso de datos no agrupados se llaman diagramas de barras.
- ◊ Tanto un diagrama de barras como un histograma de basa en una representación de las frecuencias, ya sean absolutas o relativas, frente a los datos.
- ◊ En el caso de los histogramas, si se considera igual amplitud de intervalos, la anchura de los rectángulos ha de ser la misma y las alturas coincidirán con las frecuencias de cada clase.

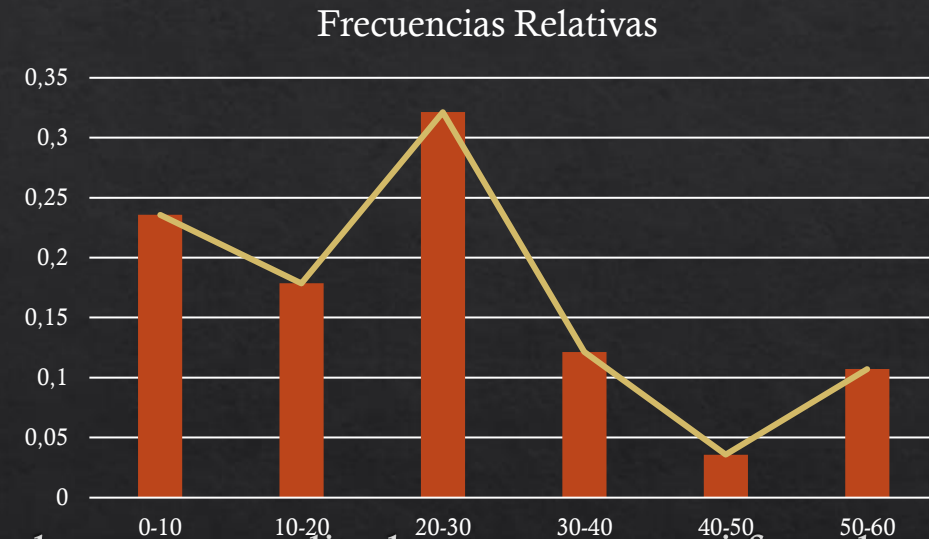


Histogramas - Ejemplo

- Los datos de la siguiente tabla muestran los aciertos que han tenido un grupo de personas en un test. Construir un histograma.

I_i	n_i	f_i
0-10	33	0,236
10-20	25	0,179
20-30	45	0,321
30-40	17	0,121
40-50	5	0,036
50-60	15	0,107
Total	140	1

- Cuando la amplitud de los intervalos no es constante, la menor amplitud se toma como si fuera la unidad y se van levantando las alturas.



$$\text{Frecuencia (área)} = \text{amplitud del intervalo} \cdot \text{altura}$$

Histogramas - Ejemplo

- ◊ Construcción de un histograma de frecuencias relativas con datos agrupados en intervalos de diferente amplitud.
- ◊ La tabla a continuación muestra los datos de resultados en un test psicológico realizado a 90 personas. Construir el histograma de frecuencias relativas:

I_i	x_i	n_i	f_i
[0-10)	5	6	1/15
[10-15)	12,5	20	2/9
[15-20)	17,5	25	5/18
[20-30)	25	7	7/90
[30-35)	32,5	23	23/90
[35-50)	42,5	9	1/10
Total		90	1,000

Histogramas - Ejemplo

- Se toma amplitud de valor 5 como la unidad. Las alturas son:

Frecuencia (área) = amplitud del intervalo · altura

$$\text{Intervalo } [10 - 15]: \quad \frac{2}{9} = 1 \cdot a_2 \rightarrow a_2 = \frac{2}{9}$$

$$\text{Intervalo } [15 - 20]: \quad \frac{5}{18} = 1 \cdot a_3 \rightarrow a_3 = \frac{5}{18}$$

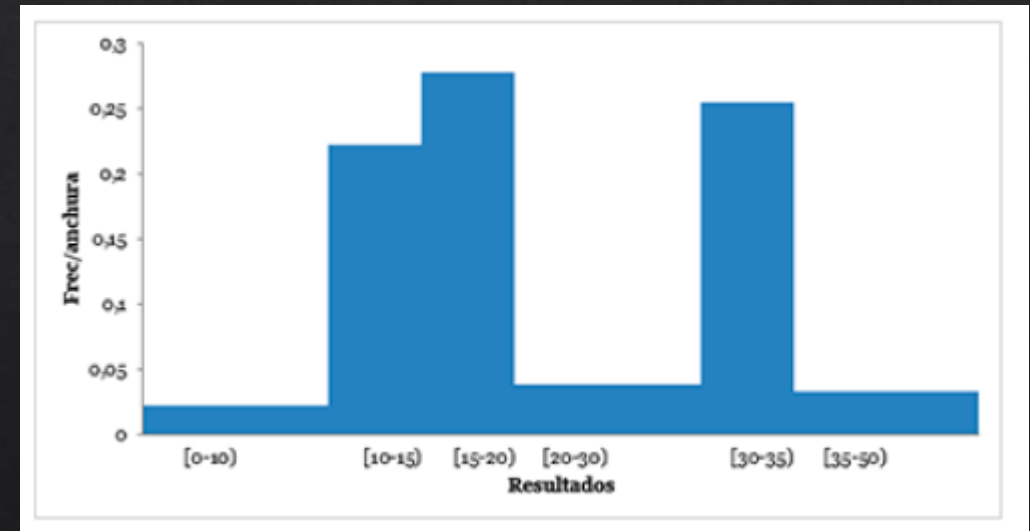
$$\text{Intervalo } [30 - 35]: \quad \frac{23}{90} = 1 \cdot a_5 \rightarrow a_5 = \frac{23}{90}$$

- Para los siguientes intervalos hay que calcular primero cuántas veces es mayor la amplitud del intervalo que el que se ha tomado como unidad y luego aplicar la fórmula. Los resultados son:

$$\text{Intervalo } [0 - 10]: \quad \frac{2}{45} = 2 \cdot a_1 \rightarrow a_1 = \frac{1}{45}$$

$$\text{Intervalo } [20 - 30]: \quad \frac{7}{90} = 2 \cdot a_4 \rightarrow a_4 = \frac{7}{180}$$

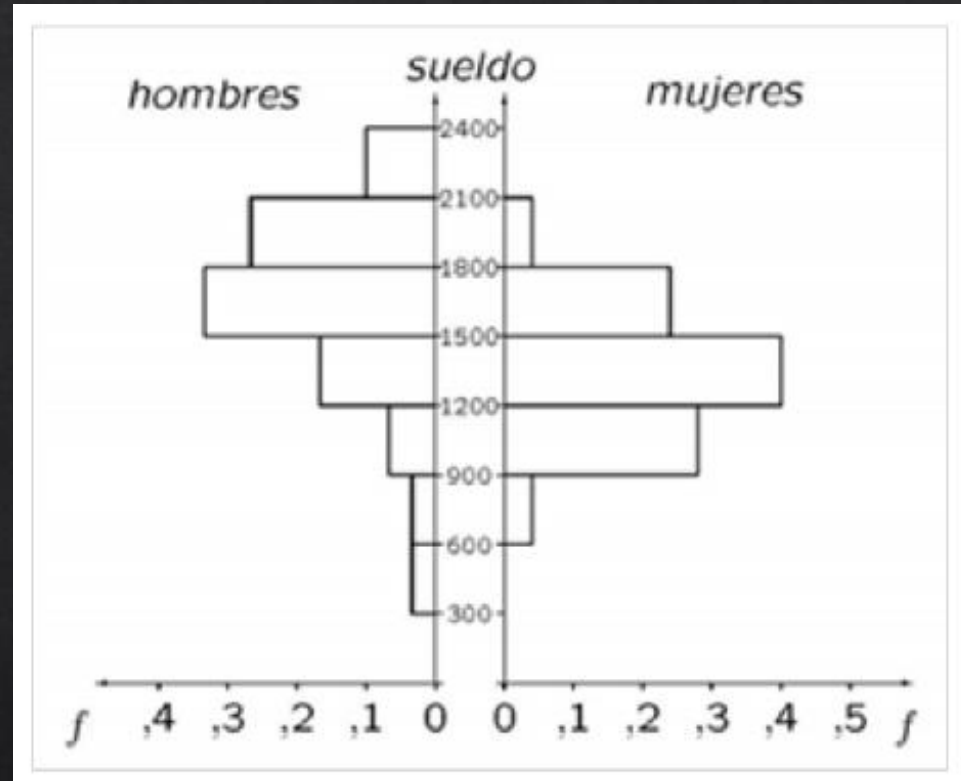
$$\text{Intervalo } [35 - 50]: \quad \frac{1}{10} = 3 \cdot a_6 \rightarrow a_6 = \frac{1}{30}$$



- Así mismo, se pueden hacer histogramas con las frecuencias acumuladas, ya sean absolutas o relativas.

Histogramas - Interpretación

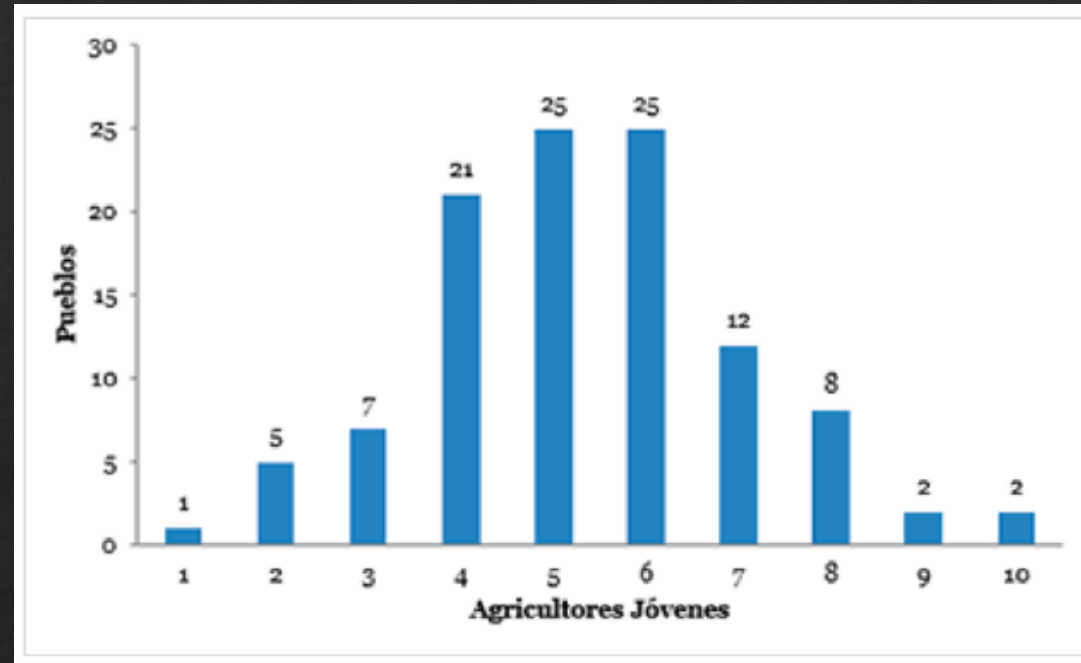
- ¿Qué se puede decir sobre el salario de los hombres en relación al de las mujeres en función de los siguientes histogramas?



- El sueldo medio de los hombres es más alto y se distribuye más asimétricamente

Histogramas - Interpretación

- ◊ El siguiente diagrama de barras muestra el número de agricultores jóvenes en unos pueblos de una comunidad autónoma. Responder a las siguientes preguntas:



- ◊ ¿Cuántos pueblos se han tomado para el estudio? 108 pueblos
- ◊ ¿Cuántos pueblos tienen como máximo 7 agricultores jóvenes? 96 pueblos
- ◊ ¿Cuál es el porcentaje de pueblos con 3 agricultores jóvenes? 6,5% (7/108)
- ◊ ¿Cuál es el número más frecuente de agricultores? 5 y 6 agricultores en 25 empresas

Histogramas - Interpretación

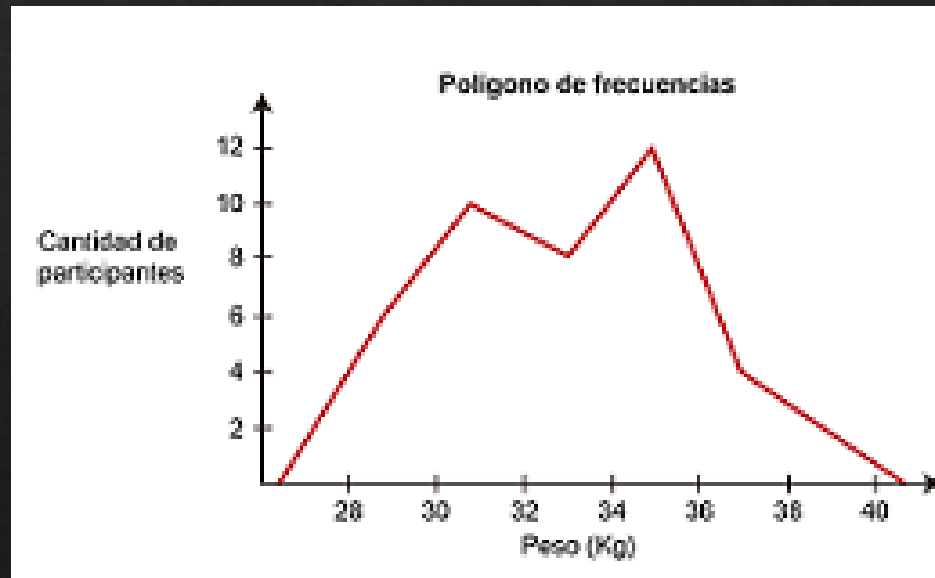
- ♦ *¿Cuál es el número máximo de agricultores jóvenes en los 25 pueblos más pequeños, es decir, con menos trabajadores? Se realiza la tabla de frecuencias acumuladas*

Agricultores jóvenes	Pueblos	Frecuencia acumulada
1	1	1
2	5	6
3	7	13
4	21	34
5	25	59
6	25	84
7	12	96
8	8	104
9	2	106
10	2	108

- ♦ En los 25 pueblos con menos agricultores el máximo número de los mismo es 4
- ♦ *Si una empresa quiere suministrar fertilizantes y sólo quiere visitar a aquellos pueblos con más de 7 agricultores jóvenes, ¿a qué porcentaje de pueblos iría? A los que tienen más de 7 agricultores ($8+2+2$), $12/108 = 11,1\%$*
- ♦ *Si el 50% de pueblos con mayor número de agricultores puede beneficiarse de una subvención, ¿qué mínimo de agricultores ha de haber en los pueblos conseguir esa subvención? El 50% de 108 es 54. Comparando con las frecuencias acumuladas, el máximo número de agricultores es 5.*

Polígono de frecuencia

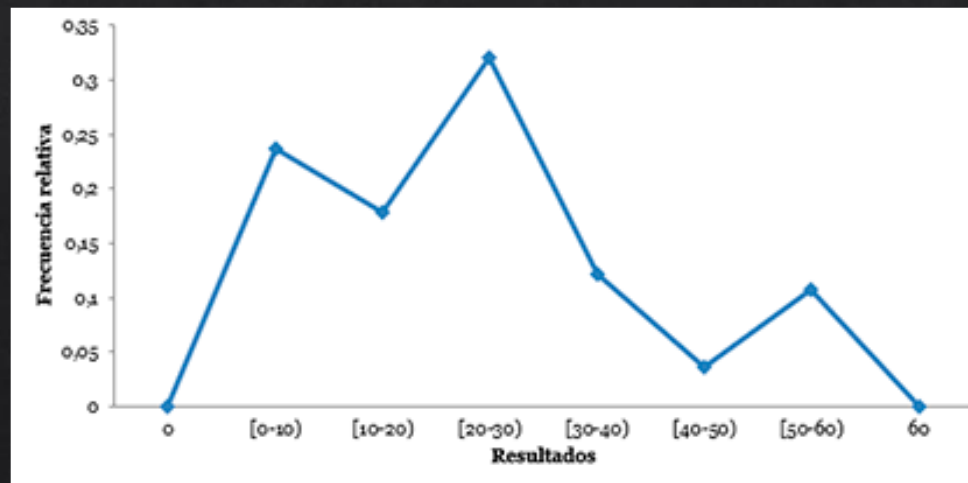
- ❖ Los polígonos de frecuencias son gráficos que proporcionan la misma información que los histogramas, pero, en vez de ser diagramas de barras, son diagramas de puntos conectados por una línea, de manera que el área que dejan debajo coincide con el área de las barras del histograma y tiene forma poligonal, o curva cuando hay muchos datos.
- ❖ En estos gráficos, se dan dos puntos en los extremos en los que las frecuencias sean cero para cerrar el polígono con el eje horizontal.



Polígono de frecuencia - Ejemplo

- ◇ Construir un polígono de frecuencias con los datos

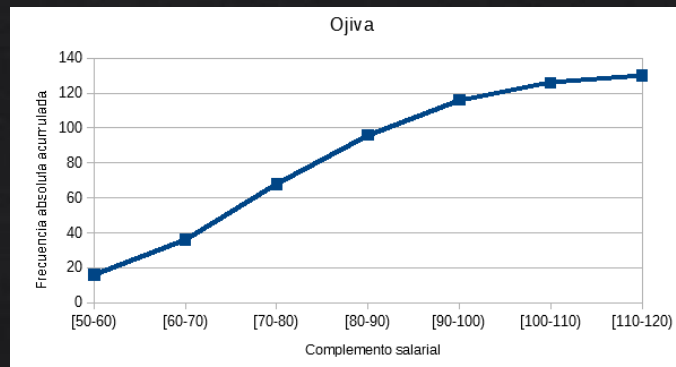
I_i	n_i	f_i
0-10	33	0,236
10-20	25	0,179
20-30	45	0,321
30-40	17	0,121
40-50	5	0,036
50-60	15	0,107
Total	140	1



- ◇ A partir del polígono de frecuencias se puede elaborar el correspondiente histograma haciendo líneas paralelas al eje horizontal que pasen por los puntos del polígono, de manera que cada punto caiga en la mitad de la barra.

Ojivas

- ◈ Las ojivas son tipos de diagramas para datos agrupados que dan información de las frecuencias acumuladas.
- ◈ Se muestran los datos que están por debajo o por encima de unos determinados valores, en vez de los datos de los intervalos o las marcas de clase.
- ◈ Hay dos tipos de ojivas: ojivas ‘menor que’ y ojivas ‘mayor que’.
- ◈ Las ojivas ‘menor que’ dan información sobre frecuencias menores en relación de la que se está comparando y se inclinan positivamente;
- ◈ Las ojivas ‘mayor que’, al revés. En la ojiva ‘menor que’ el extremo derecho no toca al eje horizontal, mientras que en la ojiva ‘mayor que’ es el extremo izquierdo. En el eje horizontal se ponen los límites de los intervalos.
- ◈ La frecuencia absoluta o relativa acumulada ‘menor que’ para la ojiva ‘menor que’, se calcula a partir de la primera frecuencia acumulada del primer intervalo y se acumula hacia abajo, mientras que la frecuencia absoluta o relativa ‘mayor que’ para la ojiva ‘mayor que’ se calcula a partir de la última frecuencia acumulada del último intervalo y se acumula hacia arriba.



Ojivas - Ejemplo

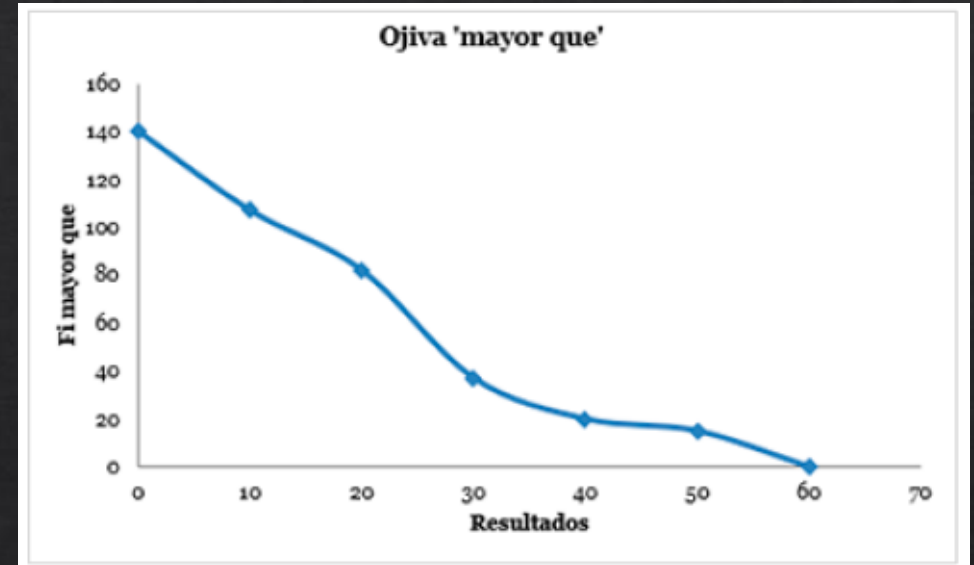
◊ A partir de los datos del ejemplo anterior hacer una ojiva 'menor que' y una ojiva 'mayor que'.

I_i	f_i	h_i
0-10	33	0,236
10-20	25	0,179
20-30	45	0,321
30-40	17	0,121
40-50	5	0,036
50-60	15	0,107
Total	140	1

Ojiva 'menor que'		Ojiva 'mayor que'	
Clase	N_i (menor que)	Clase	N_i (mayor que)
Menor que 0	0	Mayor que 0	140
Menor que 10	33	Mayor que 10	107
Menor que 20	58	Mayor que 20	82
Menor que 30	103	Mayor que 30	37
Menor que 40	120	Mayor que 40	20
Menor que 50	125	Mayor que 50	15
Menor que 60	140	Mayor que 60	0

Ojivas - Ejemplo

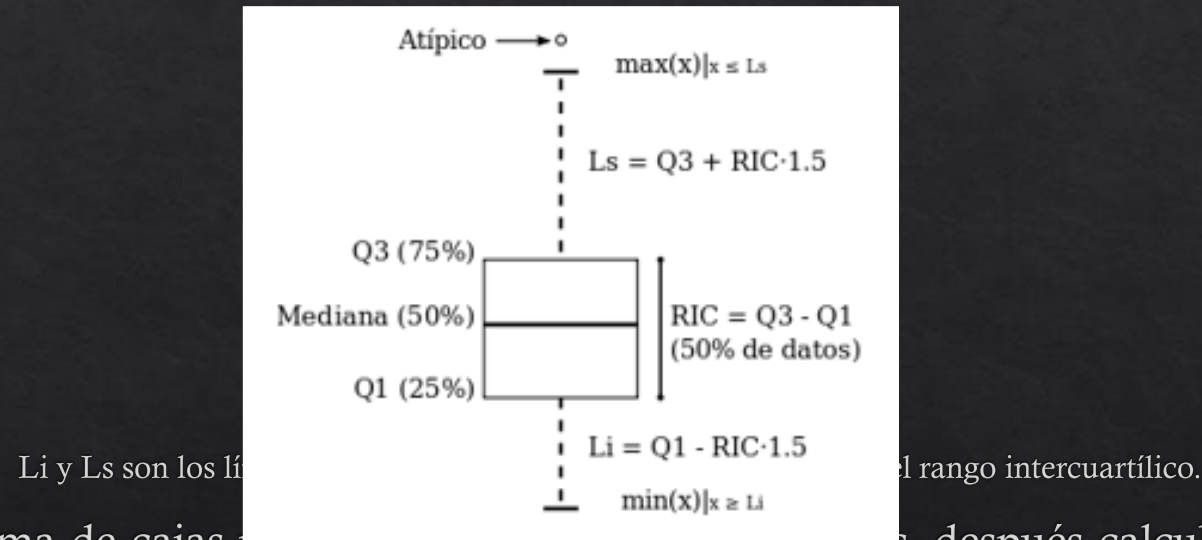
- ◇ A partir de los datos del ejemplo anterior hacer una ojiva 'menor que' y una ojiva 'mayor que'.



- ◇ Si aquellos participantes que hayan tenido menos de 35 aciertos en un test se quedan sin poder asistir a la siguiente fase, ¿qué porcentaje de participantes se quedará sin asistir? Interpolando en la ojiva 'menor que' se obtiene que son 105 participantes.
- ◇ Así mismo, se puede hacer ojivas con las frecuencias relativas acumuladas.

Diagramas de caja

- ◈ Un diagrama de caja es un tipo de representación basada en los cuartiles de una distribución.
- ◈ Su representación es una caja, donde quedan los valores de los cuartiles y unos brazos llamados bigotes que van hasta los límites de la distribución.
- ◈ Son ampliamente empleados para detectar la presencia de algún dato anómalo dentro de un conjunto de datos. En la siguiente figura, se puede ver una representación y la información que proporciona un diagrama de cajas.



- ◈ Para hacer un diagrama de cajas primero hay que ordenar los datos, después calcular los cuartiles y el rango intercuartílico, seguidamente calcular los límites superior e inferior, (tal y como se muestra en la figura) y, por último, dibujar el diagrama para lo cual hay que calcular los anchos de los rectángulos de la caja. Si algún valor cae fuera de los límites superior o inferior permitidos, se toma como dato atípico.

Diagramas de caja - Ejemplo

- ♦ A partir de los datos presentados a continuación elaborar un diagrama de cajas e interpretar los datos.

34	48	49	51	52	54	55	71	76
----	----	----	----	----	----	----	----	----

- ♦ Calculamos la mediana y los cuartiles 1 y 3.

$$\text{Posición de la Mediana} = \frac{n+1}{2} = \frac{9+1}{2} = 5$$

$$\text{Posición de } Q_1 = \frac{n+1}{4} = \frac{9+1}{4} = 2,5$$

$$\text{Posición de } Q_3 = \frac{3(n+1)}{4} = \frac{3(9+1)}{4} = 7,5$$

La mediana es el dato 52.

Q_1 es 48,5.

Q_3 es 63.

- ♦ Calculamos el rango intercuartílico

$$\text{Rango intercuartílico} = Q_3 - Q_1 = 63 - 48,5 = 14,5$$

- ♦ Calculamos los límites inferior y superior

$$L_i = Q_1 - 1,5 \cdot RIC = 48,5 - 1,5 \cdot 14,5 = 26,75$$

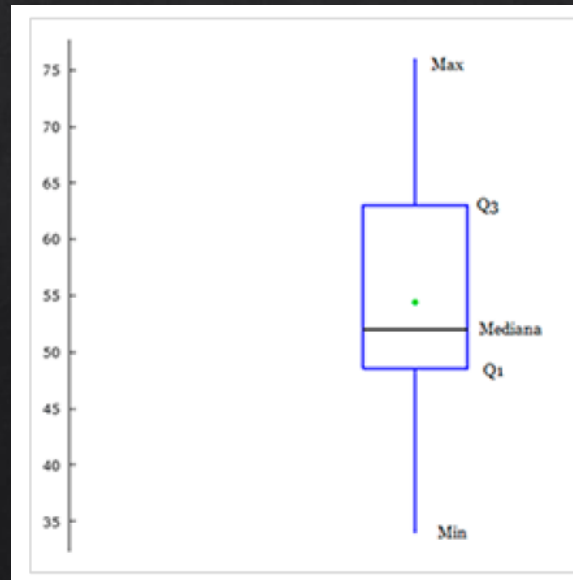
$$L_s = Q_3 + 1,5 \cdot RIC = 63 + 1,5 \cdot 14,5 = 84,75$$

Diagramas de caja - Ejemplo

- ◊ A continuación, se calcula la anchura de cada rectángulo del diagrama:

	Valores	Anchos
Mínimo	34	34
Q_1	48,5	14,5
Mediana	52	3,5
Q_3	63	11
Máximo	76	13

- ◊ Se dibuja el diagrama



- ◊ En el diagrama de cajas se puede ver que no hay ningún valor que exceda los valores de los límites superior ni inferior, por lo que no hay ningún valor atípico.

GRACIAS