

Bioestadística

Sesión 3-4: Estadística descriptiva

José Aurelio Pina Romero

Bioestadística - Grado Enfermería

UA- Departamento de Enfermería

-ÍNDICE

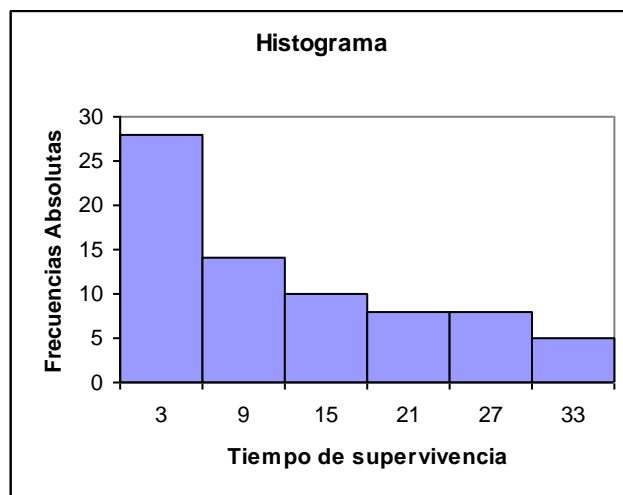
- Estadística descriptiva para datos agrupados.

Ejemplo

En un estudio sobre supervivencia tras un tratamiento con quimioterapia para cierto tipo de cáncer ha sido registrado el tiempo transcurrido desde el inicio del tratamiento hasta el fallecimiento de los individuos. Los tiempos registrados se resumen en la tabla adjunta, agrupados por intervalos de 6 meses de amplitud:

Tiempo en meses	Nº de sujetos
0 - 6	28
6 - 12	14
12 - 18	10
18 - 24	8
24 - 30	8
30 - 36	5
Total	73

Tiempo en meses	Marca de clase (X_{m_i})	Frecuencias Absolutas (f_i)	Frecuencias Absolutas Acumuladas (F_i)	Frecuencias relativas (fr_i)	Porcentaje (p_i)	Porcentaje Acumulado (P_i)
0 - 6	3	28	28	$28/73 = 0,38$	38%	38%
6 a 12	9	14	42	$14/73 = 0,19$	19%	57%
12 a 18	$15=(18+12)/2$	10	52	$10/73 = 0,14$	14%	$71=38+19+14$
18 - 24	21	8	$60=28+14+10+8$	$8/73 = 0,10$	10%	81%
24 - 30	27	8	68	$8/73 = 0,10$	10%	91%
30 - 36	33	5	73	$5/73 = 0,07$	7%	$98\% \approx 100\%$
Total		73		1,00	100%	



Media

Tiempo	Marca de clase (x_m)	f_i	F_i
0 - 6	3	28	28
6 a 12	9	14	42
12 a 18	15	10	52
18 - 24	21	8	60
24 - 30	27	8	68
30 - 36	33	5	73
Total		73	

$$\bar{x} = \frac{\sum_{i=1}^6 x m_i f_i}{n} = \frac{(3 \times 28) + (9 \times 14) + \dots + (33 \times 5)}{73} = 12,45$$

Moda = 3 meses

Mediana:

1.- Los datos por encontrarse en una tabla están ordenados

2.- Calculamos la posición:

$$r_q = \frac{n}{2} = \frac{73}{2} = 36,5$$

3.- Calculamos el percentil 50

$$Md = X_i + \left(\frac{\frac{n}{2} - F_{i-1}}{F_i - F_{i-1}} \right) \cdot a_i = 6 \left(\frac{36,5 - 28}{42 - 28} \right) \cdot 6 = 9,64 \text{ meses}$$

$$\text{Varianza} = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{n} = \frac{(3 - 12,45)^2 \times 28 + \dots + (33 - 12,45)^2 \times 5}{73} = 97,56 \text{ meses}^2$$

Desviación típica = $s = 9,88$ meses y el Coeficiente de variación = $Cv = 0,27$

Por tanto, el tiempo de supervivencia está entorno a 12,45 meses con una variabilidad entorno a este valor de 9,88 meses.

Percentil 95

Tiempo	X_{m_i}	f_i	F_i
0 - 6	3	28	28
6 a 12	9	14	42
12 a 18	15	10	52
18 - 24	21	8	60
24 - 30	27	8	68
30 - 36	33	5	73
Total		73	

Esta variable recoge el tiempo entre el inicio del tratamiento con quimioterapia y la defunción del paciente. Por tanto, buscamos aquel valor de la variable (tiempo en meses) tras el cual quedan vivos solo el 5% de la población, o lo que es lo mismo, por debajo de este valor quedarán las defunciones del 95% de la población. Calculamos por tanto el percentil del 95%:

1.- Los datos por encontrarse en una tabla están ordenados

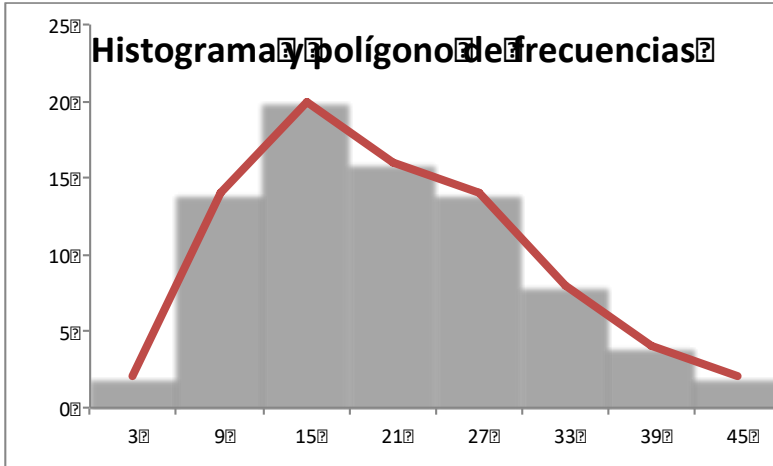
2.-Calculamos la posición: $r_q = \frac{q \cdot n}{100} = \frac{95 \cdot 73}{100} = 69,35$

luego fijándonos en las frecuencias acumuladas obtenemos el intervalo que referencia, en este caso es el último.

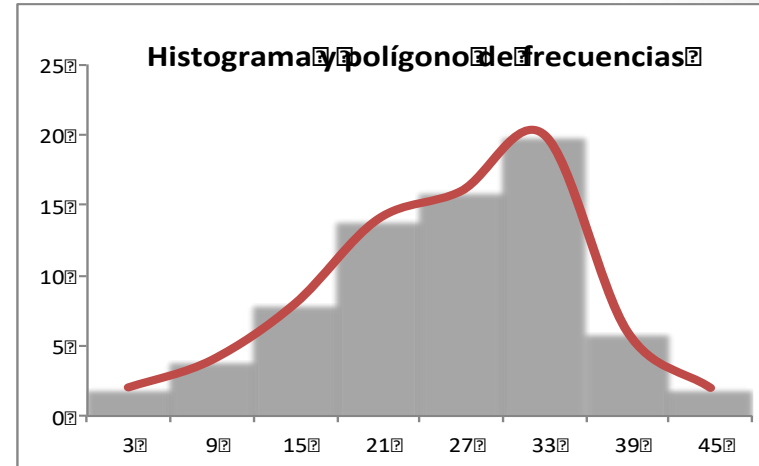
3.-Calculamos el percentil 95

$$P_{95\%} = X_i + \left(\frac{qn / 100 - F_{i-1}}{F_i - F_{i-1}} \right) a_i = 30 + \left(\frac{69,35 - 68}{73 - 68} \right) \times 6 = 31,62 \text{ meses}$$

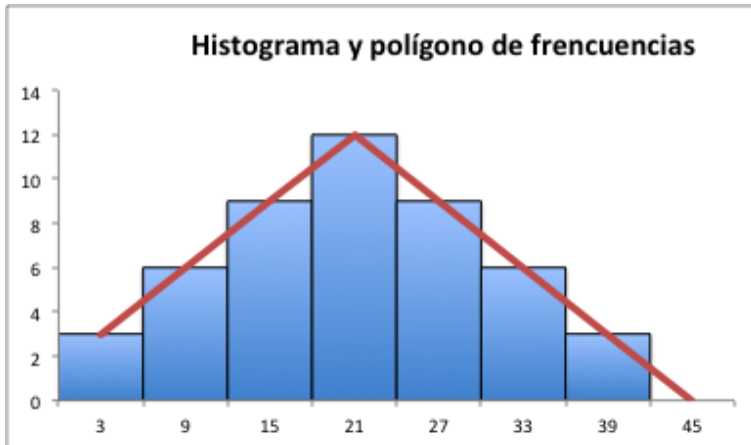
Gráficos variables cuantitativas → Polígonos de frecuencias



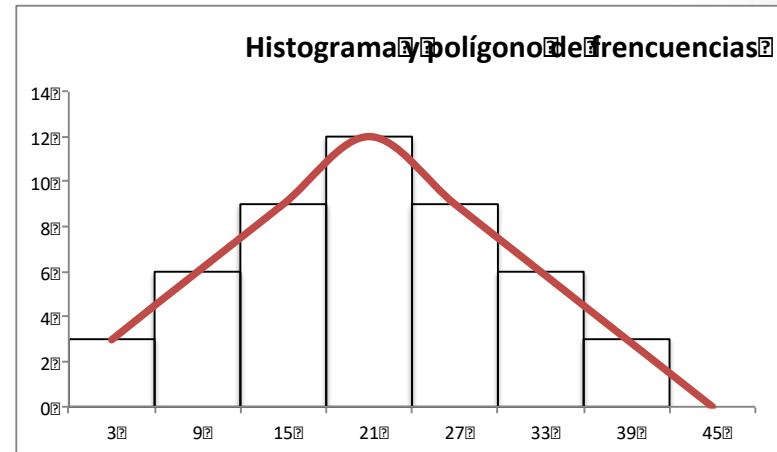
Distribución asimétrica positiva → **media > mediana**



Distribución asimétrica negativa → **media < mediana**



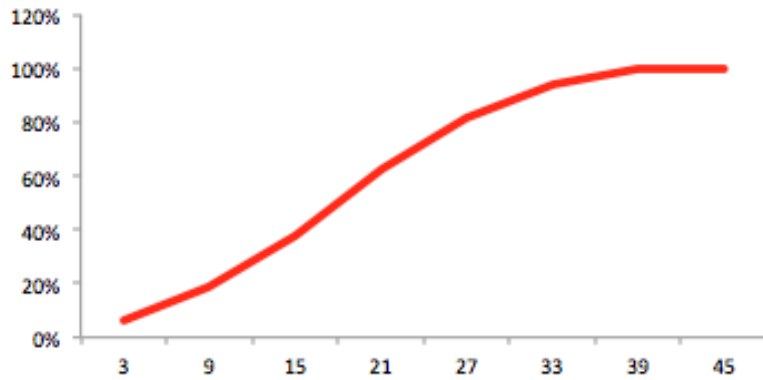
Distribución simétrica → **media = mediana**



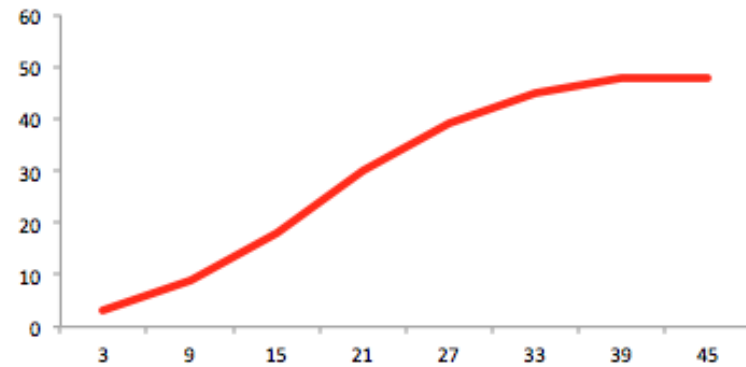
fichero:meses.xls

Gráficos variables cuantitativas → Polígonos de frecuencias acumulado

Polígono de frecuencias acumulado



Polígono de frecuencias acumulado



Medidas de forma: **datos sin agrupar** X_1, X_2, \dots, X_n

- Coeficiente de asimetría:

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

$As = 0 \rightarrow$ Simetría

$As > 0 \rightarrow$ Asimetría positiva

$As < 0 \rightarrow$ Asimetría negativa

- Coeficiente de curtosis

$$Cu = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{s^4}$$

$Cu = 3 \rightarrow$ Mesocúrtica

$Cu > 3 \rightarrow$ Leptocúrtica

$Cu < 3 \rightarrow$ Platicúrtica

Coeficiente de asimetría (según SPSS)

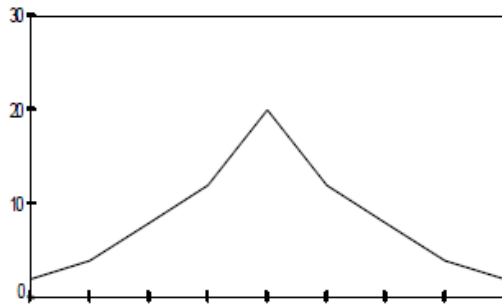
Medidas de forma: **datos sin agrupar** X_1, X_2, \dots, X_n

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

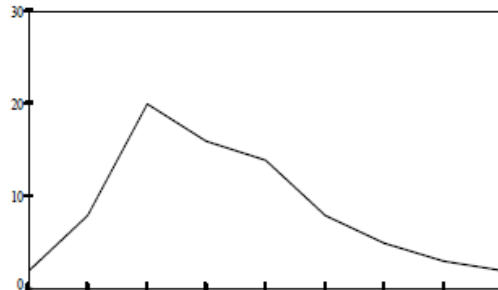
$As = 0 \rightarrow$ Simetría

$As > 0 \rightarrow$ Asimetría positiva

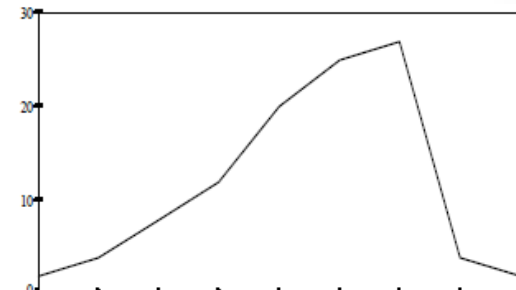
$As < 0 \rightarrow$ Asimetría negativa



Distribución simétrica



Distribución asimétrica positiva



Distribución asimétrica negativa

"ESTADÍSTICA BÁSICA EN CIENCIAS DE LA SALUD" Nolasco A, Moncho J.

Publicación electrónica en RUA: <http://hdl.handle.net/10045/60526>

Coeficiente de Curtosis

Medidas de forma: **datos sin agrupar** X_1, X_2, \dots, X_n

$$Cu = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n}}{s^4}$$

$Cu = 3 \rightarrow$ Mesocúrtica

$Cu > 3 \rightarrow$ Leptocúrtica

$Cu < 3 \rightarrow$ Platicúrtica

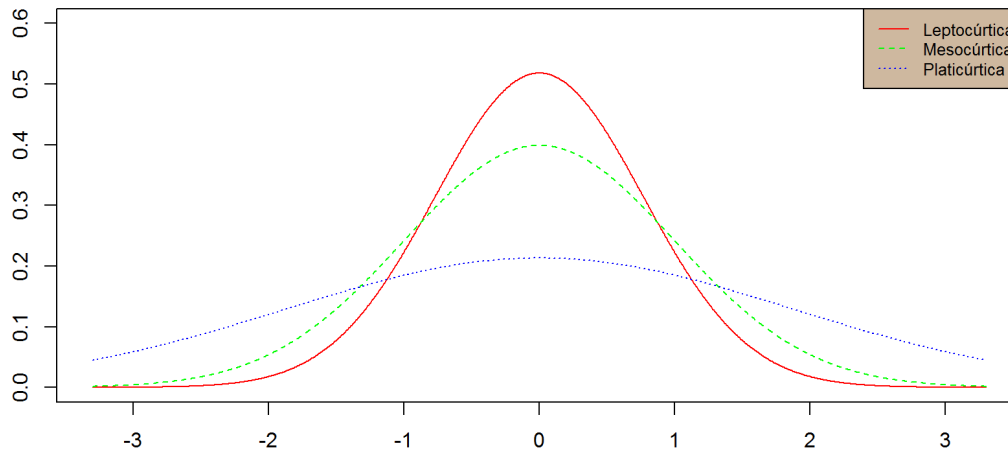
$Cu = 0 \rightarrow$ Mesocúrtica

$Cu > 0 \rightarrow$ Leptocúrtica

$Cu < 0 \rightarrow$ Platicúrtica



Distribuciones platicúrtica, mesocúrtica y leptocúrtica

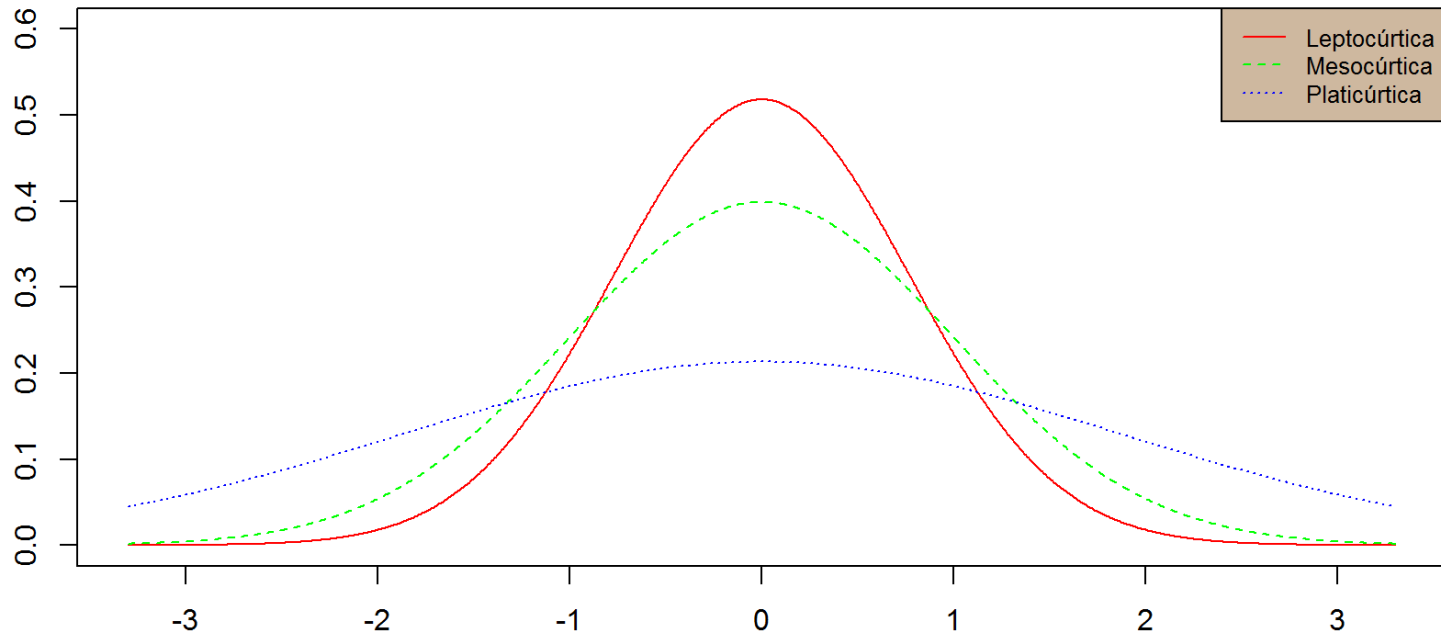


"ESTADÍSTICA BÁSICA EN CIENCIAS DE LA SALUD" Nolasco A, Moncho J.

Publicación electrónica en RUA: <http://hdl.handle.net/10045/60526>

Gráficos variables cuantitativas → Polígonos de frecuencias

Distribuciones platicúrtica, mesocúrtica y leptocúrtica



Medidas de forma: **datos agrupados por intervalos**

Variable(X_i)	xm_i	f_i	F_i
$[x_0 - X_0)$	Xm_1	f_1	F_1
$[x_1 - X_2)$	Xm_2	f_2	F_2
...			F_{i-1}
$[x_i - X_{i+1})$	xm_i	f_i	F_i
n			

- Coeficiente de asimetría:

$$As = \frac{\sum_{i=1}^n (xm_i - \bar{x})^3 \cdot f_i}{s^3}$$

$As = 0 \rightarrow$ Simetría
 $As > 0 \rightarrow$ Asimetría positiva
 $A < 0 \rightarrow$ Asimetría negativa

- Coeficiente de curtosis

$$Cu = \frac{\sum_{i=1}^n (xm_i - \bar{x})^4 \cdot f_i}{s^4}$$

$Cu = 3 \rightarrow$ Mesocúrtica
 $Cu > 3 \rightarrow$ Leptocúrtica
 $Cu < 3 \rightarrow$ Platicúrtica

