

9

Estadística

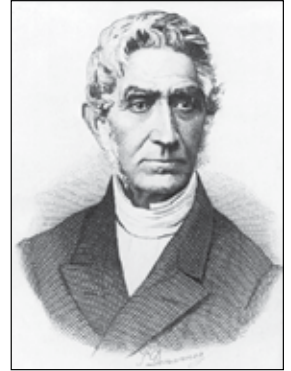
Origen y evolución de la estadística como ciencia

En el desarrollo histórico de la estadística se pueden distinguir tres grandes etapas.

Censos. Desde la Antigüedad y hasta el siglo XVI, solo se realizan recogidas de datos y, a lo sumo, una exposición ordenada y clara de estos.

Análisis de datos. Abarca los siglos XVII, XVIII y XIX. Se supera lo meramente descriptivo y los datos pasan a ser analizados científicamente con el fin de extraer conclusiones.

Se suele considerar que esta etapa comienza con los trabajos de **John Graunt** (s. XVII), quien utilizó archivos parroquiales para realizar un profundo estudio de los nacimientos y las defunciones en Londres durante 30 años: anotó el sexo de cada nacido, las enfermedades de los fallecidos y otras muchas variables. Con ello pudo extraer conclusiones válidas para el futuro e inauguró, así, la estadística demográfica.



*Adolphe Quetelet (1796-1874).
Matemático belga considerado uno
de los padres de la estadística moderna.*



*Estatua en honor de Florence Nightingale (1820-1910),
en Londres. Estadística y enfermera británica, pionera
en el uso de gráficos, que contribuyó decisivamente con
sus estudios estadísticos a la mejora de la sanidad.*

Hicieron aportaciones destacadas a esta ciencia **Neumann** (s. XVII), cuyos métodos sirvieron de base para elaborar las tablas de mortalidad utilizadas por las compañías de seguros, y **Quetelet** (s. XIX), que fue el primero que se valió de la probabilidad para aplicar la estadística a las Ciencias Sociales.



*Ronald Fisher (1890-1962), biólogo y matemático, aplicó métodos
estadísticos al diseño de experimentos científicos.*

Estadística inferencial. Se inicia a finales del siglo XIX. La esencia de esta rama de la estadística es que a partir de una muestra se extraen conclusiones válidas para toda una población. Para ello, se echa mano de la alta matemática. Son figuras destacadas en este campo **Ronald Fisher** y **Karl Pearson**.

Ejemplo

En una determinada diputación se quieren estudiar algunas características de los 2537 pueblos que la componen. Para ello, toman los datos de 300 de ellos.

Población: los 2537 pueblos.

Individuo: cada uno de los pueblos.

Muestra: los 300 pueblos que se estudian.

Ejemplo de estadística inferencial

Una editorial realiza una encuesta a 387 estudiantes de una universidad sobre sus preferencias de lectura, con el fin de extraer consecuencias válidas para todos. Esto es estadística inferencial, pues a partir de una muestra, se desea obtener información sobre algún aspecto de la población. Es decir:

Se estudia el comportamiento de una variable en una MUESTRA.



Se INFIERE el comportamiento de esa variable en la POBLACIÓN.

La estadística tiene por objeto el desarrollo de técnicas para el conocimiento numérico de un conjunto de datos empíricos (recogidos mediante experimentos o encuestas).

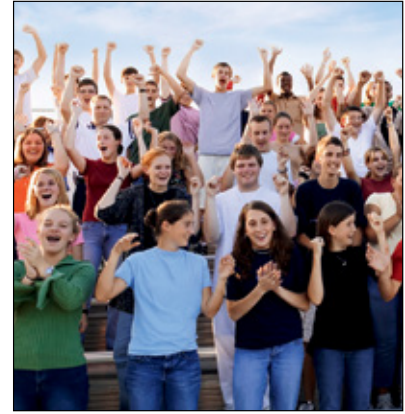
Recordemos las nociones básicas, algunas de ellas adquiridas hace años, necesarias para entender la estadística.

- **Población** es el conjunto de todos los elementos cuyo conocimiento nos interesa y que serán objeto de nuestro estudio.

Por ejemplo, una población puede ser los 1580 estudiantes matriculados en un centro de estudios.

- **Individuo** es cada uno de los elementos que forman la población.

En el caso del centro de estudios, cada estudiante es un individuo.



- **Muestra** es un subconjunto extraído de la población, cuyo estudio sirve para inferir características de toda la población.

Siguiendo con el ejemplo anterior, una muestra puede ser un grupo significativo de 50 estudiantes.

- **Caracteres** son los aspectos que deseamos estudiar en los individuos de una población. Cada carácter puede tomar distintos valores.

De los estudiantes podríamos analizar algunos caracteres como la edad, la altura, el peso, el color de pelo, cómo llegan al centro (medio de transporte), número de personas que viven en su casa, preferencias de lectura...

- **Variable estadística** es aquella que recorre todos los valores de un cierto carácter. Esta puede ser:

— **Cuantitativa**, si toma valores numéricos.

– **Discreta:** solo toma valores aislados.

– **Continua:** puede tomar cualquier valor de un intervalo.

— **Cualitativa**, si toma valores no numéricos.

El número de personas que viven en casa toma valores cuantitativos discretos. La edad, la altura y el peso toman valores cuantitativos continuos. El color de pelo y el medio de transporte para llegar al centro toman valores cualitativos.

La **estadística descriptiva** expone y analiza algunos caracteres de los individuos de un grupo dado (población) sin extraer conclusiones para un grupo mayor.

La **estadística inferencial**, por el contrario, trabaja con muestras y pretende, a partir de ellas, “inferir” características de toda la población.

Tablas con datos aislados

Tras la recogida de datos, la elaboración de una tabla de frecuencias es el siguiente paso. Cuando la variable toma pocos valores, la elaboración de la tabla es sumamente sencilla. No hay más que hacer el recuento de los resultados.

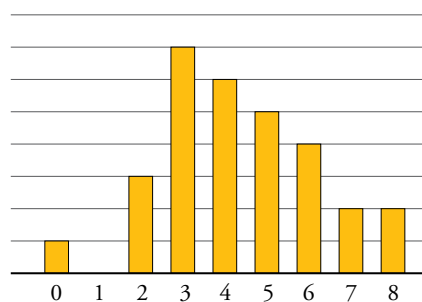
Ejemplo

A los estudiantes de una clase se les ha preguntado por el número de primos que tienen. Estos son los resultados:

5 3 2 6 4 3 6 5 8 2 4 6 3 2 0
4 3 7 5 4 3 3 6 7 5 4 8 5 4 3

Elaboramos una tabla en la que se muestra ordenadamente cada valor de la variable (número de primos) con su correspondiente frecuencia. La representamos mediante un diagrama de barras.

x_i	f_i
0	1
1	0
2	3
3	7
4	6
5	5
6	4
7	2
8	2



En la web



Recuerda: diagramas de barras e histogramas.



Tablas con datos agrupados

Cuando en una distribución estadística la variable es continua, o bien cuando, siendo discreta, el número de valores que toma la variable es muy grande, conviene elaborar una tabla de frecuencias agrupándolos en intervalos.

Los siguientes son algunos ejemplos en los que es preferible agrupar los datos en intervalos:

- Alturas, pesos, grosores, contornos, áreas, volúmenes, tiempos y, en general, las medidas físicas. Estas son variables continuas, por lo que toman cualquier valor dentro de un intervalo.
- El número de coches que pasan cierto día por cada calle de un barrio; el número de libros consultados en una biblioteca por los usuarios; el número de páginas que tienen los libros de una biblioteca... son medidas discretas que conviene agruparlas como continuas por la cantidad tan elevada de valores distintos que toman.

En la web

Refuerza la elaboración de tablas de frecuencias.

Observa

Al elaborar una tabla con datos agrupados se pierde algo de información, pues en ella se ignora cada valor concreto, que se difumina dentro de un intervalo. A cambio, se gana en claridad y eficacia.

x_i	f_i
x_1	f_1
x_2	f_2
...	...
...	...
...	...
x_n	f_n

Elaboración de una tabla con datos agrupados en intervalos

Para elaborar una tabla con datos agrupados, conviene dar los siguientes pasos:

1. Se localizan los valores extremos, a y b , y se halla su diferencia, $r = b - a$ (recorrido).
2. Se decide el número de intervalos que se quiere formar, teniendo en cuenta la cantidad de datos que se poseen. El número de intervalos no debe ser inferior a 6 ni superior a 15.
3. Se toma un intervalo, r' , de longitud algo mayor que el recorrido, r , y que sea múltiplo del número de intervalos, con objeto de que estos tengan una longitud entera.
4. Se forman los intervalos, de modo que el extremo inferior del primero sea algo menor que a y el extremo superior del último sea algo mayor que b . Es deseable que los extremos de los intervalos no coincidan con ninguno de los datos. Para ello, conviene que los extremos de los intervalos tengan una cifra decimal más que los datos.

El punto medio de cada intervalo se llama **marca de clase**. Es el valor que representa a todo el intervalo para el cálculo de algunos parámetros.

La tabla de frecuencias del margen puede corresponder a:

- Una distribución de datos aislados que toma los valores x_1, x_2, \dots, x_n .
- Una distribución de datos agrupados en intervalos, siendo x_1, x_2, \dots, x_n sus marcas de clase.

En el primer caso, la tabla refleja exactamente la distribución real. En el segundo, la tabla es una buena aproximación a la realidad.

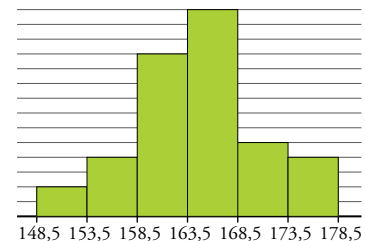
Ejercicio resuelto

Elaborar una tabla de frecuencias con las estaturas de 40 adolescentes y representarla en el gráfico correspondiente.

168	160	167	175	175
167	168	158	149	160
178	166	158	163	171
162	165	163	156	174
160	165	154	163	165
161	162	166	163	159
170	165	150	167	164
165	173	164	169	170

1. Valores extremos: $a = 149$, $b = 178$. Recorrido: $r = 178 - 149 = 29$.
2. Tomaremos solo 6 intervalos. Un múltiplo de 6 mayor que 29 y próximo a él es 30. Longitud de cada intervalo: 5.
3. Formamos los intervalos comenzando por un número algo menor que $a = 149$ y terminando en un número algo mayor que $b = 178$.
4. Repartimos los datos en los intervalos:

INTERVALOS	M. DE CLASE	FRECUENCIA
148,5-153,5	151	2
153,5-158,5	156	4
158,5-163,5	161	11
163,5-168,5	166	14
168,5-173,5	171	5
173,5-178,5	176	4



Piensa y practica

1. Reparte los cuarenta datos del ejercicio resuelto anterior en 10 intervalos con el mismo recorrido total.
2. Reparte los cuarenta datos del ejercicio resuelto anterior en 8 intervalos. Para ello, toma $r' = 32$.

Recordemos cómo se obtienen los parámetros a partir de una tabla, ya sea con datos aislados o con datos agrupados en intervalos. En este último caso, recurrimos a las marcas de clase para calcular los parámetros.

PUNTUACIONES EN UN TEST

x_j	f_j	$f_j x_j$
0	12	0
1	31	31
2	86	172
3	92	276
4	48	192
5	19	95
	288	766

Para qué sirven los parámetros?

Los parámetros estadísticos resumen la información que se da en una tabla.

¿Por qué conviene resumirla?

Si un profesor desea ver cómo van los estudiantes de una clase, en la tabla de resultados lo tiene muy claro. Pero si el director del centro quiere comparar unas clases con otras, en diversas asignaturas, incluso con los resultados de otros años... hay demasiada información. Es necesario resumirla. Para eso están los parámetros.

La media, \bar{x} , y la desviación típica, σ , se complementan. Ambas, conjuntamente, aportan una visión bastante buena de la distribución correspondiente.

En la web

- Ampliación: demostración de que las dos expresiones dadas para la varianza coinciden.
- HOJA DE CÁLCULO: aplicación para confeccionar tablas de frecuencias, representar el gráfico correspondiente y calcular \bar{x} , σ y C.V.



En la web

- Interpretación de la media.
- Interpretación de la desviación típica.

■ **MEDIA:** $\bar{x} = \frac{\sum f_j x_j}{\sum f_j}$ $\sum f_j x_j \rightarrow$ suma de todos los datos
 $\sum f_j = N \rightarrow$ n.º total de individuos

Por ejemplo, en la distribución que tenemos en el margen:

$$\sum f_j = 288. \text{ Hay 288 individuos (que han realizado el test).}$$

$$\sum f_j x_j = 766. \text{ Es la suma de las puntuaciones de todos los individuos.}$$

La media es $\bar{x} = 766/288 = 2,66$.

■ **VARIANZA:** $Var = \frac{\sum f_j (x_j - \bar{x})^2}{N}$ o bien $Var = \frac{\sum f_j x_j^2}{N} - \bar{x}^2$

Las dos expresiones coinciden (consulta la demostración en la web).

— En la primera de ellas, se ve claro el significado de la varianza: promedio de los cuadrados de las desviaciones a la media.

— La segunda es más cómoda para realizar los cálculos.

Utilizamos la segunda fórmula para obtener la varianza a partir de esta tabla:

x_j	f_j	$f_j x_j$	$f_j x_j^2$
0	12	0	0
1	31	31	31
2	86	172	344
3	92	276	828
4	48	192	768
5	19	95	475
	288	766	2446

$$Var = \frac{2446}{288} - 2,66^2 = 1,42$$

■ **DESVIACIÓN TÍPICA:** $\sigma = \sqrt{\text{varianza}}$

La desviación típica es un parámetro más razonable que la varianza, pues se expresa en la misma magnitud que los datos y que la media (por ejemplo, si los datos vienen en centímetros, la desviación típica viene en centímetros; sin embargo, la varianza se daría en centímetros cuadrados).

En el ejemplo: $\sigma = \sqrt{1,42} = 1,19$

■ **COEFICIENTE DE VARIACIÓN:** $C.V. = \frac{\sigma}{\bar{x}}$

El coeficiente de variación sirve para comparar las dispersiones de poblaciones heterogéneas, pues indica la *variación relativa*.

En el ejemplo: $C.V. = \frac{1,19}{2,66} = 0,447$. O bien 44,7%.

Ejercicio resuelto

Calcular \bar{x} , σ y C.V. en la siguiente distribución:

DISTRIBUCIÓN DE PESOS (en kg)

INTERVALOS	FRECUENCIAS
42,5-53,5	4
53,5-64,5	19
64,5-75,5	86
75,5-86,5	72
86,5-97,5	41
97,5-108,5	7

Empezamos sustituyendo los intervalos por sus marcas de clase:

$$\frac{42,5 + 53,5}{2} = 48; \quad \frac{53,5 + 64,5}{2} = 59; \quad \dots; \quad \frac{97,5 + 108,5}{2} = 103$$

La tabla queda así:

x_i	f_i	$f_i x_i$	$f_i x_i^2$
48	4	192	9 216
59	19	1 121	66 139
70	86	6 020	421 400
81	72	5 832	472 392
92	41	3 772	347 024
103	7	721	74 263
229	17 658	1 390 434	

$$N = \sum f_i = 229$$

$$\sum f_i x_i = 17 658$$

$$\sum f_i x_i^2 = 1 390 434$$

Los números de la 3.ª columna, $f_i x_i$, se obtienen multiplicando los números de las columnas anteriores ($x_i \cdot f_i = f_i x_i$). Por ejemplo, $59 \cdot 19 = 1 121$.

Análogamente, los de la 4.ª columna se obtienen multiplicando los de la 1.ª por los de la 3.ª ($x_i \cdot f_i x_i = f_i x_i^2$). Por ejemplo, $59 \cdot 1 121 = 66 139$.

Con las sumas de las columnas de la tabla, obtenemos los parámetros:

$$\text{MEDIA: } \bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{17 658}{229} = 77,1 \text{ kg}$$

$$\text{DESVIACIÓN TÍPICA: } \sigma = \sqrt{\frac{\sum f_i x_i^2}{\sum f_i} - \bar{x}^2} = \sqrt{\frac{1 390 434}{229} - 77,1^2} = 11,2 \text{ kg}$$

$$\text{COEF. DE VARIACIÓN: } \text{C.V.} = \frac{\sigma}{\bar{x}} = \frac{11,2}{77,1} = 0,145 = 14,5 \%$$

CON CALCULADORA

- Preparamos la calculadora para que trabaje en el **MODO SD**.
- Borramos los datos que pudiera haber acumulados de otras ocasiones:
- Introducimos los datos: 48 4 ; 59 19 ; ...; 103 7
- Resultados obtenidos:

N.º DE INDIVIDUOS $\sum f_i$ →

SUMA DE VALORES $\sum f_i x_i$ →

SUMA DE CUADRADOS $\sum f_i x_i^2$ →

MEDIA \bar{x} →

DESV. TÍPICA σ →

En la web

- CALCULADORA: explicación pormenorizada del uso de la calculadora en estadística.
- Refuerza el cálculo de \bar{x} , σ y C.V.

Piensa y practica

- Halla, manualmente y con calculadora, \bar{x} , σ y C.V. en la tabla obtenida en el ejercicio resuelto de la página 92:

x_i	151	156	161	166	171	176
f_i	2	4	11	14	5	4

- Halla, manualmente y con calculadora, \bar{x} , σ y C.V. en la distribución de los ejercicios 1 y 2 de la página 92:

Compara los resultados entre sí y con los del ejercicio 1 de esta página.

En la web

- Interpretación del coeficiente de variación.
- Relaciona un histograma con su media y su desviación típica.

¿Qué aportan estos parámetros?

Los parámetros de posición son otra forma de sintetizar la información. Tienen la ventaja de que son muy fáciles de interpretar por los no expertos: con una sencilla explicación todo el mundo entiende que la afirmación "tu estatura está en el percentil 78" significa que, en ese colectivo, de cada 100 hay 78 más bajos que yo y 22 más altos.

DISTRIBUCIÓN I: 16 individuos

0 1 2 2 4 5 5 5
6 7 7 8 8 9 10 10
16 : 4 = 4 individuos en cada grupo:
4, 4 · 2 = 8, 4 · 3 = 12

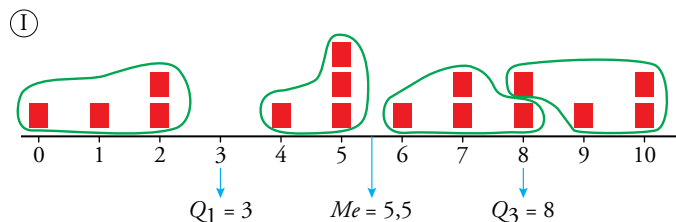
DISTRIBUCIÓN II: 19 individuos

0 1 2 2 3 4 5 5 5 6
6 7 7 8 8 9 10 10 10
19 : 4 = 4,75 individuos en cada grupo:
4,75, 4,75 · 2 = 9,5, 4,75 · 3 = 14,25
↓ ↓ ↓
ind. 5.º ind. 10.º ind. 15.º

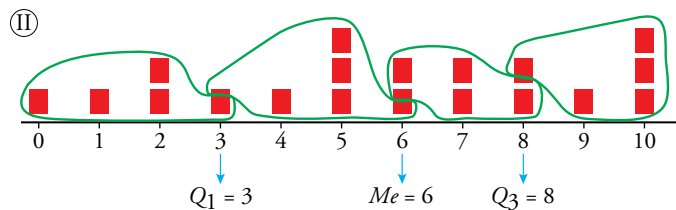
Mediana y cuartiles

Si los individuos de una población están colocados en orden creciente según la variable que se estudia y partimos la población en cuatro trozos con el mismo número de individuos, los puntos de separación son los cuartiles y la mediana:

- Q_1 (**primer cuartil**) es el valor de la variable que supera al 25% de la población y, por tanto, queda por debajo del otro 75%.
- Me (**mediana**) es el valor de la variable que deja el 50% de la población por debajo y el otro 50% por encima.
- Q_3 (**tercer cuartil**) supera al 75% de los individuos. Se llama tercer cuartil porque la mediana es el segundo cuartil.



En este caso, es muy fácil partir en 4 trozos iguales.



Cada paquete tiene un número fraccionario de individuos.

Observa que los cuartiles (Q_1 , Me , Q_3) se sitúan en los individuos que "se parten" para que los cuatro paquetes sean iguales.

Ejercicio resuelto

Calcular Me , Q_1 y Q_3 en la distribución:

1 1 2 3 4 4 5 5 5
5 6 7 7 7 8 9 10

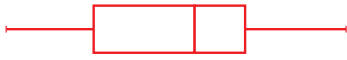
Hay 17 individuos. $17/2 = 8,5 \rightarrow$ La Me es el valor del individuo 9.º, $Me = 5$.
 $17/4 = 4,25 \rightarrow$ (5.º lugar) $Q_1 = 4$
 $17 \cdot (3/4) = 12,75 \rightarrow$ (13.º lugar) $Q_3 = 7$

Piensa y practica

1. Calcula Me , Q_1 y Q_3 en la siguiente distribución, cuyos datos están dados ordenadamente:

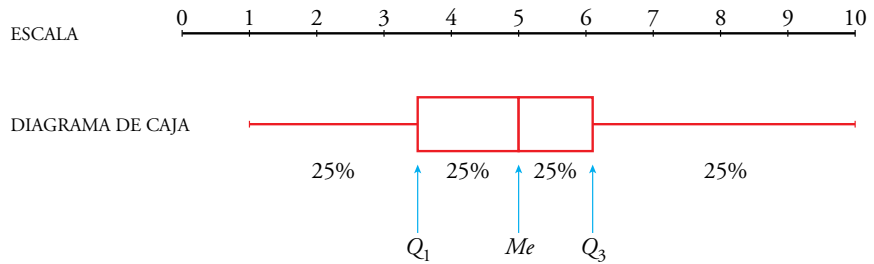
0 0 1 1 1 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 4 4 4 4 5 5 5 5 5 5 5
5 6 6 6 6 6 6 7 7 7 7 7 7 8 8 8 8 8 9 9 9 9 9 9 10 10

5 Diagramas de caja



Este diagrama se llama también de **caja y bigotes**.

Recordemos la forma de representar una distribución estadística teniendo en cuenta las principales medidas de posición:

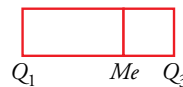


La gráfica corresponde a la distribución de notas en un cierto examen. En la parte alta se ha puesto la escala sobre la que se mueve la variable. Debajo se pone el diagrama propiamente dicho, que consiste en lo siguiente:

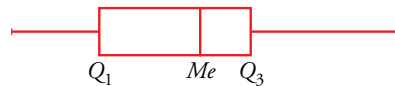
- El 50 % de los valores centrales, que se encuentran entre los cuartiles Q_1 y Q_3 , se destacan mediante un rectángulo (**caja**). Esta caja se divide en dos partes: el 25 % de los primeros valores centrales está entre Q_1 y Me y el otro 25 %, entre Me y Q_3 .
- Los valores extremos (el 25 % de los menores y el 25 % de los mayores) se representa mediante sendos segmentos (**bigotes**).

Los **diagramas de caja** (o caja y bigotes) se construyen del siguiente modo:

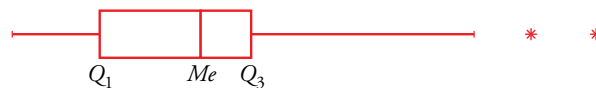
- La caja abarca el intervalo Q_1 , Q_3 (llamado **recorrido intercuartílico**) y en ella se señala expresamente el valor de la mediana, Me .



- Los bigotes se trazan hasta abarcar la totalidad de los individuos, con la condición de que cada lado no se alargue más de una vez y media la longitud de la caja.



- Si uno o más de los individuos quedaran fuera de esos tramos, el correspondiente lado del bigote se dibujaría con esa limitación ($1,5 \cdot$ longitud de la caja) y se añadirían, mediante asteriscos, los individuos en los lugares que les correspondan.



La longitud de la rama derecha es 1,5 veces la de la caja. En ella no están incluidos los dos individuos extremos que aparecen representados mediante asteriscos.

No lo olvides

La longitud de cada una de las ramas laterales (bigotes) no debe ser superior a 1,5 veces la longitud de la caja ($Q_3 - Q_1$). Si hay individuos cuyo valor se aleja de Q_1 o de Q_3 una distancia mayor que esa, se representan mediante puntos sueltos.

Ejercicios resueltos



1. Las estaturas de los 40 estudiantes de una clase son, dadas ordenadamente:

149	150	154	156	157
158	159	160	160	160
161	162	162	163	163
163	163	164	165	166
166	166	167	167	167
168	168	168	169	169
170	170	170	171	172
173	174	175	175	189

Representar la distribución mediante un diagrama de caja.

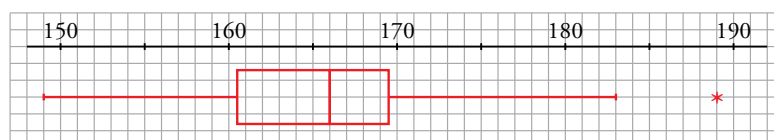
Puesto que el número de individuos es 40, Q_1 , Me y Q_3 serán los valores que hay entre los individuos 10.º y 11.º, entre 20.º y 21.º y entre 30.º y 31.º, respectivamente. Es decir:

$$Q_1 = 160,5 \quad Me = 166 \quad Q_3 = 169,5$$

La longitud de la caja es $Q_3 - Q_1 = 169,5 - 160,5 = 9$.

Una vez y media esta longitud es $1,5 \cdot 9 = 13,5$.

El altísimo estudiante que mide 189 cm se separa de Q_3 , el extremo superior de la caja, $189 - 169,5 = 19,5$. Esta distancia es mayor que una vez y media la longitud de la caja. Por eso, ponemos a la derecha un bigote de la mayor longitud posible, 13,5, y añadimos un asterisco que señala la situación del individuo excepcional, 189.



2. Representar, mediante un diagrama de caja, la siguiente distribución:

x_j	f_j
0	10
1	20
2	41
3	29
4	14
5	5
6	1

Hay 120 valores, por tanto, Q_1 está entre el 30.º y el 31.º, Me está entre el 60.º y el 61.º y Q_3 , entre el 90.º y el 91.º:

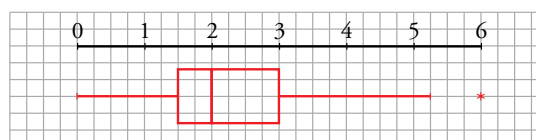
$$Q_1 = 1,5 \quad Me = 2 \quad Q_3 = 3$$

La caja abarca el intervalo $[Q_1, Q_3] = [1,5; 3]$.

La longitud del recorrido intercuartílico es $3 - 1,5 = 1,5$.

Los segmentos del bigote han de medir, como mucho, $1,5 \cdot 1,5 = 2,25$.

El bigote izquierdo mide menos de 2,25; sin embargo, el derecho, de 2,25, no abarca al elemento mayor (una familia con 6 hijos), ya que $Q_3 + 2,25 = 5,25$. Lo representamos mediante un asterisco.



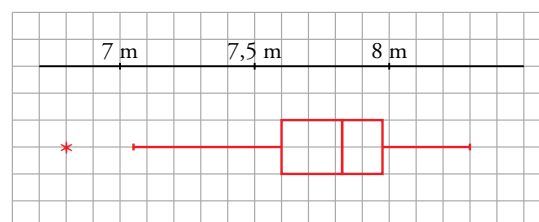
Piensa y practica

En la web Representación de diagramas de caja.

1. Haz el diagrama de caja correspondiente a esta distribución de notas:

x_j	f_j
1	6
2	15
3	22
4	24
5	33
6	53
7	22
8	16
9	8
10	1


2. Interpreta el siguiente diagrama de caja y bigotes relativo a las marcas de algunos saltadores de longitud:



Ejercicios y problemas

Practica


Tablas de frecuencias

1.  El número de faltas de ortografía que cometieron un grupo de estudiantes en un dictado fue:

0 3 1 2 0	2 1 3 0 4
0 1 1 4 3	5 3 2 4 1
5 0 2 1 0	0 0 0 2 1
2 1 0 0 3	0 5 3 2 1

Di cuál es la variable y de qué tipo es.


Haz una tabla de frecuencias y representa los datos en un diagrama adecuado.

2.  En una maternidad se han tomado los pesos (en kilogramos) de 50 recién nacidos:

2,8 3,2 3,8 2,5 2,7	3,7 1,9 2,6 3,5 2,3
3,3 2,6 1,8 3,3 2,9	2,1 3,4 2,8 3,1 3,9
2,9 3,5 3,0 3,1 2,2	3,4 2,5 1,9 3,0 2,9
2,4 3,4 2,0 2,6 3,1	2,3 3,5 2,9 3,0 2,7
2,9 2,8 2,7 3,1 3,0	3,1 2,8 2,6 2,9 3,3


- a) ¿Cuál es la variable y de qué tipo es?
 b) Construye una tabla con los datos agrupados en 6 intervalos desde 1,65 hasta 4,05.

Media, desviación típica y C.V.


3.  Halla la media, la desviación típica y el coeficiente de variación en estas distribuciones:

x_i	f_i
0	12
1	9
2	7
3	6
4	3
5	3

INTERVALO	f_i
1,65-2,05	4
2,05-2,45	5
2,45-2,85	13
2,85-3,25	17
3,25-3,65	8
3,65-4,05	3

4.  Los gastos mensuales de una empresa A tienen una media de 100 000 euros y una desviación típica de 12 500 euros. En otra empresa B, la media es 15 000 euros, y la desviación típica, 2 500 euros. Calcula el coeficiente de variación y di cuál de las dos tiene más variación relativa.


Parámetros de posición para datos aislados

5.  La altura, en centímetros, de un grupo de estudiantes de una misma clase es:

150 169 171 172 172 175 181
 182 183 177 179 176 184 158

Halla la mediana y los cuartiles y explica el significado de estos parámetros.



6.  Halla la mediana y los cuartiles de cada una de las siguientes distribuciones correspondientes al número de respuestas correctas en un test realizado por dos grupos de estudiantes:


A: 25 - 22 - 27 - 30 - 23 - 22 - 31 - 18


24 - 25 - 32 - 35 - 20 - 28 - 30

B: 27 - 32 - 19 - 22 - 25 - 30 - 21


29 - 23 - 31 - 21 - 20 - 18 - 27

Diagramas de caja

7.  Las puntuaciones obtenidas por 87 personas tienen los siguientes parámetros de posición: $Q_1 = 4,1$; $Me = 5,1$ y $Q_3 = 6,8$. Todas las puntuaciones están en el intervalo 1 a 9. Haz el diagrama de caja.

8.  En una clase de 38 estudiantes de Primaria, las estaturas de 35 de ellos están comprendidas entre 153 cm y 179 cm. Los tres restantes miden 150 cm, 151 cm y 183 cm. Sabemos que $Q_1 = 163$; $Me = 166$ y $Q_3 = 170$.

Representa los datos en un diagrama de caja.

9.  Haz el diagrama de caja correspondiente a las siguientes distribuciones.

a) La del ejercicio 5.

b) La A y la B del ejercicio 6.

