

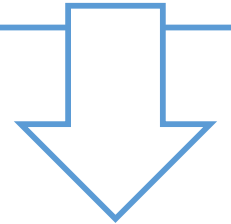


CLASE N° 3

MEDIDAS DE POSICIÓN

OA 2: Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión y probabilidades condicionales.

MEDIDAS DE ESTADÍSTICA DESCRIPTIVA



Medidas de tendencia central

Analizar los datos en torno a un valor central

Media, mediana y moda

Medidas de dispersión

Muestran la variabilidad de una distribución

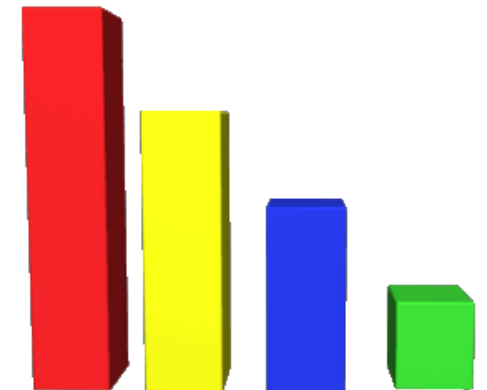
Rango, desviación media, varianza, desviación estándar y coeficiente de variación

Medidas de posición

Su intención es dividir el conjunto de observaciones en grupos con el mismo número de valores

Cuartiles y percentiles

SON VALORES NUMÉRICOS CALCULADOS A PARTIR DE LA MUESTRA Y QUE NOS RESUMEN LA INFORMACIÓN CONTENIDA EN ELLA.



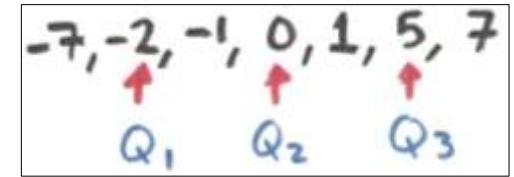
CUARTILES

Una de las **medidas de posición** son los **cuartiles** (Q_k , con $k = 1, 2, 3$), que corresponden a tres valores que dividen una distribución de datos en cuatro partes iguales.



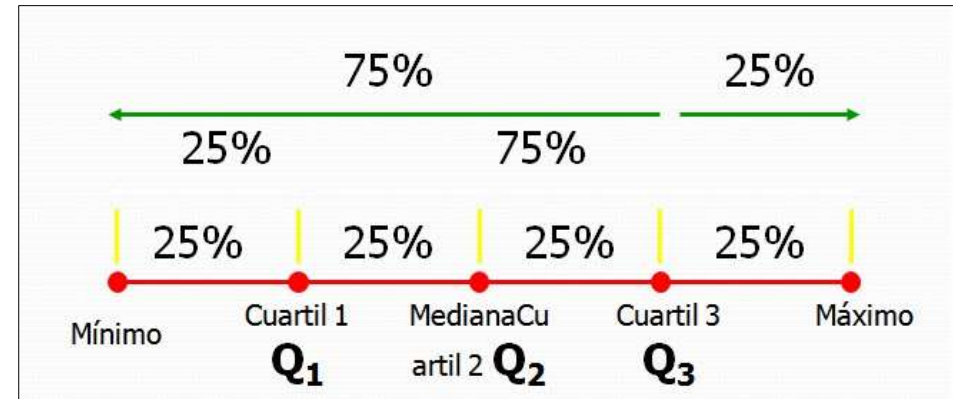
Para calcular el cuartil Q_k se deben ordenar los n datos en forma creciente y calcular $\frac{n \cdot k}{4}$.

- Si resulta un número entero, Q_k es igual al promedio entre el dato que se ubica en esa posición y el dato siguiente.
- Si resulta un número decimal, Q_k es igual al dato que ocupa la posición $\left\lceil \frac{n \cdot k}{4} \right\rceil + 1$.



Para datos agrupados

$$Q_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i \quad k = 1, 2, 3$$



EJEMPLO 1

Los siguientes datos son los puntajes obtenidos en relación con una prueba de admisión a una empresa.



100 - 121 - 134 - 123 - 142 - 118 - 123 - 142 - 126 - 127 - 131 - 98 - 116

Si para postular a la empresa se debe estar sobre el 50% de los mejores puntajes de todos los que rindieron la prueba, ¿cuál es el puntaje de corte?

SOLUCIÓN

100 - 121 - 134 - 123 - 142 - 118 - 123 - 142 - 126 - 127 - 131 - 98 - 116

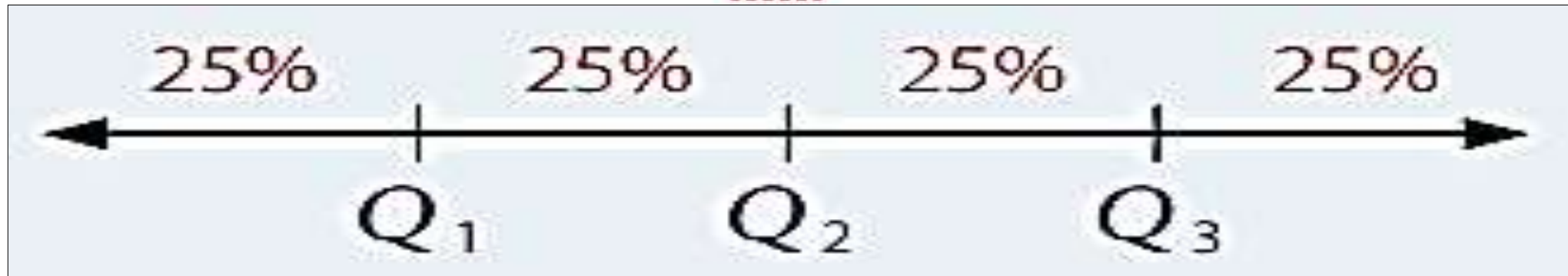
$$\frac{k \cdot N}{4}$$

- 1 Debemos calcular Q_2 , por lo que ordenamos los datos de forma creciente.

98 - 100 - 116 - 118 - 121 - 123 - 123 - 126 - 127 - 131 - 134 - 142 - 142

- 2 Identificamos el puntaje que divide a los datos en dos partes iguales.

98 - 100 - 116 - 118 - 121 - 123 - **123** - 126 - 127 - 131 - 134 - 142 - 142



- 3 El dato encerrado es el valor de Q_2 , el cual separa el 50 % de los datos de la distribución, por lo tanto para postular a la empresa se debe obtener un puntaje superior a 123.

PERCENTILES

Los **percentiles** (P_k , con $k = 1, 2, 3, \dots, 99$) corresponden a los 99 valores de una distribución que la dividen en 100 partes iguales. La diferencia entre dos percentiles consecutivos corresponde al 1% de la distribución. (Entre dos percentiles consecutivos está contenido el 1% de la muestra)

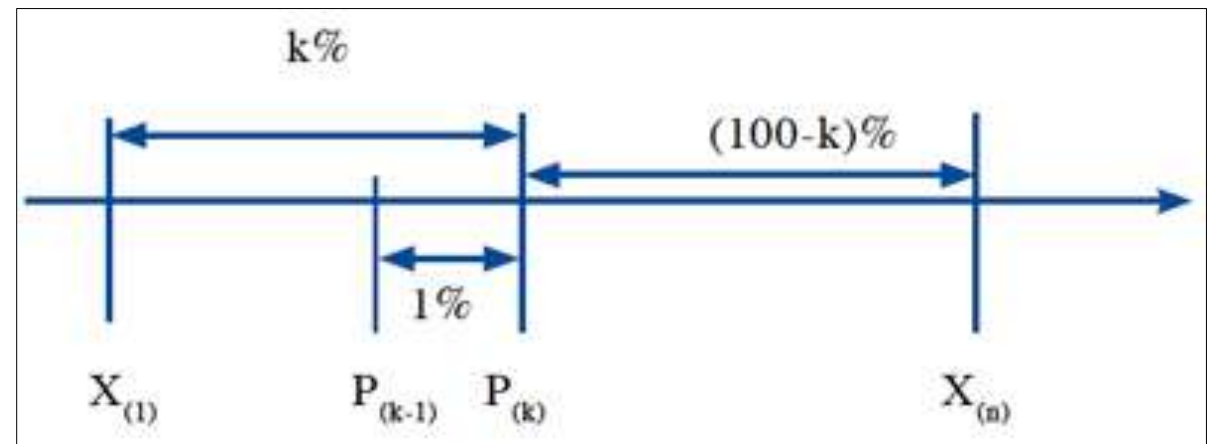
Para calcular el percentil P_k se deben ordenar los n datos en forma creciente y calcular $\frac{n \cdot k}{100}$.

- Si resulta un número entero, P_k es igual al promedio entre el dato que se ubica en esa posición y el dato siguiente.
- Si resulta un número decimal, P_k es igual al dato que ocupa la posición $\left[\frac{n \cdot k}{100} \right] + 1$.



Para datos agrupados

$$P_k = L_i + \frac{\frac{k \cdot N}{100} - F_{i-1}}{f_i} \cdot a_i$$



EJEMPLO 2

Se quiere seleccionar a un grupo de estudiantes para competir en las olimpiadas de atletismo. Las marcas (en metros) obtenidas por los estudiantes en una prueba son las siguientes:



52,4 - 56,3 - 57,5 - 65,3 - 65,3 - 66,5 - 66,8 - 67,9 - 68,7
69,3 - 70,2 - 71,4 - 72,4 - 74,7 - 74,9 - 75,5 - 75,6

Si se selecciona el 90% de las mejores marcas, ¿cuántos estudiantes no fueron seleccionados?

SOLUCIÓN

- 1 Debemos calcular P_{10} , ya que los estudiantes no seleccionados equivalen al 10%.

$$P_{10} = \frac{17 \cdot 10}{100} = \frac{170}{100} = 1,7$$

Como 1,7 es un número decimal, calculamos $[1,7] + 1 = 1 + 1 = 2$.

- 2 Como los datos ya están ordenados de forma creciente, identificamos aquel dato que ocupa la posición 2.

Posición	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Dato	52,4	56,3	57,5	65,3	65,3	66,5	66,8	67,9	68,7	69,3	70,2	71,4	72,4	74,7	74,9	75,5	75,6

10% DE LOS PARTICIPANTES TIENEN UNA MARCA MENOR A 56,3 M

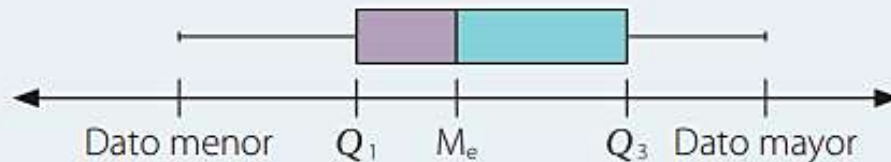
- 3 Luego, el valor de P_{10} corresponde a 56,3, por lo tanto 2 estudiantes no fueron seleccionados.

DIAGRAMA DE CAJÓN

Para **construir un diagrama de cajón** se traza una recta graduada a partir de los datos y se construye un rectángulo (cajón) cuyos extremos deben estar ubicados sobre Q_1 y Q_3 .

Así, la medida del largo de la caja es $Q_3 - Q_1 = Ric$, donde *Ric* corresponde al **recorrido intercuartil o rango intercuartil**, es decir, a la variabilidad de los datos con respecto a la mediana (*Me*).

Dentro del cajón se traza una línea vertical en el lugar de la mediana (*Me*); de esta manera, se divide el conjunto de datos en dos partes porcentualmente iguales. Luego, se trazan dos líneas, a ambos lados del cajón, desde sus extremos hasta los valores del dato menor y del mayor de la distribución.



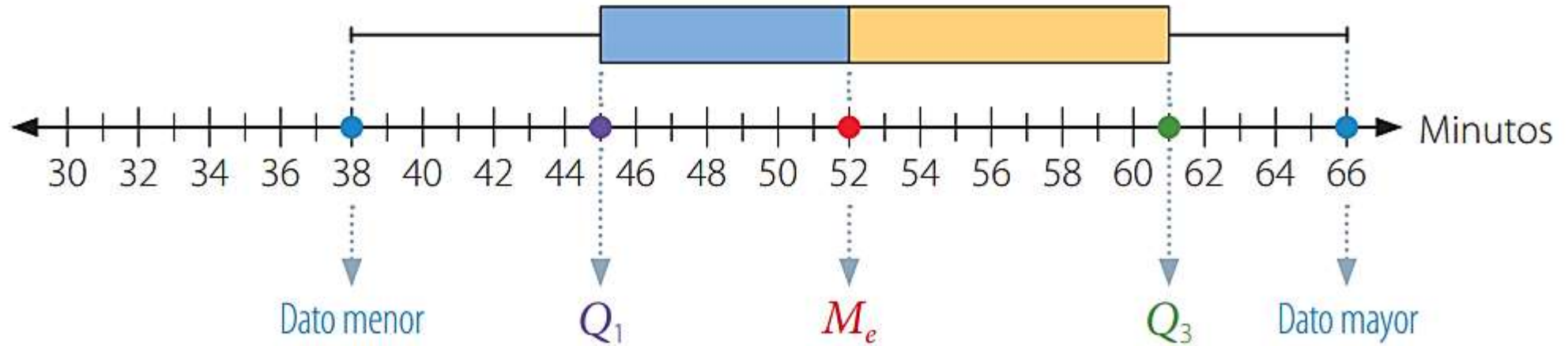
Al observar un diagrama de cajón es posible obtener conclusiones respecto de la distribución de la variable en estudio. Si uno de los cajones tiene mayor área, quiere decir que los datos que se ubican entre determinados cuartiles están más dispersos.

GUÍA 14 HAY UN VIDEO
EXPLICATIVO DE COMO
CONSTRUIR EL
DIAGRAMA DE CAJÓN



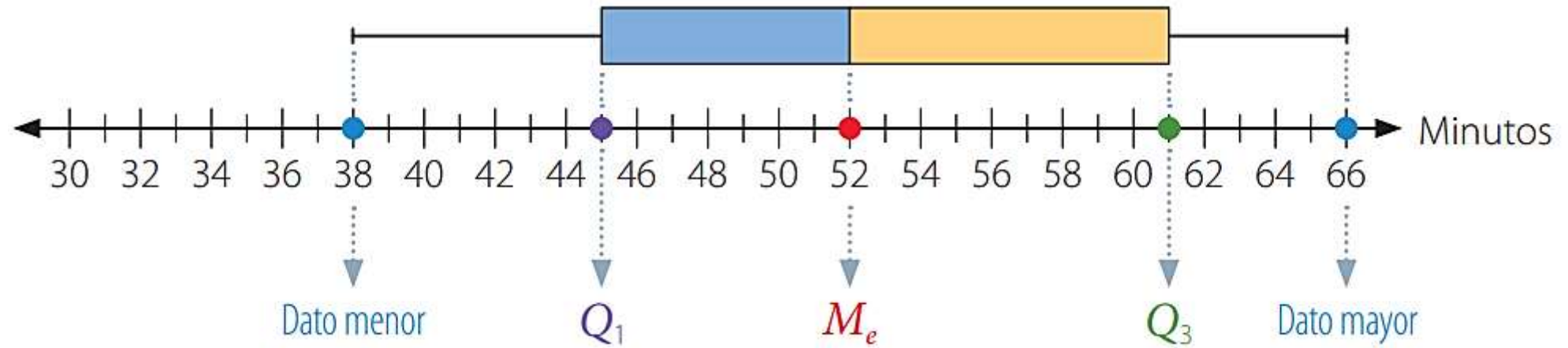
EJEMPLO 3

Los minutos que tardaron los estudiantes en responder un examen están representados en el siguiente diagrama.



¿Al cabo de cuántos minutos el 50% de los estudiantes terminó de contestar el examen?
¿Cuántos minutos tardaron en contestar el examen todos los estudiantes?

SOLUCIÓN



- 1 La mediana separa el 50% de los datos, por lo tanto a los 52 minutos la mitad de los estudiantes termina el examen.
- 2 Para determinar el tiempo que tardaron en responder el examen todos los estudiantes basta que observemos el dato mayor de la distribución de datos. Es decir, tardaron 66 minutos en responder el examen.

- Un **diagrama de cajón** es una representación que permite visualizar algunas características de la población a partir de las medidas de tendencia central y de posición.



GRACIAS