

CAPITULO 3

MEDIDAS DE POSICION

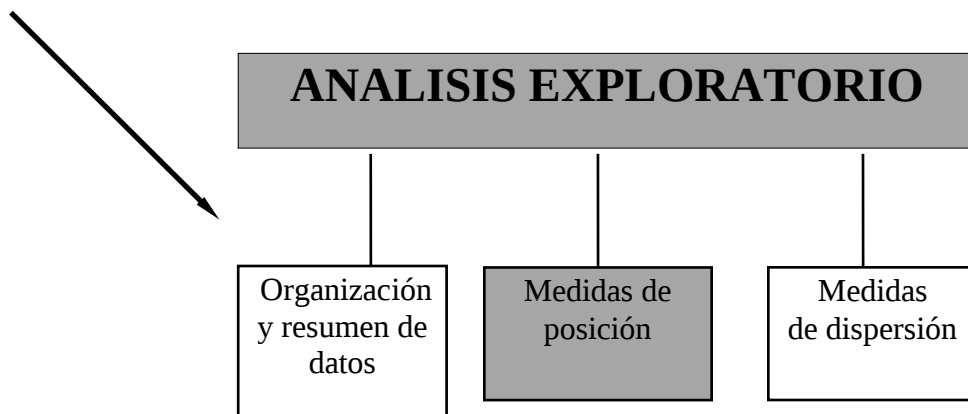
OBJETIVOS

- Identificar las medidas de posición más frecuentemente utilizadas como forma de síntesis de la información y aplicarlas en el tratamiento de datos concretos.
- Interpretar las medidas en situaciones reales para su mejor comprensión y aplicación.

CONTENIDOS

1. Introducción.....
2. Medidas de Posición.....
 - 2.1 La media aritmética.....
 - 2.1.1 Concepto.....
 - 2.1.2 Cálculo.....
 - 2.1.3 Ventajas y desventajas de la media aritmética.....
 - 2.1.4 Propiedades.....
 - 2.2 La media geométrica.....
 - 2.3 La media armónica.....
 - 2.4 La mediana.....
 - 2.4.1 Concepto.....
 - 2.4.2 Cálculo.....
 - 2.4.3 Determinación gráfica.....
 - 2.5 La moda.....
 - 2.5.1 Concepto.....
 - 2.5.2 Cálculo.....
 - 2.5.3 Determinación gráfica.....
 - 2.6 Los cuartiles.....
 - 2.6.1 Concepto.....
 - 2.6.2 Cálculo.....
 - 2.7 Los percentiles.....
 - 2.7.1 Concepto.....
 - 2.7.2 Cálculo.....
 - 2.8 Los deciles.....
 - 2.8.1 Concepto.....
 - 2.8.2 Cálculo.....
3. Indicadores de control
4. Problemas propuestos.....

Ud. está aquí



1. Introducción

En los capítulos anteriores vimos que los datos pueden presentarse por medio de una distribución de frecuencia y por medio de gráficos y cuadros. Podemos servirnos también de distintos tipos de medidas para resaltar las tendencias características de cada distribución, por ello necesitamos introducir conceptos que se expresen a través de números y que nos permitan inferir esas tendencias.

El primer grupo de medidas que veremos son las medidas de posición, dentro de ellas se encuentran las que denominamos de tendencia central. Estas medidas deben su nombre, porque tienden a ocupar posiciones centrales en una distribución de datos u observaciones. Las otras medidas de posición se ubican en distintos lugares dentro de la distribución de los datos. La unidad de estas medidas es la misma que la unidad de las observaciones.

Por ejemplo, si uno quiere caracterizar el rendimiento escolar de un niño, inmediatamente recurre al promedio de sus notas; de esta forma, aún sin saberlo, se está haciendo uso de una medida de tendencia central (la media aritmética). Un psicólogo quiere saber cual es la patología que aqueja a sus pacientes, también está usando una medida de posición (la moda). Un pediatra determinará el peso del bebé y en qué percentil se encuentra, está recurriendo a medidas de posición (los percentiles). Abundan innumerables ejemplos más sobre la importancia y su utilización en cada contexto de la realidad.

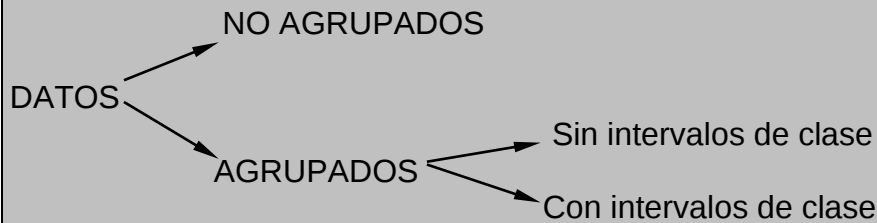
Las principales medidas de tendencia central o de posición que estudiaremos en este capítulo son:

- La media aritmética
- La media geométrica
- La media armónica
- La mediana
- La moda

Las otras medidas de posición pero que no se ubican en el centro de la distribución de datos que estudiaremos son:

- Los cuartiles
- Los percentiles
- Los deciles

Tengamos presente de acuerdo al capítulos anterior que los datos se pueden presentar en la siguiente forma:



En cada caso el cálculo de las medidas de posición será diferente, tal como veremos a continuación.

2. Medidas de Posición

2.1 La media aritmética

2.1.1 Concepto

La media aritmética es la medida de tendencia central o de posición que se utiliza con mayor frecuencia y es fácil de calcular. Muchas personas la utilizan, la conocen como “promedio” y tiene un gran uso general. En la práctica cuando se habla de “**media**” nos estamos refiriendo a la media aritmética, y entonces hablaremos de calificación media, estatura media, rendimiento medio de un alumno u obrero, etc.

La media aritmética es la suma de todos los valores que toma la variable dividida entre el número total de los mismos.

A continuación vamos a presentar la forma en que se calcula la media según como estén dispuestos los datos.

2.1.2 Cálculo de la media aritmética

2.1.2.1 Para datos no agrupados:

Sea un conjunto de n observaciones $X_1, X_2, \dots, X_n$, la media aritmética se expresa simbólicamente de la siguiente manera:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Utilizando la notación de sumatoria, la media aritmética nos queda:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Ejemplo 3.1

La producción lechera de una vaca (en litros) durante una semana fue:

10, 14, 13, 15, 16, 18 ,12

Por lo tanto la producción media de la semana fue:

$$\bar{x} = \frac{10+14+13+15+16+18+12}{7} = \frac{98}{7} = 14 \text{ litros}$$

X = 14 litros

Interpretación: La producción promedio de leche de una vaca, durante una semana fue de 14 litros.

Actividad 1:



En la guardia de un hospital público, los pacientes con cuadro gripal que concurrieron durante la semana fueron los siguientes:

Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
20	17	15	25	16	12	18

¿Cuántos pacientes en promedio con cuadro gripal concurrieron esa semana a la guardia del hospital?

Actividad 2:

\$

La recaudación tributaria (en \$) en la provincia de Jujuy entre los meses de abril y septiembre del año 2002 proporcionada por la Dirección Provincial de Rentas, se detalla a continuación:

ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE
3.784.059	4.316.839	4.228.730	5.089.305	6.082.525	6.210.358

Calcular la media aritmética e interpretarla. Realice los comentarios que le parezcan pertinentes.

2.1.2.2 Para datos agrupados sin intervalos de clase

En este caso, como las frecuencias son números indicadores de la intensidad de cada valor de la variable, ellas funcionan como factores de ponderación, lo que nos da por resultado una media aritmética ponderada.

$$\bar{X} = \frac{\sum_{i=1}^n x_i f_i}{n}$$

donde n = número de observaciones.

Ejemplo 3.2



En un barrio fueron encuestadas 34 familias, y entre las preguntas que le hicieron fue la cantidad de hijos que posee cada familia. Los resultados se observan en el cuadro siguiente:

Número de hijos (X)	Número de familias	$X_i f_i$
0	2	0
1	6	6
2	10	20
3	12	36
4	4	16
Total	34	78

Sugerencia

Un modo práctico para la obtención de la media es agregar una columna correspondiente a los productos que van en el numerador $X_i \cdot f_{ii}$

Calcular la media aritmética e interpretarla.

$$\bar{x} = \frac{78}{34} = 2.29 \cong 2.3$$

$\bar{x} = 2$ hijos.

Dado que la variable “ número de hijos” es una variable discreta, tomamos un número entero.

Interpretación: Las familias encuestadas tienen en promedio dos hijos.

Actividad 3:

Complete el esquema para el cálculo de la media aritmética de la siguiente distribución de frecuencias:

X_i	1	2	3	4	5	6
f_i	2	4	6	8	3	1

Tenemos:

X_i	f_i	$X_i f_i$
1	2	2
2	4	...
3	6	...
4	8	...
5	3	...
6	1	...
Total	(1)	(2)

$$\bar{X} = \frac{(2)}{(1)} = \dots$$

Importante: Si los datos que se tabulan son cuantitativos discretos, se puede verificar que la media aritmética del conjunto de datos originales, coincide con la media de los datos agrupados y en este caso **no hay pérdida de información**.

Comprobemos:

Datos originales

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

Supongamos que son las notas de examen de 28 alumnos. Calculamos la media:

$$\bar{x} = \frac{1+1+2+2+2+4+4+4+4+5+5+5+5+5+6+6+6+7+7+7+7+8+8+9+9+9+9+10}{28} =$$

$$= \frac{157}{28} = 5.6$$

Datos agrupados

Calificaciones X	Número de alumnos	X fi
1	2	2
2	3	6
4	4	16
5	5	25
6	3	18
7	4	28
8	2	16
9	4	36
10	1	10
TOTAL	28	157

$$\bar{x} = \frac{157}{28} = 5.6$$

Como podemos comprobar el resultado es el mismo y no hay pérdida de información. Esto no ocurre para las variables continuas agrupadas en intervalos de clase, que es de lo que nos ocupamos en el punto siguiente.

2.1.2.3 Para datos agrupados con intervalos de clase

Los datos al estar agrupados en intervalos de clase, nos permiten simplificar nuestra labor reduciendo los cálculos. Pero, por otra parte, al agruparlos en intervalos perdemos información en relación con el conjunto de datos originales. La media calculada a partir de los datos agrupados es una aproximación a la media aritmética del conjunto original de datos.

Un ejemplo de datos agrupados lo constituyen la información que usualmente se publica sobre los censos de población. Por ellos sabemos que hay cierto número de personas en edades de 0 a 4 años, o de 5 a 9 años, pero desconocemos cuántos individuos de una determinada edad exacta hay.

Al trabajar con intervalos de clase hemos de proceder a hacer ciertos supuestos simplificadores. Trataremos todos los casos como si se hallaran concentrados en los puntos medios de sus intervalos respectivos. Estas simplificaciones llevan aparejadas cierta inexactitud. En efecto, no podemos esperar obtener en esta forma exactamente los mismos resultados que nos proporcionarían los datos brutos. Pero, por otra parte, si el número de datos es grande, las distorsiones introducidas serán por lo regular insignificantes, y compensarán sobradamente el ahorro de tiempo.

NOTA



Cuanto más pequeña sea la amplitud de los intervalos, tanto menos información perderemos y tanto mayor será la exactitud.

La fórmula de la media después de haber hecho las consideraciones anteriores, se convierte en:

$$\bar{X} = \frac{\sum_{i=1}^n x_{mi} f_i}{n}$$

donde X_{mi} es el punto medio de cada intervalo de clase.

Ejemplo 3.3

El siguiente cuadro registra los salarios del personal jerárquico de una empresa transnacional. Calcular la media aritmética.

Salarios (\$) X_i	N° de Ejecutivos f_i
1950 - 2950	17
2950 - 3950	26
3950 - 4950	38
4950 - 5950	51
5950 - 6950	36
6950 - 7950	21
Total	189

Para realizar el cálculo agregamos dos columnas: una con los puntos medios de cada intervalo o marcas de clase, y otra multiplicándolos por las frecuencias respectivas.

Salarios (u\$s) X_i	N° de Ejecutivos f_i	Punto medio del intervalo	
		X_{mi}	$X_{mi}f_i$
1950 - 2950	17	2450	41650
2950 - 3950	26	3450	89700
3950 - 4950	38	4450	169100
4950 - 5950	51	5450	277950
5950 - 6950	36	6450	232200
6950 - 7950	21	7450	156450
Total	189		967050

$$\bar{X} = \frac{\sum_{i=1}^n x_{mi} f_i}{n} = \frac{967.050}{189} = \$ 5.117$$

Interpretación: El sueldo promedio del personal jerárquico de la empresa es de \$ 5.117.

Observación!!!!

Si escribiéramos todos los datos originales en una columna, observaríamos que el valor de 5.117, estaría ubicado en la parte central de la serie de datos. Siempre la media va a tender a ocupar un lugar central.

Actividad 4:

Las estaturas de 40 alumnos de una escuela pueden apreciarse en el siguiente cuadro:

Estatura (cm)	N° de Niños (fi)
150-154	4
154-158	9
158-162	11
162-166	8
166-170	5
170-174	3
Total	40

Calcular la media aritmética e interpretarla.



2.1.3 Ventajas y desventajas de la media aritmética

VENTAJAS

- ◇ Cálculo sencillo
- ◇ Fácil interpretación y concepto familiar a la mayoría de las personas
- ◇ Intervienen todos los valores de la variable

DESVENTAJAS

- ◇ Fuertemente afectada por la existencia de valores extremos que no son representativos del resto de las observaciones.
- ◇ No se puede calcular cuando hay intervalos abiertos en una serie de frecuencias

Ventajas: Hemos podido observar que tanto el cálculo como su interpretación no presentan dificultades. En todas las disciplinas u ocupaciones su uso es generalizado y se toma como referente importantísimo a esta medida para describir o analizar cualquier serie de datos u observaciones.

Desventajas: La existencia de valores extremos afecta el resultado de la media, lo cual en algunos casos que se de ésta presencia, la torna poco representativa de la serie de datos. Veamos un ejemplo que nos clarificará esta dificultad que puede presentar la media aritmética:

Tomemos dos series de datos iguales donde en la segunda cambia un extremo.

Serie 1:

100 100 120 140 200 200 300 300 350

Serie 2:

100 100 120 140 200 200 300 300 **10.000**

Supongamos que se trata de los ingresos per-cápita de la población de dos países que se utiliza como medida de desarrollo económico de un país (como medida es cuestionada por considerarla defectuosa por varias circunstancias, entre ellas el no considerar la forma en que está distribuido ese ingreso entre la población). Calculemos la media en ambos países. En el primer país nos da \$ 201.11; y en el segundo país nos da \$ 1.273,33. Vemos que la diferencia es considerable, a causa de la influencia de un solo valor extremo. De aquí surge que cuando calculamos un promedio es necesario ser cuidadoso para llegar a conclusiones determinantes. Más adelante estudiaremos otro tipo de medidas que nos indicarán cuán representativo es un promedio. En referencia a lo anterior lo ilustraremos con un hecho de la realidad:

El diario Clarín, el 20 de octubre de 1997, publicó en la tapa:



Esto surge de la encuesta nacional que realiza el INDEC. En Capital y el Gran Buenos Aires el sueldo promedio es \$ 635. Los salarios más altos están en Ushuaia, con \$ 951. Y los más bajos en Corrientes, con \$ 399.

?

De esto surge los siguientes comentarios:

- ✓ No se puede cuestionar el cálculo.
- ✓ Para el imaginario social esto es una mentira. Un empleado estatal de categoría 1, que sus ingresos no superan los \$ 150, se pregunta ¿de qué me están hablando?
- ✓ La representatividad de ese ingreso promedio es cuestionable.

- ✓ No se puede publicar una noticia de ese impacto en la gente, sin hacer las salvedades del caso. Hay que ser más cuidadoso, sobre todo cuando es un medio masivo, y no una publicación científica.

Además la otra desventaja es de que la media no se puede calcular para intervalos abiertos. Si tenemos un intervalo abierto, para el cálculo de la media necesitamos la marca de clase, cosa que en dicho intervalo es imposible determinar.

2.1.4 Propiedades de la media

Propiedad 1: Si todos los valores observados $x_1, x_2, \dots, x_n$ son iguales a d (donde d es una constante), entonces

$$\underline{X(x_i) = X(d) = d}$$

Probemos con un ejemplo:

Valores observados: 10,10,10,10,10

$d = 10$

Calculamos la media de los valores observados: $50 / 5 = 10$

Calculamos la media de d : $10/1 = 10$

Por lo tanto la media aritmética de una constante es la misma constante.

Propiedad 2: Si a cada valor de las observaciones $x_1, x_2, \dots, x_n$ se le suma o resta una constante, la media aritmética del nuevo conjunto transformado, es la media aritmética del conjunto original más o menos la constante.

$$\underline{\bar{Y}(y_i) = \bar{X}(x_i) \pm d}$$

donde $y_i = x_i \pm d$, $d = \text{constante}$

Propiedad 3: Si a cada valor de un conjunto de observaciones $x_1, x_2, \dots, x_n$ se multiplica por una constante diferente de cero, la media aritmética del nuevo conjunto transformado es la media aritmética del conjunto original multiplicado por la constante.

$$\underline{\bar{Y}(y_i) = d\bar{X}(x_i)}$$

donde $y_i = d x_i$

Propiedad 4: La suma algebraica de las desviaciones de cada valor observado $x_1, x_2, \dots, x_n$ con respecto a la media es cero.

$$\underline{\sum_{i=1}^n (x_i - \bar{x}) = 0}$$

Si los datos están tabulados se multiplica por las frecuencias correspondientes y el resultado es el mismo.

Veamos con un ejemplo:

Se tienen los valores 3, 4, 2 y 3, cuya media es igual a 3.

Calculamos los desvíos:

$$3 - 3 = 0$$

$$4 - 3 = 1$$

$$2 - 3 = -1$$

$$3 - 3 = 0$$

La suma de los desvíos (0+1-1+0) da cero por resultado.

Propiedad 5: La suma de los cuadrados de las desviaciones de cada una de las observaciones $x_1, x_2, \dots, x_n$ con respecto a su media es mínima. Esta propiedad significa que, si a es cualquier valor, entonces:

$$\sum_{i=1}^n (x_i - \bar{x})^2 \leq \sum_{i=1}^n (x_i - a)^2$$

Propiedad 6: En casos especiales, esto cuando es apropiado hacerlo, la media aritmética se puede utilizar para estimar la cantidad total de una población, esto es:

$$\text{Total} = N\bar{X}$$

donde N = tamaño de la población
 $\bar{x}$ = media aritmética de la muestra

2.2 La media geométrica

La media geométrica de un conjunto de n números positivos se define como la raíz n -ésima del producto de los n valores. Su fórmula es la siguiente:

$$Mg = \sqrt[n]{(x_1)(x_2)\dots(x_n)}$$

Para cierto tipo de variables que sigue una tendencia exponencial, se hace indispensable su uso. La media geométrica también es útil para encontrar el promedio de porcentajes, índices o tasas de crecimiento.

La media geométrica siempre será menor o igual que la media aritmética. Todos los valores deben ser positivos y si algún valor de la variable es cero, la media geométrica será cero.

Ejemplo 3.4

Una empresa obtuvo durante un año las siguientes ganancias trimestrales: 1%, 3%, 2% y 7%. Calcular la media geométrica de las ganancias de la empresa.

$$Mg = \sqrt[4]{(1).(3).(2).(7)} = \sqrt[4]{42} = 2.54\%$$

Interpretación: El valor medio geométrico de las ganancias es 2.54%.



Si calculamos la media aritmética, ¿qué obtendríamos?:

$$\bar{x} = \frac{0.01 + 0.03 + 0.02 + 0.07}{4} = 0.0325 = 0.0325 * 100 = 3.25\%$$

Observamos que el valor que se obtiene es mayor, y que esta influida por el valor extremo.

!!!! NOTA: Para utilizar la calculadora para el cálculo de la media geométrica se siguen los siguientes pasos: al resultado del producto de los valores de la variable que constituye el radicando, en el ejemplo “42” se lo teclea, luego según la calculadora se presiona la tecla “2nd ó INV”, inmediatamente la tecla $x\sqrt{y}$ después el índice de la raíz (4) y finalmente el signo =.

Otra aplicación:

Otra aplicación de la media geométrica es encontrar un aumento porcentual promedio en un intervalo de tiempo. La fórmula es la siguiente:

$$Mg = \sqrt[n]{\frac{\text{Valor al final del período}}{\text{Valor al principio del período}}} - 1$$

$\Delta \%$

Ejemplo 3.5



La población de una ciudad en 1990 fue de 30.000 habitantes. En el año 2000 pasó a tener 45.000 habitantes. ¿Cuál fue la tasa del incremento porcentual anual promedio para el período?.

$$Mg = \sqrt[10]{\frac{45.000}{30.000}} - 1 = 0.04$$

La tasa de crecimiento anual promedio es del 4%. Esto significa que cada año la población de esa ciudad aumenta en promedio un 4%.

Actividad 5

Cuatro acciones de un producto importante, tuvieron un rendimiento anual en porcentaje de 5.61, 6.33, 9.21 y 42.3.

1. Calcule la media geométrica de los rendimientos. Interprete.
2. Calcule la media aritmética de los rendimientos y compare.

2.3 La media armónica

La media armónica “ M_h ” de n términos no nulos $x_1, x_2, \dots, x_n$, es el recíproco de la media aritmética de los recíprocos de esos términos. Es decir:

$$M_h = \frac{1}{\sum_{i=1}^n \frac{1}{x_i} / n} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

La media armónica para su cálculo contiene todas las observaciones, por ello está afectada por los valores extremos. No está definida si alguno de los valores es cero.

Es muy aplicada cuando existe una relación inversa entre las variables. Ejemplos: productividad y tiempo, precio y cantidad demandada, velocidad y tiempo, etc.

Ejemplo 3.6

La media armónica de los valores 2, 4, 7 y 9 es:

$$M_h = \frac{4}{\frac{1}{2} + \frac{1}{4} + \frac{1}{7} + \frac{1}{9}} = \frac{4}{1.004} = 3.98$$

Siempre se cumple:

$$M_h \leq M_g \leq \bar{x}$$

Ejemplo 3.7

Una editorial publica un aviso en los clasificados de un diario, solicitando la incorporación a la empresa de un operador de computadora. Luego de una primera selección quedaron seis personas preseleccionadas. Para seleccionar una persona de las seis, la empresa decide darle un mismo texto a cada uno y medir la velocidad de palabras por minuto y tomar como parámetro de decisión la velocidad media. La prueba

brindó los siguientes resultados: 42, 38, 30, 35, 34 y 40 palabras por minuto. Calcular la velocidad promedio.

NOTA

Dado que la velocidad y el tiempo son variables que mantienen una relación inversa, es conveniente calcular la media armónica

Pasamos a calcular:

$$M_h = \frac{6}{\frac{1}{42} + \frac{1}{38} + \frac{1}{30} + \frac{1}{35} + \frac{1}{34} + \frac{1}{40}} = \frac{6}{0.024 + 0.026 + 0.033 + 0.028 + 0.029 + 0.025} = \frac{6}{0.165} = 36.4$$

$M_h = 36.4$ palabras por minuto.

Interpretación: La velocidad media de los operadores es de 36.4 palabras por minuto.

Observación:

- Para el cálculo de los números fraccionarios de la fórmula a través de la calculadora, cuenta con un tecla que en la parte superior figura 1/ X. Teclee el denominador y luego presione la tecla indicada e inmediatamente obtiene el resultado.

1 / X



Actividad 6

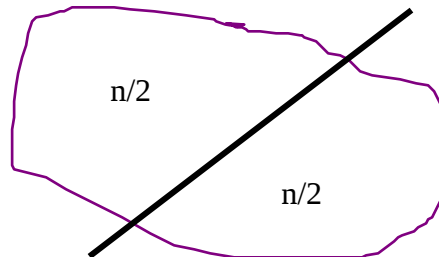
Un automovilista parte de su pueblo hacia la capital de su provincia, que se encuentra a una distancia de 60 km. Los primeros 20 km los recorre a una velocidad de 110 km /h y el resto a 140 km / h. Calcule la velocidad media del automovilista.

2.4 La Mediana

2.4.1 Concepto

Es el valor medio que separa en partes iguales a la serie de datos. Es el valor a cuya izquierda se encuentra el 50% de los datos, y a cuya derecha se encuentra el restante 50% de los datos.

Dado un conjunto de n valores de una variable, $X_1, X_2, X_3, \dots, X_n$, la mediana es una X particular, tal que supera a $n/2$ de los X_i y es superado por los $n/2$ restantes. Se simboliza como **Me**.



2.4.2 Cálculo de la mediana

Para el cálculo de la mediana, se presentan dos casos: cuando los datos no están agrupados y cuando los datos están agrupados.

2.4.2.1 Para datos no agrupados

El número de elementos de la distribución de los datos, puede ser par o impar. Para cada caso el cálculo es distinto. Veamos:

a) Con número impar de datos

Dados los valores

10, 7, 2, 8 y 16

En primer lugar se ordenan los datos en forma creciente:

2 7 8 10 16



Me

La mediana es el valor central.

Comprobemos:

$8 > 2$

$8 < 10$

$8 > 7$

$8 < 16$

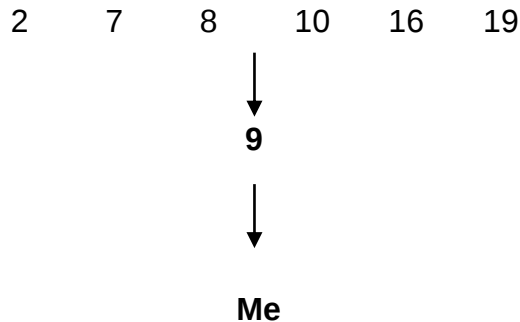
8 supera a dos valores

8 es superado por dos valores

b) Con número par de datos

Dados los valores ordenados 2, 7, 8, 10, 16 y 19, no hay un valor que ocupe la posición central, sino que hay dos valores centrales: 8 y 10. Cualquiera de los valores comprendidos entre 8 y 10 divide al conjunto de datos en dos partes iguales, cumpliendo con la definición de mediana.

Para resolver esta indeterminación se toma como mediana la media aritmética simple de los valores centrales:



$$Me = (8 + 10) / 2 = 9$$

$$9 > 2 \quad 9 < 10 \quad || \quad 9 \text{ supera a tres datos y}$$

$$9 > 7 \quad 9 < 16 \quad \text{también es superado por tres datos} \quad ||$$

$$9 > 8 \quad 9 < 19$$

Actividad 7

Calcular la mediana para las dos series de datos:

a) 5, 13, 10, 2, 18, 15, 6, 16, 9

b) 2, 6, 7, 10, 12, 13, 18, 21

Notas:

- ♦ El valor de la mediana puede coincidir o no con un elemento de la serie, como vimos. Cuando el número de elementos de la serie es impar, hay coincidencia. Lo mismo no acontece cuando ese número es par.
- ♦ La mediana y la media aritmética no tienen necesariamente el mismo valor.
- ♦ La mediana, como vimos, depende de la posición y no de los valores de los elementos en la serie ordenada. Esa es una de las diferencias fundamentales entre la mediana y la media que se deja influenciar, y mucho, por los valores extremos. Esta propiedad de la mediana puede ser constatada a través de los ejemplos a seguir:

$$5, 7, 10, 13, 15 \Rightarrow \bar{X} = 10 \text{ y } Me = 10$$

$$5, 7, 10, 13, 65 \Rightarrow \bar{X} = 20 \text{ y } Me = 10$$

Esto es, la media del segundo conjunto de valores es mayor que la del primero, por influencia de un valor extremo, mientras que la mediana permanece igual.

2.4.2.2 Para datos agrupados sin intervalos de clase

Los pasos para calcular la mediana son los siguientes:

- Calculamos $n / 2$, es decir, el total de observaciones (la suma de las frecuencias absolutas) lo dividimos por dos.
- En la tabla de distribución de frecuencias agregamos una columna y calculamos las frecuencias acumuladas “ menor que”.
- Al resultado de $n / 2$, buscamos la frecuencia acumulada “ menor que” inmediatamente superior al mismo. Ese valor de frecuencia acumulada corresponde a un determinado valor de la variable, **ese valor de la variable será la mediana.**

A través de un ejemplo observaremos en forma clara su cálculo.

Ejemplo 3.8

En una empresa se registró el número de días trabajados por los obreros durante un mes. Calcular la mediana e interpretarla.



Nº de días trabajados por cada obrero	Nº de obreros
19	5
20	15
21	30
22	35
23	10
24	5
Total	100

1º- El primer paso es calcular el total de datos dividido por dos.

$$n / 2 = 100 / 2 = 50$$

2º- Calculamos las frecuencias acumuladas de la forma “ menor que” , agregando otra columna en la tabla.

Nº de días trabajados por cada obrero	Nº de obreros	F<
19	5	5
20	15	20
21	30	50
22	35	85
23	10	95
24	5	100
Total	100	

En la columna de frecuencias acumuladas se busca la primera que supere a $n/2 = 50$, que en este ejemplo es 85 (algunos autores consideran 50 que coincide). A 85 le corresponde el valor de $X = 22$; por lo tanto, 22 es la mediana de este conjunto de datos.

Interpretación: El 50% de los obreros trabajaron 22 días o menos, y el 50% restante 22 días o más.

Actividad 8

Dada la siguiente distribución de niños de 0 a 4 años seleccionados a través de una muestra en un barrio, calcule la mediana e interprete.

X_i	f_i
0	2
1	6
2	10
3	12
4	5

2.4.2.3 Para datos agrupados con intervalos de clase

Los pasos a seguir para el cálculo son los siguientes:

- 1) Se calculan las frecuencias acumuladas.
- 2) Se calcula $n/2$.
- 3) En la columna de frecuencias acumuladas se busca la primera frecuencia que supere el valor $n/2$.
- 4) Para encontrar el valor exacto de la mediana se aplica la siguiente fórmula:

$$Me = L_i + \frac{n/2 - F_{(i-1)}}{f_i} \cdot a$$

en la cual:

- * **F (i-1)** simboliza la frecuencia acumulada correspondiente al intervalo anterior al que contiene la mediana.
- * **L_i** es el límite inferior del intervalo de clase donde se halla contenida la mediana.
- * **f_i** es la frecuencia absoluta correspondiente al intervalo que contiene a la mediana.
- * **a** es la amplitud del intervalo de clase.

Ejemplo 3.9

Dada la distribución de frecuencias que se expone en la tabla a continuación, calcularemos la mediana correspondiente:

Intervalos	f _i	F _i
0 - 10	2	2
10 - 20	7	9
20 - 30	15	24
30 - 40	9	33
40 - 50	7	40
50 - 60	2	42

Una vez calculadas las frecuencias acumuladas se procede a obtener $n/2$:

$$n / 2 = 42 / 2 = 21$$

Ahora buscamos la primera frecuencia acumulada que supere a $n/2 = 21$. En el ejemplo, la primera frecuencia acumulada que supera a 21 es 24, pero no corresponde a un valor de la variable, sino a un intervalo de clase: 20 - 30. Es decir:

$$20 \leq Me < 30$$

Una vez encontrado el intervalo donde se halla contenida la mediana, reemplazamos en la fórmula los valores, y obtenemos:

$$Me = 20 + \frac{21 - 9}{15} \cdot 10 = 28$$

NOTA :

En el caso de existir una frecuencia acumulada exactamente igual a $n/2$, la mediana será el límite superior de la clase correspondiente

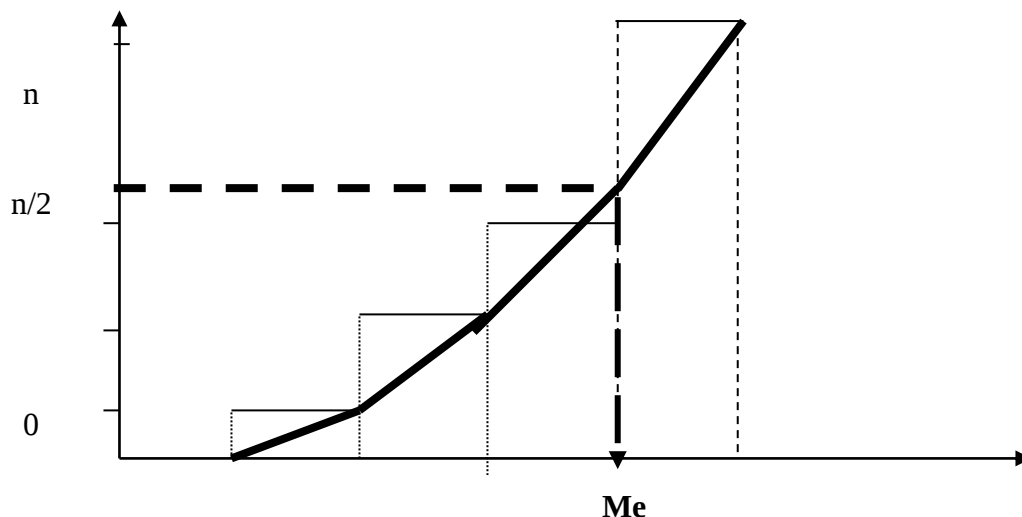
Actividad 9

Se realizó un test de inteligencia a 100 alumnos seleccionados aleatoriamente en un colegio, los resultados se presentan en el siguiente cuadro:

Puntaje	N° de alumnos
0 – 10	5
10 – 20	6
20 – 30	4
30 – 40	15
40 – 50	30
50 – 60	20
60 – 70	8
70 – 80	5
80 – 90	5
90 – 100	2
Total	100

Calcular la mediana e interpretar el resultado.

2.4.3 Determinación gráfica de la mediana



Para obtener la mediana en forma gráfica, se construye la ojiva “menor que” proveniente de las frecuencias absolutas o relativas acumuladas.

Sobre el eje de ordenada buscamos el punto $n/2$, y a partir de ese punto trazamos una recta paralela al eje de abscisa que corte a la ojiva. Desde ese punto trazamos una perpendicular al eje de abscisa. El punto que corta a dicho eje es la **mediana**.

2.5 La Moda

2.5.1. Concepto

La moda o modo se define como el valor o clase que tiene la mayor frecuencia en un conjunto de observaciones. Otra forma de definir la moda: es la categoría o valor de la variable que más se repite, según la variable que se trate, cualitativa o cuantitativa. Por lo tanto la moda se puede calcular tanto para variables cualitativas como para variables cuantitativas. Se simboliza como **Mo**.

Ejemplo 3.1.1

Supóngase que en una clase de 20 estudiantes se clasifica según el color de ojos de los mismos de la siguiente forma:

Color de ojos de los estudiantes	Azul	Marrón	Gris	Verde
Nº de estudiantes	6	8	4	2

 **Mo = ojos marrones**

Entonces, la moda será los estudiantes de ojos marrones.

!!

Importante:

A diferencia de la media y la mediana, la moda puede usarse para variables cualitativas. La moda presenta la ventaja de no estar afectada por los valores extremos, como es el caso de la media aritmética.

2.5.2. Cálculo de la moda

2.5.2.1. Para variables cualitativas

Dado que su cálculo es simple, vamos a verlo a través de un ejemplo:

Ejemplo 3.1.2

El director de personal de una empresa tomó una prueba de selección a 50 personas. La prueba consistió en un test que mide la capacidad para desempeñar una determinada tarea, midiéndose la misma como: nula, baja, media y superior. Los resultados obtenidos fueron:

B – N – M – M – N – S – N – N – N – S – B – S – B – N – M – M – B – M – B – B – S – B – M –
N – N – N – B – S – S – N – M – B – B – M – B – B – M – M – N – S – N – N – N – B – B – B –
B – S – B – N

- ¿Cuál es la variable en estudio y su escala de medición?
- Construir una tabla de distribución de frecuencias con la información brindada.
- Determinar la moda y la mediana.

Resultado:

- La variable es “la capacidad para desempeñar la tarea en la empresa”

Esta variable es cualitativa y su escala es ordinal.

b)

Capacidad	Cantidad de personas (fi)	Frec. Relativa(ri)	Frec. Acumulada “menor que”
Superior (S)	8	0,16	8
Media (M)	10	0,20	18
Baja (B)	17	0,34	35
Nula (N)	15	0,30	50
TOTAL	50	1	

c)

La moda es la capacidad de desempeño baja. Es la más frecuente en la prueba realizada.

Para calcular la mediana, procedemos a dividir el total de la muestra por dos ($n / 2$). Nos da 25 ese cociente, luego buscamos la frecuencia acumulada inmediata superior. Dicha frecuencia es 35 y corresponde a la categoría "baja". Por lo tanto, la mediana es la capacidad de desempeño baja.

Actividad 10

En una escuela se realizó una prueba de matemáticas para observar el rendimiento de los alumnos en esa materia. Para ello se seleccionó aleatoriamente una muestra de 40 alumnos, y dio los siguientes resultados:

Excelente (E) , Bueno (B), Regular (R) y Malo (M)

E	R	M	M	R	B	E	M	R	E
M	M	B	B	E	R	M	M	B	B
E	B	M	M	M	B	R	R	M	B
B	B	M	E	M	M	M	R	E	R

a) ¿Cuál es la variable en estudio y su escala de medición?

b) Construir una tabla de distribución de frecuencias con la información brindada.

c) Determinar la moda y la mediana.

2.5.2.2. Para variables cuantitativas con datos no agrupados

Cuando tenemos datos no agrupados, la moda es fácilmente reconocida; basta, de acuerdo con la definición, encontrar el valor que más se repite.

Ejemplo 3.1.4

Dada la serie de datos:

7, 8, 9, 10, 10, 10, 11, 12, 13, 15

la moda será igual a 10, que es el valor que más se repite $\Rightarrow$ **Mo** = 10

Actividad 11

Calcule la moda en la siguiente serie de datos:

3, 3, 6, 7, 8, 8, 13, 16, 19

2.5.2.3. Para variables cuantitativas con datos agrupados sin intervalos de clase

Una vez agrupados los datos, es posible determinar inmediatamente la moda: basta encontrar el valor de la variable al que le corresponde la mayor frecuencia absoluta.

Ejemplo 3.1.5

Se realizó una encuesta en un barrio carenciado, se seleccionó una muestra y entre las variables en estudio, fue el número de hijos de los hogares encuestados que se detallan a continuación:

N° de hijos	N° de hogares f_i
0	5
1	10
2	25
3	32
4	44
5	31

En la siguiente tabla la frecuencia absoluta mayor es 44, y corresponde al valor 4 de la variable. Luego, la $M_o = 4$ hijos.

Interpretación: Los hogares que más hijos tienen son de 4 hijos.

Actividad 12

Los salarios semanales obtenidos por 50 obreros de una empresa metalúrgica, se observan en el siguiente cuadro:

$X_i (\$)$	f_i
55	2
60	3
65	4
70	5
75	8
80	10
85	8
90	6
95	4

Calcular la moda e interpretarla.

2.5.2.4. Para variables cuantitativas con datos agrupados en intervalos de clase.

La clase que presenta la mayor frecuencia es denominada “clase modal”. Por definición, podemos afirmar que la moda, en este caso, es el valor más frecuente que está comprendido entre los límites de la clase modal.

Para determinar el valor de la moda dentro del intervalo modal, utilizamos la siguiente fórmula:

$$M_o = L_i + \frac{d_1}{d_1 + d_2} \cdot a$$

en la cual:

- * L_i es el límite inferior de la clase modal
- * a es la amplitud de la clase modal
- * $d_1 = f_{1^*} - f_{(i-1)}$
- * $d_2 = f_{1^*} - f_{(i+1)}$
- * f_{1^*} es la frecuencia absoluta de la clase modal
- * $f_{(i-1)}$ es la frecuencia absoluta correspondiente al intervalo anterior al de la clase modal
- * $f_{(i+1)}$ es la frecuencia absoluta correspondiente al intervalo posterior al de la clase modal

Ejemplo 3.1.6

La variable en estudio ahora es la longitud de tornillos en un proceso productivo y se tiene la siguiente tabla de distribución de frecuencias:

Longitud de los tornillos x_i	Número de tornillos f_i
6 – 7	11
7 – 8	9
8 – 9	14
9 – 10	11
10 – 11	22
11 – 12	14
12 – 13	7
13 – 14	5
14 – 15	4
15 – 16	3

En el ejemplo se tiene:

- * $f^* = 22$
- * $L_i = 10$
- * $a = 1$
- * $f_{(i-1)} = 11$
- * $f_{(i+1)} = 14$
- * $d_1 = f_{1^*} - f_{(i-1)} = 22 - 11 = 11$
- * $d_2 = f_{1^*} - f_{(i+1)} = 22 - 14 = 8$

Entonces la moda será:

$$Mo = 10 + \frac{11}{11 + 8} * 1 = 10.58 \text{ mm}$$

Este valor de la variable indica que la longitud de tornillos que produce este proceso con mayor frecuencia se ubica entre los 10 y 11 mm, más precisamente 10.58 mm. Este valor se encuentra comprendido en la clase modal, de mayor frecuencia.

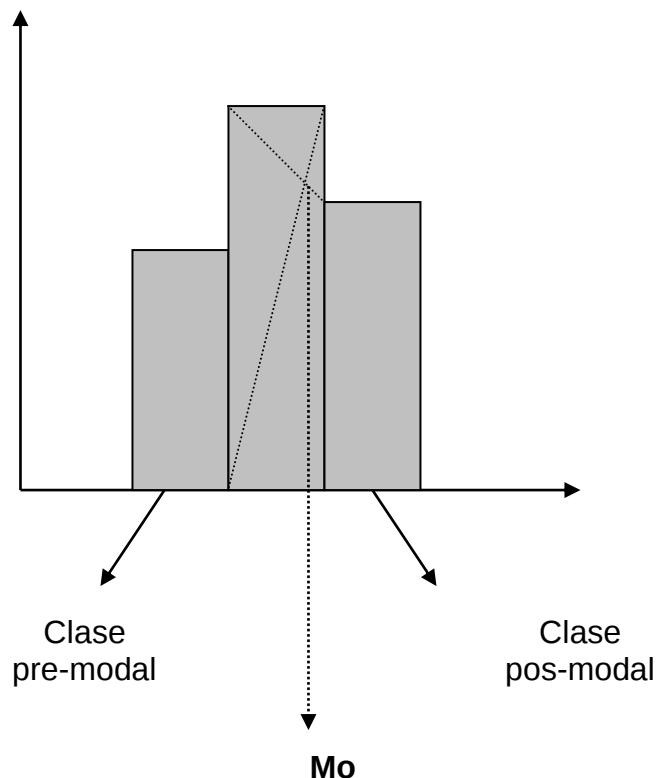
Actividad 13

En un experimento hortícola se trata de determinar el rendimiento en la producción de zapallitos debido al uso de un determinado fertilizante en el suelo. Para ello se sembraron 348 parcelas, y cumplido el ciclo hasta la cosecha, se agruparon los rendimientos de las parcelas en la siguiente tabla de frecuencias:

Rendimiento por parcela (kg de zapallitos)	Nº de parcelas f_i
21.0 - 21.5	1
21.5 - 22.0	2
22.0 - 22.5	1
22.5 - 23.0	5
23.0 - 23.5	38
23.5 - 24.0	55
24.0 - 24.5	102
24.5 - 25.0	61
25.0 - 25.5	49
25.5 - 26.0	25
26.0 - 26.5	8
26.5 - 27.0	1
Total	348

Calcular la moda e interpretarla.

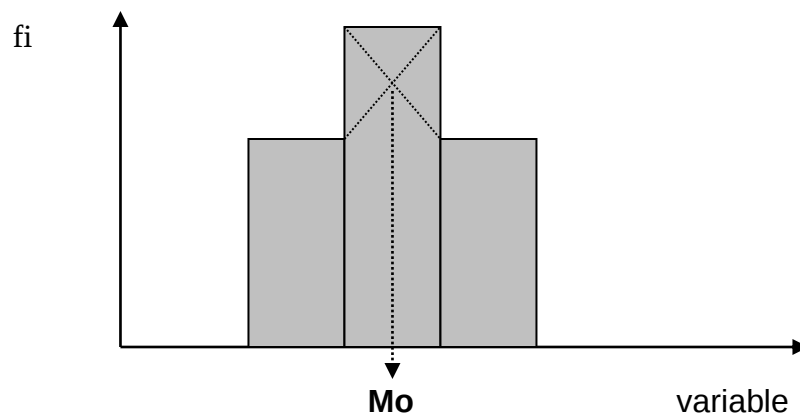
2.5.3 Determinación gráfica de la moda



Para obtener en forma gráfica la moda, se siguen los siguientes pasos:

- ✓ Construir el histograma de frecuencias absolutas.
- ✓ En el histograma el rectángulo de mayor altura corresponde al intervalo modal.
- ✓ En dicho rectángulo se unen los vértices superiores con los vértices superiores de las clases pre-modal y pos-modal. De esta forma se obtienen dos segmentos que se intersectan en un punto.
- ✓ A partir del punto de intersección se traza una perpendicular al eje de abscisas. Dicho punto es la moda.

Observación: En el histograma anterior podemos observar que la moda no coincide con el punto medio de la clase modal. Únicamente coincide con el punto medio o marca de clase modal, cuando la frecuencia absoluta de la clase pre-modal es igual a la pos-modal. Esto lo podemos ver en la siguiente figura:



Es importante destacar que la moda tiene ventajas como su uso tanto para variables cualitativas (con escalas nominales y ordinales), como cuantitativas. No está afectada por los valores extremos en una distribución de datos. Puede calcularse para intervalos de clase abiertos en los dos extremos. La desventaja que presenta es que en un conjunto de datos, la ausencia de la misma o la presencia de dos o más modas, dificulta la interpretación.

Dos conjuntos de datos:

10 15 18 22 24

no hay valor modal (porque no se repite ningún valor)

10 10 15 18 18 22 24

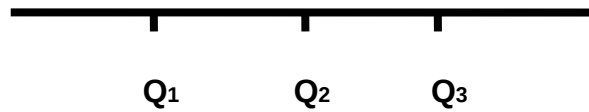
existen dos modas, se denomina esta agrupación “bimodal”

2.6 Los cuartiles

2.6.1. Concepto

Los cuartiles de una distribución, son valores de la variable que dividen al conjunto de datos (ordenados) en cuatro subconjuntos que contienen la misma cantidad de datos. El primer cuartil se simboliza como Q_1 , el segundo cuartil Q_2 y el tercero como Q_3 . Son medidas de posición.

Gráficamente :



Como se observa en la figura, los cuartiles son los valores que dividen en cuatro partes iguales a la serie. Por lo tanto tendremos tres cuartiles :

a - El **primer cuartil (Q_1)** : es el valor de la variable, situado de tal modo en la serie, que una cuarta parte (25%) de los datos es menor o igual que él, y las tres cuartas partes restantes (75%) de los datos son mayores o igual.

b - El **segundo cuartil (Q_2)** : evidentemente, coincide con la mediana. ($Q_2 = M_e$) y su interpretación es la misma.

c - El **tercer cuartil (Q_3)** : es el valor de la variable, situado de tal modo que las tres cuartas partes (75%) de los datos son menores o igual que él, y una cuarta parte (25%) es mayor o igual.

2.6.2. Cálculo de los cuartiles

2.6.2.1. Para datos agrupados sin intervalos de clase

El primer cuartil es el valor de la variable al cual le corresponde la frecuencia acumulada inmediatamente superior a $n / 4$.

El segundo cuartil es el valor de la variable que coincide con la mediana.

El tercer cuartil es el valor de la variable al cual le corresponde la frecuencia acumulada inmediatamente superior a $3n / 4$.

NOTA : En este texto los cuartiles no los calculamos para datos no agrupados por su escasa aplicación.

Ejemplo 3.1.7

Para ejemplificar el cálculo, tenemos la siguiente distribución de frecuencias :

	x_i	f_i	F_i
	1	2	2
Q_1	2	3	5
Q_2	3	4	9
Q_3	4	3	12
	5	3	15

$$\begin{aligned} n/4 &= 3.75 \\ 2n/4 &= 7.5 \\ 3n/4 &= 11.25 \end{aligned}$$

———— Cálculos

Q_1 ——— Primera frec. Acumulada que supera al valor 3.75
 Q_2 ——— Primera frec. Acumulada que supera al valor 7.5
 Q_3 ——— Primera frec. Acumulada que supera al valor 11.25

Por lo tanto tenemos :

$$Q_1 = 2$$

$$Q_2 = 3$$

$$Q_3 = 4$$

Como puede observarse, para determinar los cuartiles, usamos la misma técnica para el cálculo de la mediana.

Actividad 14

En una clase de 50 alumnos, las notas obtenidas forman la siguiente distribución :

NOTAS	N° de alumnos
2	1
3	3
4	6
5	10
6	13
7	8
8	5
9	3
10	1

Calcule el primer y tercer cuartil e interprételos.

2.6.2.2. Para datos agrupados en intervalos de clase

Un primer paso es encontrar el intervalo donde se encuentra el cuartil que queremos calcular. Para ello aplicamos la fórmula :

$$\frac{kn}{4}$$

siendo k el número de orden del cuartil . Una vez obtenido el valor, en la columna de las frecuencias absolutas o relativas acumuladas “ menor que”, buscamos el valor de la frecuencia inmediata superior al valor obtenido. A esa frecuencia le corresponde un intervalo de clase, en él está contenido el cuartil que se busca.

Una vez encontrado el intervalo, para encontrar el valor aproximado o igual, aplicamos las siguientes fórmulas para encontrar Q_1 y Q_3 .

$$Q_1 = L_i + \frac{n/4 - F_{i-1}}{f} * a$$

$$Q_3 = L_i + \frac{3n/4 - F_{i-1}}{f} * a$$

Ejemplo 3.1.8

Utilizando el problema de la Actividad 4, calculemos el primer y tercer cuartil, e interpretemos los resultados :

Estatura (cm)	Nº de Niños (fi)	Frecuencia Acumulada
150-154	4	4
154-158	9	13
158-162	11	24
162-166	8	32
166-170	5	37
170-174	3	40
Total	40	

Primer cuartil

Tenemos :

$$n_4 = 40 / 4 = 10$$

|

$$Q_1 = 154 + \frac{(10 - 4) \cdot 4}{9} = 154 + 2.66 = 156.66$$

$Q_1 = 156.66 \text{ cm.}$

Tercer cuartil

Tenemos :

$$\frac{3 \cdot n_4}{4} = \frac{3 \times 40}{4} = 30$$

|

$$Q_3 = 162 + \frac{(30 - 24) \cdot 4}{8} = 162 + 3 = 165$$

$Q_3 = 165 \text{ cm.}$

Interpretación :

Q_1 : El 25% de los estudiantes tiene una estatura menor o igual a 156.7 cm. , y el resto es mayor o igual.

Q_3 : El 75% de los estudiantes tiene una estatura menor o igual a 165 cm. , y el resto es mayor o igual.

Actividad 15

Los varones que entre los 20 y 50 años contrajeron matrimonio durante el año 1991 en un país, presentan la siguiente distribución por edades :

Edad	Varones (en miles)
20 a 25	46
25 - 30	143
30 - 35	44
35 - 40	13
40 - 45	17
45 - 50	3

Calcular el primer y tercer cuartil. Interpretarlos.

2.7. Los percentiles

2.7.1. Concepto

Denominamos percentiles a los noventa y nueve valores que dividen una serie de datos en 100 partes iguales.

Indicamos :

$$P_1, P_2, \dots, P_{32}, \dots, P_{99}$$

Es evidente que :

$$P_{50} = M_e$$

$$P_{25} = Q_1$$

$$P_{75} = Q_3$$

2.7.2. Cálculo de los percentiles

El cálculo de un percentil sigue la misma técnica del cálculo de la mediana, la fórmula $n / 2$ será substituida por :

$$\frac{k \cdot n}{100}$$

siendo **k** el número de orden del percentil.

Así, para el 27° percentil, tenemos :

$$k = 27$$

$$\frac{27 \cdot n}{100}$$

Una vez realizado el cálculo anterior, buscamos en la columna de las frecuencias acumuladas “menor que”, la frecuencia acumulada inmediata superior a ese valor. Dicha frecuencia corresponde a un determinado intervalo, en él se encuentra contenido el percentil. Lo que resta es determinar el valor aproximado o exacto del percentil 27, para ello aplicamos la siguiente fórmula:

$$P_{27} = L_i + \frac{(27n/100) - F_{i-1}}{f} a$$

Ejemplo 3.1.9

Considerando el problema de la Actividad 4, el cálculo del percentil 8 es el siguiente :

$$k = 8 \rightarrow \frac{8n}{100} = \frac{8 \times 40}{100} = 3.2$$

Luego :

$$P_8 = 150 + \frac{(3.2 - 0) 4}{4} = 150 + 3.2 =$$

$$= 153.2 \text{ cm}$$

donde :

$$P_8 = 153.2 \text{ cm.}$$

Interpretación : El 8% de los estudiantes tiene una estatura inferior o igual a 153.2 cm. y el resto igual o mayor.

Actividad 16

Del problema de la Actividad 4, calcule el percentil 20 e interprételo.

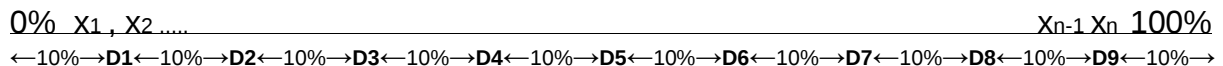
2.8. Los deciles

2.8.1. Concepto

Los valores que dividen a un conjunto de datos ordenados en forma ascendente (o descendente) en diez partes iguales se llaman **Deciles** y se denota por :

$$D_i = 1, 2, \dots, 9$$

Gráficamente se puede observar a los deciles:



Primer decil: es el valor que deja el 10% de las observaciones menores o iguales a él y el 90% superiores o igual a él. Otra forma de decirlo: es el valor que supera a no más de un décimo de las observaciones y es superado por no más de nueve décimos de ella.

Noveno decil: es el valor que deja 90% de las observaciones menores o iguales a él y el 10% superiores o igual a él. Otra forma: es el valor que supera a no más de nueve décimos de las observaciones y es superado por no más de un décimo de ellas.

2.8.2. Cálculo

Los deciles se calculan en forma similar a los cuartiles y los percentiles. Lo único que cambia es el denominador :

$$(i \times n) / 10$$

donde: **(i)** es el orden del decil.

Una vez encontrado el intervalo donde se encuentra contenido el decil buscado, se aplica la siguiente fórmula para determinar el valor aproximado o exacto del decil.

$$D_i = L_i + \frac{(i \cdot n)/10 - F_{i-1}}{f} a$$

Ejemplo 3.2.1

Consideremos el problema de la Actividad 4, calculemos el decil 6:

$$6 \times 40 / 10 = 24$$

Buscamos el intervalo que contiene a este decil, para ello buscamos la frecuencia absoluta acumulada “menor que” inmediata superior a ese valor. Dicha frecuencia es 32, que corresponde al intervalo 162 – 166. Con esto ya tenemos los elementos para calcular el valor aproximado o exacto del decil 6:

$$D_6 = 162 + \frac{(24 - 32)}{8} \cdot 4 = 158 \text{ cm}$$

Interpretación: El 60% de los estudiantes tiene una estatura menor o igual a 158cm, y el 40% restante una estatura mayor o igual.

2.9 Indicadores de control

Actividad 1:

En la guardia de un hospital público, los pacientes con cuadro gripal que concurrieron durante la semana fueron los siguientes:

Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
20	17	15	25	16	12	18

¿Cuántos pacientes en promedio con cuadro gripal concurrieron esa semana a la guardia del hospital?

Solución:

$$\bar{x} = \frac{20 + 17 + 15 + 25 + 16 + 12 + 18}{7} = \frac{123}{7} = 18 \text{ pacientes}$$

En promedio concurren 18 pacientes con cuadro gripal esa semana a la guardia del hospital.

Observación: Debido a que la variable es discreta, redondeamos a 18.

Actividad 2:

La recaudación tributaria (en \$) en la provincia de Jujuy entre los meses de abril y septiembre del año 2002 proporcionada por la Dirección Provincial de Rentas, se detalla a continuación:

ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE
3.784.059	4.316.839	4.228.730	5.089.305	6.082.525	6.210.358

Calcular la media aritmética e interpretarla. Realice los comentarios que le parezcan pertinentes.

Solución:

$$\bar{x} = \frac{3.784.059 + 4.316.839 + 4.228.730 + 5.089.305 + 6.082.525 + 6.210.358}{6} = \frac{29.711.816}{6} = 4.951.969$$

$$\bar{x} = \$ 4.951.969$$

La recaudación tributaria promedio en la provincia de Jujuy, entre los meses de abril y septiembre del año 2002 fue de \$ 4.951.969.

Comentario: Podemos observar en el cuadro que la recaudación tributaria fue creciente y un factor determinante fue el cambio de modelo a nivel nacional (del Plan de Convertibilidad se pasa al Modelo de sustitución de importación, vía devaluación). Desde el año 1998 hasta el 2002 se pasa de un período recesivo a uno de reactivación.

Actividad 3

Complete el esquema para el cálculo de la media aritmética de la siguiente distribución de frecuencias:

X_i	1	2	3	4	5	6
f_i	2	4	6	10	3	1

Tenemos:

X_i	f_i	$X_i f_i$
1	2	2
2	4	8
3	6	18
4	10	40
5	3	15
6	1	6
Total	24 (1)	89 (2)

$$\bar{X} = \frac{(2)}{(1)} = 3.71$$

Actividad 4

Las estaturas de 40 alumnos de una escuela pueden apreciarse en el siguiente cuadro:

Estatura (cm)	N° de Niños (fi)
150-154	4
154-158	9
158-162	18
162-166	8
166-170	6
170-174	3
Total	40

Calcular la media aritmética e interpretarla

Estatura (cm)	N° de Niños (fi)	Marca de clase	$X_i \times f_i$
150-154	4	152	608
154-158	9	156	1404
158-162	18	160	2880
162-166	8	164	1312
166-170	6	168	1008
170-174	3	172	516
Total	48		7728

$$\bar{x} = \frac{7728}{48} = 161 \text{ cm.}$$

La estatura promedio de los alumnos de la escuela es de 161 cm.

Actividad 5: Cuatro acciones de un producto importante, tuvieron un rendimiento anual en porcentaje de 5.61, 6.33, 9.21 y 42.3.

1. Calcule la media geométrica de los rendimientos. Interprete.
2. Calcule la media aritmética de los rendimientos y compare.

Solución:

1.

$$Mg = \sqrt[4]{(5.61) \cdot (6.33) \cdot (9.21) \cdot (42.3)} = \sqrt[4]{13834.6} = 10.85\%$$

El rendimiento promedio de las acciones fue del 10.85%.

2.

$$\bar{x} = \frac{5.61 + 6.33 + 9.21 + 42.3}{4} = 15.86\%$$

Observamos que la media aritmética es mayor dado que está influenciada por los valores extremos.

Actividad 6: Un automovilista parte de su pueblo hacia la capital de su provincia, que se encuentra a una distancia de 60 km. Los primeros 20 km los recorre a una velocidad de 110 km /h y el resto a 140 km / h. Calcule la velocidad media del automovilista.

Solución:

$$M_h = \frac{60}{\left(\frac{20}{110}\right) + \left(\frac{40}{140}\right)} = \frac{60}{0.18 + 0.29} = \frac{60}{0.47} = 127.66 \text{ km/h}$$

La velocidad media o promedio del automovilista es de 127.66 km/h.

Actividad 7:

Calcular la mediana para las dos series de datos:

a) 5, 13, 10, 2, 18, 15, 6, 16, 9

b) 2, 6, 7, 10, 12, 13, 18, 21

Solución:

a)

2, 5, 6, 9, 10, 13, 15, 16, 18

Me = 10

b)

2, 6, 7, 10, 12, 13, 18, 21

$10 + 12 / 2 = 22 / 2 = 11$

Me = 11

Actividad 8

Dada la siguiente distribución de niños de 0 a 4 años seleccionados a través de una muestra en un barrio , calcule la mediana e interprete.

X_i	f_i
0	2
1	6
2	10
3	12
4	5

Solución:

X_i	f_i	Fi
0	2	2
1	6	8
2	10	18
3	12	30
4	5	35
Total	35	

$$N / 2 = 35/2 = 17.5$$

Buscamos la frecuencia acumulada inmediata superior a ese valor, y observamos que es 18. A esa frecuencia le corresponde el valor de la variable "2". Por lo tanto la mediana es 2 años.

Me = 2 años

Interpretación: La mitad de los niños tienen 2 años o menos, y la otra mitad 2 años o más.

Actividad 9

Se realizó un test de inteligencia a 100 alumnos seleccionados aleatoriamente en un colegio, los resultados se presentan en el siguiente cuadro:

<i>Puntaje</i>	<i>Nº de alumnos</i>
0 – 10	5
10 – 20	6
20 – 30	4
30 – 40	15
40 – 50	30
50 – 60	20
60 – 70	8
70 – 80	5
80 – 90	5
90 – 100	2
Total	100

Calcular la mediana e interpretar el resultado.

Solución:

<i>Puntaje</i>	<i>Nº de alumnos</i>	<i>Fi</i>
0 – 10	5	5
10 – 20	6	11
20 – 30	4	15
30 – 40	15	30
40 – 50	30	60
50 – 60	20	80
60 – 70	8	88
70 – 80	5	93
80 – 90	5	98
90 – 100	2	100
Total	100	

$$N / 2 = 50$$

La frecuencia acumulada inmediata superior a ese valor es 60, y corresponde al intervalo de clase 40 – 50. En ese intervalo está contenida la mediana. Ahora determinamos el valor de ella a través de la fórmula:

$$Me = 40 + \frac{50 - 30}{30} - 10 = 46.66 \cong 47 \text{ puntos}$$

Me = 47 puntos

Interpretación: El 50% de los alumnos obtuvo un puntaje en el test de 47 puntos o menos y el otro 50% , 47 puntos o más.

Actividad 10:

En una escuela se realizó una prueba de matemáticas para observar el nivel en esa materia. Para ello se seleccionó aleatoriamente una muestra de 40 alumnos, y dio los siguientes resultados:

Excelente (E) , Bueno (B), Regular (R) y Malo (M)

E	R	M	M	R	B	E	M	R	E
M	M	B	B	E	R	M	M	B	B
E	B	M	M	M	B	R	R	M	B
B	B	M	E	M	M	M	R	E	R

- a) ¿Cuál es la variable en estudio y su escala de medición?
- b) Construir una tabla de distribución de frecuencias con la información brindada.
- c) Determinar la moda.

Solución:

- a) Variable: Nivel o rendimiento en Matemáticas de los alumnos de una escuela.
Escala ordinal.
- b)

Nivel de rendimiento de los alumnos en Matemáticas	Nº de alumnos f_i	r_i	F_i
E	7	0.18	7
B	10	0.25	17
R	8	0.20	25
M	15	0.37	40
Total	40	1	

- c) Calcularemos la moda:

Moda: Malo

Interpretación: Los alumnos que tienen un nivel malo en Matemáticas son los más frecuentes.

Actividad 11

Calcule la moda en la siguiente serie de datos:

3, 3, 6, 7, 8, 8, 8, 13, 16, 19

Solución:

Mo = 8

El valor que más se repite de la serie de datos es 8.

Actividad 12

Los salarios semanales obtenidos por 50 obreros de una empresa metalúrgica, se observan en el siguiente cuadro:

X_i (\$)	f_i
55	2
60	3
65	4
70	5
75	8
80	10
85	8
90	6
95	4

Calcular la moda e interpretarla.

Solución:

Mo = \$ 80.-

Interpretación: La mayoría de los obreros cobran un salario semanal de \$ 80.-

Actividad 13:

En un experimento hortícola se trata de determinar el rendimiento en la producción de zapallitos debido al uso de un determinado fertilizante en el suelo. Para ello se sembraron 350 parcelas, y cumplido el ciclo hasta la cosecha, se agruparon los rendimientos de las parcelas en la siguiente tabla de frecuencias:

Rendimiento por parcela (kg de zapallitos)	Nº de parcelas f_i
21.0 - 21.5	1
21.5 - 22.0	2
22.0 - 22.5	1
22.5 - 23.0	5
23.0 - 23.5	38
23.5 - 24.0	55
24.0 - 24.5	102
24.5 - 25.0	61
25.0 - 25.5	49
25.5 - 26.0	25
26.0 - 26.5	8
26.5 - 27.0	1
Total	348

Calcular la moda e interpretarla.

Solución:

El intervalo que tiene la mayor frecuencia es: 24.0 – 24.5. En este intervalo está contenida la moda, ahora obtendremos el valor de la moda a través de la fórmula correspondiente:

En el ejemplo se tiene:

* $f^* = 102$

* $L_i = 24$

- * $a = 0.5$
- * $f_{(l-1)} = 55$
- * $f_{(l+1)} = 61$
- * $d_1 = f_{1^*} - f_{(l-1)} = 102 - 55 = 47$
- * $d_2 = f_{1^*} - f_{(l+1)} = 102 - 61 = 41$

Entonces la moda será:

$$Mo = 24 + \frac{47}{47 + 41} * 0.5 = 24.28 \text{ mm}$$

Interpretación: El rendimiento por parcela que más se repite es de 24.28 mm.

Actividad 14:

En una clase de 50 alumnos, las notas obtenidas forman la siguiente distribución :

NOTAS	N° de alumnos
2	1
3	3
4	6
5	10
6	13
7	8
8	5
9	3
10	1

Calcule el primer y tercer cuartil e interprételos.

Solución:

NOTAS	N° de alumnos	Fi
2	1	1
3	3	4
4	6	10
5	10	20
6	13	33
7	8	41
8	5	46
9	3	49
10	1	50
Total	50	
Primer cuartil		

Tenemos :

$$\frac{n}{4} = \frac{50}{4} = 12.5$$

Tercer cuartil

Tenemos :

$$\frac{3n}{4} = \frac{3 \times 50}{4} = 37.5$$

$Q_1 = 5$ puntos.

$Q_3 = 7$ puntos.

Interpretación del primer cuartil: El 25% de los alumnos obtuvieron una nota menor o igual a 5 puntos, y el 75% de los alumnos una nota mayor o igual a 5 puntos.

Interpretación del tercer cuartil: El 75% de los alumnos obtuvieron una nota menor o igual a 7 puntos, y el 25% de los alumnos una nota mayor o igual a 7 puntos.

Actividad 15:

Los varones que entre los 20 y 50 años contrajeron matrimonio durante el año 1991 en un país, presentan la siguiente distribución por edades :

Edad	Varones (en miles)
20 a 25	46
25 - 30	143
30 - 35	44
35 - 40	13
40 - 45	17
45 - 50	3

Calcular el primer y tercer cuartil. Interpretarlos

Solución:

Calculamos las frecuencias acumulada “menor que” :

Fi
46
189
233
246
263
266

Primer cuartil

Tenemos :

$$\frac{n}{4} = \frac{266}{4} = 66.5$$

$$Q_1 = 25 + \frac{(66.5 - 46)}{143} \times 5 =$$

$$\underline{Q_1 = 26 \text{ años.}}$$

Tercer cuartil

Tenemos :

$$\frac{3n}{4} = \frac{3 \times 266}{4} = 199.5$$

$$Q_3 = 30 + \frac{(199.5 - 189)}{44} \times 5 =$$

$$\underline{Q_3 = 31 \text{ años.}}$$

Interpretación :

Q₁ : El 25% de los varones que contrajeron matrimonio tienen una edad menor o igual a 26 años. , y el resto es mayor o igual.

Q₃ : El 75% de los varones que contrajeron matrimonio tienen una edad menor o igual a 31 años , y el resto es mayor o igual.

Actividad 16:

Del problema de la Actividad 4, calcule el percentil 20 e interprételo.

Estatura (cm)	Nº de Niños (fi)	Fi
150-154	4	4
154-158	9	13
158-162	18	31
162-166	8	39
166-170	6	45
170-174	3	48
Total	48	

Solución:

Calcularemos el percentil 20, para ello determinaremos en qué intervalo se encuentra contenido:

$$k = 20 \longrightarrow \frac{20 \cdot n}{100} = \frac{20 \times 48}{100} = 9.6$$

El intervalo indicado es 154 - 158

Luego :

$$P_{20} = 154 + \frac{(9.6 - 4) \times 4}{9} = 156.49 \text{ cm}$$

$$= 156.49 \text{ cm}$$

donde :

$$P_{20} = 156.49 \text{ cm.}$$

Interpretación: El 20% de los niños tienen una estatura de 156.49 cm o menos, y el resto una estatura mayor o igual.

4. Problemas propuestos



PRACTICA 3

Problema 1

El número de accidentes ocurridos con 70 choferes de una empresa de taxis durante 2 años es el siguiente:

N° de accidentes	N° de choferes
0	20
1	10
2	16
3	9
4	6
5	5
6	2
7	2

- Definir la variable y clasificarla.
- Calcular la media aritmética, la mediana y la moda. Interpretarlos.
- Calcular el primer y tercer cuartil. Interpretélos.

Problema 2

Considerando las notas de un test de inteligencia aplicado a 100 alumnos:

64	78	66	88	74	103	78	86	103	87
73	95	82	89	73	92	85	80	81	90
78	86	78	101	85	98	75	73	90	86
86	84	86	76	76	83	103	86	84	85
76	80	92	102	73	87	70	85	79	93
82	90	83	81	85	72	81	96	81	85
68	96	86	70	72	74	84	99	81	89
71	73	63	105	74	98	78	78	83	96
95	94	88	62	91	83	98	93	83	76
94	75	67	95	108	98	71	92	72	101

- Forme una distribución de frecuencias con intervalos de clase, de una amplitud igual a 5.
- Calcule las frecuencias absolutas y relativas. Grafique.
- Calcule y grafique las frecuencias acumuladas de las formas “mayor y menor que”.
- Calcule las principales medidas de posición . Interpretélas.

Problema 3

La siguiente distribución de frecuencias corresponde a los pesos (en kg) de 50 terneros de una explotación pecuaria :

Peso (kg)	N° de terneros
145-151	10
151-157	9
157-163	8
163-169	6
169-175	13
175-181	3
181-187	1

Total	50

- a) Definir la variable y clasificarla.
- b) Construir el histograma de frecuencias absolutas.
- c) Calcular la media aritmética e interpretarla.
- d) Calcular la mediana e interpretarla.
- e) Calcular la moda e interpretar el resultado.

Problema 4

Se entrevista a 40 trabajadores de una empresa y se les pregunta su salario. En base a sus respuestas se obtiene la siguiente información:

Salarios (\$)	N° de Trabajadores
7.5 - 12.5	2
12.5 - 17.5	5
17.5 - 22.5	19
22.5 - 27.5	6
27.5 - 32.5	3
32.5 - 37.5	5
Total	40

- a) Definir la variable y clasificarla. Indique la escala de medición.
- b) Construir el histograma de frecuencias absolutas.
- c) Calcular la media aritmética e interpretarla.
- d) Calcule la mediana, la moda y los cuartiles e interpréte los.

Problema 5

Los siguientes datos corresponden a las mediciones de los 50 terremotos más recientes en la escala de Richter.

Escala	N° de Terremotos
2.25 - 2.75	4
2.75 - 3.25	2
3.25 - 3.75	5
3.75 - 4.25	8
4.25 - 4.75	12
4.75 - 5.25	8
5.25 - 5.75	5
5.75 - 6.25	3
6.25 - 6.75	2
6.75 - 7.25	1
Total	50

- Definir la variable y clasificarla. Indicar la escala de medición.
- Calcular la media aritmética, la mediana y la moda. Interpretarlos.
- Calcular el percentil 30, el decil 4 e interpretarlos.
- Calcular el primer y tercer cuartil. Interpretélos.

Problema 6

Se aplicó una prueba de inteligencia a 50 niños que de una clase de preescolar, obteniéndose los siguientes resultados:

Resultados	N° de Niños
67.5 - 72.5	2
72.5 - 77.5	5
77.5 - 82.5	8
82.5 - 87.5	15
87.5 - 92.5	10
92.5 - 97.5	7
97.5 - 102.5	3
Total	50

- Definir la variable y clasificarla. Indicar la escala de medición.
- Calcular la media aritmética, la mediana y la moda. Interpretarlos.
- Calcular el percentil 40, el decil 6 e interpretarlos.
- Calcular el primer y tercer cuartil. Interpretélos.

Problema 7

Se observó el estado civil de los empleados de una determinada empresa. Los resultados obtenidos se muestran en la tabla siguiente.

Estado civil	Número de
--------------	-----------

	empleados
Soltero	35
Casado.	19
Viudo.	14
Divorciado	18

- a- Reconozca y clasifique la variable en estudio.
b- ¿Qué medidas de posición propone para ella?

Problema 8

Los siguientes datos corresponden a los costos unitarios de producción para la fabricación de un mismo artículo de 65 firmas diferentes:

Costos Unitarios (\$)	Nº de Firmas
450-550	8
550-650	10
650-750	11
750-850	16
850-950	13
950-1050	5
1050-1150	2
Total	65

- Definir la variable y clasificarla. Indicar la escala de medición.
- Calcular la media aritmética, la mediana y la moda. Interpretarlos.
- Calcular el percentil 40, el decil 6 e interpretarlos.
- Calcular el primer y tercer cuartil. Interpretélos.

Problema 9

Si la población de San Salvador de Jujuy fue de 124.487 en 1980 y de 179.887 en 1991 (según los Censos Nacionales respectivos):

- ¿Cuál fue el promedio de crecimiento anual?
- ¿Cuál es la población estimada para 1997, suponiendo que se mantenga esta tasa de crecimiento?

Problema 10

Se aplicó una prueba de conocimientos a un grupo de personas (30 en total), obteniéndose las siguientes puntuaciones:

39 53 44 45 32 48 46 23 36 43 60 44 40 26 44
21 40 64 45 41 43 59 50 34 37 42 47 46 36 31

- Define y clasifique la variable.
- Construya una tabla de distribución de frecuencias listando los valores y otra agrupando los datos en 9 intervalos de amplitud 5.
- Calcule todas las medidas de posición que la caracterizan.

d- ¿Qué valor de la variable le corresponde a cada cuartil?

e- ¿Qué proporción de personas tiene una puntuación que está por debajo del primer cuartil?