Probabilidad y estadística

Docente a Cargo:
Ing. Pablo Vega
Mail: pablorvega@gmail.com

Material extracto del apunte teórico del Ing. Fugiglando

UNIDADES TEMATICAS

1. Metodología estadística
2. Medias de posición y dispersión
3. Algebra de probabilidades
4. Variable aleatoria
5. Modelos especiales de probabilidad
6. Elementos de muestreo
7. Teoría de la estimación estadística
8. Contraste, prueba o docimasia de hipótesis

UNIDADES TEMATICAS

1. Metodología estadística
2. Medias de posición y dispersión
3. Algebra de probabilidades
4. Variable aleatoria
5. Modelos especiales de probabilidad
6. Elementos de muestreo
7. Teoría de la estimación estadística
8. Contraste, prueba o docimasia de hipótesis

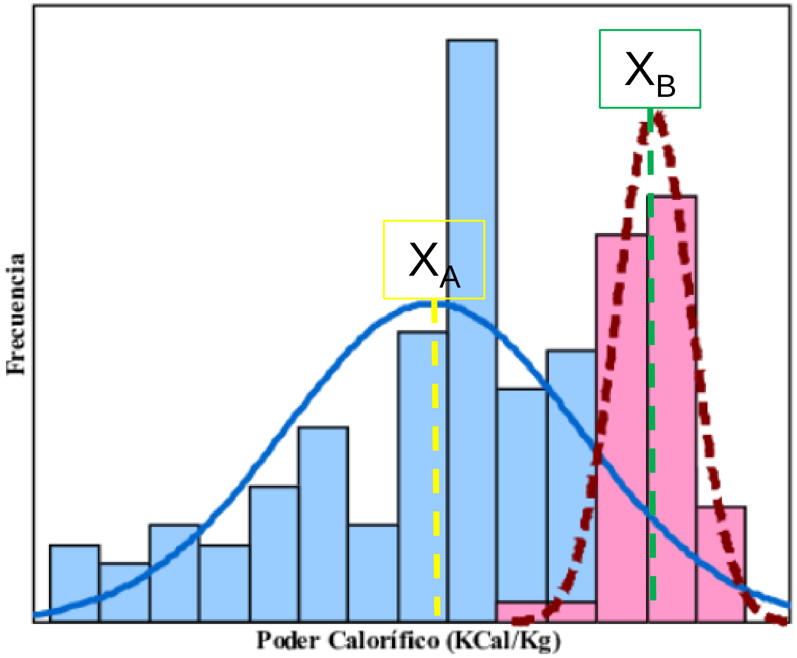
PARAMETROS Y ESTADISTICOS

Principales **medidas estadísticas descriptivas** de resumen de un conjunto de datos, dadas por:

1. Medias de posición
2. Medidas de dispersión
3. Medidas de forma, asimetría o sesgo
4. Mediadas de puntiagudez o curtosis

- **Medidas de posición**

Son medidas descriptivas de resumen, representativas del valor de la variable en torno a la cual se concentraran las observaciones. Así por ejemplo, si se tiene a las distribuciones A y B.



Se observa que ambas tienen la misma forma, concentrándose los datos en torno al valor de la variable X_A en la distribución A, y en torno a X_B en la distribución B, siendo X_A y X_B valores de una **medida de posición.**

MEDIA ARITMETICA

a) **Serie simple:** dada la serie simple correspondiente a una muestra de **n** observaciones de una variable: $x_1, x_2, x_3, \ldots\ldots\ldots, x_n$, la **Media Aritmética o promedio MUESTRAL**, simbolizada con M(x) o $\overline{x}$, surge de la suma de los valores de la variable observada, dividido por el numero de los mismos, dando lugar a la denominada **«Formula de la media simple»,** es decir:

$$M(x) = \overline{X} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Media simple Muestral

En el caso de que la serie simple se corresponda con todos los valores de la variable de la población bajo estudio (N observaciones): $x_1, x_2, x_3, \ldots\ldots\ldots, x_N$, la **Media Aritmética o promedio POBLACIONAL,** denotada con µ va a estar dada por:

$$M(x) = \mu = \frac{x_1 + x_2 + x_3 + \cdots + xn}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Media simple Poblacional

MEDIA ARITMETICA

b) **Media Ponderada:** surge de la suma de los productos entre cada valor de la variable o marca de clase (según sea distribución de frecuencia de variable discreta o de variable continua, respectivamente) y las respectivas frecuencias absolutas, dando lugar a la denominada **«Formula de la media ponderada»**, pudiendo ser **muestral** o **poblacional** según se considere los datos de una muestra o de toda la población, respectivamente, es decir:

$$M(y) = \overline{y} = \frac{y_1 n_1 + y_2 n_2 + y_3 n_3 + \cdots + y_n n_n}{n} = \frac{\sum_{i=1}^{n} y_i n_i}{n}$$

Media ponderada Muestral

$$M(y) = \mu = \frac{y_1 n_1 + y_2 n_2 + y_3 n_3 + \cdots + y_n n_n}{N} = \frac{\sum_{i=1}^{N} y_i n_i}{N}$$

Media simple Poblacional

y_i : valor de la variable (DISCRETA) o marca de clase (CONTINUA)

Ejemplo 1: el numero de empleados que faltaron por día en una empresa en los últimos 20 días laborales, esta dado por:

4, 5, 2, 2, 1, 1, 1, 6, 2, 2, 2 ,2, 2, 5, 4, 4, 1, 1, 1, 2

Calcular el promedio diario de empleados que faltaron considerando:

1) Serie Simple
2) Distribución de frecuencia de variable discreta.

Ejemplo 2: la antigüedad de sus empleados (en años) de los empleados de un hipermercado esta dado por: Cual es la antigüedad promedio?

y_{i-1}	y_i	N° de empleados
3	5	50
5	7	70
7	9	90
9	11	40
11	13	30
13	15	20
	Σ	300

Propiedades de la media aritmética

1. «La media aritmética de una constante, es la misma constante»

$$\boxed{M(k) = k;\ k \rightarrow \text{constante}}$$

2. «La media aritmética del producto entre una constante y una variable es igual al producto de la constante por la media aritmética de la variable»

$$\boxed{M(k*x) = k*M(x);\ k \rightarrow \text{constante}}$$

3. «La media aritmética de la suma de una constante y una variable es igual a la suma de la constante y la media aritmética de la variable»

$$\boxed{M(k+x) = k+M(x);\ k \rightarrow \text{constante}}$$

4. La media aritmética de una suma de variables **expresadas en igual unidad de medida** es igual a la suma de las medias aritméticas de cada una de las variables consideradas»

$$\boxed{M(x+y) = M(x)+M(y);\ x,y \rightarrow \text{variables}}$$

Propiedades de la media aritmética

1. «La suma de los desvíos o diferencia entre cada valor de la variable y su media aritmética, ponderadas en el caso de datos agrupados, <u>es igual a cero</u>»

a) Serie simple → $\Sigma(x_i - \overline{x}) = 0$

a) Datos agrupados → $\Sigma(y_i - \overline{y}).n_i = 0$

2. «La suma de los desvíos o diferencia entre cada valor de la variable y su media aritmética elevadas al cuadrado y ponderadas en el caso de datos agrupados, <u>es un mínimo</u>»

a) Serie simple → $\Sigma(x_i - \overline{x})^2 < \Sigma(x_i - k)^2$; $k \neq \overline{x}$

a) Datos agrupados → $\Sigma(y_i - \overline{y})^2.n_i < \Sigma(y_i - k)^2.n_i$; $k \neq \overline{y}$

c) Media general o total

Dado un conjunto de **n** observaciones, las que se subdividen en **k** grupos distintos y se calcula en cada uno de ellos la media aritmética, la **«media total o general»** se obtiene de la suma de los productos entre la media de cada grupo y su respectivo tamaño, dividido todo por el total de observaciones.

Grupo	Media aritmética	Tamaño
1	y_1	n_1
2	y_2	n_2
3	y_3	n_3
.	.	.
k	y_k	n_k

$$M(y) = \overline{y} = \frac{\sum_{i=1}^{n} y_i n_{ii}}{n}$$

Media general o total

MEDIANA

La media aritmética como medida de posición resulta útil *siempre* que los valores de la variable observada sean homogéneos. En caso contrario, resulta mas adecuada como medida de posición la **Mediana**, dada por **el valor de la variable que supera a no mas de la mitad de las observaciones y es superada por no mas de la mitad de las mismas, <u>siendo menos sensible que la media aritmética</u> ante la presencia de valores extremos.** Es decir, la **mediana** es el valor de la variable ubicado en la mitad de un conjunto de valores ordenados de la variable, o sea el *valor central*. Para que un valor de la mediana tenga sentido, el nivel de medida de los datos debe ser por lo menos ordinal. Se lo denota con M_e.

Calculo de la mediana

A. **Serie simple**

i. Se ordenan los datos en sentido creciente o decreciente.

ii. Se determina la mediana que corresponde con el valor de la variable ubicado en el valor central. Es decir:

$$M_e = x_{.5} = X_{\frac{n+1}{2}}$$

Observaciones:

- Si el numero de datos **n** es **par**, el valor de la Mediana se obtiene de la media aritmética de los valores centrales de las variables, siendo su valor coincidente con uno observado de la variable, solo si son iguales los valores centrales.
- Si el numero de datos **n** es **impar**, el valor de la mediana será igual a un valor realmente observado de la variable.

Ejemplo 1: las temperaturas observadas en un cierto lugar en grados centígrados (°C), están dadas por:

-10, 2, 1, 0, -1, -3, -4, 1, -1, 0, 3

Calcular: temperatura media y temperatura mediana

Ejemplo 2: las temperaturas observadas en un cierto lugar en grados centígrados (°C), están dadas por:

-10, 2, 1, 0, -1, -3, -4, 1, -1, 0, 3, 45

Calcular: temperatura media y temperatura mediana

B. Datos agrupados de variable discreta:

En el caso de una distribución de frecuencias o datos agrupados de variable discreta, el valor de la mediana surge de la siguiente manera:

i. Calcular la mitad de las observaciones, es decir: **n/2**, siendo **n** el total de observaciones.

ii. Determinar N_j que es la **menor de las frecuencias absolutas acumuladas que supera a la mitad de las observaciones (n/2)**

iii. Relacionar N_{j-1} que es la **frecuencia absoluta acumulada inmediata anterior a N_j**, con **n/2**, pudiéndose presentar dos situaciones.

a) $N_{j-1} = n/2 \rightarrow M_e = \frac{y_{i-1} + y_i}{2}$ →Promedio aritmético de los valores de las variable relacionadas con N_{j-1} y N_j respectivamente

b) $N_{j-1} < n/2 \rightarrow M_e = y_j$ → y_j es el valor de la variable asociada a N_j

Ejemplo 1: calcular la Mediana de la siguiente distribución de frecuencia de variable discreta:

y_i	n_i	N_i
1	9	9
3	12	**21**
7	15	**36**
9	10	46
20	5	51
100	3	54
Σ	54	

$M_e = y_j$ → 7

N_{j-1} → 21

N_j → 36

Solución: procediendo según a lo antes expuesto

1. $n/2 = ? \rightarrow n/2 = 54/2 =$ **27**

2. $N_j = ? \rightarrow N_j =$ **36**

3. $N_{j-1} = 21 \rightarrow N_{j-1} < n/2 \rightarrow M_e = y_j \rightarrow M_e = 7$

Ejemplo 2: calcular la Mediana de la siguiente distribución de frecuencia de variable discreta:

	y_i	n_i	N_i	
	1	8	8	
	3	12	20	
y_{j-1}	**7**	7	**27**	N_{j-1}
y_j	**9**	13	**40**	N_j
	20	8	48	
	100	6	54	
	Σ	54		

Solución: procediendo en forma análoga al problema anterior.

1. $n/2 = ? \rightarrow n/2 = 54/2 = \mathbf{27}$

2. $N_j = ? \rightarrow N_j = 40$

3. $N_{j-1} = 27 \rightarrow N_{j-1} = n/2 \rightarrow M_e = \frac{y_{j-1}+y_j}{n} \rightarrow M_e = \frac{7+9}{2} = 8$

C. Datos agrupados de variable continua:

En el caso de una distribución de frecuencias o datos agrupados de variable continua, el valor de la mediana surge de la siguiente manera:

i. Calcular la mitad de las observaciones, es decir: **n/2**, siendo **n** el total de observaciones.

ii. Determinar N_j que es la **menor de las frecuencias absolutas acumuladas que supera a la mitad de las observaciones (n/2)**

iii. Relacionar N_{j-1} que es la **frecuencia absoluta acumulada inmediata anterior a** $\mathbf{N_j}$, con **n/2**.

Se demuestra que la **formula de calculo de la mediana** en una distribución de frecuencia de variable continua, esta dada por:

$$M_e = y_{j-1} + c_j \frac{\frac{n}{2} - N_{j-1}}{n_j}$$

Siendo:

y_{j-1} → extremo izquierdo del intervalo correspondiente a **N_j**, denominado **intervalo mediano.**

c_j → amplitud del intervalo mediano.

$n/2$ → mitad de las observaciones.

N_{j-1} → frecuencia absoluta acumulada inmediata anterior a N_j.

n_j → frecuencia absoluta correspondiente al intervalo mediano.

Observación: si → $N_{j-1} = n/2$ → $Me = y_{j-1}$ → extremo izquierdo del intervalo de clase mediano.

Ejemplo 1: calcular la Mediana de la siguiente distribución de frecuencia de variable continua:

yi_{-1}	y_i	n_i	N_i
4	8	5	5
8	12	9	**14**
12	16	20	**34**
16	20	6	40
20	24	10	50
	Σ	50	

y_{j-1}

N_{j-1}

N_j

Solución: procediendo en forma análoga al problema anterior.

1. n/2 = ? → n/2 = 50/2 = **25**

2. N_j = ? → N_j = 34 → N_{j-1} = **14**

3. Relacionar $\mathbf{N_{j-1}}$ con **n/2** → 14<25 → $M_e = y_{j-1} + c_j \frac{\frac{n}{2} - Nj\_{1}}{n_j}$

4. $Me = 12 + 4\,\frac{25-14}{20} = 14{,}2$

Ejemplo 2: calcular la Mediana de la siguiente distribución de frecuencia de variable continua:

yi_{-1}	y_i	n_i	N_i
4	8	5	5
8	12	9	14
12	16	11	**25**
16	20	16	**41**
20	24	9	50
	Σ	50	

$M_e = y_{j-1}$ → **16**

25 ← N_{j-1}

41 ← N_j

Solución: procediendo en forma análoga al problema anterior.

1. $n/2 = ? \rightarrow n/2 = 50/2 = \mathbf{25}$

2. $N_j = ? \rightarrow N_j = 41 \rightarrow N_{j-1} = \mathbf{25}$

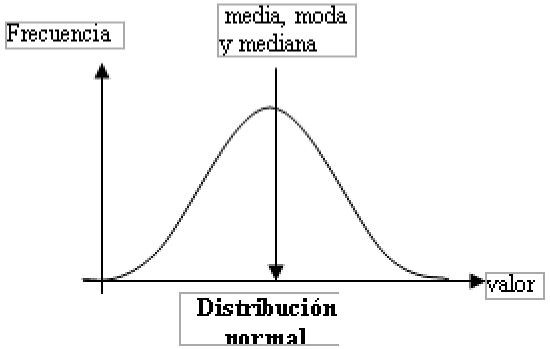
3. Relacionar $\mathbf{N_{j-1}}$ con $\mathbf{n/2}$ → $25=25 \rightarrow M_e= y_{j-1} \rightarrow M_e= 16$

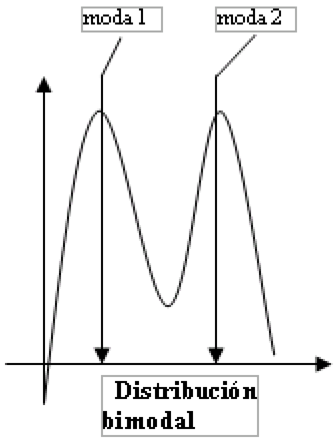
MODO, MODA, O VALOR MODAL

Es otra medida de posición definida **por el valor de la variable con mayor frecuencia de presentación.** También se dice que es el **valor típico** de un conjunto de datos y resulta aplicable tanto a variable cuantitativa como a variable cualitativa.

Si de un conjunto de datos un valor de la variable tiene mayor frecuencia que los demás, se dice que la distribución es **unimodal**, mientras que si son dos los valores que tienen mayor frecuencia, se dice que es **bimodal.** Se denota a la moda M_o.

Si existen 2 valores que son dominantes en relación a los otros, incluso sin un empate exacto para la moda, el conjunto es bimodal.





Calculo de la moda

a) **Serie Simple:** el **Modo**, si existe, será el valor de la variable que se presenta con mas frecuencia en la serie de datos.

Ejemplos: calcular en cada caso, si existe, el **Modo:**

1. El N° de artículos fallados por día en una línea esta dado por: 6, 8, 10, 8, 8, 6, 5, 3

2. El N° de empleados que faltaron a su trabajo, por dia, esta dado por: 3, 1, 1, 1, 2, 2, 2, 2, 1, 4, 5, 3

3. El estado civil de un conjunto de empleados es: casado, soltero, casado, casado, soltero, soltero, soltero

4. El medio de movilidad escogido por un grupo de empleados para ir a su lugar de trabajo: moto, auto, colectivo, caminando, moto, colectivo, caminando, auto

5. El color de los autos vendidos en una semana en una concesionaria es: blanco, rojo, gris, azul, negro

Calculo de la moda

b) **Datos agrupados de variable discreta:** en este caso el **Modo** esta dado por $\mathbf{y_j}$ que es la variable cuantitativa o cualitativa asociada a $\mathbf{n_j}$, que es el **mayor** de las frecuencias absolutas **que supera simultáneamente** a la inmediata anterior y posterior. Es decir:

$$\text{Modo} = \mathbf{y_j} \rightarrow \mathbf{n_j} \text{ siendo: } \mathbf{n_{j-1} < n_j > n_{j+1}}$$

En forma análoga a una serie simple, si existe mas de un valor de la variable que satisface la condicion anterior, la distribución podrá ser *bimodal, trimodal*, etc.

Ejemplos: calcular en cada caso, si existe, el **Modo: Gaseosa (120), otros 1 (270), Fernet (270), otros 2 (310), otros 3 (250), cerveza (390)**

Calculo de la moda

c) **Datos agrupados de variable continua:** para la determinación del modo en una distribución de frecuencias de variable continua existen distintos métodos, entre los que se encuentran:

1) El que utiliza la marca de clase del intervalo de clase modal (que es aquel con mayor valor de la frecuencia absoluta $\mathbf{n_j}$).
2) El de interpolación mediante formula, que establece que el modo surge de la aplicación de la formula:

$$M_o = y_{j-1} + c_j \frac{d_1}{d_2 + d_1}$$

Siendo:

d_1: $n_j - n_{j-1}$

d_2: $n_j - n_{j+1}$

n_j: mayor de las frecuencias absolutas que supera simultáneamente a la inmediata anterior y superior.

y_{j-1}: extremo izquierdo del intervalo de clase modal (donde se encuentra n_j)

c_j: amplitud del intervalo de clase modal.

n_{j-1}: frecuencia absoluta inmediata anterior a $\mathbf{n_j}$

N_{j+1}: frecuencia absoluta inmediata posterior a $\mathbf{n_j}$

Ejemplo 1: calcular la Mediana de la siguiente distribución de frecuencia de variable continua:

yi_{-1}	y_i	n_i
4	8	5
8	12	9
12	16	20
16	20	6
20	24	4

n_{j-1}

n_j

N_{j+1}

Solución: procediendo en forma análoga al problema anterior.

1. Cual es el intervalo de clase modal? Como se lee a dicho intervalo?
2. Calcular el **Modo** de la distribución aplicando la «marca de clase» del intervalo de clase modal.
3. Calcular el **Modo** de la distribución aplicando interpolación mediante formula.

PERCENTILES O CENTILES

Son medidas de posición o tendencia central que dividen a un grupo de datos en 100 partes. Hay 99 **Percentiles,** porque se necesita 99 divisores para separar un grupo de datos en 100 partes. El **n - esimo Percentil** es el valor de la variable tal que **«al menos n por ciento de los datos están bajo de ese valor y a lo sumo (100 – n) por ciento superan a ese valor».** Así por ejemplo, el percentil 67 es un valor de la variable tal que al menos el 67 % de los datos están por debajo de ese valor y no mas del 33% están sobre ese valor.

Determinación de la ubicación de un percentil

1) Ordenar los valores de la variable en forma creciente.

2) Calcular la ubicación del **percentil (i)** de la forma:

$$i = \frac{P}{100} n$$

Siendo:

i: ubicación del Percentil

P: el percentil de interés

n: numero de datos

3) Determinar la ubicación de **a)** o de **b)**

a) Si **i=a es un numero entero**, el *P-esimo* Percentil es el promedio de los valores de las variables localizadas entre la i-esima e (i+1)-esima ubicación. Por ejemplo, si **i=16**→ el percentil buscado será el **promedio** entre los valores de las variables ubicadas en los lugares **16 y 17.**

b) Si **i≠a un numero entero**, el valor del P-esimo percentil esta dado por el valor de la variable ubicado en la parte entera del numero resultante de sumarle **1** a **i**, es decir, lugar **(i + 1).** Por ejemplo, si **i=8,6,** el percentil buscado será el valor de la variable ubicado en el **lugar 9**, (parte entera del numero 8,6+1=9,6)

Ejemplo 1: calcular el Percentil 30 de la siguiente serie:

22, 27, 38, 23, 33, 15, 24, 29

Solución:

1) Ordenar los datos en forma creciente:

15, 22, 23, 24, 27, 29, 33, 38

i → 1 2 3 4 5 6 7 8

3 → P_{30}

2) Aplicar la formula:

$$i = \frac{P}{100} n$$

i=(30*8)/100=2,4→i=2,4→i+1=3,4→P_{30}→posicion 3=23

Ejemplo 2: calcular el Percentil 50 de la siguiente serie:

22, 27, 38, 23, 33, 15, 24, 29

Solución:

1) Ordenar los datos en forma creciente:

15, 22, 23, 24, 27, 29, 33, 38

i → | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

P_{50}

2) Aplicar la formula:

$$i = \frac{P}{100} n$$

M_e

i=(50*8)/100=4→i=4→i+(i+1)=(4+5)/2=4,5→P_{50}→posicion 4,5→(24+27)/2=25,5

Observación: si consideramos la definición de **Mediana** y de **Percentiles**, podemos concluir que **siempre** se va a verificar que: **P_{50}=M_e** (valor de la variable que <u>supera a no mas de la mitad</u> de las observaciones y <u>es superado por no mas de la mitad de las mismas)</u>

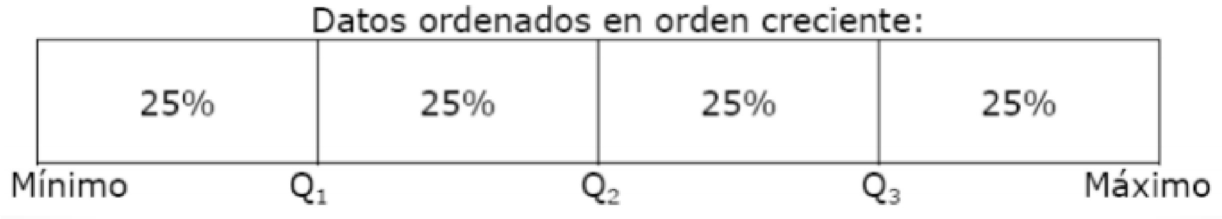
CUARTILES

Los **cuartiles** son medidas de posición o tendencia central que dividen a un conjunto de datos en cuatro partes iguales, teniendo por lo tanto 3 cuartiles:

Q_1→Primer Cuartil: representativo del valor de la variable que supera a no mas del 25% de las observaciones y es superado por no mas del 75% de las mismas. Es decir, separa el cuarto mas bajo de los datos de los tres cuartos mas altos, siendo por lo tanto igual al **Percentil 25**

Q_2→Segundo Cuartil: representativo del valor de la variable que supera a no mas del 50% de las observaciones y es superado por no mas del 50% de las mismas. Es decir, separa la mitad mas baja de los datos de la mitad mas alta, siendo igual al **Percentil 50,** e igual a la **Mediana.**

Q_3→Tercer Cuartil: representativo del valor de la variable que supera a no mas del 75% de las observaciones y es superado por no mas del 25% de las mismas. Es decir, separa los tres cuartos mas bajos del cuartos mas altos, siendo por lo tanto igual al **Percentil 75.**



Determinación la ubicación de un cuartil

Se procede en forma análoga a la determinación de la ubicación de un **Percentil** considerando la relación existente entre **Percentil** y **Cuartil**, de la forma:

1) Ordenar los datos en sentido creciente.
2) Calcular la ubicación del **cuartil** (i) aplicando la formula para calcular la ubicación de un **Percentil**, dada por:

$$i = \frac{P}{100} n$$

Considerando que, **i=ubicación del Cuartil**:

$$Q_1 = P_{25} \rightarrow i = \frac{Q_1 = P_{25}}{100} n$$

$$Q_2 = P_{50} \rightarrow i = \frac{Q_2 = P_{50}}{100} n$$

$$Q_3 = P_{75} \rightarrow i = \frac{Q_3 = P_{75}}{100} n$$

1) Determinar la ubicación de a) o de b), procediendo en forma análoga al caso de determinación de Percentiles. Es decir:

a) Si i=a **es un numero entero**, el **Q-esimo** Cuartil es el promedio de los valores de las variables localizadas entre la **i-esima** e **(i+1)-esima** ubicación.

b) Si i≠a **es un numero entero**, el valor del **Q-esimo Cuartil** esta dado por el valor de la variable ubicado en la parte entera del numero decimal resultante de sumarle **1** a **i**, es decir **lugar (i+1)**

Ejemplo 1: 225, 206, 214, 216, 229, 222, 221, 209

Ejemplo 2: 225, 206, 214, 216, 229, 222, 221, 209, 325

Determinar:

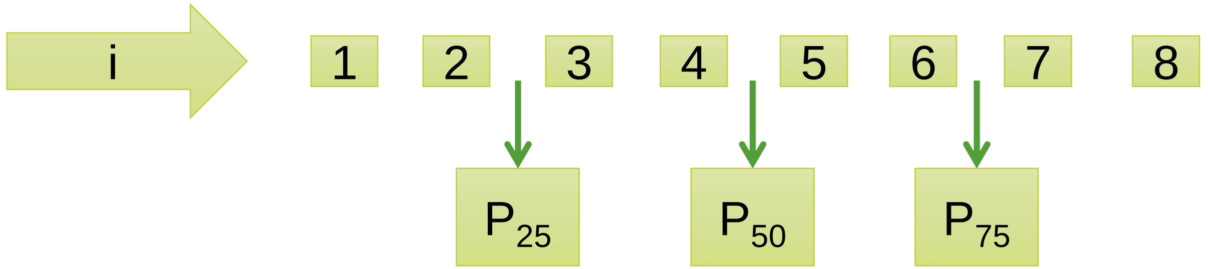
1. Q_1
2. Q_2
3. Q_3

Solución 1: para obtener los cuartiles en forma análoga a la determinación de los Percentiles, recordando la relación existente entre ambas medidas de posición. En consecuencia se deberá:

Solución:

1) Ordenar los datos en forma creciente:

206, 209, 214, 216, 221, 222, 225, 229



2) Aplicar la formula:

$i=(25*8)/100=2 \rightarrow i=2 \rightarrow i+(i+1)=(2+3)/2=2,5 \rightarrow P_{25} \rightarrow (209+214)/2=211,5$

$i=(50*8)/100=4 \rightarrow i=4 \rightarrow i+(i+1)=(4+5)/2=4,5 \rightarrow P_{50} \rightarrow (216+221)/2=218,5$

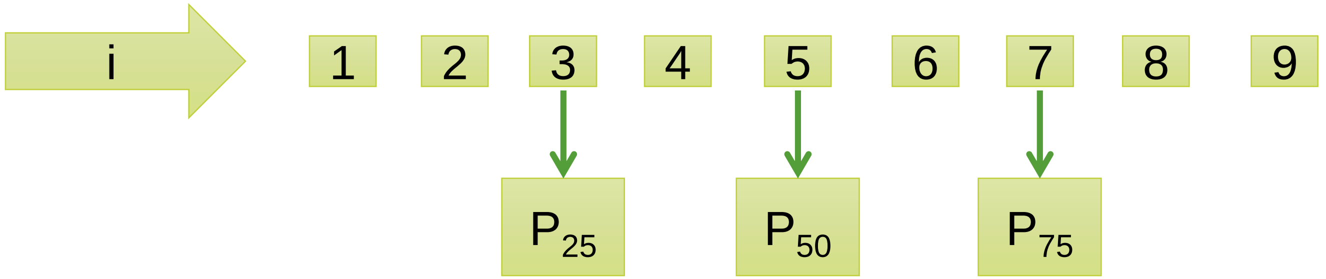
$i=(75*8)/100=6 \rightarrow i=6 \rightarrow i+(i+1)=(6+7)/2=6,5 \rightarrow P_{75} \rightarrow (222+225)/2=223,5$

Solución 2: para obtener los cuartiles en forma análoga a la determinación de los Percentiles, recordando la relación existente entre ambas medidas de posición. En consecuencia se deberá:

Solución:

1) Ordenar los datos en forma creciente:

206, 209, 214, 216, 221, 222, 225, 229, 325



2) Aplicar la formula:

$i=(25*9)/100=2{,}25 \rightarrow (i+1)=2{,}25+1=3{,}25 \rightarrow P_{25} \rightarrow$ posicion $3 \rightarrow 214$

$i=(50*9)/100=4{,}5 \rightarrow (i+1)=4{,}5+1=5{,}5 \rightarrow P_{50} \rightarrow$ posicion $5 \rightarrow 221$

$i=(75*9)/100=6{,}75 \rightarrow (i+1)=6{,}75+1=7{,}75 \rightarrow P_{75} \rightarrow$ posicion $7 \rightarrow 225$

DECILES

Se procede en forma análoga a la determinación de la ubicación de un **Decil** considerando la relación existente entre **Percentil y Deciles** de la forma:

1) Ordenar los valores de la variable en forma creciente.

2) Calcular la ubicación del **decil (i)** aplicando la formula para calcular la ubicación de un **Percentil**, considerando la relación existente entre ambas medidas de posición, dada por:

$$i = \frac{P}{100} n$$

Siendo:

i: ubicación del Cuartil

P o D: el Decil de interés

n: numero de datos

$$D_1 = P_{10} \rightarrow i = \frac{D_{10} = P_{10}}{100} n,\ D_2 = P_{20} \rightarrow i = \frac{D_{20} = P_{20}}{100} n, \ldots..,\ D_9 = P_{90} \rightarrow i = \frac{D_{90} = P_{90}}{100} n$$

Ejemplo 1 dada la serie:

243, 232, 225, 206, 214, 216, 229, 222, 221, 209

Determinar: Q_1, Q_5, Q_8

Solución:

1) Ordenar los datos en forma creciente:

206, 209, 214, 216, 221, 222, 225, 229, 232, 243

2) Aplicar la formula:

$i=(10*10)/100=1 \rightarrow i=1 \rightarrow i+(i+1)=(1+2)/2=1,5 \rightarrow P_{10} \rightarrow (206+209)/2=207,5$

$i=(50*10)/100=5 \rightarrow i=5 \rightarrow i+(i+1)=(5+6)/2=5,5 \rightarrow P_{50} \rightarrow (222+221)/2=221,5$

$i=(80*10)/100=8 \rightarrow i=8 \rightarrow i+(i+1)=(8+9)/2=8,5 \rightarrow P_{80} \rightarrow (229+232)/2=230,5$

Actividad 1: Se registraron las siguientes mediciones para el tiempo de secado (en minutos) de cierta marca de pintura esmaltada.

60	90	77	58	72	47	77	39	78	86
82	86	90	89	50	95	58	63	44	94
88	64	68	55	92	72	95	97	85	80

Calcule:

1. Media
2. Mediana
3. Moda
4. Q_1,Q_2 y Q_3

MEDIA GEOMETRICA

Cuando se quiere obtener una medida de posición de cocientes, razones, tasas o variables que tienen un comportamiento de una progresión geométrica, se debe utilizar la **Media Geométrica** dada por **«la raíz n-esima del producto de las variables consideradas, siendo n el numero de las mismas».** Suponiendo una variable x, se la denota de la forma: Mg(x) o G(x)

a) **Serie simple:** dada la serie simple: x_1, x_2, x_3,………, x_n, la **Media Geométrica** se obtiene de la forma:

$$\mathrm{Mg(x)} = \sqrt[n]{\mathrm{x}_1 \times \mathrm{x}_2 \times \mathrm{x}_3 \times \cdots \times \mathrm{x}_\mathrm{n}} = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

a) **Datos agrupados:** en el caso de una distribución de frecuencias de variable discreta o de variable continua, la **Media Geométrica** denotada de la forma **Mg(y)** o **G(y),** surge de: «la raíz n-esima del producto de las potencias con base igual a cada valor de la variable o marca de clase y con exponente igual a las respectivas frecuencias absolutas», es decir:

$$\mathrm{Mg(y)} = \sqrt[n]{{y_1}^{n_1} \times {y_2}^{n_2} \times \cdots \times {y_m}^{n_m}} = \sqrt[n]{\prod_{i=1}^{m} {y_i}^{n_i}}$$

MEDIA ARMONICA

Cuando se quiere calcular promedios en problemas en los que los valores de la variable son proporciones, que son relaciones que se pueden expresar en forma reciproca, como es el calculo de la velocidad promedio, o de productividad media, se utiliza como media de posición a la Media Armónica dada por **«la inversa de la media Aritmética de los inversos de los valores de la variable»**. Suponiendo una variable x, se la denota con H(x) o Ma(x)

a) **Serie simple:** dada la serie simple: x_1, x_2, x_3,………, x_n, la **Media Armónica** se obtiene de la forma:

$$Ma(x) = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_{ii}}}$$

a) **Datos agrupados:** en el caso de una distribución de frecuencias de variable discreta o de variable continua, la **Media Geométrica** denotada de la forma **Mg(y)** o **G(y),** surge de: «la raíz n-esima del producto de las potencias con base igual a cada valor de la variable o marca de clase y con exponente igual a las respectivas frecuencias absolutas», es decir:

$$Ma(y) = \frac{n}{\frac{n_1}{y_1}+\frac{n_2}{y_2}+\cdots+\frac{n_m}{y_m}} = \frac{n}{\sum_{i=1}^{m}\frac{n_i}{y_i}}$$