UNIDAD III. MEDIDAS DE TENDENCIA CENTRAL PARA DATOS AGRUPADOS

depositphotos.com, 2019

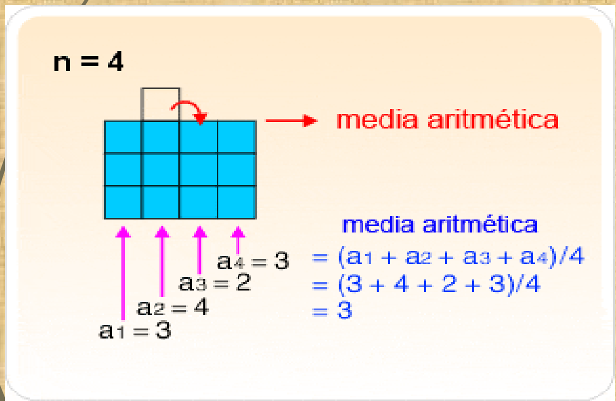
PROPÓSITO

- Razonamiento lógico y sistémico
- Búsqueda y análisis de información
- Reflexión critica
- Aplicación , análisis y solución de problemas

Media Aritmética

Es la forma mas común de sintetizar un conjunto de datos con una medida representativa ya que es el conjunto de cifrar sumadas y dividida entre numero de datos contemplados.

Esta medida es totalmente numérica o sea sólo puede calcularse en datos de características cuantitativas.



Fuente: RPDP 2019

Según Lind, Marchal, Wathen (2012, 60) se determina de la siguiente manera:

$$\overline{X} = \frac{\Sigma X}{n}$$

Donde:

$\overline{X}$=es la media de la muestra; se lee: X barra.

n= es el número de valores de la muestra.

X= representa cualquier valor particular.

Σ= es la letra mayúscula griega sigma e indica la operación de suma.

ΣX= es la suma de X valores de la muestra.

EJEMPLO:

La altura en cm de los jugadores de un equipo de baloncesto está en la siguiente tabla. Calcular la media.

Intérvalo	x_i	f_i	$x_i \cdot f_i$
$[160, 170)$	165	1	165
$[170, 180)$	175	2	350
$[180, 190)$	185	4	740
$[190, 200)$	195	3	585
$[200, 210)$	205	2	410
		12	2250

Calculamos la media para datos agrupados:

$$\overline{x} = \frac{165 \cdot 1 + 175 \cdot 2 + 185 \cdot 4 + 195 \cdot 3 + 205 \cdot 2}{1 + 2 + 4 + 3 + 2} =$$

$$= \frac{2250}{12} = 187.5$$

Fuente: sangakoo 2018

Mediana

Según Lind, Marchal, Wathen (2012, 64) es definido como el Punto medio de los valores una vez que se han ordenado de menor a mayor o de mayor a menor.

Fuente: Universo Formulas 2015

La definición de mediana se replantea así (Sweeney, William, Anderson, 2008,84)

a) Si el número de observaciones es impar, la mediana es el valor de en medio

b) Si el número de observaciones es par, la mediana es el promedio de las dos observaciones de en medio.

Para la primera suposición el ejemplo seria:

5 7 9 10 15

Media

Fuente: Elaboración propia

Segunda suposición

$$123 \quad 225 \quad \underbrace{302 \quad 444}_{\text{Valores de la mediana}} \quad 467 \quad 707$$

302+444= 746

746/2= 373 Mediana= 373

Fuente: Elaboración propia

La tendencia a todo esto es que es recomendable para no inflar demasiado la demanda de un ingresos anuales o valores de propiedades muy altos ya que no contempla cantidades de extremos solo partes centrales y proporcionales de la misma

Moda

Otra medida de tendencia central es la moda, tal y como se considera en el ámbito social : es lo que mas se usa o prefiere como ropa, teléfonos, los actores de grupos musicales. Por tal motivo en estadística es considerado como el valor mas frecuente.

Fuente: Bcn Cool Hunter 2019

La moda es una medida de fácil obtención, aun que no siempre hay un dato cuya frecuencia este claramente por encima de las frecuencias de otros datos.

Hay conjunto de datos que tienen dos modas (bimodales) o incluso mas (multimodales).

Por esta razón, la moda se utiliza se utiliza solo para hacerse una idea provisional de la tendencia central de un conjunto de datos (Arreola, Robles, 2016, 56)

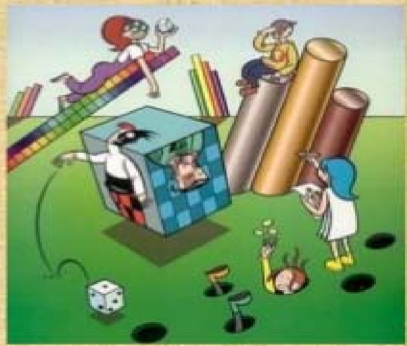
Ejemplo.

En esta serie de datos identificar cual es la moda.

1; **2**; 8; 10; 11; 15; 2; 2; **16; 18; 15;**
2; 19; 2; 20, 2; 12; 16; 21; 22; 25.

Mo= 2

Porque es el numero q mas se repite en este grupo de datos



Fuente: Slideshare 2017

Media Geométrica

La media geométrica resulta útil para determinar el cambio promedio de porcentajes, razones, índices o tasas de crecimiento. Posee amplias aplicaciones en la administración y la economía, ya que con frecuencia hay interés en determinar los cambios porcentuales de ventas, salarios o cifras económicas, como el producto interno bruto, los cuales se combinan o se basan unos en otros. La media geométrica de un conjunto de **n** números positivos se define como la raíz enésima de un producto de **x** variables.(lind, Marchal, Wathen, 2012, 72)

Media Geométrica

- Es la raíz N-ésima del producto de N valores.

$$MG = \sqrt[N]{X_1 \cdot X_2 \cdot \ldots \cdot X_{N-1} \cdot X_N}$$

- Ejemplo, la media geométrica de 2,4,9 es

$$MG = \sqrt[3]{2 \cdot 4 \cdot 9} = \sqrt[3]{72} = 4.1602$$

Fuente: Brainly 2016

Donde:

MG= Media Geométrica.
N= Numero de datos.
X= representa cualquier valor particular.

Fuente: Elaboración propia

Media Armónica

La media armónica **H** de un conjunto de **N** números X1, X2, X3, . . . , XN es el recíproco de la media aritmética de los recíprocos de los números: (Spiegel, Stephens, 2009, 65)

Los elementos del conjunto deben ser necesariamente no nulos. Esta media es poco sensible a los valores grandes y los infravalora respecto a la media aritmética, pero muy sensible a los valores próximos a cero, ya que los recíprocos 1/Xi son muy altos, Si algún valor fuese cero, la media armónica quedaría indeterminada.

$$H = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_N}}$$

Ejemplo

Un tren realiza un trayecto de 400km. La vía tiene en mal estado que no permitían correr. Los primeros 100 km los recorre a 120km/h, los siguientes 100km la vía está en mal estado y va a 20km/h, los terceros a 100km/h y los 100 últimos a 130km/h. Para calcular el promedio de velocidades, calculamos la media armónica.

La media armónica es de H=52,61km/h.

$$H = \frac{4}{\frac{1}{120} + \frac{1}{20} + \frac{1}{100} + \frac{1}{130}} = \frac{4}{\frac{593}{7800}} = \frac{4 \cdot 7800}{593} = 52,61$$

Fuente: Universo Formula 2015

Cuartiles

Son 3 valores de la variable que dividen a un conjunto de datos ordenados en 4 partes iguales. Según vitutor.net (2019)

Los cuales se representan con valores 25%, 50%, 75% de los datos

Q1 =primer cuartil 25

Q2 =segundo cuartil 50

Q3 =tercer cuartil 75

FORMULA: **basado en (Sweeney, William, Anderson, 2008,87)**

$$i = \left(\frac{P}{100}\right) n$$

i= Cuartil

P= Porcentaje del cuartil (25, 50, 75)

n= Es el número de valores de la muestra.

EJEMPLO:

330 | 350 390 458 | 485 512 534 | 550

Q1 Q2 Q3

$$i = \left(\frac{25}{100}\right) 8 = 2$$

$$i = \left(\frac{50}{100}\right) 8 = 4$$

$$i = \left(\frac{75}{100}\right) 8 = 6$$

$$350 + 390/\,2 = 370$$

Q1

$$458 + 485/\,2 = 471.5$$

Q2

$$512 + 534/\,2 = 523$$

Q3

Fuente: Elaboración propia

Deciles

Según Vitutor-A (2019), son los nueve valores que dividen la serie de datos ordenados en diez partes porcentualmente iguales. Son los nueve valores que dividen al conjunto de datos ordenados en diez partes iguales.

Los deciles se denotan D1, D2,..., D9.

Los deciles dan los valores correspondientes al 10%, al 20%... y al 90% de los datos.

El D_5 coincide con la mediana.

El decil 5 coincide con el cuartil 2

Cálculo De Los Deciles

Basado en Superprof (2019):

- En primer lugar buscamos el intervalo donde se encuentra el decil 1,2…y 9, utilizando la fórmula: $\frac{k * N}{a_i}$

Donde:

K: número de decil que queremos encontrar

N: suma de frecuencias absolutas

ai: amplitud de la clase

Fórmula

Basado en Superprof (2019):

$$D_k = L_i + \frac{\frac{k \cdot N}{10} - F_{i-1}}{f_i} \cdot a_i$$

Donde:

Li: limite inferior de la clase donde se encuentra el decil

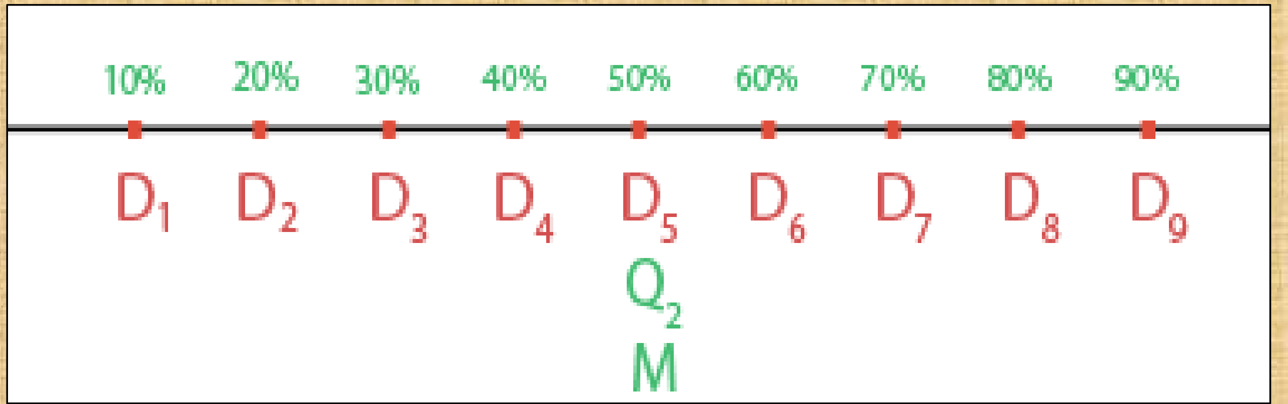
K: número de decil que queremos encontrar

N: suma de frecuencias absolutas

Fi-1: frecuencia acumulada anterior de a la clase del decil

ai: amplitud de la clase

fi: frecuencia absoluta

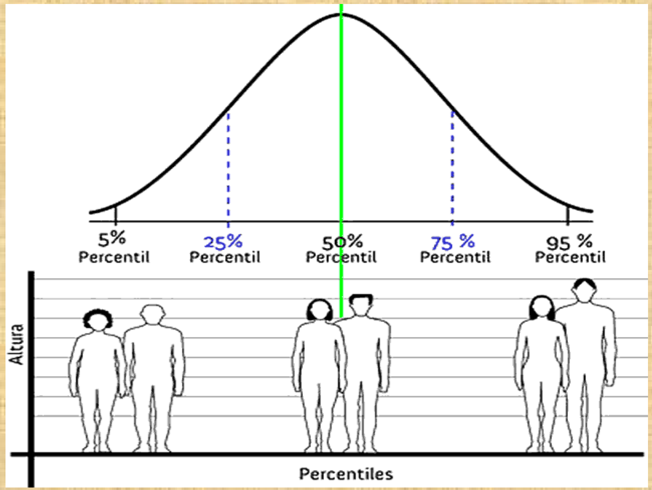


Fuente: Deciles/ Superprof

Percentiles

Según superprof (2019), Los percentiles son los 99 valores que dividen una serie de datos ordenados en 100 partes iguales.

Los percentiles dan los valores correspondientes al 1%, al 2%... y al 99% de los datos.



Fuente: curiosoando

FORMULA Y EJEMPLO SEGÚN (Sweeney, William, Anderson, 2008,87)

CÁLCULO DEL PERCENTIL p

Paso 1. Ordenar los datos de menor a mayor (colocar los datos en orden ascendente).
Paso 2. Calcular el índice i

$$i = \left(\frac{p}{100}\right)n$$

donde p es el percentil deseado y n es el número de observaciones.

Paso 3. (a) Si *i no es un número entero, debe redondearlo.* El primer entero mayor que i denota la posición del percentil p.
(b) Si *i es un número entero,* el percentil p es el promedio de los valores en las posiciones i e $i + 1$.

Para ilustrar el empleo de este procedimiento, determine el percentil 85 en los sueldos mensuales iniciales de la tabla 3.1.

Paso 1. Ordenar los datos de menor a mayor

3310 3355 3450 3480 3480 3490 3520 3540 3550 3650 3730 3925

Paso 2.

$$i = \left(\frac{p}{100}\right)n = \left(\frac{85}{100}\right)12 = 10.2$$

Paso 3. Como i no es un número entero, se debe *redondear*. La posición del percentil 85 es el primer entero mayor que 10.2, es la posición 11.

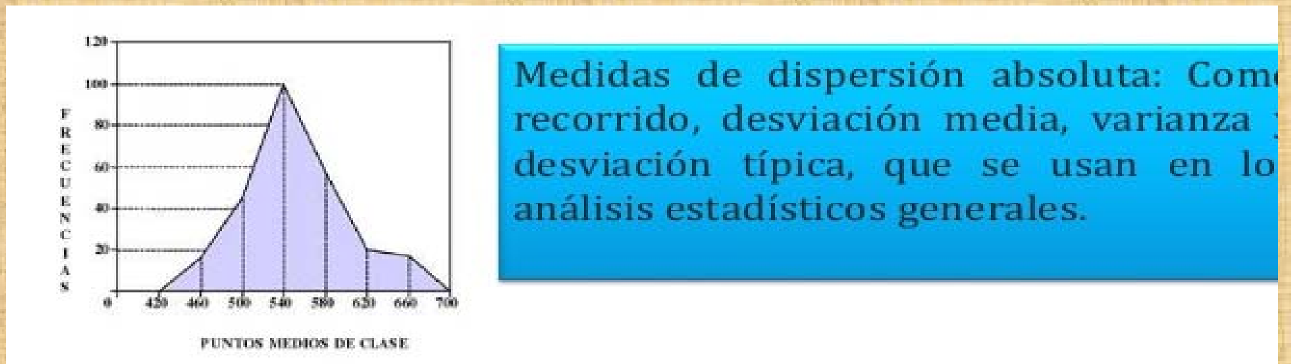
Observe ahora los datos, entonces el percentil 85 es el dato en la posición 11, o sea 3730.

MEDIDAS DE DISPERSIÓN Y DISPERSIÓN RELATIVA:

PARA DATOS AGRUPADOS

Medidas De Dispersion

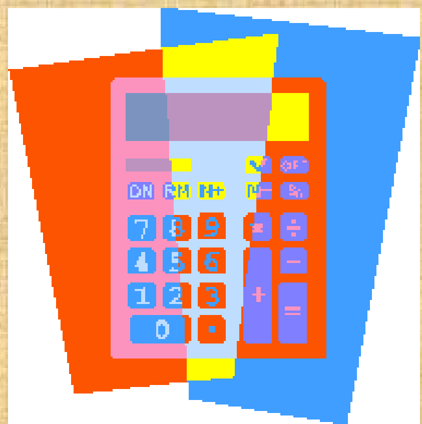
Según ecured (2019) Las medidas de dispersión muestran la variabilidad de una distribución, indicándolo por medio de un número, si las diferentes puntuaciones de una variable están muy alejadas de la media.



Fuente: slideshare.net

- Cuanto mayor sea ese valor, mayor será la variabilidad, cuanto menor sea, más homogénea

- Las medidas de dispersión son números reales no negativos, su valor es igual a cero cuando los datos son iguales y este se incrementa a medida que los datos se vuelven más diversos.



Fuente: matematicas.reduaz.m

Fuente: educame.com

Las medidas de dispersión más utilizadas son el rango, la desviación estándar y la varianza.

- **Rango**

Indica la dispersión entre los valores extremos de una variable. se calcula como la diferencia

entre el mayor y el menor valor de la variable. Se denota como R.

Para datos ordenados se calcula como:

R = x(n) - x(1)

- **Desviación media**

Es la media aritmética de los valores absolutos de las diferencias de cada dato respecto a la

media. Según ecured (2019)

- **Desviación estándar**

La desviación estándar mide el grado de dispersión de los datos con respecto a la media, se

denota como s para una muestra o como σ para la población. Se define como la raíz cuadrada

de la varianza según la expresión:

- **Varianza**

Es otro parámetro utilizado para medir la dispersión de los valores de una variable respecto a la

media. Corresponde a la media aritmética de los cuadrados de las desviaciones respecto a la

media. Según ecured (2019)

Medidas De Dispersión Relativas

Según ucemestadistica.word(2019) Son indicadores de la dispersión de la distribución que se han relativizado, para que no afecten

las unidades de medida de la variable y para que puedan hacerse comparaciones entre las

dispersiones de conjuntos de datos dispares.

Fuente: slideshare.net

Se obtiene dividiendo la medida de dispersión elegida entre la media del conjunto de datos

Se trata de parámetros

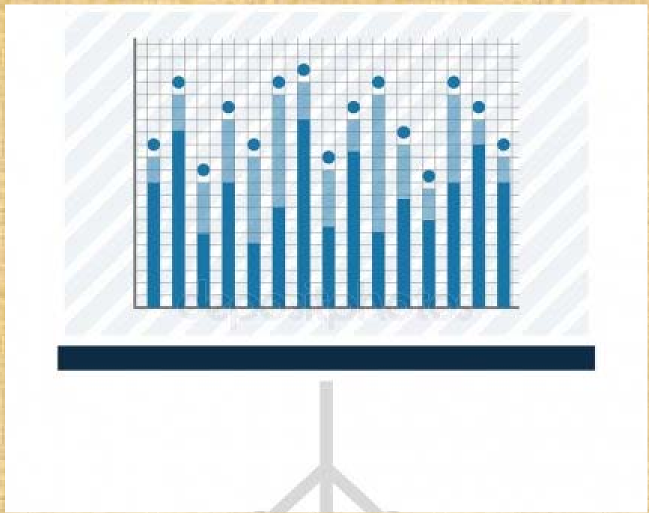
más técnicos y utilizados en estudios específicos, y entre ellas se encuentran los coeficientes

De apertura, el recorrido relativo, el coeficiente de variación (índice de dispersión de Pearson) y

el índice de dispersión mediana. Según ucemestadistica.word(2019)

MEDIDAS DE TENDENCIA CENTRAL PARA DATOS NO AGRUPADOS

Media Aritmética



depositphotos.com, 2019

Según Anderson, (2008, 113), la medida de localización más importante es la media, o valor promedio, de una variable. La media proporciona una medida de localización central de los datos. Si los datos son datos de una muestra, la media se denota $\bar{x}$.

- En las fórmulas estadísticas se acostumbra denotar el valor de la primera observación de la variable x con x1, el valor de la segunda observación de la variable x con x2 y así con lo siguiente. En general, el valor de la i-ésima observación de la variable x se denota xi. La fórmula para la media muestral cuando se tiene una muestra de n observaciones es la siguiente.

$$\overline{X} = \frac{\sum x_i}{n}$$

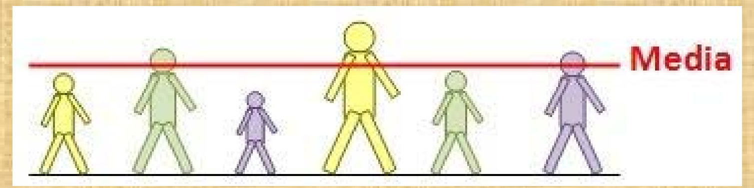
En la fórmula anterior el numerador es la suma de los valores de las n observaciones. Es decir:

$$\sum x_i = x_1 + x_2 + \cdots + x_n$$

Ejemplo:

Para ilustrar el cálculo de la media muestral, considere los siguientes datos que representan el tamaño de cinco grupos de una universidad.

- 46
- 54
- 42
- 46
- 32



UniversoFormulas.com, 2018

- Se emplea la notación $x_1, x_{2,} x_{3,} x_4, x_{5,}$ para representar el número de estudiantes en cada uno de los cinco grupos.

$x_1 = 46$

$x_2 = 54$

$x_3 = 42$

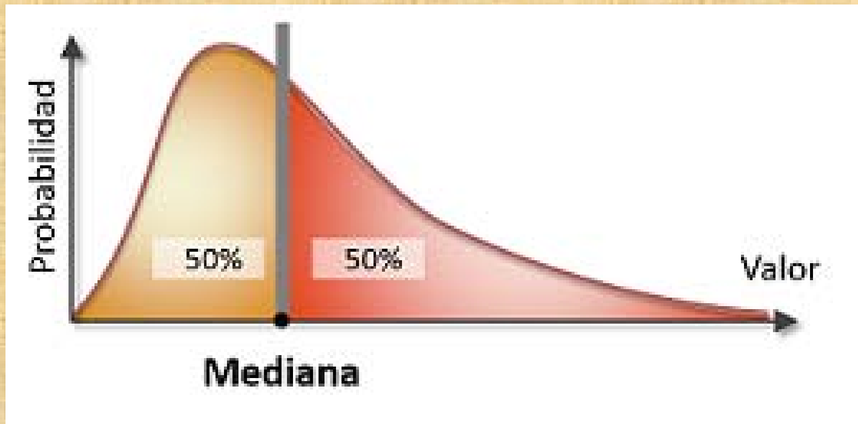
$x_4 = 46$

$x_5 = 32$

$$\overline{X} = \frac{\sum x_i}{n} = \frac{46+54+42+46+32}{5} = 44$$

Mediana

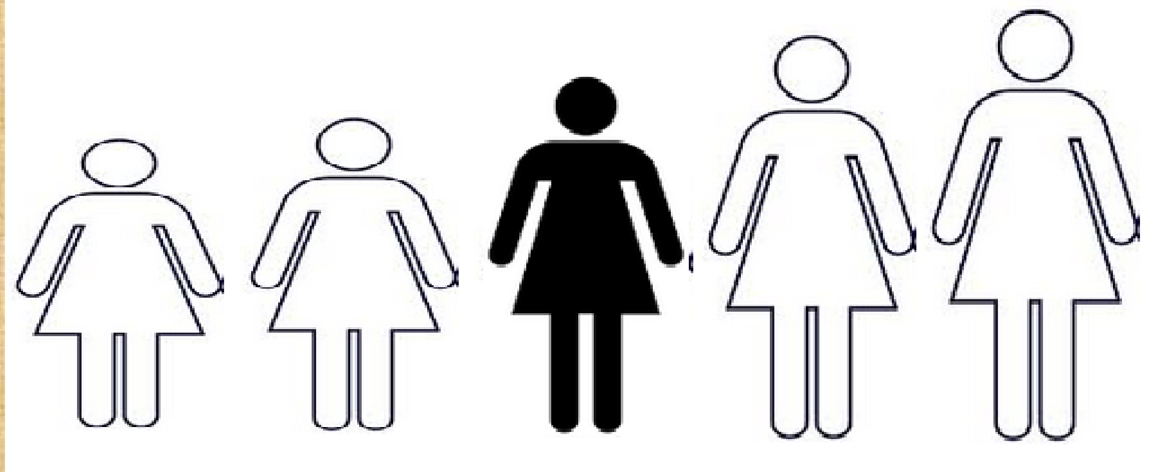
La mediana es la observación que ocupa el lugar central de un conjunto de observaciones ordenadas en sentido ascendente o descendente. (Newbold, 2008, 50).



Matematicas10.net, 2019

- Si el tamaño de la muestra, n, es un número impar, la mediana es la observación que se encuentra en el medio.

- Si el tamaño de la muestra, n, es un número par, la mediana es la media de las dos observaciones que se encuentran en el medio.



ProyectoDescartes.org, 2019

Ejemplos. (Murray, 2009, 64)

- La mediana del conjunto de números: 3, 4, 5, <u>6</u>, 8, 8 ,10

Me: 6

- La mediana del conjunto de números: 5, 5, 7, <u>9, 11</u>, 12, 15, 18

Me: (9,11) En este caso la mediana se calcula obteniendo la media:

$$\text{Me} = \frac{9+11}{2} = \frac{20}{2} = 10$$

Moda

La Moda es el valor de la observación que aparece con mayor frecuencia.(Douglas,2012, 65)

Ventaja

- No influyen en ella valores extremadamente grandes o pequeños.

Desventaja

- En el caso de muchos conjuntos de datos no existe la moda, porque ningún valor se presenta más de una vez.

Elaboración propia

depositphotos.com, 2019

Según VITUTOR.net (2015) La moda se representa por Mo.

Se puede hallar la moda para variables cualitativas y cuantitativas.

Hallar la moda de la distribución:
2, 3, 3, 4, 4, 4, 5, 5
Mo= 4

Otros ejemplos:

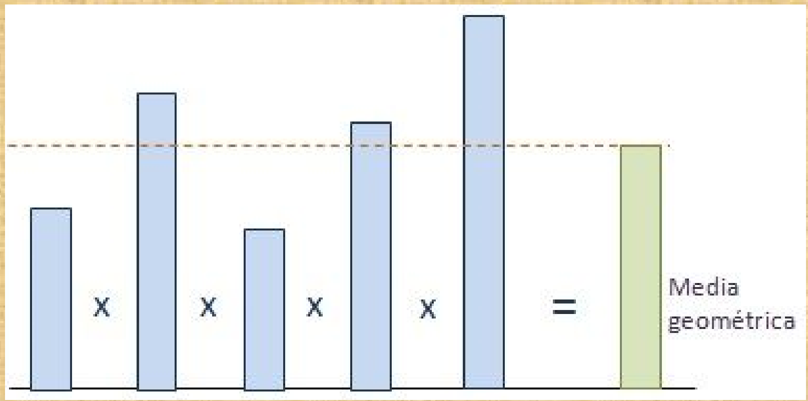
- Si en un grupo hay dos o varias puntuaciones con la misma frecuencia y esa frecuencia es la máxima, la distribución es bimodal o multimodal, es decir, tiene varias modas.

1, 1, 1, 4, 4, 5, 5, 5, 7, 8, 9, 9, 9

Mo= 1, 5, 9

- Cuando todas las puntuaciones de un grupo tienen la misma frecuencia, no hay moda.
  2, 2, 3, 3, 6, 6, 9, 9

Media Geométrica



UniversoFormulas.com, 2017

La media geométrica de un conjunto de n números positivos se define como la raíz enésima de un producto de n variables. La fórmula de la media geométrica se escribe de la siguiente manera:(Douglas, 2012, 72)

$$MG=\sqrt[n]{(x_1)(x_2)\ldots(x_n)}$$

Ejemplo: (Murray, 2009, 83)

Encontrar la media geométrica de los números: 3, 5, 6, 6, 7, 10, 12

- MG= $\sqrt[7]{(3)(5)(6)(6)(7)(10)(12)}$

  MG= $\sqrt[7]{453600}$. Empleando logaritmos comunes, log MG= $\frac{1}{7}$ log 453600 = $\frac{1}{7}(5.6567) = 0.8081$ y MG =6.43 (a la centésima más cercana).

Otro método:

Log MG= $\frac{1}{7}(\log 3 + \log 5 + \log 6 + \log 6 + \log 7 + \log 10 + \log 12)$

$= \frac{1}{7}(0.4771 + 0.6990 + 0.7782 + 0.7782 + 0.8451 + 1.0 + 1.0792)$

= 0.8031

MG= 6.43

MEDIA ARMÓNICA

Media Armónica

- Se utiliza cuando se desea promediar velocidades, cuando el tiempo es constante y las distancias varían o se mantienen y la velocidad varia (Elorza, 1999).
- Su formula es:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots \frac{1}{x_n}} = \frac{n}{\sum_{i=1} n \frac{1}{x_i}}$$

(Elorza, 1999, 84)

Cuartiles

En un conjunto de datos en el que éstos se hallan ordenados de acuerdo con su magnitud, el valor de en medio (o la media aritmética de los dos valores de en medio), que divide al conjunto en dos partes iguales, es la mediana. Continuando con esta idea se puede pensar en aquellos valores que dividen al conjunto de datos en cuatro partes iguales. Estos valores, denotados Q1, Q2 y Q3 son el primero, segundo y tercer cuartiles, respectivamente; el valor Q2 coincide con la mediana (Spiegel y Stephens, 2009).

Con frecuencia es conveniente dividir los datos en cuatro partes; así, cada parte contiene una cuarta parte o 25% de las observaciones. A los puntos de división se les conoce como cuartiles (ANDERSON SWEENEY WILLIAMS, 2008).

- Q1 = primer cuartil, o percentil 25
- Q2 = segundo cuartil, o percentil 50
- Q3 = tercer cuartil, o percentil 75

- **Cómo se calculan los cuartiles**
- Esta calculadora utiliza el siguiente método para calcular los cuartiles:
- n es el número total de valores. x_1, x_2 ... x_n son los valores ordenados de menor a mayor.
- **Fórmulas para calcular el primer cuartil**
- Si $1/4(n+1)$ es igual a un numero entero, la formula seria: $x_{\frac{1}{4}(n+1)}$
- Si no es numero enteros entonces seria

MEDIDAS DE DISPERSIÓN Y DISPERSIÓN RELATIVA: PARA DATOS NO AGRUPADOS.

Medidas de dispersión y dispersión relativa: para datos no agrupados.

- Las medidas de dispersión muestran la variabilidad de una distribución indicando por medio de un numero si las diferentes puntuaciones de una variable están muy alejadas de la media. Cuando sea eses valor mayor será la variabilidad. Cuando sea menor será la media (Villegas Alemán, 2012).
- Varianza

Es el resultado de la división de la sumatoria de las distancias existentes entre cada dato y su media aritmética elevadas al cuadrado y el numero total de datos (elaboración propia).

$$(s^2) = \frac{\sum_i^n = 1\ (x_i - \underset{x}{\rightarrow})^2}{n}$$

- Desviación estándar:
- Es una medida de dispersión usada en estadística que nos dice cuanto tienden a alejarse los valores puntuales del promedio en una distribución, es el promedio de la distancia de cada punto respecto del promedio (Quintana, 1992). si se conoce la varianza se calcula:
- $DE\ (s) = \sqrt{\frac{\sum_{i}^{n}=1\ (x_i - \underset{x}{\rightarrow})^2}{n}}$

RANGO

Se mide como la diferencia entre el valor mayor y el menor (elaboración propia).

$$R = X\ max - Xmenor$$

- Coeficiente de variación
- Indica la importancia de la desviación estándar en relación al promedio aritmético y cuya definición puede representarse de la siguiente manera.
- $(c.v) = \frac{desvicion\ estandar}{media\ aritmetica}(100)$
- Desviación media
- Es la medida aritmética de los valores absolutos de las desviaciones respecto a la media y se representa (elaboración propia).
- $DM = \frac{\sum_{i}^{n}=1\ (x_i - \underset{x}{\rightarrow})}{n}$

BIBLIOGRAFIA

1. Elorza, al (2000),Estadística para las Ciencias Sociales y del Comportamiento, 2ed, México.
2. Anderson, Sweeney, Williams, al (2008), Estadística para Administración y Economía, 10ed, México D.F.
3. Newbold, Carlson, Thorne, al (2008), Estadística para Administración y Economía, 6ed, Madrid (España).
4. Spiegel, Sthepens, al (2009), Estadística, 4ed, México D.F.
5. Lind, Wathen, Marchal, al (2012), Estadistica Aplicada a los Negocios y la Economia, 15ed, Mexico D.F.

BIBLIOGRAFÍA

6. Arreola, Robles, al (2016), La probabilidad y Estadística, 1ed, México.

REFERENCIAS

1. Sangaku Maths (2019), acezado el 25 de agosto del 2019 en: https://www.sangakoo.com/es/temas/media-geometrica
2. Universo formulas (2015), acezado el 25 de agosto del 2019 en: https://www.universoformulas.com/estadistica/descriptiva/media-armonica/
3. Vitutor.net(2019), accesado el 25 de agosto en https://www.vitutor.net/2/11/cuartiles_percentiles.html
4. Superprof(2019), accesado el 25 de agosto del 2019 en:https://www.superprof.es/apuntes/escolar/matematicas/estadistica/descriptiva/deciles.html

REFERENCIAS

5. Ecured (2019), accesado el 26 de agosto del 2019 en: https://www.ecured.cu/Medidas_de_dispersi%C3%B3n
6. Ceaces (2019), accesado el 26 de agosto del 2019 en: https://www.uv.es/ceaces/base/descriptiva/disrelati.htm
7. Slideshare (2019), accesado el 26 de agosto del 2019 en:https://es.slideshare.net/edidpanccaapaza1/medidas-de-dispersion-absolutas-y-relativas-55214966