Descriptive statistics (II): Measures of central tendency and dispersion

Estadística descriptiva (II): Medidas de tendencia central y de dispersión

Juan Luis Soto Espinosa[1]

[1] Especialización en Salud en el Trabajo, FES Zaragoza, UNAM, PAPIME PE216019
Correo electrónico de contacto: soej@gmail.com

Fecha de envío: *01/06/2020*
Fecha de aprobación: *17/07/2020*

Medidas de tendencia central

Las Medidas de Tendencia Central son cálculos estadísticos cuyo propósito es resumir las características de un conjunto de datos en unos pocos valores "representativos" del conjunto. Las medidas de tendencia central definen un punto central en torno al cual se concentra el conjunto de los datos.

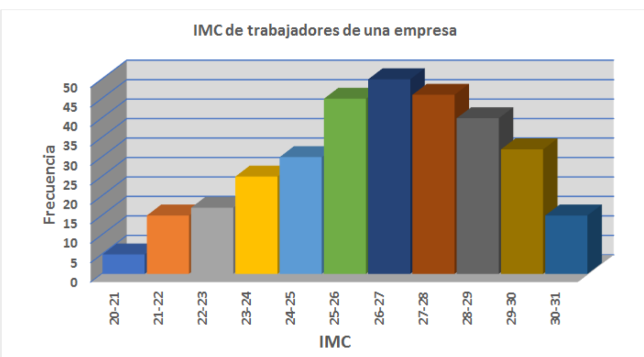
Tres son las medidas de tendencia central más utilizadas:

a) Media aritmética o promedio estadístico

b) Mediana

c) Moda

Considerando la gráfica siguiente:



Es posible observar que los datos tienden a concentrarse hacia las clases que ocupan la región central de la distribución de frecuencias, mientras que los datos que ocupan las clases más externas tienden a tener menos valores.

Media

La Media (que se denota por μ en el caso de poblaciones y como $\bar{x}$ en el caso de muestras) es el promedio aritmético del conjunto de datos, siempre y cuando se trabaje con variables cuantitativas.

Para obtener la media debemos obtener la suma de todos los valores numéricos asociados (xi) a cada dato y dividirlos entre el número total de datos. Supongamos que tenemos n datos y x1, x2, x3,…,xn son los valores numéricos correspondientes a cada uno de esos datos, entonces matemáticamente la media aritmética se expresa como:

$$\mu = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N}$$

(en poblaciones)

Dónde:

μ = Media poblacional

X_N = Valor en la posición N

N = Tamaño total de la población

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

(en muestras)

Dónde:

$\bar{x}$ = Media muestral

xn = Valor en la posición n

n = Tamaño total de la muestra

La expresión anterior se puede representar de forma compacta utilizando la notación sigma, es decir,

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

donde:

$\bar{x}$ = Media aritmética

$\sum$ = Sumatoria de los valores numéricos desde el primero (i = 1) hasta el n-ésimo valor (i = n).

X_i = Valor en la posición i

n = Número de valores en la muestra

Con base en las fórmulas anteriores, tomemos los datos de pesos (kg) de una muestra de 20 personas y obtengamos la media

167.2	74.8
83.9	102.5
102.5	74.8
83.9	83.9
83.9	74.8
74.8	90.85
74.8	90.85
90.85	102.5
128.7	74.8
74.8	83.9

Sustituyendo los valores en la expresión matemática de la media, se tiene

$$\bar{X} = \frac{167.2 + 83.9 + \cdots + 102.5 + 74.8 + 83.9}{20}$$

Por lo tanto, la media es

$$\bar{x} = \frac{1\,819.05}{20} = 90.95$$

Es posible obtener la media partiendo de una tabla de frecuencias, para ello debemos realizar el siguiente procedimiento

Cada valor x se multiplica por su frecuencia, lo que corresponde a la frecuencia absoluta de la clase y se suma al producto de las clases subsecuentes. En el ejemplo:

Peso (x)	**Frecuencia (f)**	**x*f**
74.8	7	523.6
83.9	5	419.5
90.85	3	272.55
102.5	3	307.5
128.7	1	128.7
167.2	1	167.2
Total	20	1819.05

Empleando el procedimiento para cálculo de la media a partir de una tabla de frecuencia organizada en clases, se tiene

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i * f(x_i)}{N}$$

Donde:

xi = valor de clase o centro de clase

f(xi) = Frecuencia de clase

Sustituyendo valores, tenemos:

$$\bar{x} = \frac{(7)(74.8) + (5)(83.9) + (3)(90.85) + (3)(102.5) + (1)(128.7) + (1)(167.2)}{20}$$

$$\bar{x} = \frac{1\,819.05}{20} = 90.95$$

Una de las LIMITACIONES de la media es su sensibilidad a la presencia de valores extremos en el conjunto de datos, pongamos un ejemplo:

Si tenemos el siguiente conjunto de datos:

Paciente	Peso
1	45.5
2	52.5
3	40.0
4	50.0
5	51.5
6	55.5
7	60.0
8	42.5
9	57.0
10	56.5
11	58.0

Al obtener la media tenemos:

$$\bar{x} = \frac{45.5 + 52.5 + 40.0 + 50.0 + 51.5 + 55.5 + 60 + 42.5 + 57.0 + 56.5 + 58.0}{10}$$

$$\bar{x} = \frac{569}{11} = 51.73$$

La media del conjunto de datos es 51.73 kg

Ahora supongamos que el tercer valor en lugar de **40.0 kg** fuera **150 kg**, en ese caso tendríamos:

$$\bar{x} = \frac{45.5 + 52.5 + 150.0 + 50.0 + 51.5 + 55.5 + 60 + 42.5 + 57.0 + 56.5 + 58.0}{10}$$

$$\bar{x} = \frac{679}{11} = 61.73$$

En este caso, la media del conjunto sería **61.73 kg**

Ahora supongamos que, en lugar de ser **150 kg**, fuera de **20 kg**, en ese caso tendríamos:

$$\bar{x} = \frac{45.5 + 52.5 + 20.0 + 50.0 + 51.5 + 55.5 + 60 + 42.5 + 57.0 + 56.5 + 58.0}{10}$$

$$\bar{x} = \frac{549}{11} = 49.91$$

Ahora, la media es de **49.91 kg**

Como se puede observar, el hecho de que UN SOLO VALOR sea extremo hacia arriba (muy alto) o hacia abajo (muy bajo) provoca que la media del conjunto de datos se desplace. En estos casos el investigador debe decidir si la media es la medida de tendencia central más conveniente o debe tomar como referencia otra medida.

Mediana

La mediana se define como el dato central de un conjunto ordenado de valores (ascendente o descendente).

Retomando el ejemplo anterior, procedemos a ordenar de forma ascendente el conjunto de datos, con lo que tenemos

Valor	40.0	42.5	45.5	50.0	51.5	52.5	55.5	56.5	57.0	58.0	60.0
Posición	1	2	3	4	5	6	7	8	9	10	11

En este ejemplo, el dato que se ubica en el centro del conjunto de datos es el número 52.5, que se ubica en la sexta posición.

Ahora supongamos que el primer valor el lugar de ser 40.0 kg es 20.0 kg

Valor	20.0	42.5	45.5	50.0	51.5	52.5	55.5	56.5	57.0	58.0	60.0
Posición	1	2	3	4	5	6	7	8	9	10	11

La mediana sigue siendo el dato que ocupa la sexta posición, es decir 52.5

Ahora supongamos que el valor es ahora 150 kg

Valor	42.5	45.5	50.0	51.5	52.5	55.5	56.5	57.0	58.0	60.0	150
Posición	1	2	3	4	5	6	7	8	9	10	11

La mediana sigue siendo el dato que ocupa la sexta posición, aunque ahora el valor es 55.5

Considerando los ejemplos anteriores, los datos pueden variar, pero la mediana sigue correspondiendo al dato que ocupa **LA POSICIÓN CENTRAL DEL ARREGLO** (la sexta posición en el ejemplo), es decir, la mediana depende de la posición central del arreglo y no del valor de los datos, con lo que podemos afirmar que no es sensible a valores extremos.

Para determinar la mediana existen dos posibilidades:

1) Si la cantidad de datos es impar, como en el ejemplo anterior, la mediana es el valor intermedio que queda después de ir descartando los valores que van quedando en los extremos.

2) Si la cantidad de datos es par, la mediana es el promedio de los dos valores que ocupan las posiciones centrales del conjunto ordenado de datos, si tuviéramos solamente 10 datos tendríamos:

Valor	40.0	42.5	45.5	50.0	51.5	52.5	55.5	56.5	57.0	58.0
Posición	1	2	3	4	5 ↑	6 ↑	7	8	9	10

En este caso, los datos que ocupan las posiciones centrales son 51.5 y 52.5 (posiciones 5 y 6 en el arreglo ordenado), por lo que se procede a calcular el promedio de los dos para obtener la mediana, es decir:

$$Mediana = \frac{51.5 + 52.5}{2} = \frac{104}{2} = 52$$

La mediana en este ejemplo sería 52 kg.

Para facilitar la determinación de la mediana, tomemos en cuenta el siguiente algoritmo:

$$posición\ de\ la\ mediana = \frac{n}{2}$$

Donde

n = Número de datos en el arreglo

Al resolver la ecuación tenemos dos posibles resultados:

a) Que el resultado sea un número entero: en este caso, el arreglo tiene un número par de elementos, por lo que el valor de la mediana estaría dado por:

$$mediana = \frac{dato\ en\ posición\ \frac{n}{2} + dato\ en\ posición\ (\frac{n}{2} + 1)}{2}$$

Si tiene un conjunto de 10 datos, tendremos:

$$posición\ de\ la\ mediana = \frac{10}{2} = 5$$

Al obtener un número entero, el valor de la mediana está dado por:

$$mediana = \frac{dato\ en\ posición\ \frac{10}{2} + dato\ en\ posición\ (\frac{10}{2} + 1)}{2}$$

$$mediana = \frac{dato\ en\ posición\ 5 + dato\ en\ posición\ (5 + 1)}{2}$$

$$mediana = \frac{dato\ en\ posición\ 5 + dato\ en\ posición\ 6}{2}$$

En este caso, debemos localizar el dato que ocupa la posición 5, sumarle el dato que ocupa la posición 6 y al resultado dividirlo entre dos

.

b) Que el resultado sea un número fraccionario: Esto quiere decir que tenemos un número impar de elementos, por lo que el valor de la mediana estaría dado:

$$mediana = dato\ en\ posición\ (\frac{n}{2} + 0.5)$$

Si tenemos un conjunto de 11 datos, se tiene:

$$mediana = dato\ en\ posición\ (\frac{11}{2} + 0.5)$$

$$mediana = dato\ en\ posición\ (5.5 + 0.5)$$

$$mediana = dato\ en\ posición\ (6)$$

Lo que indica que debemos consultar el dato que ocupa la sexta posición para conocer el valor de la mediana.

Como se puede observar, el cálculo de la mediana depende solamente de los **VALORES CENTRALES** del conjunto de datos, si hay valores atípicamente altos o bajos, estos no afectan el valor de la mediana, es decir, la mediana **NO ES SENSIBLE A VALORES EXTREMOS**.

Si se Repite la operación tomando los datos de la siguiente Tabla PESO a fin de obtener la mediana. Lo primero que debemos hacer es ordenarlos de forma ascendente:

Ordenados en forma ascendente

Ordenados en forma ascendente
74.8
74.8
74.8
74.8
74.8
74.8
74.8
83.9
83.9
83.9
83.9
83.9
90.85
90.85
90.85
102.5
102.5
102.5
128.7
167.2

En este caso la cantidad de datos es par, por lo que:

$$posición\ de\ la\ mediana = \frac{20}{2} = 10$$

Como obtuvimos un valor entero, tenemos:

$$mediana = \frac{dato\ en\ posición\ \frac{20}{2} + dato\ en\ posición\ (\frac{20}{2} + 1)}{2}$$

la mediana será el promedio de los dos valores centrales (dato en posición 10 y dato en posición 11), sustituyendo valores tenemos:

74.8	1	
74.8	2	
74.8	3	
74.8	4	
74.8	5	
74.8	6	
83.9	7	
83.9	8	
83.9	9	
83.9	10	←
83.9	11	←
83.9	12	
90.85	13	
90.85	14	
90.85	15	
102.5	16	
102.5	17	
102.5	18	
128.7	19	
128.7	20	

Sustituyendo valores en la ecuación:

$$Mediana = \frac{83.9 + 83.9}{2} = 83.9$$

En este caso, el valor de la mediana es **83.9 kg**.

La importancia de la mediana radica en el hecho de que por encima y por debajo de ese valor se distribuyen equitativamente los datos, es decir, por encima y por debajo de la mediana se encuentra en 50 % de los datos.

Moda

En ESTADÍSTICA, se conoce como MODA al dato que tiene la mayor frecuencia, es decir, el dato que más se repite en el conjunto de datos. Cabe señalar que, coloquialmente, cuando existe una canción que es escuchada por un amplio sector de la población, una prenda o color de ropa que es utilizado por la mayor parte de la gente o existe un dispositivo que usa un amplio sector de la población, se utiliza la expresión "está de moda", haciendo referencia a que es un evento de alta frecuencia en la población.

Con base en la tabla de frecuencias obtengamos la moda

Peso(x)	**Frecuencia (f)**
74.8	7
83.9	5
90.85	3
102.5	3
128.7	1
167.2	1
Total	20

Observa que para obtener la moda se debe determinar el dato con mayor frecuencia, en este caso la frecuencia más alta es de 7 y su dato asociado es de 74.8 kg. Por lo tanto, la moda = 74.8.

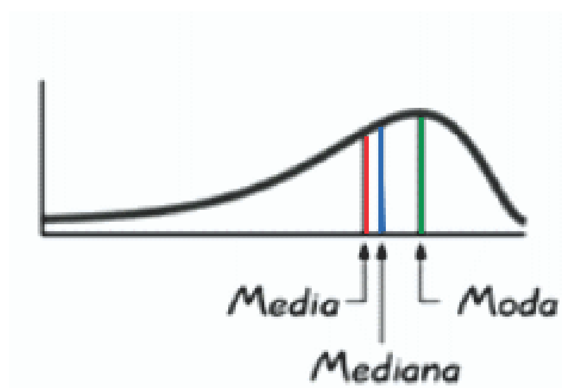
Frecuentemente podemos encontrar conjuntos de datos en los que se presentan varios datos que se repiten más que los demás, cuando esto se presenta, se dice que se tienen poblaciones o muestras MULTIMODALES, es decir, que tienen varias modas. Cuando un conjunto de datos tiene dos modas, se dice que es BIMODAL, si tiene tres, se le llama TRIMODAL, con cuatro modas tendremos un TETRAMODAL, si tiene 5 el conjunto es PENTAMODAL y así, sucesivamente. Cabe señalar que cuando se tienen más de tres modas, se llama al conjunto de datos MULTIMODAL de forma genérica.

Medidas de tendencia central y la forma de una distribución de frecuencia

Además de la utilidad que ha sido descrita para cada una de las medidas de tendencia central, las medidas de tendencia central nos permiten conocer la forma que esperaríamos de la distribución de frecuencia de un conjunto de datos al ser representado en una gráfica.

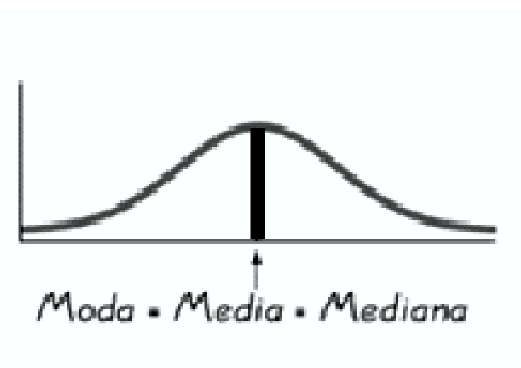
Una distribución de datos está sesgada si no es simétrica respecto a la media, es decir, si se extiende más hacia un lado que hacia el otro. Es posible conocer la forma de una distribución de datos conociendo os valores de la MODA, la MEDIA y la MEDIANA. A este respecto tendremos tres casos:

a) La media es menor que la moda y la mediana, lo cual indica que la parte más alta de la distribución de frecuencia se encuentra desplazada hacia la derecha. A una curva donde la parte más elevada se encuentra desplazada hacia la derecha se le conoce como CURVA ASIMÉTRICA NEGATIVA o CURVA CON SESGO NEGATIVO, En esta distribución los datos se concentran hacia los valores más altos; la apariencia de la gráfica tiene la forma:



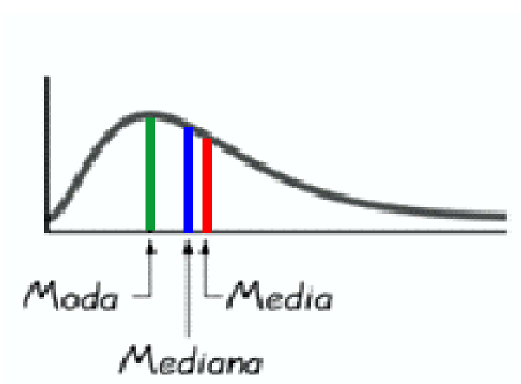
Gráfica con sesgo negativo (sesgo a la izquierda). La media y la mediana se encuentran a la izquierda de la moda

b) La media, la mediana y la moda tienden a tener el mismo valor (son prácticamente iguales), en este caso se dice que la distribución de frecuencia es simétrica, es decir, cumple con tres características: las tres medidas ocupan prácticamente la misma posición, a la derecha y ala izquierda de ellas podemos encontrar el 50 % de los datos y exactamente al centro de la distribución se localiza el dato con mayor frecuencia. Su apariencia gráfica es:



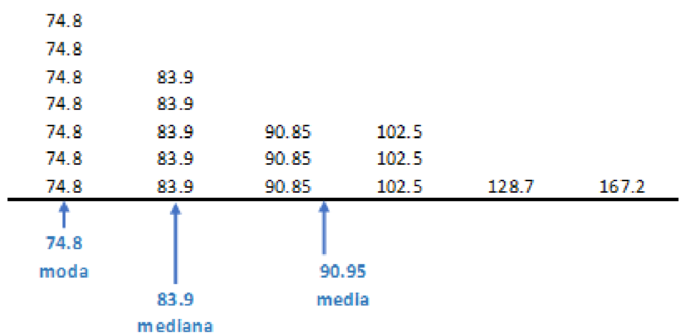
Gráfica simétrica (Sesgo neutro). La media, la moda y la mediana Son iguales o aproximadamente iguales.

c) La media es mayor que la moda y la mediana, lo cual indica que la parte más alta de la distribución de frecuencia se encuentra desplazada hacia la izquierda. A una curva donde la parte más elevada se encuentra desplazada hacia la derecha se le conoce como **CURVA ASIMÉTRICA POSITIVA o CURVA CON SESGO POSITIVO**, cuya apariencia gráfica tiene la forma:



Gráfica con sesgo positivo (Sesgo a la derecha) La media y la mediana están a la derecha de la moda

Si tomamos los datos de la Tabla 1, tenemos las tres medidas de tendencia central representadas en la siguiente gráfica:



En este caso nuestra media y mediana están a la derecha de la moda, por lo que nuestros datos están sesgados a la derecha, sesgo positivo. Note que las medidas están en donde se agrupan la mayor concentración de datos, por eso su nombre de medidas de tendencia central

Medidas de dispersión

Así como existen medidas estadísticas que nos indican la forma en que los datos se concentran en torno a un valor central de la distribución, existen medidas que nos indican la forma en que los datos se dispersan del valor de la media o de la median a; a estas medidas se les conoce como MEDIDAS DE DISPERSIÓN. Estas medidas nos indican que tan alejados están los datos respecto del centro de la distribución. Las medidas de dispersión más comúnmente utilizadas en estadística descriptiva son:

1. Rango
2. Varianza
3. Desviación estándar o desviación típica

Rango

El término Rango, en este punto del curso, nos es un concepto conocido, pues como se revisó en el tema de frecuencia de datos agrupados, el rango de un conjunto de valores es la diferencia entre el valor máximo y el valor mínimo existentes en la distribución. Si consideramos que x_máx es el valor máximo y x_mínes el valor mínimo, entonces el rango se expresa matemáticamente como:

rango=$x_{máx}$ - $x_{mín}$

El rango tiene como principal característica que para determinarlo sólo considera los valores extremos, es debido a esto que, al igual que la media, es sensible a la presencia de valores extremos (mediciones atípicas inusualmente altas o bajas).

Si recurrimos al ejemplo de Peso en Kg, tenemos los siguientes datos:

Peso en Kg	
74.8	83.9
74.8	83.9
74.8	90.85
74.8	90.85
74.8	90.85
74.8	102.5
74.8	102.5
83.9	102.5
83.9	128.7
83.9	167.2

Donde

$x_{máx}$=167.2

$x_{mín}$=74.8

Sustituyendo valores en la ecuación, tenemos:

rango=167.2 – 74.8=92.4

La distancia que separa al valor más alto del más bajo en este conjunto de dos, es decir, el rango es igual a 92.4 kg

Varianza

La varianza de un conjunto de valores es una medida de variación total del conjunto de datos respecto de la media. Como hemos podido apreciar, la variación de los datos respecto de la media se da en dos sentidos, los que son menores y los que son mayores que la media. Las diferencias menores tienen signo negativo, mientras que las mayores, tienen signo positivo.

Si sumáramos las diferencias de forma directa, se presentaría un efecto de amortiguamiento pues cantidades positivas se verían reducidas por cantidades negativas, lo que impediría apreciar la variabilidad total del conjunto de datos.

Para evitar la presencia del efecto de amortiguamiento, se elevan las diferencias de cada valor **al cuadrado**, con lo que tenemos varias ventajas:

1. El resultado SIEMPRE ES POSITIVO.
2. Las diferencias pequeñas al elevarse al cuadrado se vuelven más pequeñas.
3. Las diferencias grandes, al elevarse al cuadrado, se vuelven más evidentes.
4. Permite obtener una medición de la variabilidad total del conjunto de datos; mientras más grande sea la varianza, el conjunto de datos será más disperso. Por el contrario, si la varianza es más pequeña, el conjunto de datos tenderá a ser homogéneo y a presentar poca variabilidad.

La varianza puede obtenerse a partir de los datos de una población (en cuyo caso se representa con s^2) o a partir de los datos de una muestra (en cuyo caso se representa con una s^2).

El cálculo de la Varianza poblacional se realiza a través de la siguiente ecuación:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Dónde:

s2 = Varianza poblacional

N = Número de individuos o elementos en la población

Xi = Valor del dato en la posición i.

μ = Media poblacional

El cálculo de la varianza muestral esta dado por:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

Dónde:

s^2 = Varianza muestral

n = Tamaño de la muestra

Xi = Valor del dato en la posición i.

$\bar{X}$ = Media muestral

Pongamos un ejemplo:

Si consideramos el siguiente conjunto de datos:

Alumno	Peso (kg)
1	76
2	98
3	74
4	87
5	115
6	119
7	94
8	54
9	104
10	119

¿Cuál será la varianza?

- Primer cuestionamiento, ¿se trata de una población o una muestra?

Por el número de elementos, suponemos que se trata de una muestra a menos que exista un planteamiento o indicación de que se trata de una población.

- Segunda cuestión ¿cuál es la media y el valor de n?

n=10

Determinemos el valor de la media:

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

$$\bar{X} = \frac{76 + 98 + 74 + 87 + 115 + 119 + 94 + 54 + 104 + 119}{10}$$

$$\bar{X} == \frac{940}{10} = 94$$

La media muestral es igual a 94 kg

La fórmula a utilizar es:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

Sustituyendo valores tenemos:

$$S^2 = \frac{(76-94)^2 + (98-94)^2 + \cdots + (104-94)^2 + (119-904)^2}{(10-1)}$$

$$S^2 = \frac{(-18)^2 + (4)^2 + (20)^2 + \cdots + (-40)^2 + (10)^2 + (25)^2}{(10-1)}$$

$$S^2 = \frac{324 + 16 + 400 + 49 + 441 + 625 + 0 + 1600 + 100 + 625}{9}$$

$$S^2 = \frac{4180}{9} = 464.44$$

Por lo tanto, la varianza es S^2= 464.44 kg^2

Sin embargo, a pesar de que la varianza es una forma matemáticamente precisa de evaluar la dispersión de los datos respecto de la media, los valores obtenidos serán siempre dados en las unidades de medida de la variable elevadas al cuadrado (kg^2 en el ejemplo anterior). Esa unidad al cuadrado puede dificultar su comprensión pues carece de interpretación física en el mundo real, debido a lo anterior, se recurre a una de las medidas de dispersión más ampliamente utilizada: la desviación estándar.

Desviación estándar (desviación típica).

La desviación estándar de un conjunto de valores (muestrales) x_1, x_2, x_3,…,x_n, es una medida de variación de los valores respecto a la media aritmética. Se denota por la letra s y se calcula como:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

donde

x_i = i-ésimo valor en el conjunto de datos

$\bar{x}$ = media aritmética,

$(x_i-\bar{x})$ = valor de desviación de x_i respecto de la media

n = total de datos de la muestra.

Si comparamos la fórmula de cálculo de la desviación estándar con la de la varianza, podemos notar que son muy parecidas, la diferencia radica en que en el caso de la desviación estándar se saca la raíz cuadral resultado de la suma de las diferencias entre el total de elementos, es por ello por lo que la desviación estándar se define como la **"Raíz cuadrada de la Varianza".**

Con base en los datos calculamos la Desviación estándar,

PESO12_2
167.2
83.9
102.5
83.9
83.9
74.8
74.8
90.85
128.7
74.8
74.8
102.5
74.8
83.9
74.8
90.85
90.85
102.5
74.8
83.9

Calculando la media en forma abreviada, tenemos

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

$$\bar{x} = \frac{1\,819.05}{20} = 90.95$$

Calculando la Desviación estándar

$$s = \sqrt{\frac{1}{20-1}((167.2-90.95)^2 + \cdots + (83.9-90.95)^2)}$$

$$s = \sqrt{\frac{1}{19}(9\,713.63)} = \sqrt{511.24} = 22.61$$

Por lo tanto, la desviación estándar es S=22.61 kg

Propiedades de la desviación estándar.

1. Es la medida de dispersión o variación más importante y útil.
2. La desviación estándar adquiere valores desde 0 hasta infinito (positivos), nunca encontraremos desviaciones estándar negativas.
3. La desviación estándar es igual cero (s=0) cuando todos los valores de los datos son iguales (valores constantes).
4. Un valor grande de s implica que hay una mayor variación o dispersión.
5. El valor de s se ve afectada de forma drástica con la inclusión de valores extremos atípicos.
6. Las unidades de la desviación estándar deben ser las mismas que las de los datos originales.
7. Al tener las mismas unidades que los dato, ´se puede dar una interpretación física y directa al resultado

PRÓXIMO TEMA DE LA SERIE:

Estadística descriptiva (III): Medidas de tendencia central y de dispersión en datos agrupados

Referencias

Anderson, D. R., sweeney, D., & Williams, T. A. (1999). Estadística para la Administración y Economía. México DF, México: International Thompson Editores.

Departamento de Didáctica de la Matemática. (2011). Estadística con proyectos. (C. Batanero, & C. Díaz, Eds.) Granada, España: Facultad de Ciencias de la Educación, Universidad de Granada.

García Pérez, A. (2008). Estadística aplicada: conceptos básicos (2a edición ed.). Madrid, España: Educación permanente / Universidad Nacional de Educación a Distancia.

Kazmier, L. J., Díaz Mata, A., & Eslava Gómez, G. (1991). Estadística Aplicada a Administración y Economía. Naucalpan, Estado de México, México, Atlacomulco, México: McGraw Hill.

Pérez López, C. (1999). Control estadístico de la calidad. Madrid, España: ALFAOMEGA

Wackerly, D. D., Mendenhall III, W., & Scheaffer, R. (2010). Estadística Matemática con aplicaciones. México, D.F., México: Cengage Learning Editores, S.A

Declaración de conflicto de intereses

El autor declara no tener ningún interés comercial o asociativo que represente un conflicto de intereses en relación con el trabajo presentado.