MEDIDAS DE DISPERSIÓN

Elaborado por:
MSc. Nohora Báez Sánchez

MEDIDAS DE DISPERSIÓN

La variabilidad trata de descubrir las irregularidades que puedan existir en un conjunto de datos.

Además establece la medida en que los datos se concentran o se dispersan alrededor de este valor. Entre sus medidas están:

1. RANGO
2. VARIANZA
3. DESVIACIÓN TÍPICA
4. COEFICIENTE DE VARIACIÓN DE PEARSON

EL RANGO

Es la diferencia entre el mayor valor y el menor valor de un conjunto de datos.

POR EJEMPLO: Se tiene una muestra con los siguientes datos: 3, 10, 2, 8, 7.

El Rango corresponde a: R= 10-2 = 8

Por otro lado la **varianza** indica cuánto se alejan los datos de la media aritmética.. Es la medida de dispersión más usada en estadística.

La desviacón corresponde a la raíz de la varianza y suele usarse como medida del riesgo relacionado con una inversión en acciones o en fondos de acciones (Bussines-Week, 7 de enero de 2000). Proporciona una medida de cómo fluctúa la rentabilidad mensual respecto de la rentabilidad promedio a largo plazo.

Para datos no agrupados se emplea alguna de las siguientes fórmulas, las dos primeras para cuando se trabaja con muestras y la tercera para cuando se trabaja poblaciones:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$ **FÓRMULA 1**

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$$ **FÓRMULA 2**

$$s^2 = \frac{\sum (x_i - \mu)^2}{N}$$ **FÓRMULA 3**

Por ejemplo: obtenga la varianza y desviación típica de la siguiente muestra 6, 3, 8, 5, 3.

PASOS

PASO 1: se obtiene la media aritmética.

$$\bar{x} = \frac{6 + 3 + 8 + 5 + 3}{5} = \frac{25}{5} = 5$$

PASO 2: se obtiene las desviaciones y sus cuadrados para cada dato.

DATO	DESVIACIÓN	CUADRADO
6	6 - 5 = 1	1
3	3 - 5 = -2	4
8	8 - 5 = 3	9
5	5 - 5 = 0	0
3	3 - 5 = -2	4
TOTAL		**18**

Tabla 1. Datos para obtener desviaciones y cuadrados. Fuente: Gómez, M. (2011).

PASO 3: se reemplaza en la fórmula 1.

$$s^2 = \frac{\sum(x - \overline{x})^2}{n - 1}$$

$$s^2 = \frac{18}{5 - 1} = \frac{18}{4} = 4{,}5$$

Para la desviación, se calcula la raíz de la varianza

$$s = \sqrt{4{,}5} = 2{,}12$$

SOLUCIÓN CON LA FÓRMULA 2

PASO 1: se obtiene el cuadrado de los datos y se suma la columna de los datos y de sus cuadrados.

DATO	CUADRADO
6	36
3	9
8	64
5	25
3	9
25	**143 TOTALES**

PASO 2: se reemplaza en la fórmula 2.

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$$

$$s^2 = \frac{143 - \frac{(25)^2}{5}}{4}$$

$$s^2 = \frac{143 - \frac{625}{5}}{4}$$

$$s^2 = \frac{143 - 125}{4} = \frac{18}{4} = 4{,}5$$

FÓRMULAS PARA DATOS AGRUPADOS

$$s^2 = \frac{\sum (x - \bar{x})^2 fi}{n - 1}$$

FÓRMULA 1

$$s^2 = \frac{\sum x^2 fi - \frac{(\sum x)^2 fi}{n}}{n - 1}$$

FÓRMULA 2

$$s^2 = \frac{\sum x^2 fi - \frac{(\sum x\, fi)^2}{N}}{N}$$

FÓRMULA 3

EJEMPLO: Calcule la varianza y desviación estándar de la siguiente muestra.

CLASES	FA	XI
22 - 32	1	27
33 - 43	2	38
44 - 54	5	49
55 - 65	2	60
66 - 76	9	71
77 - 87	9	82
88 - 98	10	93
99 - 109	5	104
110 - 120	3	115
121 - 131	4	126
TOTAL	**50**	

Tabla 2. Datos para encontrar varianza y desviación. Fuente: Gómez, M. (2011).

PARA LA FÓRMULA 1

PASO 1: Obtener la media aritmética:

CLASES	FA	XI	XIFI
22 - 32	1	27	27
33 - 43	2	38	76
44 - 54	5	49	245
55 - 65	2	60	120
66 - 76	9	71	639
77 - 87	9	82	738
88 - 98	10	93	930
99 - 109	5	104	520
110 - 120	3	115	345
121 - 131	4	126	504
TOTAL	**50**		**4144**

Tabla 3. Datos para obtener media aritmética. Fuente: Gómez, M. (2011).

X= 4144/50 = 82,88

PASO 2: encontrar las desviaciones y sus cuadrados

CLASES	DEVIACIONES	CUADRADOS
22 - 32	27 - 82,88 = -55,88	23122,57
33 - 43	38 - 82,88 = -44,88	2014,22
44 - 54	49 - 82,88 = -33,88	1147,85
55 - 65	60 - 82,88 = -22,88	523,49
66 - 76	71 - 82,88 = -11,88	141,13
77 - 87	82 - 82,88 = -0,88	0,77
88 - 98	93 - 82,88 = 10,12	102,41
99 - 109	104 - 82,88 = 21,12	446,05
110 - 120	115 - 82,88 = 32,12	1031, 69
121 - 131	126 - 82,88 = 43,12	1859,33
TOTAL		**10389,44**

Tabla 4. Datos para encontrar desviaciones y cuadrados. Fuente: Gómez, M. (2011).

PASO 3: multiplicar cada cuadrado por su frecuencia.

CLASES	FA	CUADRADOS	(XI - X)^2 FI
22 - 32	1	23122,57	3122,57
33 - 43	2	2014,22	4028,43
44 - 54	5	1147,85	5739,27
55 - 65	2	523,49	1046,99
66 - 76	9	141,13	1270,21
77 - 87	9	0,77	6,97
88 - 98	10	102,41	1024,14
99 - 109	5	446,05	2230,27
110 - 120	3	1031, 69	3095,08
121 - 131	4	1859,33	7437,34
TOTAL		**10389,44**	**29001,28**

Tabla 5. Obtención de (Xi-X)^2 Fi. Fuente: Gómez, M. (2011).

PASO 4: Se reemplaza en fórmula 1

$$s^2 = \frac{\sum (x - \bar{x})^2 fi}{n - 1}$$

$$s^2 = \frac{29001,28}{50 - 1} = 591,86$$

SOLUCIÓN CON FÓRMULA 2

Paso 1: encontrar x^2, x^2f, xf

CLASES	FA	XI	XIFI	X^2	X^2F
22 - 32	1	27	27	729	729
33 - 43	2	38	76	1444	2888
44 - 54	5	49	245	2401	12005
55 - 65	2	60	120	3600	7200
66 - 76	9	71	639	5041	45369
77 - 87	9	82	738	6724	60516
88 - 98	10	93	930	8649	86490
99 - 109	5	104	520	10816	54080
110 - 120	3	115	345	13225	39675
121 - 131	4	126	504	15876	63504
TOTAL	**50**		**4144**	**68505**	**372456**

Tabla 6. Datos para encontrar x^2, x^2f,xf. Fuente: Gómez, M. (2011).

PASO 2: reemplazar los valores en la fórmula 2

$$s^2 = \frac{\sum x^2 fi - \frac{(\sum x\, fi)^2}{n}}{n-1}$$

$$s^2 = \frac{372456 - \frac{4144^2}{50}}{50-1} = \frac{372456 - \frac{17172736}{50}}{49}$$

$$s^2 = \frac{372456 - 343454{,}7}{49} = \frac{29001{,}28}{49} = 591{,}86$$

CUALQUIERA DE LAS FÓRMULAS QUE SE UTILICEN PARA OBTENER LA VARIANZA, DARÁ EN MISMO RESULTADO, COMO SE PUDO OBSERVAR CON LOS EJEMPLOS REALIZADOS TANTO PARA DATOS AGRUPADOS COMO PARA DATOS SIN AGRUPAR.

COEFICIENTE DE VARIACIÓN DE PEARSON (CV)

Indica la importancia de la desviación estándar en relación con el promedio aritmético. Representa el número de veces que la desviación típica contiene a la media aritmética. Por lo tanto, cuanto mayor es el CV, mayor es la dispersión y menor la representatividad de la media.

$$C.V.(x) = \frac{s}{R} * 100$$

$$C.V.(x) = \frac{s}{\mu} * 100$$

EJEMPLO: Un conjunto de valores dado por una muestra tiene como promedio 36,5 y su desviación estándar es 28,98. Determine el coeficiente de variación.

$$C.V.(x) = \frac{28{,}98}{36{,}5} * 100 = 79{,}40$$

En este caso el coeficiente es bastante alto, por lo que la muestra no es representativa con respecto a la media, por lo que para tomar decisiones debe tener en cuenta otra muestra o buscar la representatividad con otra medida.

PROBABILIDAD

Corresponde a un valor numérico que debe cumplir con ciertas condiciones o propiedades matemáticas. Se asocia a un evento o suceso determinado para expresar el grado de confianza que se tiene en la verificación futura de dicho evento.

EJEMPLO: Considérese los resultados que se obtienen al lanzar un dado. Intuitivamente la probabilidad correspondiente a cada uno de los puntos es la misma e igual a 1/6.

DEFINICIÓN CLÁSICA

Los matemáticos Pascal, Fermat y Laplace formularon una primera definición de la probabilidad, basada en los juegos de azar.

La definición clásica de la probabilidad indica:

> **Si un suceso puede ocurrir de n maneras mutuamente excluyentes e igualmente posibles, y si n(a) de ellas poseen un atributo a, la probabilidad de a es: p(a)=n(a)/n. (Gómez, 2011, p.30).**

PROPIEDADES

- La probabilidad es un número positivo o nulo: P(A) 0.
- La suma de las probabilidades correspondientes a cada uno de los eventos simples es igual a 1:

$$P(A1)+ P(A2)+ \dots +P(An) =1.$$

Siguiendo el ejemplo de los dados 1/6+1/6+1/6+1/6+1/6+1/6=1

- Cuando el evento es imposible: P(A)=0 Por ejemplo, que nieve en Costa Rica.
- Si el evento es seguro o cierto: P(A)=1.

EJEMPLO: Se lanza un dado, compruebe que la probabilidad total es igual a 1:

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6)$$

$$1/6 + 1/6 + 1/6 + 1/6 + 1/6 +1/6 = 1$$

Calcule la probabilidad de obtener menos de 4:

$$P(1) + P(2) +P(3) = 1/6 +1/6 +1/6 = 3/6 = 1/2$$

Si se lanzan dos dados, uno de color rojo (X) y otro de color azul (Y) y se toma la suma de puntos. Para ello, se debe obtener el espacio muestral:

El espacio muestral corresponde a cada una de las posibles combinaciones obtenidas; para este caso:

(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)

La probabilidad de que X + Y = 8

Las posibilidades son: (2,6), (3,5), (4,4), (5,3), (6,2)

Por lo tanto, la probabilidad es 5/36.

LEY DE LA SUMA

P(A o B) = P(A) + P(B) – P(AB)

P(A o B): Probabilidad de que sucede alguno de los eventos o los dos

P(A) : Probabilidad de que suceda A

P(B) : Probabilidad de que suceda B

P(AB): Probabilidad de que sucedan A Y B

EJEMPLO: Siguiendo con los dados

Evento A: X+Y = número par

Evento B: X+Y <5

Probabilidad de A: 18/36

Probabilidad de B: 6/36

Probabilidad de AB: 4/36

P(AoB)= 18/36 + 6/36 – 4/36= 20/36

Cuando los eventos son mutuamente excluyentes, P(AB)=0. Por lo que

P(AoB) = P(A)+ P(B)

EJEMPLO: Al lanzar un dado, ¿cuál es la probabilidad de obtener 3 ó 5?

P(A o B)= P(A) + P(B)

= 1/6 + 1/6 = 2/6

COMPLEMENTO

Formado por todos los eventos que no cumplen el atributo.

P(A)= 1- P(-A)

PROBABILIDAD CONDICIONAL

Se presenta cuando un evento sucede, porque otro ya sucedió.

P(B/A)= P(AB)/P(A)

EJEMPLO: Se lanzan dos dados y la suma de sus caras resulta par. ¿Cuál es la probabilidad de que la suma sea menor que 5?

Evento A: suma par

Evento B: suma menor que 5

Se desea saber P(B/A)

Espacio muestral nuevo= 18 puntos

Cantidad de puntos con suma menor que 5: (1,3), (1,1), (2,2), (3,1)

Por lo tanto P(A/B) = 4/18

Otra forma: espacio muestral= 36 PUNTOS

P(B/A) = 4/36 = 4/18= 2/9

18/36

REGLA DEL PRODUCTO

B= evento bola blanca

N= evento bola negra

P(BN)=P(B) P(N/B)

P(B)= 7/12

P(N/B) = 5/11

P(N/B) = 7/12 * 5/12 = 35/144

REFERENCIAS BIBLIOGRÁFICAS

Gómez, M. (2011). *Elementos de Estadística Descriptiva*. (24ª Reimpresión). Costa Rica: UNED.