2 HERRAMIENTAS DE ESTADÍSTICA

La estadística puede ser considerada como un método para tratar datos numéricos. Es la rama de las matemáticas que se ocupa de recoger, analizar y extraer información relevante y útil del conjunto de datos obtenidos; esta información aparece en forma de números, porcentajes, o a través de gráficos [1].

2.1 Media aritmética

La media es la suma de los valores de una variable dividida por el número de datos que se tiene. Expresada en forma algebraica [2,3]:

$$\mu = \frac{(X_1 + X_2 + \ldots + X_N)}{N} = \sum_{X1}^{XN} \frac{X}{N} \tag{1}$$

Donde: μ= Media aritmética

N= Num. de valores

∑= Verbo matemático que nos ordena sumar

2.2 Mediana

En distribuciones de frecuencias agrupadas, la mediana es un valor que divide a la distribución en dos partes iguales. Supongamos los siguientes dos casos [2,3]:

Caso I. Supongamos 4 personas con los ingresos siguientes: \$5, \$6, \$7, \$8.

En casos como este, se suele considerar como mediana el valor medio (6+7)/2= 6.5.

Caso II. Supongamos 5 personas con los siguientes ingresos \$2, \$3, \$4, \$5, \$6.

En este grupo la mediana es \$4.

2.3 Desviación típica o estándar

La desviación estándar representa dispersión de valores por lo que la variabilidad de diferentes distribuciones puede compararse en términos de la desviación estándar. Para disponer de una medida de la dispersión para los valores se encuentra la raíz cuadrada de la varianza que se llama desviación típica de la variable o desviación estándar. Ahora si:

$$\mathrm{var}\,ianza(X) = \sum_{X1}^{XN} \frac{(X-\mu)^2}{N} \qquad (2)$$

Entonces desviación Estándar=Varianza (X) [3,5]. Expresada en forma algebraica:

$$\sigma = \sqrt{\frac{\sum_{X1}^{XN} (X-\mu)^2}{N}} \qquad (3)$$

2.4 Tipos de error

Sea *x* una medida [6].

- Error máximo de medición = $x - X = \Delta x$ = Δx, donde: *x*= medida real
  X= medida aproximada
- Error relativo = $\frac{\Delta x}{x}$
- Error porcentual = $\frac{\Delta x}{x}(100\%)$

2.5 Mínimos cuadrados

El método de mínimos cuadrados consiste en encontrar la única curva que posea la propiedad de que la suma de los cuadrados de las desviaciones de los valores observados de X con respecto a los valores correspondientes sobre la curva (valores calculados o estimados) sea mínima [3, 4].

Para la ecuación de la recta los mínimos cuadrados se aplican como lo indica la siguiente ecuación:

$$y = mx + b \tag{4}$$

Dados m y b, la desviación puede ser calculada como sigue:

$$\Delta y_i = y_i - mx - b \tag{5}$$

Para cualquier valor de $x=x_i$, la probabilidad PP_i para obtener la medición observada y_i con distribución gausseana y desviación estándar σ_i las cuales su valor real $y(x_i)$ es:

$$PP_i = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left[-\frac{1}{2}\left[\frac{y_i - y(x_i)}{\sigma_i} \right]^2 \right] \tag{6}$$

Por lo tanto para N valores de y_i, tenemos:

$$PP(m,b) = \prod_i^N PP_i = \prod_i^N \left(\frac{1}{\sigma_{i\sqrt{2\Pi}}} \right) \exp\left[-\frac{1}{2}\sum_i^N \left[\frac{y_i - y(x_i)}{\sigma_i} \right]^2 \right] \tag{7}$$

Utilizando el método de mínimos cuadrados, la suma de dicho cuadrados queda de la siguiente forma:

$$\overline{X^2} = \sum_i^N \frac{[y_i - y(x_i)]}{\sigma_i} = \sum_i^N \frac{1}{\sigma_i^2}\left[(y_i - mx_i - b)^2 \right] \tag{8}$$

Derivamos parcialmente la ecuación respecto a m y b:

$$\frac{\partial \overline{X}^2}{\partial m} = 0 = \frac{\partial}{\partial m}\left[\frac{1}{\sigma^2}\sum\left(y_i - mx_i - b\right)^2\right] = \frac{-2}{\sigma^2}\sum x_i\left(y_i - mx_i - b\right) \tag{9}$$

$$\frac{\partial \overline{X}^2}{\partial b} = 0 = \frac{\partial}{\partial b}\left[\frac{1}{\sigma^2}\sum\left(y_i - mx_i - b\right)^2\right] = \frac{-2}{\sigma^2}\sum\left(y_i - mx_i - b\right) \tag{10}$$

El coeficiente de correlación lineal mide la fuerza de la relación entre las variables. El coeficiente tiene el signo que tiene b y su valor estará entre -1≤ $\overline{R}_{cc}$ ≥1. El coeficiente se obtiene como a continuación se describe:

Sean $m = \frac{1}{m'}$, $b = -\frac{b'}{m'}$ y $\overline{\sigma_2}' = \frac{N\sum x_i y_i - \sum x_i y_i}{N\sum y_i^2 - (\sum y_i)^2}$; entonces:

$$\overline{R}_{cc} = \sqrt{mm'} = \frac{N\sum x_i y_i - \sum x_i y_i}{\left[N\sum x_i^2 - (\sum x_i)^2\right]^{\frac{1}{2}}\left[N\sum y_i^2 - (\sum y_i)^2\right]^{\frac{1}{2}}} \tag{11}$$