1

Introducción y estadística descriptiva

1.3 Medidas numéricas descriptivas

Definición 1.1 La media de las observaciones $x_1, x_2, \ldots, x_n$ es el promedio aritmético de éstas y se denota por

$$\overline{x} = \sum_{i=1}^{n} \frac{x_i}{n} \tag{1.1}$$

Definición 1.2 La mediana de un conjunto de observaciones es el valor para el cual, todas las observaciones se ordenan de manera creciente, la mitad de éstas es menor que este valor y la otra mitad mayor.

Definición 1.3 La moda de un conjunto de observaciones es el valor de la observación que ocurre con mayor frecuencia en el conjunto.

$$\overline{x} = \sum_{i=1}^{k} \frac{f_i x_i}{n} \tag{1.2}$$

$$Mediana = L + c\left(\frac{j}{f_m}\right) \tag{1.3}$$

Definición 1.4 La varianza de las observaciones $x_1, x_2, \ldots, x_n$ es, en esencia, el promedio del cuadrado de las distancias entre cada observación y la media del conjunto de observaciones. La varianza se denota por

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1} \tag{1.4}$$

Definición 1.5 La raíz cuadrada positiva de la varianza recibe el nombre de la desviación estándar y se denota por

$$s = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}} \tag{1.5}$$

El uso de la ecuación (1.4) puede dar origen a errores grandes por redondeo. Con un poco de álgebra se obtiene, a partir de (1.4), una fórmula computacional más exacta para esas condiciones:

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1} = \frac{\sum \left(x_i^2 - 2\overline{x}x_i + \overline{x}^2\right)}{n-1} = \frac{\sum x_i^2 - 2\overline{x}\sum x_i + n\overline{x}^2}{n-1} = \frac{\sum x_1^2 - \frac{2\left(\sum x_i\right)\left(\sum x_1\right)}{n} + \frac{n\left(\sum x_i\right)^2}{n^2}}{n-1} =$$

$$s^2 = \frac{\sum\limits_{i=1}^{n} x_i^2 - \frac{\left(\sum\limits_{i=1}^{n} x_i\right)^2}{n}}{n-1} \tag{1.6}$$

$$s = \sqrt{\frac{\sum\limits_{i=1}^{n} x_i^2 - \frac{\left(\sum\limits_{i=1}^{n} x_i\right)^2}{n}}{n-1}} \tag{1.7}$$

Para datos agrupados, puede calcularse el valor aproximado de la varianza mediante el uso de la fórmula

$$s^2 = \frac{\sum\limits_{i=1}^{k} f_i(x_i - \overline{x})^2}{n-1} \tag{1.8}$$

ó

$$s^2 = \frac{\sum\limits_{i=1}^{k} f_i x_i^2 - \frac{\left(\sum\limits_{i=1}^{k} f_i x_i\right)^2}{n}}{n-1} \tag{1.9}$$

La fórmula para la desviación estándar es

$$s = \sqrt{\sum_{i=1}^{k} \frac{f_i(x_i - \overline{x})^2}{n-1}} \tag{1.10}$$

Definición 1.6 La desviación media es el promedio de los valores absolutos de las diferencias entre cada observación y la media de las observaciones. La desviación media está dada por

$$D.M. = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n}$$

Para datos agrupados, el valor de la desviación media se aproxima por

$$D.M. = \frac{\sum_{i=1}^{k} f_i |x_i - \overline{x}|}{\sum_{i=1}^{k} f_i} \tag{1.11}$$

Definición 1.7 La desviación mediana es el promedio de los valores absolutos de las diferencias entre cada observación y la mediana de éstas. Esta dada por:

$$D.M. = \frac{\sum_{i=1}^{n} |x_i - D.Md|}{n} \tag{1.12}$$