# My first Latex files

SC

August 28, 2022

## 1 Text 1

Let $x$ be a variable that can take the values: 1, 5, 2, 5, 4, i.e $x_1 = 1, x_2 = 5, x_3 = 2, x_4 = 5, x_5 = 4$. The sum of these observations is given by $S = x_1 + x_2 + x_3 + x_4 + x_5 = 1 + 5 + 2 + 5 + 4 = 17$. The arithmetic mean is given by $\bar{x} = \dfrac{x_1 + x_2 + x_3 + x_4 + x_5}{5}$. A short hand way of adding these terms is to use the summation sign ($\Sigma$). Thus we can write $S = \sum_{i=1}^{5} x_i$ and $\bar{x} = \dfrac{S}{5}$. The inline summation sign is not very visually appealing. A better way of presentation is to write it as $S = \displaystyle\sum_{i=1}^{5} x_i$.
Median is the middle most value in the ordered set. The ordered data is written as $x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq x_{(4)} \leq x_{(5)}$. Clearly $x_{(1)} = x_1, x_{(2)} = x_3, x_{(3)} = x_4, x_{(4)} = x_2, x_{(5)} = x_4$. Here median is $x_{(3)} = 4$.
Mode is that value of $x$ which is repeated the maximum number of times. Here 5 is repeated twice and all other values are repeated once. So mode=5.

## 2 Text 2

Two distributions may be close in their central tendency (location), but markedly different in their scatter. Among the different measures of dispersion we have Range (**R**), standard deviation (**S**), mean deviation (**MD**) and quartile deviation (**QD**). Given a set of observations $x_1, x_2, .., x_n$ we define these measures.

$$R = x_{(n)} - x_{(1)} \tag{1}$$

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

$$MD_A = \frac{1}{n}\sum_{i=1}^{n}|x_i - A| \tag{3}$$

$$QD = (Q_3 - Q_1)/2 \tag{4}$$

Here in $MD_A$ , $A$ is any constant specifically an appropriate measure of central tendency and in QD, $Q_3$, $Q_1$ are the first and third quartiles respectively, i.e $Q_3$ is that value of the variable $x$ below which 75% of the observations lie. Note that $Q_1$, $Q_2$ and $Q_3$ divide the entire distribution into four equal parts , area of each part being 25%. Therefore $Q_2$ is in fact the median.