Measure of variation / Dispersion :→

In addition to locating the center of the observed values of the variable in the data, another important aspect of a descriptive study of the variable is numerically measuring the extent of variation about the center. Two data sets of the same variable may exhibit similar positions of center but may be remarkably different with respect to variability.

Just as there are several different measure of center, there are also several different measure of variation are used mostly only for quantitative variables.

Range :-

The simple range is obtained by computing the difference b/w the largest observed value of the variable in a data set and smallest one.

or

The simple range of the variable is the difference b/w its maximum and minimum values in a data set.

$$\text{Range} = \text{max} - \text{min}.$$

The sample range of the variable is quite easy to compute. H/w, in using the range, a great deal of information is ignored. i.e, only the largest and smallest values of the variable are considered; the other value are disregarded. range always increase as the data added.

Ex:- 7 participant in bike race had the following finishing times in minutes 28, 22, 26, 29, 21, 23, 24

What is the range?

$$\text{range} = 29 - 21 = 8$$

Ex:- when participant are 8; 28, 22, 26, 29, 21, 23, 24, 50

Quartiles:- Let n denote the number of observations in a data set. Arranged the observed values of variable in a data in increasing order

1. The first quartile Q_1 is at Position $\frac{n+1}{4}$,
2. The second quartile Q_2 (the Median) is at Position $\frac{n+1}{2}$
3. The Third quartile Q_3 is at Position $\frac{3(n+1)}{4}$

in the ordered list.

Next we define the Sample interquartile range. Since the interquatile range is defined using quartiles, it is Befered Measure of variation when the Median is used as the Measure of center (i.e in case of Skewed distribution).

or

Definition:- The Sample interquartile range of the variable, denoted IQR, is the difference b/w the first and third quartiles of the variable, that is

$$IQR = Q_3 - Q_1$$

Roughly speaking, the IQR gives the range of the middle 50% of the observed values.

Ex:- Find IQR of 28, 22, 26, 29, 21, 23, 24

→ 21, 22, 23, 24, 26, 28, 29

Q_1 Q_2 Q_3

$IQR = Q_3 - Q_1 = 28 - 22 = 6$

Find IQR of 28, 22, 26, 29, 21, 23, 24, 50

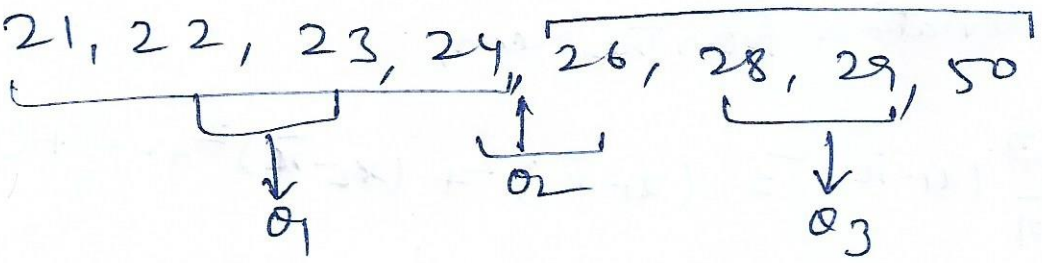
$IQR = Q_3 - Q_1 \Rightarrow 28.5 - 22.5$

$Q_1 = \frac{22+23}{2}$

$Q_2 = \frac{24+26}{2}$

$Q_3 = \frac{23+24}{2}$

Ascending order



Five Number Summary :→

The five number summary of the variable consists of Minimum, maximum, and quartiles written in increasing order:

$$Min, Q_1, Q_2, Q_3, Max$$

Standard deviation :—

The sample standard deviation is the most frequently used measure of variability, it can be considered as a kind of average of the absolute deviations of observed values from the mean of the variable.

Definition :— for a variable x, the sample standard deviation denoted by s, is

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

Since the standard deviation is defined using the sample mean $\bar{x}$ of the variable x, it is preferred measure of variation when the mean is used as the measure of center. Note that the standard deviation is always positive number i.e. $s \geq 0$

In a formula of the standard deviation, the sum of the square deviations from the mean,

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2$$

is called sum of squared deviations and provides a measure of total deviation from the mean for all the observed values of the variable. once the sum of squred deviations is divided by $n-1$, we get

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

which is called the sample variance or variance. the sample standard deviation has the following more formulas

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

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \left\{\left(\sum_{i=1}^{n} x_i\right)^2 / n\right\}}{n-1}}$$

Ex:— 7 participants in bike race had the following times in minutes 28, 22, 26, 29, 21, 23, 24 what is SD?

Sol:— $\bar{x} = \frac{173}{7} = 24.7$

$$\Rightarrow \sqrt{\frac{1}{6}\left(28-24.7\right)^2 + (22-24.7)^2 + (26-24.7)^2 + (29-24.7)^2 + (21-24.7)^2 + (23-24.7)^2 + (24-24.7)^2}$$