Measures of Dispersion

Dispersion: Dispersion indicates the measures the extent to which individual items differs. It indicates lack of uniformity in the size of items. It is also called variation or variability. The measures of dispersion are called average of second order.
Following are the measures of dispersion:

A. Absolute Measures:

1. Range
2. Mean Deviation(MD)
3. Standard Deviation(SD)
4. Quartile Deviation(QD)

B. Relative Measures:

1. Coefficient of range
2. Coefficient of mean deviation
3. Coefficient of Standard Deviation
4. Coefficient of quartile deviation

A. Absolute Measure: When we measure the dispersion and express it in terms of original data, it is called an absolute measure.

Range:

Range is the difference between highest and lowest observation in a set of data.

Calculation of Range:

Ungrouped Data: Let, $x_1, x_2, ..., x_n$ are n observations. x_H is the highest value and x_L is the lowest value.
So, Range $= x_H - x_L$

Grouped Data: Let, $x_1, x_2, ..., x_n$ are the middle value of the class intervals and their corresponding frequencies are $f_1, f_2, ..., f_n$.Then-
Range=Upper limit of last class interval –lower limit of first class interval.

Example: Find range in the series:
80, 90, 60, 63, 68, 61, 67, 65, 100, 75, 89, 84, 86
Uses: The range is most widely used in statistical process control (SPC) applications because it is very easy to calculate and understand.

Standard deviation:

The arithmetic mean of the square deviations of observations from their arithmetic mean is known as variance. The positive square root of variance is known as standard deviation.

Calculation of standard deviation:

For ungrouped sample data: Let $x_1, x_2, ..., x_n$ be n observations with mean $\bar{x}$, then standard deviation is given by-

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

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

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

For grouped sample data:

Let, $x_1, x_2, ..., x_n$ are the middle value of the class intervals and their corresponding frequencies are $f_1, f_2, ..., f_n$ and their mean is $\bar{x}$ .Then standard deviation is given by-

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

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

Population Standard deviation: $\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$,

Where population mean, $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$

Population variance, $\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$

Advantages: i) It utilizes all observations.

ii) It is less erratic.

iii) Suitable foe algebraic manipulation.

iv) It is used for higher statistical operations.

v) Less affected by fluctuation of sampling

Disadvantages:

i) Its calculation demands greater time and labour.

ii) It gives greater weights to extreme items.

Uses of Standard Deviation:

As a measure of dispersion standard deviation is the most important. An important application of SD is to judge how much representative the AM is as a measure of central tendency. Thus if two or more comparable distribution have the same mean, Then the mean is the most representative in the distribution with smallest SD. It is useful for locating the place of an individual item in the distribution. It is a key stone in sampling and correlation and also used in interpretation of curves.

Quartile Deviation:

Quartile deviation or semi-interquartile range is obtained by dividing the difference between the upper and lower quartile by 2.

i.e. Quartile Deviation or Q.D= $\frac{Upper quartile - lower quartile}{2}$

$$= \frac{Q_3 - Q_1}{2}$$

Where, **For ungrouped data-**

$Q_1 = \frac{n+1}{4}$ th observation in a series arranged in ascending order

$Q_3 = \frac{3(n+1)}{4}$ th observation in a series arranged in ascending order.

For grouped Data-

First quartile group= size of $\frac{N}{4}$ th item

$Q_1 = L_1 + \frac{N/4 - c.f}{f_1} \times c$, Where L_1 =lower limit of the first quartile group

N =Total number of observation

$c.f$ = Cumulative frequency of the group just preceding the 1st quartile group.

f_1 =Frequency of the first quartile group

$c =$ Class interval

Third quartile group= size of $\frac{3N}{4}$ th item

$Q_3 = L_3 + \frac{3N/4 - c.f}{f_3} \times c$, Where L_3 =lower limit of the Third quartile group

N =Total number of observation

$c.f$ = Cumulative frequency of the group just preceding the 3rd quartile group.

f_3 =Frequency of the third quartile group

$c =$ Class interval

Second quartile, Q_2 =Median

Relative Measure: Absolute measure cannot be used for comparison purposes if expressed in different units. In order to compare two series, if absolute measure of dispersion of each series is expressed as a ratio or percentage of the average then itg is called relative measure of dispersion.
Now we briefly show the measure as follows:

Coefficient of range:

For ungrouped data:

Coefficient of range= $\frac{Highestvalue - lowestvalue}{Highestvalue + lowestvalue}$ (if percentage wanted then multiply by 100)

For grouped data:
Coefficient of range=(upper limit of last class interval-lower limit of first class interval)/ (upper limit of last class interval +lower limit of first class interval)

Coefficient of standard deviation or coefficient of variation:

The ratio of the standard deviation to the arithmetic mean expressed as a percentage.
It is a very useful measure when-

i) The data are in different units (such as dollar and days absent)

ii) The data are in the same unit, but the mean are far apart.(The income of the top executives and the income of the unskilled employees).

Now- Coefficient of variation, C.V= $\frac{s}{\bar{x}} \times 100$ (for sample)

C.V= $\frac{\sigma}{\mu} \times 100$ (for sample) Where, σ =population standard deviation

Population mean, $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$

Coefficient of Quartile Deviation:

Coefficient of Quartile Deviation= $\frac{Q_3 - Q_1}{Q_3 + Q_1}$