Measures of Dispersion, Moments and Skewness

There are possibilities that while finding out the average (mean, median, mode) of two or more than two data sets, we may have the same result for the average (mean, median, mode), but their individual observations may differ considerably from the average. It means that a value of central tendency does not adequately describe the data.
We need some additional information concerning with how the data are dispersed about the average.
This is done by measuring the dispersion by which we mean the extent to which the observations in a sample or population vary about their mean.
A quality that measures this characteristics is called a "measure of dispersion, scatter or variability."
The main measures of dispersion are as follows.

i) The Range
ii) The semi-Interquartile Range or Quartile Deviation
iii) The Mean Deviation or the Average Deviation
iv) The Variance and Standard Deviation

1) The Range:

The Range "R" is defined as the difference between the largest and the smallest observation in a data set.
Symbolically range is given by

$$R = x_m - x_0$$

x_m = largest value

x_0 = Smallest value

R = Range

Page.

Disadvantages of Range:

→ It ignores all the information available from the intermediate observations.

→ Its value is based only on two extreme observations ie usually large and small

Co-efficient of Dispersion:

Co-efficient of Dispersion is given by

$$\text{Co-efficient of Dispersion} = \frac{x_m - x_0}{x_m + x_0} \quad \text{OR}$$

$$\text{Co-efficient of Range} = \frac{x_m - x_0}{x_m + x_0} \times 100$$

Example: The marks obtained by 9 students are given below

45, 32, 37, 46, 39, 36, 41, 48, 36

Find Range and Co-efficient of dispersion.

Sol:=

$X_m = 48 \qquad x_0 = 32$

$$R = x_m - x_0$$
$$= 48 - 32$$
$$= 16$$

$$\text{Co-efficient of Dispersion} = \frac{x_m - x_0}{x_m + x_0}$$
$$= \frac{48 - 32}{48 + 32}$$
$$= \frac{16}{80}$$
$$= 0.2$$

2) The Semi- InterQuartile Range :

The interquartile range or sometimes called the quartile deviation is a measure of dispersion defined by the difference between the third and first quartiles; and half of this range is called the semi-interquartile Range.

Symbolically

$$Q.D = \frac{Q_3 - Q_1}{2}$$

Q_1 = First Quartile

Q_3 = Third Quartile

Its relative measure is called the Co-efficient of the Quartile Deviation or of Semi-Interquartile Range and is given by

$$\text{Co-efficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

OR

$$\text{Co-efficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

Example:

Find the quartile deviation and the co-efficient of quartile deviation

45, 32, 37, 46, 39, 36, 41, 48, 36

Sol => First arrange the date to find the Quartiles

32, 36, 36, 37, 39, 41, 45, 46, 48

$$Q_1 = \left[\frac{n}{4}+1\right]^{th} = \left[\frac{9}{4}+1\right]^{th} = \left[\frac{13}{4}\right]^{th} = [3.25]^{th}$$

$$\Rightarrow [3] = 36$$

$$Q_3 = \left[\frac{3n}{4}+1\right]^{th}$$

$$= \left[\frac{3\times 9}{4}+1\right]^{th}$$

$$= \left[\frac{27+4}{4}\right]^{th}$$

$$= \left[\frac{31}{4}\right]^{th}$$

$$= [7.75]^{th}$$

$$= 7^{th}$$

$$= 45$$

$$Q.D = \frac{45-36}{2} \Rightarrow \frac{Q_3 - Q_1}{2}$$

$$= 4.5$$

$$\text{Co-efficient of } Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{45-36}{45+36} = \frac{9}{81} = 0.11$$

OR

$$\frac{9}{81} \times 100 = 0.11 \times 100 \Rightarrow 11.11$$

The Mean Deviation:

The mean deviation of a set of data is defined as the arithmetic mean of the deviations measured either from the mean or from the median, all deviations being counted as positive. The reason to count the deviations as positive is to avoid the difficulty arising from the property that the sum of deviations of the observations from their mean is zero.

For Grouped Data:

When the data are grouped into a frequency distribution having K classes with midpoints $x_1, x_2, \ldots, x_k$ and the corresponding frequencies $f_1, f_2, \ldots, f_k$, the sample variance and standard deviation are given by

$$s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{n} \quad \text{and}$$

$$s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{n}}$$

where as, Co-efficient is given by

$$\text{Co-efficient of S.D} = \frac{\text{Standard Deviation}}{\text{Mean}} \text{ OR } \frac{\text{Standard Deviation}}{\text{Mean}} \times 100$$

$$= \frac{\delta}{A.M} \text{ or } \frac{\delta}{A.M} \times 100$$

Corresponding Formula's: (Ungrouped)

Sample Variance $s^2 = \frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2$

Standard Deviations $\delta = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x_i}{N}\right)^2}$

$$s = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

Grouped:

Sample variance $s^2 = \frac{\sum f x^2}{n} - \left(\frac{\sum f x}{n}\right)^2 \quad \therefore \sum f = n$

Standard Deviation $\delta = \sqrt{\frac{\sum f x^2}{n} - \left(\frac{\sum f x}{n}\right)^2} \quad \therefore n = \sum f$

Example:

Consider the following observations 7, 8, 10, 13, 14, 19, 20, 25, 26 and 28. Find variance and standard deviation. The data collected is of population having size N=10

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} = \frac{534}{10} = 53.4$$

and

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} = \sqrt{53.4} = 7.31$$

x_i	$x_i - \mu$	$(x_i - \mu)^2$	x_i^2
7	7−17 = −10	100	49
8	8−17 = −9	81	64
10	10−17 = −7	49	100
13	13−17 = −4	16	169
14	14−17 = −3	9	196
19	19−17 = 2	4	361
20	20−17 = 3	9	400
25	25−17 = 8	64	625
26	26−17 = 9	81	676
28	28−17 = 11	121	784

$$\mu = \frac{\sum x}{N} = \frac{170}{10} = 17 \qquad \sum (x_i - \mu)^2 = 534 \qquad \sum x_i^2 = 3424$$

OR

Using Alternative method

$$\sigma^2 = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2$$

$$\left(\frac{3424}{10}\right) - \left(\frac{170}{10}\right)^2 = 342.4 - 289 = 53.4$$

and

$$\sigma = \sqrt{\frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2} = \sqrt{53.4} = 7.31$$