Measures of dispersion

❖ Concept of Dispersion

Measures of central tendency help us to represent the entire mass of the data by a single value.
Can the central tendency describe the data fully and adequately?
In order to understand it, let us consider an example.
The daily income of the workers in two factories are :
Factory A : 35 45 50 65 70 90 100
Factory B : 60 65 65 65 65 65 70
Here we observe that in both the groups the mean of the data is the same, namely, 65.

(i) In group A, the observations are much more scattered from the mean.
(ii) In group B, almost all the observations are concentrated around the mean.

Certainly, the two groups differ even though they have the same mean.
Thus, there arises a need to differentiate between the groups. We need some other measures**,** which concern with the measure of scatteredness (or spread).
To do this, we study what is known as measures of **dispersion**.

❖ Dispersion

Dispersion means the scatteredness of observations from some central values.

❖ Types of Measures of Dispersion

There are two main types of dispersion method, which are

- Absolute measures of dispersion
- Relative measures of dispersion

➢ Absolute Measures of Dispersion

An absolute measure of dispersion contains the same unit as the original data set. Absolute dispersion method expresses the variations in terms of the average of deviations of observations. Such measures are

1. Range
2. Mean deviation
3. Variance
4. Standard deviation
5. Quartile deviation

➢ Relative Measures of Dispersion

The relative measures of depression are used to compare the distribution of two or more data sets. This measure compares values without units. Some common relative measures of dispersion are

1. Coefficient of range
2. Coefficient of mean deviation
3. Coefficient of variation
4. Coefficient of standard deviation
5. Coefficient of quartile deviation

❖ Range

Range is the difference between highest value (H) and lowest value (L) in a data set.
i.e. Range = Highest value – Lowest value
Example: 2, 3, 9, 8   Range = 9-2 = 7.

➢ Coefficient of Range

Coefficient of range is defined as
Coefficient of range = $\frac{H-L}{H+L} \times 100$; where, H = Highest value and L = Lowest value.
Example: 2, 3, 9, 8   Coefficient of range = $\frac{9-2}{9+2} \times 100 = 63.64\%$

❖ Mean Deviation

Let $x_1, x_2, x_3, \cdots\cdots\cdots, x_n$ be a set of n measurements of a variable x. Mean deviation is computed from mean, median and mode.

Then mean deviation about mean is defined as

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

Then mean deviation about median is defined as

$$MD\ (Me) = \frac{\sum_{i=1}^{n} |x_i - Me|}{n}$$

Then mean deviation about mode is defined as

$$MD\ (Mo) = \frac{\sum_{i=1}^{n} |x_i - Mo|}{n}$$

Where $\bar{x}$, Me and Mo denotes the mean, median and mode respectively.
For continuous frequency distribution, if $x_1, x_2, x_3, \cdots\cdots\cdots, x_k$ $(i = 1,2,3, \ldots\ldots, k)$be the mid value of i^{th} class and $f_1, f_2, f_3, \cdots\cdots\cdots, f_k$ be their corresponding frequency.

$$MD\ (\bar{x}) = \frac{\sum_{i=1}^{k} f_i|x_i - \bar{x}|}{N}$$

$$MD\ (Me) = \frac{\sum_{i=1}^{k} f_i|x_i - Me|}{N}$$

$$MD\ (Mo) = \frac{\sum_{i=1}^{k} f_i|x_i - Mo|}{N}$$

Where $N = f_1 + f_2 + f_3 + \cdots\cdots\cdots + f_k$.

➢ Coefficient of Mean Deviation

Let $x_1, x_2, x_3, \cdots\cdots\cdots, x_n$ be a set of n measurements of a variable x.

Coefficient of mean deviation from mean = $\frac{MD(\bar{x})}{\bar{x}} \times 100$

Coefficient of mean deviation from median = $\frac{MD(Me)}{Me} \times 100$

Coefficient of mean deviation from mode = $\frac{MD(Mo)}{Mo} \times 100$

- **Example #1:** The marks of 5 students in a class are shown here: 40, 55, 75, 65, 80. Find mean deviation about mean. Also find it coefficient of mean deviation.

Solution: We know that

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$
$$= \frac{40+55+75+65+80}{5}$$
$$= \frac{315}{5} = 63$$

The mean deviation about mean is

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

$$= \frac{|40-63| + |55-63| + |75-63| + |65-63| + |80-63|}{5}$$

$$= \frac{23 + 8 + 12 + 2 + 17}{5}$$

$$= \frac{62}{5} = 12.4$$

Therefore, the mean deviation about mean is 12.4 .

We know that

Coefficient of mean deviation from mean = $\frac{MD(\bar{x})}{\bar{x}} \times 100$

$$= \frac{12.4}{63} \times 100 = 19.68\%$$

Example #2: The following is a distribution of the final examination scores, which 200 students obtained in a three –week course in Statistics.

Scores	0-20	20-40	40-60	60-80	80-100	Total
Number of students	24	55	76	32	13	200

Find (i) Mean deviation about mean and (ii) Coefficient of mean deviation.

Solution:

Scores	**Number of students (f_i)**	**Mid value (X_i)**	f_iX_i	$f_i\|X_i - \bar{X}\|$
0-20	24	10	240	852
20-40	55	30	1650	852.5
40-60	76	50	3800	342
60-80	32	70	2240	784
80-100	13	90	1170	578.5
Total	200		9100	3409.00

We know that

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i X_i}{N}$$

$$= \frac{9100}{200} = 45.5$$

We have that the mean deviation about mean is

$$MD\ (\bar{X}) = \frac{\sum_{i=1}^{k} f_i |X_i - \bar{X}|}{N}$$

$$= \frac{3409}{200} = 17.045$$

And coefficient of mean deviation for mean $= \frac{MD(\bar{X})}{\bar{X}} \times 100$

$$= \frac{17.045}{45.5} \times 100 = 37.46\%$$

- **Quartile Deviation**

The Quartile Deviation is a simple way to estimate the spread of a distribution about a measure of its central tendency (usually the mean). So, it gives you an idea about the range within which the central 50% of your sample data lies.

Formally, the Quartile Deviation is equal to the half of the Inter-Quartile Range and thus we can write it as

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

Where Q_1 and Q_3 denotes the first and third quartile respectively. Therefore, we also call it the *Semi Inter-Quartile Range.*

- **Coefficient of Quartile Deviation**

Based on the quartiles, a relative measure of dispersion, known as the Coefficient of Quartile Deviation is defined as

Coefficient of quartile deviation $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

Since it involves a ratio of two quantities of the same dimensions, it is unit-less. Thus, it can act as a suitable parameter for comparing two or more different datasets, which may or may not involve quantities with the same dimensions.

- **Example #3:** Find the quartile deviation and coefficient of quartile deviation of the series: 2, 6, 7, 9, 5, 4, and 11.

Solution:

By arranging the series in ascending order, we have that

2, 4, 5, 6, 7, 9 and 11

Since the number of observation $(n = 7)$ is odd , therefore

$$Q_i = \text{Value of } \left(\frac{i(n+1)}{4}\right) th \text{ observation, where } i = 1, 2, 3.$$

$$Q_1 = \text{Value of } \left(\frac{(n+1)}{4}\right) th \text{ observation}$$
$$= \text{Value of } \left(\frac{7+1}{4}\right) th \text{ observation}$$
$$= \text{Value of } 2nd \text{ observation} = 4$$

$$Q_3 = \text{Value of } \left(\frac{3(n+1)}{4}\right) th \text{ observation}$$
$$= \text{Value of } \left(\frac{3(7+1)}{4}\right) th \text{ observation}$$
$$= \text{Value of } 6th \text{ observation} = 9$$

Therefore quartile deviation is

$$QD = \frac{Q_3 - Q_1}{2} = \frac{9-4}{2} = 2.5$$

And, coefficient of quartile deviation $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{9-4}{9+5} \times 100 = \frac{5}{14} \times 100 = 35.71\%$

- **Example #4:** The following is a distribution of the final examination scores, which 200 students obtained in a three –week course in Statistics.

Scores	0-20	20-40	40-60	60-80	80-100	Total
Number of students	24	55	76	32	13	200

Find (i) quartile deviation and (ii) coefficient of quartile deviation.

Solution:

Scores	**Number of students (f)**	**Cumulative frequency**
0-20	24	24
20-40	55	79
40-60	76	155
60-80	32	187
80-100	13	200
Total	200	

We know that the quartiles for grouped data is

$$Q_i = L_i + \frac{\frac{iN}{4} - F_i}{f_i} \times c \qquad ; i = 1, 2, 3.$$

Where,

Quartile class = which class have $cf \geq \frac{iN}{4}$.

L_i is the lower limit of ith quartile class

N is the total frequency

f_i is the frequency of ith quartile class

F_i is the cumulative frequency preceding of the ith quartile class

c is the class width of quartile class

1st quartile:

$$Q_1 = L_1 + \frac{\frac{N}{4} - F_1}{f_1} \times c$$

$$= 20 + \frac{\frac{100}{4} - 24}{55} \times 20$$

$$= 20 + 0.36 = 20.36$$

Where,
1st quartile class is 20-40
$L_1 = 20$
$F_1 = 24$
$f_1 = 55$
c=20

3rd quartile:

$$Q_3 = L_3 + \frac{\frac{3N}{4} - F_3}{f_3} \times c$$

$$= 40 + \frac{\frac{3 \times 200}{4} - 79}{76} \times 20$$

$$= 40 + 18.68$$

$$= 58.68$$

Where,
3rd quartile class is 40-60
$L_3 = 40$
$F_3 = 79$
$f_3 = 76$
c = 20

We know that the quartile deviation (QD) is

$$QD = \frac{Q_3 - Q_1}{2} = \frac{58.68 - 20.36}{2} = 19.16$$

And, coefficient of quartile deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{58.68 - 20.36}{58.68 + 20.36} \times 100 = \frac{38.32}{79.04} \times 100 = 48.48\%$

- **Variance and Standard Deviation**

The variance is actually the average of the square of the distance that each value is from the mean. Therefore, if the values are near the mean, the variance will be small. In contrast, if the values are far from the mean, the variance will be large. The symbol of population variance is σ^2.
The formula of population variance is

$$V(X) = \sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Where, X = individual value
μ = population mean
N = population size

And the formula of sample variance, denoted by s^2 is

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

Where,
$\bar{X}$ = sample mean
n = sample size
The standard deviation is the square root of the variance. The symbol for the population standard deviation is σ.
The formula of population standard deviation is

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$$

And the formula for sample standard deviation, denoted by s is

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

For continuous frequency distribution, if $x_1, x_2, x_3, \cdots\cdots\cdots, x_k$ $(i = 1,2,3, \ldots\ldots, k)$ be the mid value of i^{th} class and $f_1, f_2, f_3, \cdots\cdots\cdots, f_k$ be their corresponding frequency.

The formula of population variance is

$$V(X) = \sigma^2 = \frac{\sum_{i=1}^{k} f_i (X_i - \mu)^2}{N}$$

Where, X = individual value
μ = population mean
N = population size
The formula of population standard deviation is

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k} f_i (X_i - \mu)^2}{N}}$$

❖ **Example #5:** Find the variance and standard deviation for the data set: 10, 60, 50, 30, 40, 20.

Solution:

X	$(X-\mu)$	$(X-\mu)^2$
10	-25	625
60	25	625
50	15	225
30	-5	25
40	5	25
20	-15	225
$\sum X = 210$		$\sum(X-\mu)^2 = 1750$

We know that the arithmetic mean is

$$\mu = \frac{\sum X}{N} = \frac{210}{6} = 35$$

The variance of X is

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} = \frac{1750}{6} = 291.67$$

And the standard deviation of X is

$$\sigma = \sqrt{291.67} = 17.08$$

❖ **Example #6:** The following is a distribution of the final examination scores, which 200 students obtained in a three –week course in Statistics.

Scores	0-20	20-40	40-60	60-80	80-100	Total
Number of students	24	55	76	32	13	200

Find the variance and standard deviation for the above frequency distribution.

Solution:

Scores	Number of students (f_i)	Mid value (X_i)	f_iX_i	$f_i(X_i-\bar{X})^2$
0-20	24	10	240	30246
20-40	55	30	1650	13213.75
40-60	76	50	3800	1539
60-80	32	70	2240	19208
80-100	13	90	1170	25743.25
Total	200		9100	89950.00

We know that

$$\bar{X} = \frac{\sum_{i=1}^{k} f_iX_i}{N}$$

$$= \frac{9100}{200} = 45.5$$

The variance is

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i(X_i-\bar{X})^2}{N} = \frac{89950}{200} = 449.75$$

And standard deviation is

$$\sigma = \sqrt{449.75} = 21.21$$

Therefore the test scores of students vary by ± 21.21 marks from the average marks of the students (45.5).

❖ **Uses of Variance and Standard Deviation**

1) Variances and standard deviations can be used to determine the spread of the data. If the variance or standard deviation is large, the data are more dispersed. This information is useful in comparing two (or more) data sets to determine which is more (most) variable.
2) The measures of variance and standard deviation are used to determine the consistency of a variable. For example, in the manufacture of fittings, such as nuts and bolts, the variation in the diameters must be small, or the parts will not fit together.
3) The variance and standard deviation are used to determine the number of data values that fall within a specified interval in a distribution. For example, Chebyshev's theorem shows that, for any distribution, at least 75% of the data values will fall within 2 standard deviations of the mean.
4) Finally, the variance and standard deviation are used quite often in inferential statistics.

➢ **Coefficient of Variation**

Whenever two samples have the same units of measure, the variance and standard deviation for each can be compared directly. For example, suppose an automobile dealer wanted to compare the standard deviation of miles driven for the cars she received as trade-ins on new cars. She found that for a specific year, the standard deviation for Buicks was 422 miles and the standard deviation for Cadillacs was 350 miles. She could say that the variation in mileage was greater in the Buicks. But what if a manager wanted to compare the standard deviations of two different variables, such as the number of sales per salesperson over a 3-month period and the commissions made by these salespeople?

A statistic that allows you to compare standard deviations when the units are different, as in this example, is called the *coefficient of variation.*

The coefficient of variation, denoted by CV, is the standard deviation divided by the mean. The result is expressed as a percentage. Mathematically

$$CV = \frac{\sigma}{\overline{X}} \times 100$$

Example #7: The grade point average (GPA) in different semesters of two students are shown below:

Student	**GPA in semesters**							
	1	2	3	4	5	6	7	8
A	2.5	2.5	3.0	3.5	3.5	4.0	3.5	3.5
B	2.5	3.0	4.0	4.0	4.0	2.0	2.5	4.0

Which students would you consider better throughout the courses of studies?

Solution:

Let X_A and X_B denotes the GPA in different semester of student A and B respectively.

$\boldsymbol{X_A}$	$\boldsymbol{X_B}$	$\boldsymbol{(X_A - \overline{X_A})^2}$	$\boldsymbol{(X_B - \overline{X_B})^2}$
2.5	2.5	0.5625	0.5625
2.5	3	0.5625	0.0625
3	4	0.0625	0.5625
3.5	4	0.0625	0.5625
3.5	4	0.0625	0.5625
4	2	0.5625	1.5625
3.5	2.5	0.0625	0.5625
3.5	4	0.0625	0.5625
$\sum X_A = 26$	$\sum X_B = 26$	$\sum(X_A - \overline{X_A})^2 = 2$	$(X_B - \overline{X_B})^2 = 5$

<u>For Student A</u>

<u>Mean</u>

$$\bar{X}_A = \frac{\sum X_A}{n_A} = \frac{26}{8} = 3.25$$

<u>Standard deviation</u>

$$\sigma_A = \sqrt{\frac{\sum(X_A - \bar{X}_A)^2}{n_A}} = \sqrt{\frac{2}{8}} = 0.5$$

<u>Coefficient of variation</u>

$$CV(A) = \frac{\sigma_A}{\bar{X}_A} \times 100 = \frac{0.5}{3.25} \times 100 = 15.38\%$$

<u>For Student B</u>

<u>Mean</u>

$$\bar{X}_B = \frac{\sum X_B}{n_B} = \frac{26}{8} = 3.25$$

<u>Standard deviation</u>

$$\sigma_B = \sqrt{\frac{\sum(X_B - \bar{X}_B)^2}{n_B}} = \sqrt{\frac{5}{8}} = 0.79$$

<u>Coefficient of variation</u>

$$CV(B) = \frac{\sigma_B}{\bar{X}_B} \times 100 = \frac{0.79}{3.25} \times 100 = 24.32\%$$

Decision: Since CV(A) < CV(B) therefore we can say that student A is better than student B throughout the courses of studies.

❖ Difference between Absolute and Relative Measures of Dispersion

There are the following differences between absolute and relative measures of dispersion:

Absolute measures of dispersion	Relative measures of dispersion
(i) When dispersion is measured in original units then it is known as absolute measures of dispersion.	(i) A relative measure of dispersion is the ratio of absolute measures of dispersion to an appropriate average.
(ii) Absolute measures of dispersion are not expressed in terms of ratio, percentage etc.	(ii) Relative measures of dispersion are expressed in terms of ratio, percentage etc.
(iii) Absolute measures of dispersion are as follows: • Range • Mean deviation • Standard deviation • Quartile deviation.	(iii) Relative measures of dispersion are as follows: • Coefficient of range • Coefficient of mean deviation • Coefficient of variation • Coefficient of quartile deviation.
(iv) It is calculated from raw data.	(iv) It is calculated from the absolute measures of dispersion
(v) Absolute measures of dispersion are not pure number.	(v) Relative measures of dispersion are a pure number.
(vi) If the two sets of data are expressed in different units, the absolute measures of dispersions are not comparable.	(vi) If the two sets of data are expressed in different units, the relative measures of dispersions are comparable.

❖ Difference between Variance and Coefficient of Variance

The difference between variance and coefficient of variance are given below:

Variance	Coefficient of Variation
(i) Variance is the arithmetic mean of the squared deviations from mean of the distribution.	(i) Coefficient of variation is 100 times of a ration of the standard deviation to the arithmetic mean.
(ii) Mathematically, variance is $\sigma^2 = \frac{\sum_{i=1}^{k} f_i(x_i - \bar{x})^2}{N}$	(ii) Mathematically, coefficient of variance is $CV = \frac{\sigma}{\bar{x}} \times 100$
(iii) It is an absolute measure of dispersion.	(iii) It is a relative measure of dispersion.
(iv) Variance is not pure number.	(iv) Coefficient of variation is a pure number.
(v) Variance is not expressed in terms of percentage.	(v) Coefficient of variation is expressed in terms of percentage.
(vi) It is used to measure the dispersion of a distribution.	(vi) It is used to compare two or more distributions.

- ❖ **Requisites of a Good Measures of Dispersion**

The requirements of ideal measure of dispersion are:

- It should be rigidly defined.
- It should be easy to understand and calculate.
- It should be based on all the observations.
- It should be suitable for further mathematical treatment.
- It should be affected as little as possible by fluctuations of sampling.
- It should not be affected much by extreme observations.

- ❖ Which is the best measure of dispersion in your opinion? Give reasons for your answer. OR Why standard deviation is considered to be the best measure of dispersion.

Solution:

I think that standard deviation is the best measures of dispersion. Standard deviation satisfies the following properties:

- Standard deviation is rigidly defined.
- It is based on all the values of a variable.
- It is capable for further algebraic manipulation.
- It is less affected by sampling fluctuations.

Since standard deviation satisfies almost all the properties of an ideal measure. So I can say that standard deviation is the best measure of dispersion.

- ❖ **Theorem #1:** Find the variance and standard deviation of first n natural numbers.

Solution:

We know that, the first n natural numbers are $1, 2, 3, \dots \dots \dots, n$.

The arithmetic mean of first n natural numbers are

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$= \frac{1 + 2 + 3 + \cdots \cdots \cdots + n}{n}$$

$$= \frac{n(n+1)}{2n}$$

$$= \frac{n+1}{2}$$

Therefore, the variance of first n natural numbers are

$$V(X) = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} X_i{}^2 - n\bar{X}^2\right)$$

$$= \frac{1}{n}\ (1^2 + 2^2 + 3^2 + \cdots\cdots\cdots + n^2 - n\left(\frac{n+1}{2}\right)^2)$$

$$= \frac{1}{n}\ (\frac{n(n+1)(2n+1)}{6} - n\frac{(n+1)^2}{4})$$

$$= \frac{n+1}{2}\ (\frac{2n+1}{3} - \frac{n+1}{2})$$

$$= \frac{n+1}{2}\ (\frac{4n+2-3n-3}{6})$$

$$= \frac{n+1}{2}\ (\frac{n-1}{6})$$

$$= \frac{n^2-1}{12}$$

Therefore, the standard deviation of first n natural numbers are

$$SD = \sqrt{\frac{n^2-1}{12}} = \frac{\sqrt{n^2-1}}{2\sqrt{3}}$$

- **Theorem #2:** Show that variance and hence standard deviation is independent of origin but dependent on scale of measurement.

Solution:

Let $x_1, x_2, x_3, \cdots\cdots\cdots, x_n$ be a set of n measurements of a variable x. Then arithmetic mean (AM) is

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

And the variance is

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

Let us consider a new variable $u_i = \frac{x_i - a}{c}$, $where\ (i = 1, 2, 3, \ldots\ldots\ldots, n)$ and a and c are arbitrary constant denotes the origin and scale respectively.

$$x_i = a + cu_i$$

$$\bar{x} = \frac{\sum_{i=1}^{n} a + cu_i}{n}$$

$$= \frac{\sum_{i=1}^{n} a + c\sum_{i=1}^{n} u_i}{n}$$

$$= \frac{na}{n} + c\frac{\sum_{i=1}^{n} u_i}{n}$$

$$= a + c\bar{u} \quad [\text{Since } \bar{u} = \frac{\sum_{i=1}^{n} u_i}{n}]$$

We have that

$$V(x) \text{ or } \sigma_x{}^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$= \frac{\sum_{i=1}^{n}(a + cu_i - a - c\bar{u})^2}{n}$$

$$= c^2\frac{\sum_{i=1}^{n}(u_i - \bar{u})^2}{n}$$

$$= c^2V(u) \text{ or } c^2\sigma_u{}^2 \quad \text{Since } V(u) \text{ or } \sigma_u{}^2 = \frac{\sum_{i=1}^{n}(u_i - \bar{u})^2}{n}$$

The standard deviation is

$$SD(x) \text{ or } \sigma_x = \sqrt{V(x)} = c\sigma_u$$

Therefore, we can say that variance and hence standard deviation is independent of origin but dependent on scale of measurement. (Showed)

- **Theorem #3:** For two unequal observations show that mean deviation and standard deviation is equal to the half of the range.

Solution:

Let x_1 and x_2 be two observations of a variable x $(x_1 > x_2)$.

We know that the mean of two observations are

$$\bar{x} = \frac{x_1 + x_2}{2}$$

The range (R) of two observations are

$$R = x_1 - x_2$$

We know that mean deviation (MD) about mean is

$$MD = \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n}$$

$$= \frac{\left|\left(x_1 - \frac{x_1 + x_2}{2}\right) + \left(x_2 - \frac{x_1 + x_2}{2}\right)\right|}{2}$$

$$= \frac{1}{2}\left|\left(\frac{2x_1 - x_1 - x_2}{2}\right) + \left(\frac{2x_1 - x_1 - x_2}{2}\right)\right|$$

$$= \frac{1}{2}\left|\left(\frac{x_1 - x_2}{2}\right) + \left(\frac{x_2 - x_1}{2}\right)\right|$$

$$= \frac{x_1 - x_2}{2} = \frac{R}{2}\ (Showed)$$

We know that the standard deviation (SD) is

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

$$= \sqrt{\frac{(x_1 - \frac{x_1 + x_2}{2})^2 + (x_2 - \frac{x_1 + x_2}{2})^2}{2}}$$

$$= \sqrt{\frac{(\frac{2x_1 - x_1 - x_2}{2})^2 + (\frac{2x_2 - x_1 - x_2}{2})^2}{2}}$$

$$= \sqrt{\frac{(\frac{x_1 - x_2}{2})^2 + (\frac{x_2 - x_1}{2})^2}{2}}$$

$$= \sqrt{\frac{(\frac{x_1 - x_2}{2})^2 + (\frac{x_2 - x_1}{2})^2}{2}}$$

$$= \sqrt{(\frac{x_1 - x_2}{2})^2}$$

$$= \frac{x_1 - x_2}{2} = \frac{R}{2}\ \ (Showed)$$