Unit-4
Measures of Dispersion & Skewness

CONTENTS

Measures of Dispersion
Definition:

It is hardly fully representative of a mass, unless we know the manner in which the individual items scatter around it. A further description of the series is necessary if we are to gauge how representative the average is.

> *"Dispersion is the measure of the variation of the items."*
> —A.L. Bowley
>
> *"Dispersion is a measure of the extent to which the individual items vary."* —L.R. Connor
>
> *"Dispersion or spread is the degree of the scatter or variation of the variables about a central value."*
> —B.C. Brooks and W.F.L. Dicks
>
> *"The degree to which numerical data tend to spread about an average value is called the variation or dispersion of the data."* —Spiegal

Importance or Significance of Measures of Dispersion:

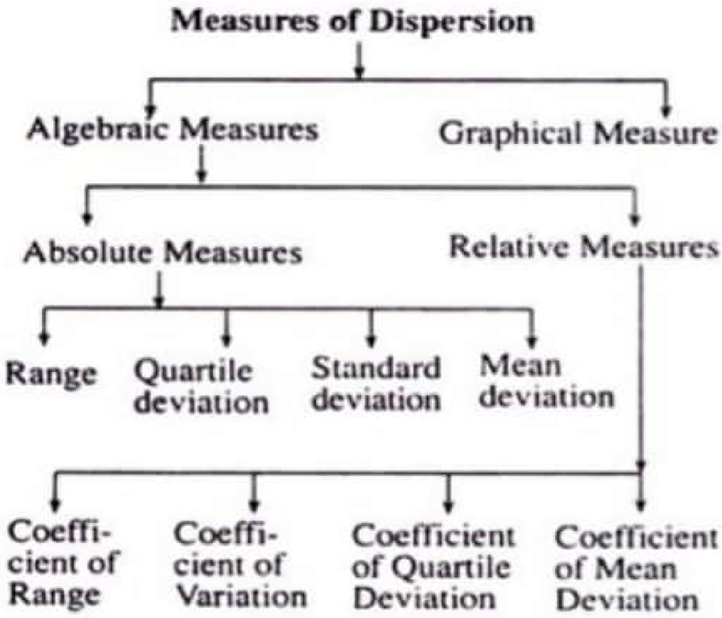
Scanned with CamScanner

The reliability of a measure of central tendency is known.

1. Measures of Dispersion provide a basis for the control of variability.
2. They help to compare two or more sets of data with regard to their variability.
3. They enhance the utility and scope of statistical techniques.

The Usual Measures of Dispersion:

The usual measures of dispersion, very often suggested by the statisticians, are exhibited with the aid of the following chart:



Difference between the Absolute measure and Relative Measure:

Absolute measure	Relative measure
Range	Co-efficient of Range
Quartile deviation	Co-efficient of Quartile Deviation
Mean Deviation (about Mean)	Co-efficient of Mean Deviation (about Mean)
Median Deviation (about Median)	Co-efficient of Median Deviation (about Median)
Mode Deviation (about Mode	Co-efficient of Mode Deviation (about Mode)
Standard deviation and Variance	Co-efficient of Variation

Absolute Measures of Dispersion:

Range

Definition: Range is the difference between the greatest (largest) and the smallest of

the values.

Scanned with CamScanner

In Symbols, Range = L-S.

L- Largest Value

S- Smallest Value

In individual observations and discrete series, L and S are easily identified. In Continuous series, the following two method are followed.

Method-I

L- Upper boundary of the highest class

S - Lower boundary of the lowest class

Method-2

L - Mid value of the highest class

S - Mid value of the lowest class

□−□□+□

Uses of Range:

1. Range is used in finding the control limits of mean chart and Range chart in S.Q.C.

2. While Quoting the prices of shares, bonds, gold, etc. on daily basis or yearly basis, the minimum and the maximum prices are mentioned.

3. The minimum and the maximum temperature likely to prevail on each day are forecasted.

Merits:

1. It is simple to understand and easy to calculate.
2. It can be calculated in no time.

Demerits:

1. Its definition does not seem to suit continuous series.
2. It is based on the two extreme items. It does not consider the other items.
3. It is usually affected by the extreme items.
4. It cannot be manipulated algebraically. The Range of combined set cannot be found from the range of the individual sets.
5. It does not have sampling stability.
6. It cannot be calculated from open-end class intervals. It is a very rarely used measure. Its Scope is limited.

Quartile Deviation (Q.D)

Definition : Quartile Deviation is half of the difference between the first and the third quartiles. Hence it is called Semi Inter Quartile Range.

Scanned with CamScanner

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

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

Merits:

1. It is simple to understand and easy to calculate.
2. It is not affected by extreme items.
3. It can be calculated for data with open and classes also.

Demerits:

1. It is not based on all items. It is based on two positional values Q_1 and Q_3 and ignores the extreme 50% of the items.
2. It cannot be manipulated algebraically.
3. It is affected by sampling fluctuations.
4. Like range, it does not measure the deviation about any measure of central tendency.

Mean Deviation or Average Deviation:

Definition: Mean deviation is the arithmetic mean of the absolute deviations of the values about their arithmetic mean or median or mode.

M.D. is the abbreviation for Mean Deviation. There are three kinds of mean deviations, Viz.,

1. mean deviation or mean deviation about mean
2. mean deviation about median
3. mean deviation about mode.

Mean deviation about median is the least. it could be easily verified in individual observations and discrete series where the actual values are considered.

The relative measures are the following:

Coefficient of Mean deviation (about Mean)

Coefficient of Mean deviation about Median

Coefficient of Mean deviation about Mode

Individual Observations:

Scanned with CamScanner

Mean Deviation (about Mean) = $\frac{\sum |X - \overline{X}|}{N}$

Mean Deviation about Median = $\frac{\sum |X - M|}{N}$

Mean Deviation about Mode = $\frac{\sum |X - Z|}{N}$

Discrete Series:

Mean Deviation (about Mean) = $\frac{\sum f |X - \overline{X}|}{N}$

Mean Deviation about Median = $\frac{\sum f |X - M|}{N}$

Mean Deviation about Mode = $\frac{\sum f |X - Z|}{N}$

Continuous Series:

Mean Deviation (about Mean) = $\frac{\sum f |m - \overline{X}|}{N}$

Mean Deviation about Median = $\frac{\sum f |m - M|}{N}$

Mean Deviation about Mode = $\frac{\sum f |m - Z|}{N}$

□.□.□□□□□ □□□□□□□□□

□.□.□□□□□ □□□□□□□□□□□□□

□.□.□□□□□ □□□□□□□□□

Uses of Mean Deviation:

Mean deviation provides an opportunity to calculate deviation, absolute deviation, total deviation and average of the deviations. Standard deviation is the most important absolute measure of dispersion. knowledge of the principle of mean deviation facilities understanding the concept of standard deviation. Standard deviation is a part of almost all the theories of Statistics, Viz., skewness, kurtosis, correlation, regression, sampling, estimation, inference, S.Q.C., etc. Mean deviation is preferred when a particular discussion is not carried to other spheres. It is found to be much useful in forecasting business cycles and a few other statistical activities connected with business, economic and sociology.

Merits:

1. Mean deviations are rigidly defined.

Scanned with CamScanner

2. They are based on all the items.
3. They are affected less by extreme items than standard deviation. Among the three mean deviations, mean deviation about median is the least.
4. They are simple to understand and not difficult to calculate.
5. They do not vary much from sample to sample.
6. They provide choice. among the three mean deviations, the one that is suitable to a particular situation can be used.
7. Formation of different distributions can be compared on the basis of a mean deviation.

Demerits:

1. Omission of negative sign of deviations makes them non-algebraic. It is pointed out as a great drawback.
2. They could not be manipulated. Combined mean deviation could not be found.
3. it is not widely used in business or economics.

Standard Deviation:

Definition: Standard Deviation is the root mean square deviation of the values from their arithmetic mean. S.D denoted by **σ** (read , sigma). **Variance** is denoted by **σ^2**. S.D. is the positive square root of variance. Karl Pearson introduced the concept of standard deviation in 1893. S.D is also called **root mean square deviation**. The corresponding relative measure is **Coefficient of Variation**.

Individual Observation:

Method-1

$$\text{Standard deviation } \sigma = \sqrt{\frac{\sum x^2}{N}}$$

Method-2

$$\text{Standard deviation } \sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2}$$

Discrete Series:

Method-1

and N = ∑f

Method – 2

$$\text{Standard deviation } \sigma = \sqrt{\frac{\sum fX^2}{N} - \left(\frac{\sum fX}{N}\right)^2}$$

Scanned with CamScanner

Continuous Series:

Method – 1

Standard deviation $\sigma = \sqrt{\frac{\sum f(m-\overline{X})^2}{N}}$

Method – 2

Standard deviation $\sigma = \sqrt{\frac{\sum fm^2}{N} - \left(\frac{\sum fm}{N}\right)^2}$

Combined Standard Deviation:

When two or three groups merge, the mean and standard deviation of the combined group are calculated as follows.

Case 1. Merger of Two Groups

	Size	Mean	S.D.
Group I	N_1	$\overline{X}_1$	σ_1
Group II	N_2	$\overline{X}_2$	σ_2

That is,

N_1 – Number of items in the first group.

N_2 – Number of items in the second group.

$\overline{X}_1$ - Mean of items in the first group

$\overline{X}_2$ - Mean of items in the second group.

σ_1 – Standard Deviation of items in the first group.

σ_2 - Standard Deviation of items in the second group.

The Mean of the combined group,

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

The Standard Deviation of the combined group,

Scanned with CamScanner

$$\sigma_{12} = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_1d_1^2 + N_2d_2^2}{N_1 + N_2}}$$

where $d_1 = \overline{X}_1 - \overline{X}_{12}$ and $d_2 = \overline{X}_2 - \overline{X}_{12}$

Case 2. Merger of Three Groups

	Size	Mean	S.D.
Group I	N_1	$\overline{X}_1$	σ_1
Group II	N_2	$\overline{X}_2$	σ_2
Group III	N_3	$\overline{X}_3$	$\sigma 3$

That is

N_1 – Number of items in the first group.

N_2 – Number of items in the second group.

$\overline{X}_1$ - Mean of items in the first group

$\overline{X}_2$ - Mean of items in the second group.

σ_1 – Standard Deviation of items in the first group.

σ_2 - Standard Deviation of items in the second group.

The Mean of the combined group,

$$\overline{X}_{123} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

The Standard Deviation of the combined group,

$$\sigma_{12} = \sqrt{\frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_3\sigma_3^2 + N_1d_1^2 + N_2d_2^2 + N_3d_3^2}{N_1 + N_2 + N_3}}$$

where $d_1 = \overline{X}_1 - \overline{X}_{123}$, $d_2 = \overline{X}_2 - \overline{X}_{123}$ and $d_3 = \overline{X}_3 - \overline{X}_{123}$

Difference between the Mean Deviation and standard Deviation:

Mean Deviations	**Standard Deviations**
Deviations are calculated from Mean or Median or Mode.	Deviations are always calculated from Mean.

Scanned with CamScanner

While finding the deviations, negative sign is omitted.	Deviations are squared. Finally square root is taken.
M.D.'s are simple to understand and not difficult to calculate.	It is not simple to under standard and it is not easy to calculate.
Omission of negative sign is considered to be deficient mathematically.	It has many desirable mathematically properties.

Uses

Standard deviation is the best absolute measure of dispersion. It is a part of many statistical concepts such as Skewness, Kurtosis, Correlation, Regression, Estimation, sampling, tests of Significance and Statistical Quality Control. Not only in statistics but also in Biology, education, Psychology and other disciplines standard deviation is of immense use.

Merits:

1. Standard deviation is rigidly defined.
2. It is calculated on the basis of the magnitudes of all the items.
3. It could be manipulated further. The combined S.D. can be calculated.
4. Mistakes in its calculation can be corrected. The entire calculation need not be redone.
5. Coefficient of variation is based on S.D.. It is the best and most widely used relative measure of dispersion.
6. It is free from sampling fluctuations. This property of sampling stability has brought it an indispensable place in tests of significance.
7. It reduces the complexity in the approach of normal distribution by providing standard normal variable.
8. It is the most important absolute measure of dispersion. It is used in all the areas of statistics. It is widely used in other disciplines such as Psychology, Education and Biology as well.
9. Scientific calculators show the standard deviation of any series.
10. Different forms of the formula are available.

Demerits:

1. Compared with other absolute measures of dispersion, it is difficult to calculate.
2. It is not simple to understand.
3. It gives more weight age to the items away from the mean than those near the mean as the deviations are squared.

Coefficient of Variation

Definition: Coefficient of variation is the most widely used relative measure of dispersion. It is based on the best absolute measure of dispersion and the best measure of central tendency. It is a percentage. While comparing two or more groups, the group which has less coefficient of variation is less variable or more consistent or

Scanned with CamScanner

more stable or more uniform or more homogeneous. Coefficient of Variation is denoted by the C.V.

OR C.V. = $\frac{\sigma}{\overline{X}} \times 100$

Variance

Variance is the mean square deviation of the values from their arithmetic mean. It is denoted by σ2. Standard deviation is the positive square root of variance and is denoted by σ. The term of variance was introduced by R.A. fisher in the year 1913. It is used much in sampling, analysis of variance, etc., In analysis of variance, total variation is split into a few components. Each component is ascribable to one factor of variation. The significance of the variation is then tested.

Individual Observation:

$$\text{Variance, } \sigma 2 = \frac{\sum (x - \overline{x})}{N}$$

Discrete Series:

$$\text{Variance, } \sigma 2 = \frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2$$

Continuous Series:

$$\text{Variance, } \sigma 2 = c^2 \left[\frac{\sum d'^2}{N} - \left(\frac{\sum fd'^2}{N}\right)^2\right]$$

Combined Variance:

Based on the notations used in combined mean and combined variance,

Combined variance of two groups,

$$\sigma_{12}^2 = \frac{N_1\sigma_1^2 + N_2\sigma_2^2 + +N_1d_1^2 + N_2d_2^2}{N_1 + N_2}$$

Combined variance of three groups,

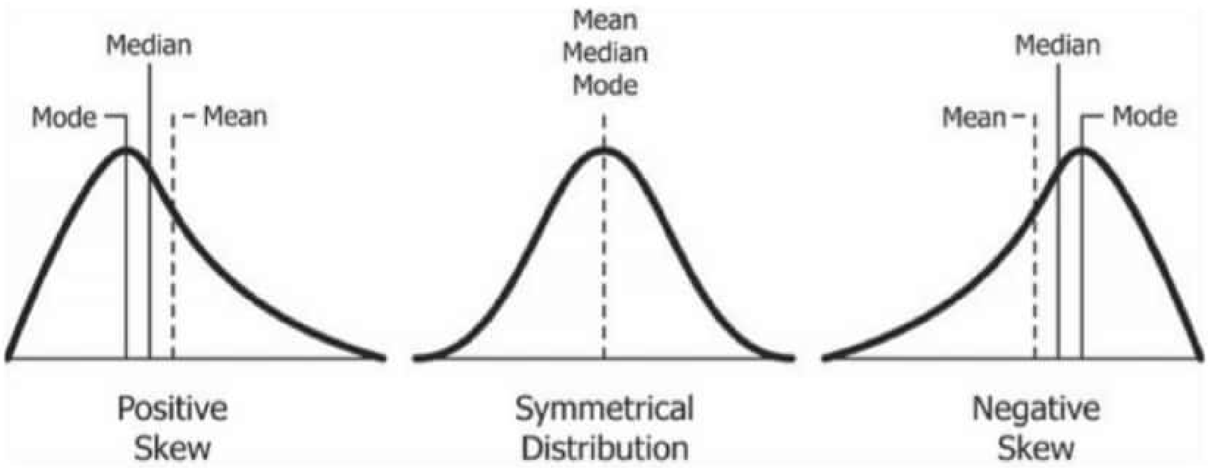
$$\sigma_{123}^2 = \frac{N_1\sigma_1^2 + N_2\sigma_2^2 + N_3\sigma_3^2 + N_1d_1^2 + N_2d_2^2 + N_3d_3^2}{N_1 + N_2 + N_3}$$

Skewness:

Scanned with CamScanner

Definition: Skewness is the degree of asymmetry, or departure from symmetry, of a distribution.

There are two types of Skewness: Positive and Negative



Positive Skewness means when the tail on the right side of the distribution is longer or fatter. The mean and median will be greater than the mode.

Negative Skewness is when the tail of the left side of the distribution is longer or fatter than the tail on the right side. The mean and median will be less than the mode. Karl – Pearson (1867- 1936) was great British Biometric and Statistician. He introduced the formula given below.

1. Karl- Pearson's Coefficient of Skewness,

$$SK_p = \frac{Mean - Mode}{S\tan dardDeviat\ ion} \quad \textbf{OR} \quad Sk_p = \frac{\overline{X} - Z}{\sigma}$$

Theoretically, no limit can be found for this measure. This is found mostly to vary between -1 and +1. Based on the interrelation between mean, median and mode in a moderately skewed distribution, his second formula:

Karl- Pearson's Coefficient of Skewness,

$$Sk_p = \frac{3(Mean - Mode)}{S\tan dardDeviat\ ion} \quad \textbf{OR} \quad Sk_p = \frac{3(\overline{X} - M)}{\sigma}$$

2. The following formula is given by Prof. Bowley,

Bowley's Coefficient of Skewness,

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

Scanned with CamScanner

This is quartile measure of skewness and the value of this between -1 and +1. This method is useful where there is an open-end class interval or extreme values are present.

Scanned with CamScanner

Measures of Dispersion (Deviation or variation)

Range

Quartile deviation

Mean deviation

Standard deviation

Their coefficients

Range and Co-efficient of Range

Range = L - S [Largest value - Smallest value]

coefficient of Range = $\frac{L-S}{L+S}$

1). Indidual series

Range = 9-1 = 8

Coefficient of Range = $\frac{9-1}{9+1} = \frac{8}{10} = 0.8$

2. Discret series

Range = 8 - 3 = 5

coefficient of Range = $\frac{8-3}{8+3} = \frac{5}{11}$

3. Continuous series

Range = 15 - 0 = 15

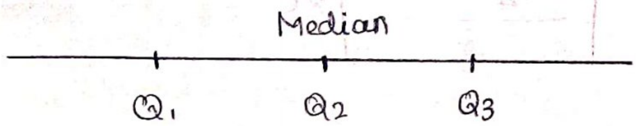
coefficient of Range = $\frac{15-0}{15+0} = \frac{15}{15} = 1$

Scanned with CamScanner

Range - Definition

Range means the difference between the largest value and smallest value from the given data.

Quartile deviation



Indidual series

Notes : Ascending order, Decending order

Formula for 3 quartors

$$Q_1 = 1\left[\frac{N+1}{4}\right]^{th} \text{ value}$$

$$Q_2 = 2\left[\frac{N+1}{2}\right]^{th} \text{ value}$$

$$= \left[\frac{N+1}{2}\right]^{th} \text{ value} = \text{Median}$$

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

Discret series

$$Q_1 = 1\left[\frac{\Sigma f+1}{4}\right]^{th} \text{ value}$$

$$Q_2 = 2\left[\frac{\Sigma f+1}{4}\right]^{th} \text{ value}$$

$$= \left[\frac{\Sigma f+1}{2}\right]^{th} \text{ value} = \text{Median}$$

$$Q_3 = 3\left[\frac{\Sigma f+1}{4}\right]^{th} \text{ value}$$

Scanned with CamScanner

Continuous series

$$Q_1 = L + \left[\frac{\frac{\Sigma f}{4} - Pcf}{f}\right] \times C$$

$$Q_2 = L + \left[\frac{\frac{\Sigma f}{2} - Pcf}{f}\right] \times C$$

$$Q_3 = L + \left[\frac{\frac{3\Sigma f}{4} - Pcf}{f}\right] \times C$$

Quartile deviation and coefficient of QD

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

$$\text{Coefficient of } QD = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

1) Individual series

2, 6, 8, 5, 3, 2, 1, 7, 9, 6

Ascending order = 1, 2, 2, 3, 5, 6, 6, 7, 8, 9

$QD = \frac{Q_3 - Q_1}{2}$ Coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

$Q_1 = \left(\frac{N+1}{4}\right)^{th}$ value $= \frac{11}{4}^{th}$ value $= 2.75^{th}$ value

$= 2^{th}$ value $+ 0.75$ [3^{rd} value $- 2^{nd}$ value]

$= 2 + 0.75(2-2)$

$= 2$

$Q_3 = 3\left(\frac{N+1}{4}\right)^{th}$ value $= 3\left(\frac{11}{4}\right)^{th}$ value $= 8.25^{th}$ value

$= 8^{th}$ value $+ 0.25$ (9^{th} value $- 8^{th}$ value)

$= 7 + 0.25(8-7)$

$= 7.25$

Scanned with CamScanner

$QD = \frac{Q_3 - Q_1}{2} = \frac{7.25-2}{2} = \frac{5.25}{2} = 2.63$

coefficient of QD $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{7.25-2}{7.25+2}$

$= \frac{5.25}{9.25} = 0.57$

Universal Rule

$\Sigma(x - \bar{x}) = 0$

2, 5, 8

$\bar{x} = \frac{2+5+8}{3} = \frac{15}{3} = 5$

$x - \bar{x}$

$2-5 = -3$

$5-5 = 0$

$8-5 = 3$

$\Sigma(x-\bar{x}) = 0$

Universal rule :- Sum of the dispersion from arithmetic mean is always zero.

03.10.2020

2) Discret series

x	F	C.F
3	8	8
4	12	20
5	9	29
6	10	39
8	6	45
	45	

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

Coefficient of $Q.D = \frac{Q_3 - Q_1}{Q_3 + Q_1}$

Scanned with CamScanner

$Q_1 = 1\left[\frac{\Sigma f+1}{4}\right]^{th}$ value $= 1\left[\frac{45+1}{4}\right]^{th}$ value

$= \left[\frac{46}{4}\right]^{th}$ value $= 11.5^{th}$ value

$= 11^{th}$ value $+ 0.5\,[12^{th}$ value $- 11^{th}$ value$]$

$= 4 + 0.5\,(4-4)$

$= 4$

$Q3 = 3\left[\frac{\Sigma f+1}{4}\right]^{th}$ value $= 3\left[\frac{45+1}{4}\right]^{th}$ value

$= 3\left[\frac{46}{4}\right]^{th}$ value $= 3\,(11.5)^{th}$ value

$= 34.5^{th}$ value

$= 34^{th}$ value $+ 0.5\,(35^{th}$ value $- 34^{th}$ value$)$

$= 6 + 0.5\,(6-6)$

$= 6$

$QD = \frac{Q_3 - Q_1}{2} = \frac{6-4}{2} = \frac{2}{2} = 1$

Coefficient of $QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{6-4}{6+4} = \frac{2}{10} = 0.2$

3. Continuous series

X	F	CF
0-3	8	8
3-6	3	11
6-9	5	16
9-12	6	22
12-15	9	31
	31	

$\frac{\Sigma f}{4} = \frac{31}{4} = 7.75$

$3\left(\frac{\Sigma f}{4}\right) = 3\left(\frac{31}{4}\right) = 23.25$

Scanned with CamScanner

$$Q_1 = L + \left[\frac{\frac{\Sigma f}{4} - Pcf}{f}\right] \times C$$

$$= 3 + \left[\frac{7.75 - 8}{3}\right] \times 3$$

$$= 3 + \left[\frac{-0.25}{3}\right] \times 3$$

$$= 3 - 0.25$$

$$= 2.75$$

$$Q_3 = L + \left[\frac{\frac{3\Sigma f}{4} - Pcf}{f}\right] \times C$$

$$= 12 + \left[\frac{23.25 - 22}{9}\right] \times 3$$

$$= 12 + \left[\frac{1.25}{9}\right] \times 3$$

$$= 12 + 0.41$$

$$= 12.41$$

$$QD = \frac{Q_3 - Q_1}{2} = \frac{12.41 - 2.75}{2} = \frac{9.66}{2} = 4.83$$

coefficient of $QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{12.41 - 2.75}{12.41 + 2.75} = \frac{9.66}{15.16}$

$$= 0.637$$

Mean deviation (MD)

Mean deviation hu matum kettal meanku mattum ~~mean~~ Mean deviation from Mean kandu puducho pothum

Scanned with CamScanner

Mean Deviation

From Mean ($\bar{x}$)	From Median (M)	From Mode (Z)
For IS $MD = \frac{\Sigma \lvert x-\bar{x} \rvert}{n}$	$MD = \frac{\Sigma \lvert x-M \rvert}{n}$	$MD = \frac{\Sigma \lvert x-Z \rvert}{n}$
For DS $MD = \frac{\Sigma f \lvert x-\bar{x} \rvert}{\Sigma f}$	$MD = \frac{\Sigma f \lvert x-M \rvert}{\Sigma f}$	$MD = \frac{\Sigma f \lvert x-Z \rvert}{\Sigma f}$
For CS $MD = \frac{\Sigma f \lvert m-\bar{x} \rvert}{\Sigma f}$	$MD = \frac{\Sigma f \lvert m-M \rvert}{\Sigma f}$	$MD = \frac{\Sigma f \lvert m-Z \rvert}{\Sigma f}$

Calculate Mean deviation from mean, from median, from mode for the following data.

14, 12, 8, 22, 28, 15, 25, 28, 26, 24

Mean deviation

from Mean

x	$\lvert x-\bar{x} \rvert$
14	6.2
12	8.2
8	12.2
22	1.8
28	7.8
15	5.2
25	4.8
28	7.8
26	5.8
24	3.8

$\bar{x} = \frac{\Sigma x}{n} = \frac{202}{10} = 20.2$

$\Sigma \lvert x-\bar{x} \rvert = 63.6$

$MD = \frac{\Sigma \lvert x-\bar{x} \rvert}{n}$

$= \frac{63.6}{10}$

$= 6.36$

Scanned with CamScanner

from Median

A.O = 8, 12, 14, 15, 22, 24, 25, 26, 28, 28

$M = \left[\frac{n+1}{2}\right]^{th}$ value $= \left[\frac{10+1}{2}\right]^{th}$ value $= \frac{11}{2}^{th}$ value

$= 5.5^{th}$ value $= \left[\frac{5^{th} \text{ value} + 6^{th} \text{ value}}{2}\right]$

$= \frac{22+24}{2} = \frac{46}{2} = 23$

x	$\lvert x-M \rvert$
14	9
12	11
8	15
22	1
28	5
15	8
25	2
28	5
26	3
24	1

$\Sigma \lvert x-M \rvert = 60$

$MD = \frac{\Sigma \lvert x-M \rvert}{n}$

$= \frac{60}{10}$

$= 6$

from Mode

x	$\lvert x-Z \rvert$
14	14
12	16
8	20
22	6
28	0
15	13
25	3
28	0
26	2
24	4

$Z = 28$

$\Sigma \lvert x-Z \rvert = 78$

$MD = \frac{\Sigma \lvert x-Z \rvert}{n}$

$= \frac{78}{10}$

$= 7.8$

Scanned with CamScanner

S.No	1	2	3	4	5
Value	20	30	50	60	90

That is indidual series.

Find Mean deviation for the following data.

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	8	23	46	84	19

X	F	m	Fm	$\|m-\bar{x}\|$	$f\|m-\bar{x}\|$
0-10	8	5	40	24.61	196.88
10-20	23	15	345	14.61	336.03
20-30	46	25	1150	4.61	212.06
30-40	84	35	2940	5.39	452.76
40-50	19	45	855	15.39	292.41
	180		5330		1490.14

$$\bar{x} = \frac{\Sigma fm}{\Sigma f} = \frac{5330}{180} = 29.61$$

$$MD = \frac{\Sigma f\,|m-\bar{x}|}{\Sigma f}$$

$$= \frac{1490.14}{180}$$

$$= 8.278$$

Wages	200	250	300	350	400	450
No. of workers	5	7	13	14	6	5

Determine MD from $\bar{x}$, M and Z

Scanned with CamScanner

Mean deviation

from Mean

x	f	fx	$\lvert x-\bar{x} \rvert$	$f\lvert x-\bar{x} \rvert$
200	5	1000	124	620
250	7	1750	74	518
300	13	3900	24	312
350	14	4900	26	364
400	6	2400	76	456
450	5	2250	126	630
	50	16200		2900

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{16200}{50} = 324$$

$$MD = \frac{\Sigma f\lvert x-\bar{x} \rvert}{\Sigma f} = \frac{2900}{50} = 58$$

from Median

x	F	C.F	$\lvert x-M \rvert$	$f\lvert x-M \rvert$
200	5	5	125	625
250	7	12	75	525
300	13	25	25	325
350	14	39	25	350
400	6	45	75	450
450	5	50	125	625
	50			2900

$$M = \left(\frac{\Sigma f+1}{2}\right)^{th} \text{value} = \left[\frac{50+1}{2}\right]^{th} \text{value} = \left[\frac{51}{2}\right]^{th} \text{value}$$

$$= 25.5^{th} \text{ value} = \frac{25^{th} \text{ value} + 26^{th} \text{ value}}{2}$$

$$= \frac{300+350}{2} = \frac{650}{2} = 325$$

Scanned with CamScanner

$$MD = \frac{\Sigma f|x-M|}{\Sigma f}$$

$$= \frac{2900}{50}$$

$$= 58$$

From Mode

x	F	|x-z|	f|x-z|
200	5	150	750
250	7	100	700
300	13	50	650
350	14	0	0
400	6	50	300
450	5	100	500

$Z = 350$

$$MD = \frac{\Sigma f|x-z|}{\Sigma f}$$

$$= \frac{2900}{50}$$

$$= 58$$

10.10.2020

Coefficient of Mean deviation

From Mean $= \frac{\text{MD from Mean}}{\text{Mean}}$

From Median $= \frac{\text{MD from Median}}{\text{Median}}$

From Mode $= \frac{\text{MD from Mode}}{\text{Mode}}$

Calculate co-efficient of MD from Mode for the following data.

x	3	5	7	9	10	14
f	4	7	9	4	3	2

Mode (z) = 7

Scanned with CamScanner

X	F	$\lvert X - Z \rvert$	$f\lvert X - Z \rvert$
3	4	4	16
5	7	2	14
7	9	0	0
9	4	2	8
10	3	3	9
14	2	7	14
	29		61

$$MD = \frac{\Sigma f \lvert X - Z \rvert}{\Sigma f}$$

$$= \frac{61}{29}$$

$$= 2.103$$

$$\text{Coefficient MD from Mode} = \frac{\text{MD from Mode}}{\text{Mode}}$$

$$= \frac{2.103}{7}$$

$$= 0.3$$

Calculate co-efficient of MD from Mean for the following data.

Marks	0-20	20-40	40-60	60-80	80-100
No. of students	4	15	17	11	3

Marks (x)	No. of students (f)	m	fm	$\lvert m - \bar{x} \rvert$	$f\lvert m - \bar{x} \rvert$
0-20	4	10	40	37.6	150.4
20-40	15	30	450	17.6	264
40-60	17	50	850	2.4	40.8
60-80	11	70	770	22.4	246.4
80-100	3	90	270	42.4	127.20
	50		2380		828.8

$$\text{Mean } (\bar{x}) = \frac{\Sigma fm}{\Sigma f}$$

$$= \frac{2380}{50}$$

$$= 47.6$$

Scanned with CamScanner

MD from Mean $= \frac{\Sigma f|m-\bar{x}|}{\Sigma f}$

$= \frac{828.8}{50}$

$= 16.576$

Coefficient of MD from Mean $= \frac{\text{MD from Mean}}{\text{Mean}}$

$= \frac{16.576}{47.6}$

$= 0.348$

Calculate coefficient of MD from Median for the following data.

20, 38, 15, 12, 9, 10, 17, 8, 4, 5

A.O = 4, 5, 8, 9, 10, 12, 15, 17, 20, 38

Median $= \left[\frac{N+1}{2}\right]^{th}$ value $= \frac{11}{2}^{th}$ value

$= 5.5^{th}$ value

$= \left[\frac{10+12}{2}\right]$

$= 11$

x	4	5	8	9	10	12	15	17	20	38	
|x-M|	7	6	3	2	1	1	4	6	9	27	66

MD from Median $= \frac{\Sigma|x-M|}{n}$

$= \frac{66}{10}$

$= 6.6$

Coefficient of MD from Median $= \frac{\text{MD from Median}}{\text{Median}} = \frac{6.6}{11} = 0.6$

Scanned with CamScanner

Standard deviation [SD] σ

Best measures for finding Mean

Indidual series

$$SD(\sigma) = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

Discret series

$$SD = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{\Sigma f}}$$

continuous series

$$SD = \sqrt{\frac{\Sigma f(m-\bar{x})^2}{\Sigma f}}$$

$$\text{Coefficient of variation (CV)} = \frac{SD}{Mean} \times 100$$

$$= \frac{\sigma}{\bar{x}} \times 100$$

Calculate SD and coefficient of variation

20, 38, 15, 12, 9, 10, 17, 8, 4, 5

x	$(x-\bar{x})^2$
20	38.44
38	585.64
15	1.44
12	3.24
9	23.04
10	14.44
17	10.24
8	33.64
4	96.04
5	77.44
	883.6

$$\bar{x} = \frac{\Sigma x}{n} = \frac{138}{10} = 13.8$$

$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$= \sqrt{\frac{883.6}{10}}$$

$$= \sqrt{88.36}$$

$$= 9.40$$

coefficient of variation

$$= \frac{SD}{Mean} \times 100$$

$$= \frac{9.40}{13.8} \times 100 = 68.11$$

Scanned with CamScanner

Find SD, CV for the following data.

x	3	5	7	9	10	14
f	4	7	9	4	3	2

x	f	fx	$(x-\bar{x})^2$	$f(x-\bar{x})^2$
3	4	12	16.2409	64.9636
5	7	35	4.1209	28.8463
7	9	63	0.0009	0.0081
9	4	36	3.8809	15.5236
10	3	30	8.8209	26.4627
14	2	28	48.5809	97.1618
	29	204		232.9661

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{204}{29} = 7.03$$

$$SD = \sqrt{\frac{\Sigma f(x-\bar{x})^2}{\Sigma f}} = \sqrt{\frac{232.9661}{29}}$$

$$= \sqrt{8.033}$$

$$= 2.834$$

$$CV = \frac{SD}{Mean} \times 100 = \frac{2.834}{7.03} \times 100$$

$$= 40.31$$

Find SD, CV for the following data

x	0–20	20–40	40–60	60–80	80–100
f	4	15	17	11	3

Scanned with CamScanner

x	f	m	fm	$(m-\bar{x})^2$	$f(m-\bar{x})^2$
0-20	4	10	40	1413.76	5655.04
20-40	15	30	450	309.76	4646.4
40-60	17	50	850	5.76	97.92
60-80	11	70	770	501.76	5519.36
80-100	3	90	270	1797.76	5393.28
	50		2380		21312

$$\text{Mean } (\bar{x}) = \frac{\Sigma fm}{\Sigma f} = \frac{2380}{50}$$

$$= 47.6$$

$$SD = \sqrt{\frac{\Sigma f(m-\bar{x})^2}{\Sigma f}}$$

$$= \sqrt{\frac{21312}{50}}$$

$$= 20.64$$

~~CV = 20.64 / 47.6~~

$$CV = \frac{SD}{Mean} \times 100 = \frac{20.64}{47.6}$$

$$= 43.36.$$

Scanned with CamScanner

Calculate Bowley's coefficient of skewness

Annual sales (in Rs)	0-20	20-50	50-100	100-250	250-500	500-1000
No. of items	20	50	69	30	25	19

Sol

Annual sales (in Rs)	No of items (f)	cf
0-20	20	20
20-50	50	70
50-100	69	139
100-250	30	169
250-500	25	194
500-1000	19	213
	N = 213	

$Q_1 = \frac{N}{4} = \frac{213}{4} = 53.25$

$Q_1 = L_1 + \left[\frac{\left(\frac{N}{4} - cf\right)}{f} \times i\right]$

$= 20 + \left[\frac{53.25 - 20}{50} \times 30\right]$

$= 39.95$

Scanned with CamScanner

$M = N/2 = \frac{213}{2} = 106.5$

$$M = L + \left[\frac{N/2 - cf}{f} \times i\right]$$

$$= 50 + \left[\frac{106.5 - 70}{69} \times 50\right]$$

$$= 76.45$$

$$Q_3 = \frac{3N}{4} = \frac{3(213)}{4} = 159.75$$

$$= 100 + \left[\frac{159.75 - 139}{30} \times 150\right]$$

$$= 203.75$$

Bowley's coeff

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

$$= \frac{203.75 + 39.95 - 2 \times 76.45}{203.75 - 39.95}$$

$$= 0.5543.$$

Scanned with CamScanner

Ex: 6 Calcul KPCOS

Wage for Item Rs	12	15	20	25	30	40	50
No of Items	10	25	40	70	32	13	10

Sol:-

	x	f	$d = x - A$, $A = 25$	fd	fd^2	fm
	12	10	-13	-130	1690	120
	15	25	-10	-250	2500	375
	20	40	-5	-200	1000	800
A&M	25	70	0	0	0	1750
	30	32	5	160	800	960
	40	13	15	195	2925	520
	50	10	25	250	6250	500
		N=200		$\Sigma fd = 25$	$\Sigma fd^2 = 15165$	5025

$$\text{Mean } \bar{x} = \frac{\Sigma fm}{N} = \frac{5025}{200} = 25.13$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} = \sqrt{\frac{15165}{200} - \left(\frac{25}{200}\right)^2} = 8.71$$

$$\text{Mode } z = 25 \qquad Sk_p = \frac{\bar{x} - z}{\sigma} = \frac{25.13 - 25}{8.71} = 0.0149.$$

Scanned with CamScanner