
بسم الله الرحمن الرحيم

Probability and Statistics (MS-301) Software Engineering

LECTURE # 4 (MEASURE OF DISPERSION)

PREPARED BY

DR. ASMA ZAFFAR

Measures of Dispersion

Measure of central tendency give us good information about the scores in our distribution. However, we can have very different shapes to our distribution, yet have the same central tendency. Measures of dispersion or variability will give us information about the spread of the scores in our distribution. Are the scores clustered close together over a small portion of the scale, or are the scores spread out over a large segment of the scale?

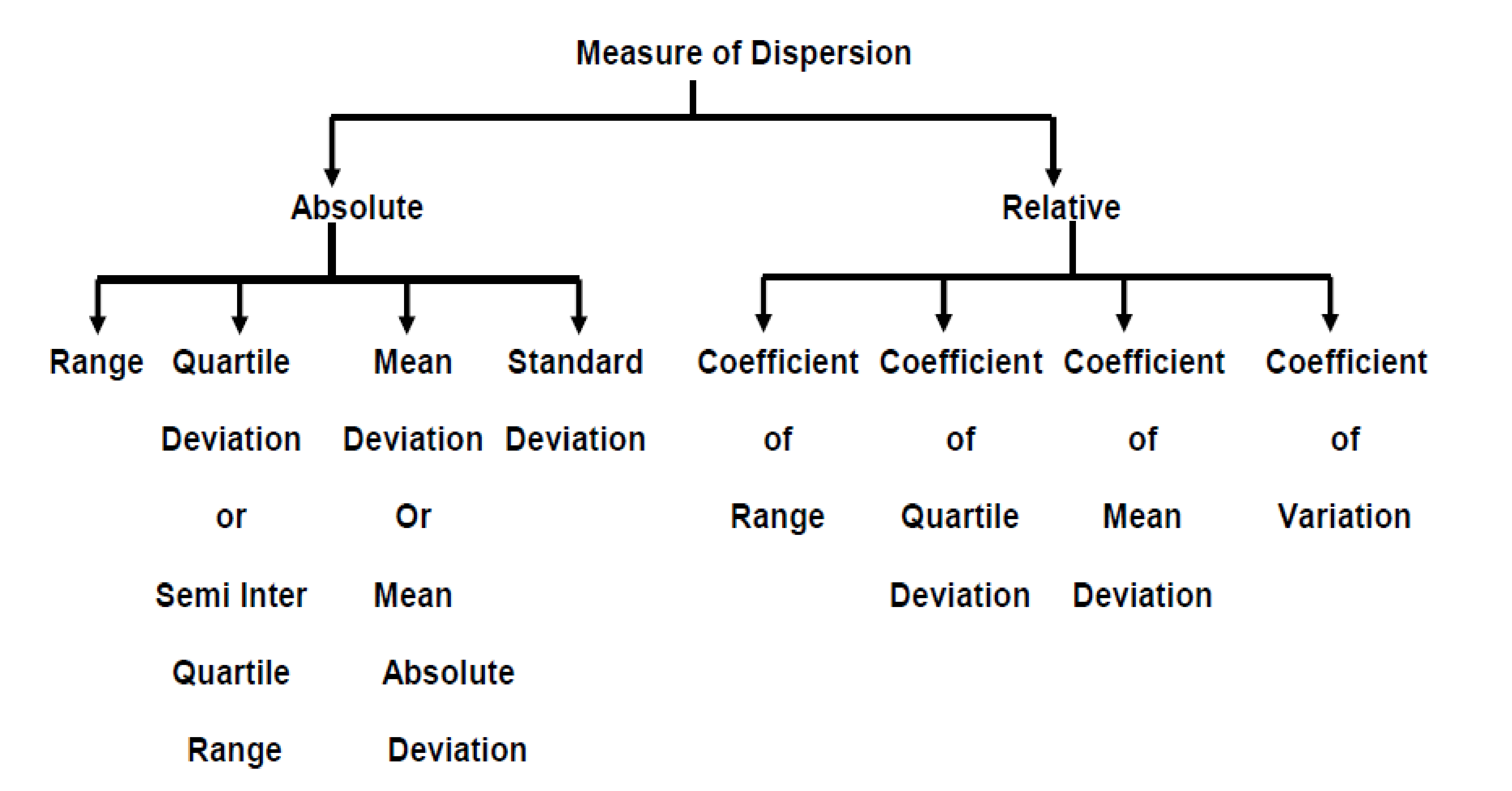
The most common measures of dispersion are

1. Range
2. Quartile Deviation
3. Mean Deviation (or Mean Absolute deviation)
4. Variance
5. Standard Deviation

Types of Measures of Dispersion

There are two types of measures of dispersion,

Absolute Measure	Relative Measure
a. These measures of dispersion will have the same units as those of the variables	a. These are usually expressed as ratios or percentages and hence unit free
b. Absolute measures are related to the distribution itself.	b. Relative measures are used i) to compare variability between two or more series. ii) To check the relative accuracy of the data


Measure of Dispersion
Absolute
Relative
Range
Quartile Deviation or Semi Inter Quartile Range
Mean Deviation Or Mean Absolute Deviation
Standard Deviation
Coefficient of Range
Coefficient of Quartile Deviation
Coefficient of Mean Deviation
Coefficient of Variation

Range:

Range is the difference between the largest (Max) and the smallest (Min) values.

$$\text{Range } (R) = \text{Max} - \text{Min}$$

Example:

Find the range for the sample values: 26, 25, 35, 27, 29, 29.

Solution:

.max = 35

.min = 25

Range (R) = 35 – 25 = 10 (unit)

Notes:

1. The unit of the range is the same as the unit of the data.
2. The range is not useful as a measure of the variation since it only takes into account two of the values (it is not good), however, it plays a significant role in many applications.

- For grouped frequency distribution:
  - Range = Upper boundary of last class – Lower boundary of 1st class
  - Range = Upper Limit of last class – Lower limit of 1st class + 1
- **Co-efficient of Range (Relative Range)=** $\frac{H - L}{H + L} \times 100$

EXAMPLE 1 Marks of ten students are recorded below:

51, 50, 40, 90, 75, 60, 44, 30, 23, 20

Calculate range and coefficient of range.

SOLUTION We have: L = 90 marks and s = 20 marks

$$\text{Range} = L - S \quad \text{Where: } L = 90 \text{ and } S = 20$$

$$\text{Range} = 90 \text{ marks} - 20 \text{ marks} = 70 \text{ marks}$$

$$\text{Coefficient of Range} = \frac{\text{Range}}{L + S} = \frac{70 \text{ marks}}{(90 \text{ marks} + 20 \text{ marks})}$$

$$= \frac{70 \text{ marks}}{110 \text{ marks}} = \frac{70}{110} = \frac{7}{11}$$

EXAMPLE 2 Calculate range and coefficient of Range from the following, adopting both the methods.

Age group (years)	No. of workers
10 – 20	10
20 – 30	20
30 – 40	40
40 – 50	30
50 – 60	5

SOLUTION

METHOD I: We have: L = 60 and s = 10

$$\text{Range} = L - S = 60 - 10 = 50 \text{ years}$$

$$\text{Coefficient of Range} = \frac{\text{Range}}{L + S} = \frac{50 \text{ years}}{(60 + 10) \text{ years}}$$

$$= \frac{50 \text{ years}}{70 \text{ years}} = \frac{50}{70} = \frac{5}{7}$$

EXAMPLE 3

Calcualte Range and coefficient of Range from the following data:

Farm size (Acres)	No. of farms	Mid – Points (x)
0 – 40	94	20
41 – 80	61	60.5
81 – 120	91	100.5
121 – 160	34	140.5
161 – 200	69	180.5
201 – 240	13	220.5

SOLUTION METHOD:

$$\text{Range} = L - S = 220.5 - (20) = 200.5 \text{ acres}$$

$$\text{Coefficient of Range} = \frac{\text{Range}}{L + S} = \frac{200.5}{220.5 + 20} = 0.83$$

picture of variabl...

4.2.2.2 MEAN DEVIATION

Mean Devaition & Coefficient of Mean Deviation is another absolute meausre of dispersion. It is defined in following terms:

"Mean deviation is arithmetic mean of the absolute deviations from some central value (Mean, Median or Mode)".

However, arithmetic mean or median is usually taken as an average to calculate mean deviation. The term "absolute deviation" means the value of the deviation without regard to its sign (i.e. disregarding a negative sign).

Mean deviations is abbreviated as M.D and mathematically defined as follows:

Nature of Data	Mean Deviation From	
	Arithmetic Mean	Median
Ungroup data	$M.D. = \frac{\sum \lvert x - \bar{x} \rvert}{n}$	$M.D = \frac{\sum \lvert x - Median \rvert}{n}$
Group data	$M.D = \frac{\sum f \lvert x - \bar{x} \rvert}{\sum f}$	$M.D = \frac{\sum f \lvert x - Median \rvert}{\sum f}$

4.2.2.2.1 Coefficient of Mean Deviation

It relative measure correspond to Mean Deviation. ***It is defined a the ratio of Mean Deviation and used measure of central tendency.***

Hence: Coefficient of M.D. $= \frac{\text{M.D. from Mean}}{\text{Mean}}$ or

Coeffiocient of M.D. $= \frac{\text{M.D from Median}}{\text{Median}}$

Example 1

Find Mean Deviation : (i) about Mean (ii) about Median, in the following data:

2, 7, 9

Solution :

i) $M.D(\bar{x}) = \dfrac{\sum |x - \bar{x}|}{n}$

where $\bar{x} = \dfrac{\sum x}{n} = \dfrac{18}{3} = 6$

	x	$x-6$	$\lvert x-6 \rvert$
	2	−4	4
	7	1	1
	9	3	3
Total	18		8

then $M.D(\bar{x}) = \dfrac{\sum |x - 6|}{3} = \dfrac{8}{3} = 2.67$

ii) First we find median i.e. Median $= \tilde{x} = 7$.

Therefore,

$M.D(\tilde{x}) = \dfrac{\sum |x - \tilde{x}|}{n}$

$M.D(\tilde{x}) = \dfrac{\sum |x - 7|}{3} = 7$

	x	$x-7$	$\lvert x-7 \rvert$
	2	-5	5
	7	0	0
	9	2	2
Total			7

$M.D(\tilde{x}) = 233$

EXAMPLE 3 **Calculate Mean Deviation from (i) arithmetic mean, and (ii) median from the marks of ten students given as follows; 40, 50, 45, 40, 65, 75, 80, 79, 25 and 45. Also complete coefficient of M.D. in each case.**

SOLUTION CASE – 1:

Mean deviation from arithmetic mean.

The data given is ungroup data and the involved variable is marks obtained by a student which is treated as continuous variable.

First we find arithmetic mean of the data by using following formula:

$$\overline{X} = \frac{\sum x}{n} = \frac{544}{10} = 54.4 \text{ marks}$$

Where: $\sum x = 544$ and $n = 10$

The necessary computations for the calculation of MD are given in following table.

x	$\lvert x - \overline{x} \rvert$	$\lvert x - \text{Median} \rvert$
40	$\lvert -14.4 \rvert = 14.4$	$\lvert -7.5 \rvert = 7.5$
50	$\lvert -4.4 \rvert = 4.4$	$\lvert 2.5 \rvert = 2.5$
45	$\lvert -9.4 \rvert = 9.4$	$\lvert -2.5 \rvert = 2.5$
40	$\lvert -14.4 \rvert = 14.4$	$\lvert -7.5 \rvert = 7.5$
65	$\lvert 10.6 \rvert = 10.6$	$\lvert 17.5 \rvert = 17.5$
75	$\lvert 20.6 \rvert = 20.6$	$\lvert 27.5 \rvert = 27.5$
80	$\lvert 25.6 \rvert = 25.6$	$\lvert +32.5 \rvert = 32.5$
79	$\lvert 24.6 \rvert = 24.6$	$\lvert 31.5 \rvert = 31.5$
25	$\lvert -29.4 \rvert = 29.4$	$\lvert -22.5 \rvert = 22.5$
45	$\lvert -9.4 \rvert = 9.4$	$\lvert -2.5 \rvert = 2.5$
$\sum x = 544$	$\sum \lvert x - \overline{x} \rvert = 162.8$	$\sum \lvert x - \text{Median} \rvert = 154$

$$\text{M.D.} = \frac{\sum \lvert x - \overline{x} \rvert}{n} = \frac{162.8}{10} = 16.28 \text{ marks}$$

Where: $\sum|x - \bar{x}| = 162.8$ and $n = 10$

$$\text{Coefficient of M.D.} = \frac{\text{M.D. from Mean}}{\text{Mean}} = \frac{16.28 \text{ marks}}{54.4 \text{ marks}}$$

$$= \frac{16.28}{54.4} = 0.30$$

CASE – II: Mean Deviation From Median

First we find median of the given data.

Data in an array: 25, 40, 40, 45, 45, 50, 65, 75, 79, 80

Median = The value of $\left(\frac{n+1}{2}\right)$th term in an array data

= The value of $\left(\frac{10+1}{2}\right)$th terms in an array data

= The value of 5.5th term in an array data

= $45 + 0.5(50 - 45)$ = 47.5 marks

The necessary computation for the calculations of M.D. are given in above table.

$$\text{M.D.} = \frac{\Sigma|x - \text{Median}|}{n} = \frac{154}{10} = 15.4 \text{ marks}$$

Where: $\Sigma|x - \text{Median}| = 154$ and $n = 10$

$$\text{Coefficient of M.D} = \frac{\text{M.D. from Median}}{\text{Meian}} = \frac{15.4 \text{ marks}}{47.5 \text{ marks}}$$

$$= \frac{17.4}{47.5} = 0.32$$

EXAMPLE 4

The following data indicate no. of patient arrived at a rural health centre in a particular month.

No. of patients:	5	6	7	8	9	10
No. of days:	4	8	10	5	2	1

Find the mean deviation from the median of the following data:

SOLUTION In this problem we are interested to find mean deviation from median. The variable involved in the problem is No. of patients arrived which is a discrete variable, so we find median first by using following formula:

Median = The value of $\left(\frac{n+1}{2}\right)$th term in an array data

= The value of $\left(\frac{30+1}{2}\right)$th term in an array data

= The value of 15.5th term in an array data

= 7 patients

The necessary computations are given in the following table:

No. of paints arrived (x)	No. of days (f)	c.f.	|x–Median|	f|x–Median|
5	4	4	|–2| = 2	8
6	8	12	|–1| = 1	8
7	10	22	|0| = 0	0
8	5	27	|1| = 1	5
9	2	29	|2| = 2	4
10	1	30 = n	|3| = 3	3
Total	$\Sigma f = 30 = n$	—	—	$\Sigma f\|x-\text{median}\| = 28$

$$\text{M.D.} = \frac{\Sigma f|x - \text{Median}|}{\Sigma f} = \frac{28}{30} = 0.93 = 1$$

patient

EXAMPLE 5 Calcualte mean deviation by using Mean and Median as an average for the following data:

Class Interval	Frequency
6.5 – 7.5	5
7.5 – 8.5	12
8.5 – 9.5	25
9.5 – 10.5	48
10.5 – 11.5	32
11.5 – 12.5	6
12.5 – 13.5	2

SOLUTION **CASE I:**

Mean Deviation From Mean

$$A.M. = \overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{1286}{130} = 9.89$$

Where: $\Sigma fx = 1286$ and $\Sigma f = 130$

Class Interval	Frequency (f)	x	f x	$f\lvert x - \overline{x}\rvert$
6.5 – 7.5	5	7	35	14.45
7.5 – 8.5	12	8	96	22.68
8.5 – 9.5	25	9	225	22.25
9.5 – 10.5	48	10	480	5.28
10.5 – 11.5	32	11	352	35.52
11.5 – 12.5	6	12	72	12.66
12.5 – 13.5	2	13	26	6.22
Total	$\Sigma f = 130$		$\Sigma fx = 1286$	$\Sigma f\lvert x - \overline{x}\rvert = 119.06$

$$M.D. = \frac{\Sigma f\lvert x - \overline{x}\rvert}{\Sigma f} = \frac{119.06}{130} = 0.92$$

Where: $\Sigma f\lvert x - \overline{x}\rvert = 119.06$ and $\Sigma f = 130$

where $\Sigma f|x - \bar{x}| = 119.00$ and $\Sigma f = 130$

CASE II: Mean Deviation From Median

Class Interval	Frequency (f)	c.f.	x	f|x−Median|
6.5 – 7.5	5	5	7	14.90
7.5 – 8.5	12	17	8	23.76
8.5 – 9.5	25	42	9	24.50
9.5 – 10.5	48	90	10	0.96
10.5 – 11.5	32	122	11	32.64
11.5 – 12.5	6	128	12	12.12
12.5 – 13.5	2	130=n	13	6.04
Total	$\Sigma f = 130 = n$			$\Sigma f\|x-\text{Median}\| = 114.92$

$$\text{Median} = L + \frac{h}{f}\left(\frac{n}{2} - c\right) = 9.5 + \frac{1}{48}(65 - 42) = 9.98$$

Where: $\frac{n}{2} = \frac{130}{2} = 65$, $L = 9.5$, $h = 1$, $f = 48$ and $c = 22$

$$\text{M.D.} = \frac{\sum f|x - \text{Median}|}{\sum f} = \frac{114.92}{130} = 0.884$$

Example 5

Calculate Mean Deviation from Mean and Mean Deviation from Median in the following data.

Classes	2 - 4	4 - 6	6 - 8	8 - 10	10 - 12
Freq.	2	3	6	2	1

Solution :

First we compute Mean ($\bar{x}$) and Median ($\tilde{x}$) i.e.

C.B.	*x*	*f*	*f.x*	*C.f.*
2 - 4	3	2	6	2
4 - 6	5	3	15	5
6 - 8	7	6	42	11
8 - 10	9	2	18	13
10 - 12	11	1	11	14
Total		14	92	

$$\bar{x} = \frac{\sum f x}{\sum f} = \frac{92}{14} = 6.57$$

$$\tilde{x} = l + \frac{h}{f}\left(\frac{\sum f}{2} - C.f.\right) = 6 + \frac{2}{6}(7-5) = 6 + \frac{2}{3} = 6.7$$

Then $M.D(\bar{x}) = \dfrac{\sum f|x - \bar{x}|}{\sum f}$ and $M.D(\tilde{x}) = \dfrac{\sum f|x - \tilde{x}|}{\sum f}$

Both are computed in the following table.

Quartile Deviation (or Semi Interquartile Range)

The quartile deviation is defined as ;

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

where Q_1 = First or lower quartile

Q_3 = Third or upper quartile

Example 1

Find Quartile Deviation for the following data.

10, 13, 9, 6, 4, 9, 18, 8, 7, 5, 14

Solution :

First we Array the data i.e. 4, 5, 6, 7, 8, 9, 9, 10, 13, 14, 18

Here n = 11 then $Q_1 = \left(\frac{n+1}{4}\right)$th value

$$Q_1 = \left(\frac{11+1}{4}\right)\text{th value} = \text{3rd value} = 6$$

$$Q_1 = 6$$

$$Q_3 = 3\left(\frac{n+1}{4}\right)\text{th value}$$

$$Q_3 = 3\left(\frac{11+1}{4}\right)\text{th value} = \text{9th value} = 13$$

$$Q_3 = 13$$

Then, Quartile Deviation $= Q.D. = \frac{Q_3 - Q_1}{2} = \frac{13-6}{2} = 3.5$

Example 2

Find Quartile Deviation for the following data:

14, 10, 17, 5, 9, 20, 8, 24, 22, 13, 26, 32, 27, 0, 4

Solution :

First we array the data i.e.

0, 4, 5, 8, 9, 10, 13, 14, 17, 20, 22, 24, 26, 27 32 Here n = 15

Then $Q_1 = \left(\frac{n+1}{4}\right)$th value $= \left(\frac{15+1}{4}\right)$th value = 24th value

$Q_1 = 4$

$Q_3 = 3\left(\frac{n+1}{4}\right)$th value $= 3\left(\frac{15+1}{4}\right)$th value = 12th value

$Q_3 = 24$

Then, Quartile Deviation $= Q.D. = \frac{Q_3 - Q_1}{2} = \frac{24 - 8}{2}$

$Q.D = 8$

Example 3

Find Q.D. in the following frequency distribution:

x	0	1	2	3	4	5	6
f	10	18	40	25	10	8	2

Solution :

Since $n = \sum f = 115$

$Q_1 = \left(\frac{n+1}{4}\right)$th value $=$ 29th value

$Q_1 = 2$

$Q_3 = 3\left(\frac{n+1}{4}\right)$th value $=$ 87th value

$Q_3 = 3$

x	f	C.f.
0	10	10
1	18	28
2	40	68
3	25	93
4	12	105
5	8	113
6	2	115

2 2

Example 4

Compute Quartile Deviation of the following data:

Wages	100	140	180	200	240	300	320
Workers	5	7	10	16	14	8	4

Solution :

$$Q_1 = \left(\frac{n+1}{4}\right) \text{th value}$$

$$= \left(\frac{64+1}{4}\right) \text{th value} = 16.25\text{th value}$$

$$Q_1 = 180$$

$$Q_3 = 3\left(\frac{n+1}{4}\right) \text{th value}$$

$$= 3\left(\frac{64+1}{4}\right) \text{th value} = 48.75\text{th value}$$

x	f	$C.f.$
100	5	5
140	7	12
180	10	22
200	16	38
240	14	52
300	8	60
320	4	64
Total	64	

16.25th val

48.75th val

$$Q_3 = 240$$

$$\text{Quartile Deviation} = Q.D. = \frac{Q_3 - Q_1}{2} = \frac{240 - 180}{2} = 30$$

$$Q.D. = 30$$

Example 5

From the following data compute Quartile Deviation:

Weight in Kg.	45-49	50-54	55-59	60-64	65-69	70-74
Freq.	62	272	314	75	22	5

Example 6

Find Quartile Deviation for the following distribution:

Classes	2 - 4	4 - 6	6 - 8	8 - 10	10 - 12	Total
Freq.	10	25	18	12	5	70

Solution :

C.B.	f	C.f.
2 - 4	10	10
4 - 6	25	35
6 - 8	18	53
8 - 10	12	65
10 - 12	5	70

$$Q_1 = l + \frac{h}{f}\left(\frac{\sum f}{4} - C.f.\right)$$

$$Q_1 = 4 + \frac{2}{25}(175 - 10)$$

$$Q_1 = 4.6$$

$$Q_3 = l + \frac{h}{f}\left(\frac{3\sum f}{4} - C.f.\right)$$

$$Q_3 = 6 + \frac{2}{18}(525 - 35) \qquad Q_3 = 7.9$$

$$\text{Quartile Deviation} = Q.D. = \frac{Q_3 - Q_1}{2} = \frac{79 - 46}{2} = 1.65$$

4.2.2.3 Standard Deviation

It is the most important and widely used absolute measure of dispersion. Its significance lies in the fact that it is free from those shortcomings from which the earlier methods suffer. It is defined in the following terms:

"Standard deviation is positive square root of the mean of squared deviations taken from arithmetic mean of the data".

Standard deviation is also known as the root mean squared deviation. It is usually denoted by the Greek letter σ (read as sigma) or S.D. or simply by S.

Symbolically, standard deviation is defined as follows:

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

Variance & Standard Deviation

If $x_1, x_2, x_3, \ldots\ldots\ldots x_n$ are the n-values of the variable X then Variance is defined as

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n} \quad \text{where } \bar{x} = \frac{\sum x}{n}$$

OR

May be computed as $\boxed{\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$

In case of a frequency distribution

$$\sigma^2 = \frac{\sum f(x - \bar{x})^2}{\sum f} \quad \text{where } \bar{x} = \frac{\sum fx}{\sum f}$$

OR

May be computed as $\boxed{\sigma^2 = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$

EXAMPLE 7 **Compute the standard deviation form the following data using all the methods: 2, 4, 8, 6, 10, 12.**

SOLUTION

CASE – I: Actual Mean Method:

x	$(x-\bar{x})^2$
2	$(2-7)^2 = 25$
4	$(4-7)^2 = 9$
8	$(8-7)^2 = 1$
6	$(6-7)^2 = 1$
10	$(10-7)^2 = 9$
12	$(12-7)^2 = 25$
$\Sigma x = 42$	$\Sigma(x-x)^2 = 70$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{42}{6} = 7$$

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

$$= \sqrt{\frac{70}{6}} = 3.42$$

EXAMPLE 8 Compute standard deviation from the following distribution of marks by using all the methods.

Marks	No. of Students
1 – 3	40
3 – 5	30
5 – 7	20
7 – 9	10

SOLUTION

METHOD I: Actual mean method.

Marks	f	x	fx	$(x - x)^2$	$f(x - \bar{x})^2$
1 – 3	40	2	80	4	160
3 – 5	30	4	120	0	0
5 – 7	20	6	120	4	80
7 – 9	10	8	80	16	160
Total	$\Sigma f = 100$	—	$\Sigma f = 400$	—	$\Sigma f(x - x)^2 = 400$

$$\bar{x} = \frac{\Sigma fx}{n} = \frac{400}{100} = 4$$

$$\sigma = \sqrt{\frac{\Sigma f(x - \bar{x})^2}{6}} = \sqrt{\frac{400}{100}} = \sqrt{4} = 2$$

EXAMPLE 9

From the following information about the accidents on a road in 150 days, calculate standard deviation of accidents:

No. of accidents per day :	0	1	2	3	4	5
No. of days :	76	38	20	10	4	2

SOLUTION

The given data is group data and the variable involved is No. of accidents per day. Let us denote No. of accidents by x and No. of days by f. We prefer to use the following formulae becaue the values of the variable are small and one can verbally do the computations.

$$S.D. = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$$

x	f	fx	fx^2
0	76	0	0
1	38	38	38
2	20	40	80
3	10	30	90
4	4	16	64
5	2	10	50
Total	$\Sigma f = 150$	$\Sigma fx = 134$	$\Sigma fx^2 = 322$

$$\text{S.D.} = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2} = \sqrt{\frac{322}{150} - \left(\frac{134}{150}\right)^2}$$

$$= \sqrt{2.1467 - 0.7980}$$

$$= \sqrt{1.3487} = 1.16$$

Questions?