1. The measures of dispersion are (1) Range (2) Quartile deviation (3) Mean deviation (4) Variance (5) Standard deviation

2. In a given series of values (data), the difference of maximum (greatest) value and minimum (least) value is called **range**

3. The arithmetic average of absolute values of the deviations of the variates measured from an average mean or median (or mode) is called **mean deviation about mean or median** (or mode)

4. Mean deviation for ungrouped data :

   Mean deviation about mean $= \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \bar{x}\right|$, where $\bar{x}$ = mean

   Mean deviation about median $= \frac{1}{n}\sum_{i=1}^{n}\left|x_i - M\right|$, where M = median

5. Mean deviation for grouped data

   Mean deviation about mean $= \frac{\sum_{i=1}^{n} f_i\left|x_i - \bar{x}\right|}{\sum_{i=1}^{n} f_i} = \frac{1}{N}\sum_{i=1}^{n} f_i\left|x_i - \bar{x}\right|$, where $N = \sum_{i=1}^{n} f_i$

   Mean deviation about median $= \frac{1}{N}\sum_{i=1}^{n} f_i\left|x_i - M\right|$, where $N = \sum_{i=1}^{n} f_i$

6. The mean of the squares of the deviations of the variates from their arithmetic mean is called variance. It is denoted by σ^2. The positive square root of variance is called standard deviation and it is denoted by σ.

7. Variance and standard deviation for ungrouped data :

   Variance, $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} x^2{}_i - \bar{x}^2$

   Standard deviation $\sigma = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2\right]} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n} x^2{}_i - \bar{x}^2\right]}$

8. Variance and standard deviation of a discrete frequency distribution:

   Variance, $\sigma^2 = \frac{1}{N}\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^2$ where $N = \sum_{i=1}^{n} f_i$

   Standard deviation, $\sigma = \sqrt{\left[\frac{1}{N}\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^2\right]}$ or $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n}\left[f_i x^2{}_i - \left(f_i x_i\right)^2\right]}$

9. Standard deviation for continuous frequency distribution:

   Variance, $\sigma = \frac{1}{N}\sqrt{N\,\Sigma\, f_i x_i^2 - (\Sigma\, f_i x_i)^2}$

   $\sigma = \frac{h}{N}\sqrt{N\,\Sigma\, f_i y_i^2 - (\Sigma\, f_i y_i)^2}$ where $y_i = \frac{x_i - A}{h}$ and h is the length of class interval.

10. Coefficient of variation $= \frac{\sigma}{(\bar{x})} \text{x}100$.

11. If each of the observations $x_1, x_2, \ldots\ldots\ldots\ldots, x_n$ is increased by k, where k is a positive or negative number, then the variance remains unchanged.

12. If each observation in a data is multiplied by a constant k, then the variance of the resulting observations is k^2 times that of the variance of original observations.

* * * * * *

LEVEL - I (VSAQ)

1. Find the mean deviation from the mean of the following discrete data : 3, 6, 10, 4, 9, 10.

A: Mean of the data 3, 6, 10, 4, 9, 10 is

$$\bar{x} = \frac{3+6+10+4+9+10}{6}$$

$$= \frac{42}{6}$$

$$= 7$$

$\therefore$ Mean deviation from the mean

$$= \frac{\sum_{i=1}^{6} |x_i - \bar{x}|}{n}$$

$$= \frac{4+1+3+3+2+3}{6}$$

$$= \frac{16}{6}$$

$$= 2.67.$$

2. Compute the mean deviation about the median of the data 6, 7, 10, 12, 13, 4, 12, 16.

A:Ascending order of the given data is 6, 7, 10, 12, 12, 13, 16.

$$\text{Median } M = \frac{x_4 + x_5}{2}$$

$$= \frac{10+12}{2}$$

$$= 11.$$

$\therefore$ Mean deviation from the median

$$= \frac{\sum_{i=1}^{8} (x_i - M)}{8}$$

$$= \frac{7+5+4+1+1+1+2+5}{8}$$

$$= \frac{26}{8}$$

$$= 3.25.$$

3. Find the variance and standard deviation of the data 5, 12, 3, 18, 6, 8, 2, 10.

$$\text{Mean } \bar{x} = \frac{5+12+3+18+6+8+2+10}{8}$$

$$\bar{x} = \frac{64}{8} = 8$$

Variance $\sigma^2 = \frac{\sum_{i=1}^{8}(x_i - \bar{x})^2}{n}$

$= \frac{9+16+25+100+4+36+4}{8}$

$= \frac{194}{8}$

= 24.25.

Standard deviation $\sigma = \sqrt{24.25}$

= 4.95.

4. The coefficient of variation of two distributions are 60 and 70 and their standard deviations are 21 and 16 respectively. Find their arithmetic means.

A: Let $\bar{x}$ and $\bar{y}$ be the means of given two distributions

Coefficient of variation $C.V. = \frac{\sigma}{\bar{x}} \times 100$

$60 = \frac{21}{\bar{x}}(100)$

$\Rightarrow \bar{x} = 35$.

For the second distribution $C.V. = \frac{\sigma}{\bar{y}} \times 100$

$70 = \frac{16}{y} \times 100$

$\Rightarrow \bar{y} = 22.85$.

5. The variance of 20 observations is 5. If each of the observations is multiplied by 2, find the variance of the resulting observations.

A: We know that if each observation in a data multiplied by a constant k, then the variance of the resulting observations is k^2 times that of the variance of original observations.

Here each of the observation is multiplied by 2 .

$\therefore$ Variance of resulting observations

$= 2^2(5)$

$= 4(5)$

$= 20$.

6. If each of the observations $x_1, x_2, ...x_n$ is increased by k, where k is a positive or negative number, then show that the variance remains unchanged.

A: For the observations $x_1, x_2, ...x_n$,

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

Variance $\sigma_1^2 = \frac{\sum(x_i - \bar{x})^2}{n}$

Mean of new observations $\bar{y} = \frac{\sum_{i=1}^{n}(x_i + k)}{n}$

$$= \frac{\sum_{i=1}^{n} x_i}{n} + \frac{kn}{n}$$

$$= \bar{x} + k$$

$\therefore$ Variance of new observations $\sigma_2^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}$

$$= \frac{\sum_{i=1}^{n}\left[x_i + k - (x_i + k)\right]^2}{n}$$

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

$$= \sigma_1^2 .$$

Thus the variance of new observations is the same as that of the original observations.

LEVEL - I (LAQ)

1. Find the mean deviation about the mean for the data :

x_i	2	5	7	8	10	35
f_i	6	8	10	6	8	2

x_i	f_i	$f_i\ x_i$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
2	6	12	6	36
5	8	40	3	24
7	10	70	1	10
8	6	48	0	0
10	8	80	2	16
35	2	70	27	54
	40	320		140

Arithemetic mean $\bar{x} = \frac{\sum f_i\ x_i}{\sum f_i} = \frac{320}{40} = 8$

Mean deviation about the mean = $\frac{\sum f_i \lvert x_i - \bar{x}\rvert}{\sum f_i} = \frac{140}{40} = 3.5$

2. **Find the mean deviation from the median for the following data :**

x_i	6	9	3	12	15	13	21	22
f_i	4	5	3	2	5	4	4	3

Arranging x_i's in ascending order, given table can be rewritten as

x_i	f_i	$\lvert x_i - M \rvert$	$f_i \lvert x_i - M \rvert$
3	3	10	30
6	4	7	28
9	5	4	20
12	2	1	2
13	4	0	0
15	5	2	10
21	4	8	32
22	3	9	27
	N = 30		$\sum f_i \lvert x_i - \bar{x} \rvert = 149$

Here median is the average of $\frac{N}{2}, \frac{N}{2}+1^{th}$ observations

$$M = \frac{\sum f_i \lvert x_i - \bar{x} \rvert}{N}$$

$$= \frac{149}{30} = 4.97.$$

3. **Find the mean deviation from the mean of the following data, using the step deviation method.**

Marks	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 70
No. of students	6	5	8	15	7	6	3

Class interval	No. of students f_i	Midvalue x_i	$d_i = \frac{x_i - A}{h}$	$f_i d_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
0 - 10	6	5	- 3	- 18	28.4	170.4
10 - 20	5	15	- 2	- 10	18.4	92.0
20 - 30	8	25	- 1	- 8	8.4	67.2
30 - 40	15	35	0	0	1.6	24.0
40 - 50	7	45	1	7	11.6	81.2
50 - 60	6	55	2	12	21.6	129.6
60 - 70	3	65	3	9	31.6	94.8
	N = 50			$\sum f_i d_i = -8$		659.2

5. Calculate the variance and standard deviation for the discrete frequency distribution :

x_i	4	8	11	17	20	24	32
f_i	3	5	9	5	4	3	1

We shall construct the following table :

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
4	3	12	- 10	100	300
8	5	40	- 6	36	180
11	9	99	- 3	9	81
17	5	85	3	9	45
20	4	80	6	36	144
24	3	72	10	100	300
32	1	32	18	324	324
	30	420			1374

Here $\sum f_i x_i = 420$, N = 30

Mean $\bar{x} = \frac{\sum f_i x_i}{N} = \frac{420}{30} = 14$

Variance $\sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{N} = \frac{1374}{30} = 45.8$.

Standard deviation $\sigma = \sqrt{45.8} = 6.77$.

6. Calculate the variance and standard deviation of the following continuous frequency distribution :

C.I.	30 - 40	40 - 50	50 - 60	60 - 70	70 - 80	80 - 90	90 - 100
Frequency	3	7	12	15	8	3	2

Now we shall construct the following table with the given data :

Class Interval	Frequency f_i	Midpoint x_i	$y_i = \frac{x_i - A}{h}$	y_i^2	$f_i y_i$	$f_i y_i^2$
30 - 40	3	35	- 3	9	- 9	27
40 - 50	7	45	- 2	4	- 14	28
50 - 60	12	55	- 1	1	- 12	12
60 - 70	15	65	0	0	0	0
70 - 80	8	75	1	1	8	8
80 - 90	3	85	2	4	6	12
90 - 100	2	95	3	9	6	18
	50				- 15	105

Take assumed origin A as 65

h = length of the class = 10

Mean $\bar{x} = A + \left(\frac{\sum f_i y_i}{N}\right) h = 65 + \frac{(-15)(10)}{50} = 62$

Variance $\sigma^2 = \frac{h^2}{N^2}\left[N \sum f_i y_i^2 - \left(\sum f_i y_i\right)^2\right]$

$= \frac{100}{2500}\left[50(105) - (-15)^2\right]$

$= \frac{1}{25}[5250 - 225] = 201$

Standard deviation $\sigma = \sqrt{201} = 14.18.$

7. **The following table gives the daily wages of workers in a factory. Compute the standard deviation and the coefficient of variation of the wages of the workers.**

Wages	125-175	175 - 225	225-275	275 - 325	325 - 375	375 - 425	425 - 475	475 - 525	525 - 575
No. of workers	2	22	19	14	3	4	6	1	1

We shall construct the following table with the given data :

Class interval	Midpoint	Frequency f_i	$y_i = \frac{x_i - A}{h}$	y_i^2	$f_i y_i$	$f_i y_i^2$
125 - 175	150	2	- 4	16	- 8	32
175 - 225	200	22	- 3	9	- 66	198
225 - 275	250	19	- 2	4	- 38	76
275 - 325	300	14	- 1	1	- 14	14
325 - 375	350	3	0	0	0	0
375 - 425	400	4	1	1	4	4
425 - 475	450	6	2	4	12	24
475 - 525	500	1	3	9	3	9
525 - 575	550	1	4	16	4	16
		72			- 103	373

Taking assumed origin A as 350, h = 50

Mean $\bar{x} = A + \left(\frac{\sum f_i y_i}{N}\right) h = 350 + \left(\frac{-103}{72}\right)(50) = 278.47$.

Variance $\sigma^2 = \frac{h^2}{N^2}\left[N\Sigma f_i y_i^2 - (\Sigma f_i y_i)^2\right]$

$= \frac{2500}{72 \times 72}\left[72(373) - (-103)^2\right]$

$\sigma = 88.52.$

Coefficient of variation = $\frac{\sigma}{\bar{x}} \times 100 = \frac{88.52}{278.47} \times 100 = 31.79$ ▪

8. **The scores of two cricketers A and B in 10 innings are given below. Find who is better run getter and who is a more consistent player :**

Scores of A : x_i	**40**	**25**	**19**	**80**	**38**	**8**	**67**	**121**	**66**	**76**
Scores of B : y_i	**28**	**70**	**31**	**0**	**14**	**111**	**66**	**31**	**25**	**4**

A: For cricketer A : Mean $\bar{x} = \frac{540}{10} = 54.$

For cricketer B : Mean $\bar{y} = \frac{380}{10} = 38.$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
40	- 4	196	28	- 10	100
25	29	841	70	32	1024
19	- 35	1225	31	- 7	49
80	26	676	0	- 38	1444
38	- 16	256	14	- 24	575
8	- 46	2116	111	73	5329
67	13	169	66	28	784
121	67	4489	31	- 7	49
66	12	144	25	- 13	163
76	22	484	4	- 34	1156
$\Sigma x_i = 540$		10596	$\Sigma y_i = 380$		10680

Standard deviation of scores of A, $\sigma_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}} = \sqrt{\frac{10596}{10}} = \sqrt{1059.6} = 32.55$

Standard deviation of scores of B, $\sigma_y = \sqrt{\frac{\Sigma(y_i - \bar{y})^2}{n}} = \sqrt{\frac{10680}{10}} = \sqrt{1068} = 32.68$

Coefficient of Variation of A = $\frac{\sigma_x}{\bar{x}} x100 = \frac{32.55}{54} x100 = 60.28$.

Coefficient of Variation of B = $\frac{\sigma_y}{\bar{y}} x100 = \frac{32.68}{38}$ x $100 = 86$.

Since $\bar{\mathbf{x}} > \bar{\mathbf{y}}$, cricketers A is a better run getter.

Also C.V. of A < C.V. of B, Cricketer A is also a more consistent player.

9. The mean of 5 observations is 4.4. Their variance is 8.24. If three of the observations are 1, 2 and 6. Find the other two observations.

A: Let the missing two observations be x, y.

Now, mean = 4.4

$\Rightarrow \frac{1+2+6+x+y}{5} = 4.4$

$\Rightarrow x + y + 9 = 22$

$\Rightarrow x + y = 13$

$\Rightarrow y = 13 - x$ (1)

Also variance $\frac{\sum (x_i - \bar{x})^2}{n} = 8.24$

$\Rightarrow \frac{(1-4.4)^2 + (2-4.4)^2 + (6-4.4)^2 + (x-4.4)^2 + (13-x-4.4)^2}{5} = 8.24$

$\Rightarrow (-3.4)^2 + (-2.4)^2 + 1.6^2 + (x-4.4)^2 + (8.6-x)^2 = 41.20$

$\Rightarrow 11.56 + 5.76 + 2.56 + x^2 - 8.8x + 19.36 + 73.96 - 17.2x + x^2 = 41.20$

$\Rightarrow 2x^2 - 26x + 72 = 0.$

$\Rightarrow x^2 - 13x + 36 = 0$

$\Rightarrow x^2 - 4x - 9x + 36 = 0$

$\Rightarrow x(x-4) - 9(x-4) = 0$

$\Rightarrow (x-4)(x-9) = 0.$

$\Rightarrow x = 4$ or 9.

So the missing two observations are 4,9.

LEVEL - II (VSAQ)

1. Find the mean deviation from the mean of the following discrete data : 6, 7, 10, 12, 13, 4, 12, 16.

Mean of the data = $\frac{6+7+10+12+13+4+12+16}{8}$.

$\bar{x} = \frac{80}{8}$

$= 10.$

$\therefore$ Mean deviation from the mean $= \frac{\sum_{i=1}^{8} |x_i - \bar{x}|}{8}$

$$= \frac{4+3+0+2+3+6+2+6}{8}$$

$$= \frac{26}{8} = 3.25.$$

2. Find the mean deviation about the median for the data : 4, 6, 9, 3, 10, 13, 2.

A: The ascending order of the data is 2, 3, 4, 6, 9, 10, 13.

Median M = x_4 = 6.

$\therefore$ Mean deviation from the median = $\dfrac{\sum_{i=1}^{7}|x_i - M|}{7}$

$$= \frac{4+3+2+0+3+4+7}{7}$$

$$\frac{23}{7}$$

= 3.29.

3. Find the variance for the discrete data : 6, 7, 10, 12, 13, 4, 8, 12.

A: Mean = $\dfrac{6+7+10+12+13+4+8+12}{8}$

$$= \frac{72}{8}$$

= 9.

Variance $\sigma^2 = \dfrac{\sum_{i=1}^{8}(x_i - \bar{x})^2}{8}$

$$= \frac{9+4+1+9+16+25+1+9}{8}$$

$$= \frac{74}{8}$$

= 9.25.

* * * * * *