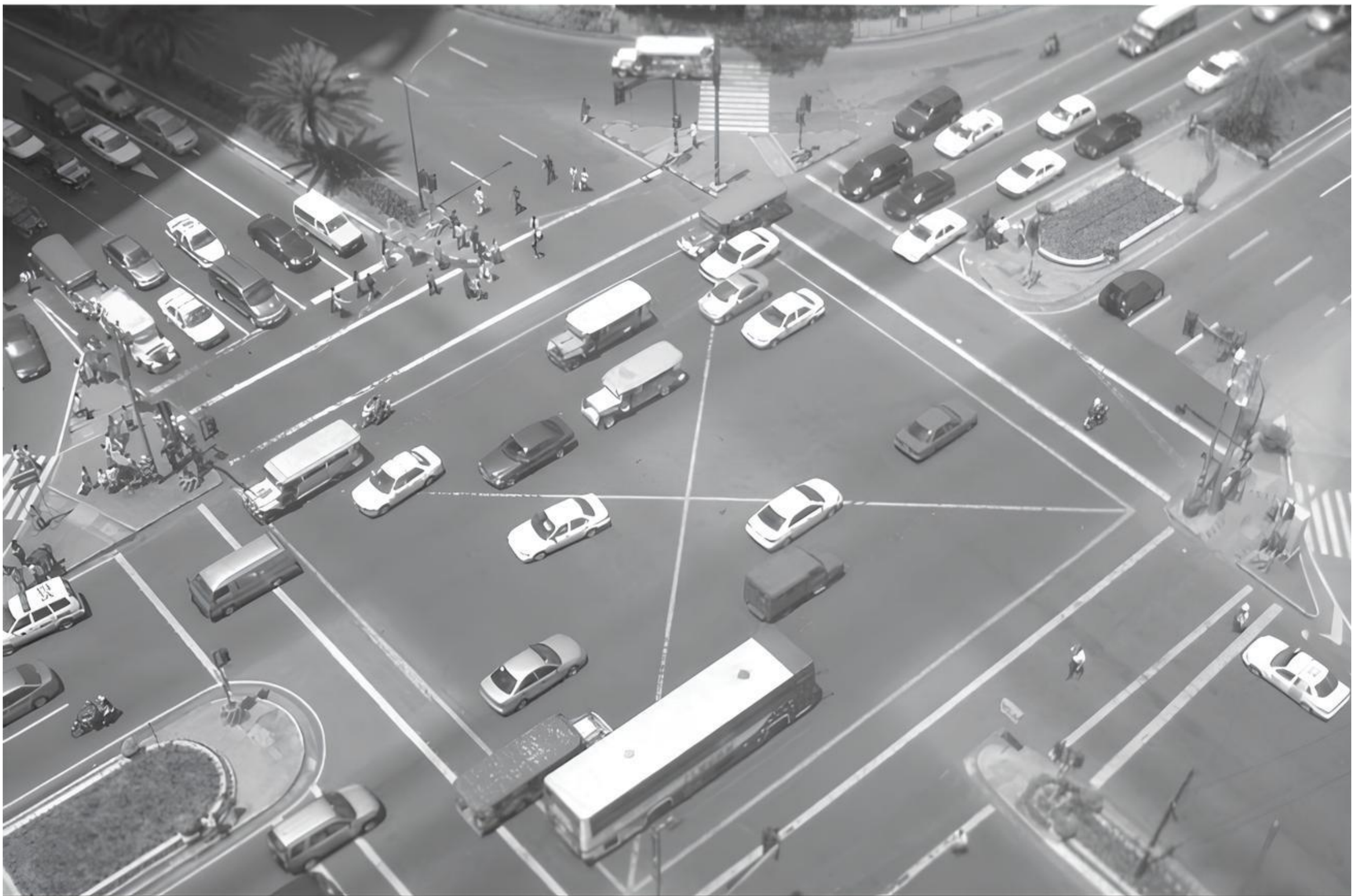
13 Statistics



Traffic engineers regularly use statistics to monitor total traffic in different areas of a city, which allows them to decide whether or not they should add or remove roads to optimize traffic flow.

Topic Notes

- *Measures of Dispersion*
- *Variance and Standard Deviation*

MEASURES OF DISPERSION 1

TOPIC 1

MEASURE OF CENTRAL TENDENCY

These are three measures of central tendency.

1. Arithmetic Mean
2. Median
3. Mode

Range

The measure of dispersion which is easiest to understand and easiest to calculate is the range.

Range is defined as the difference between two extreme observations of the distribution.

Range of distribution = Largest observation – Smallest observation.

Mean of Ungrouped Data

Ungrouped data is the data you first gather from an experiment or study. The data is raw — that is, it is not sorted into categories, classified, or otherwise grouped. An ungrouped set of data is basically a list of numbers.

Consider the ungrouped data $x_1, x_2, ..., x_n$ consisting of n values. There are two methods for computing the mean (denoted by $\bar{x}$) of the given data.

Direct Method

The mean of the given data is

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$$

Deviation Method (or Short-cut Method)

This method involves the following steps:

Step I: Take an assumed mean a. (Any value can be taken as assumed mean, but for simpler calculations, a central value is chosen.)

Step II: Compute $d_i = x_i - a$, i.e., the deviations of the values of data from assumed mean.

Step III: The mean of the given data is

$$\bar{x} = a + \frac{\Sigma d_i}{n}.$$

Example 1.1: Find the median of the following data:

36, 72, 46, 42, 60, 45, 53, 416, 51, 49 [NCERT]

Ans. Given data in ascending order is 36, 42, 45, 46, 46, 49, 51, 53, 60, 72.

Here, number of terms, n = 10, which is even.

Hence, median = $\dfrac{\left(\frac{n}{2}\right)^{th} \text{term} + \left(\frac{n}{2}+1\right)^{th} \text{term}}{2}$

$$= \frac{5^{th}\text{term} + 6^{th}\text{term}}{2}$$

$$= \frac{46 + 49}{2} = 47.5$$

Mean of Discrete Distribution

The mean of a discrete random variable X is a weighted average of the possible values that the random variable can take. Unlike the sample mean of a group of observations, which gives each observation equal weight, the mean of a random variable weights each outcome x_i.

Consider discrete frequency distribution consisting of m distinct values $x_1, x_2 ... x_m$ with frequencies $f_1, f_2, ..., f_m$ respectively. There are two methods for computing the mean (denoted by $\bar{x}$) of the given data.

Direct Method

The mean of the given data is

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

Deviation Method (or Short-cut Method)

This method involves the following steps:

Step I: Take an assumed mean a.

Step II: Compute, $d_i = x_i - a$, i.e., the deviations of the values of data from assumed mean.

Step III: The mean of the given data is

$$\bar{x} = a + \frac{\Sigma f_i d_i}{\Sigma f_i}$$

Mean of Continuous Distribution

Consider a continuous frequency distribution consisting of class intervals with their corresponding frequencies as f_r. We compute the mid-point of the class intervals and denoted them by x_r. There are three methods for computing the mean (denoted by $\bar{x}$) of the given data.

Direct Method

The mean of the given data is

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Deviation method

This method involves the following steps:

Step I: Take an assumed mean a and class-width h.

Step II: Compute $u_i = \frac{x_i - a}{h}$, i.e., the step - deviations of the mid-points of class-intervals from assumed mean.

Step III: The mean of the given data is

$$\bar{x} = a + \frac{h\sum f_i u_i}{\sum f_i}$$

Median of Ungrouped Data

Consider the ungrouped data $x_1, x_2, \ldots, x_n$ consisting of n values. In order to find the median, we have the following working rule:

Step I: Arrange the given up grounded data in ascending order.

Step II: Calculate median of the given data using the formula

(1) Median = $\left(\frac{n+1}{2}\right)^{th}$ term, if n is odd.

(2) Median = $\dfrac{\left(\frac{n}{2}\right)^{th}\text{term} + \left(\frac{n}{2}+1\right)^{th}\text{term}}{2}$, if n is even.

Median of Discrete Distribution

Consider a discrete frequency distribution consisting of m distinct values $x_1, x_2, \ldots, x_3$ with frequencies, $f_1, f_2, \ldots f_3$ respectively. In order to find the median, we have the following working rule.

Step I: Arrange the given data (i.e., values of x_i) in ascending order.

Step II: Find the cumulative frequency, where cumulative frequency represents the sum of all previous frequencies up to the current values.

Step III: Denote $n = \mp\sum f_i$

Step IV: Calculate median of the given data using the formula.

(1) Median = $\left(\frac{n+1}{2}\right)^{th}$ term, if n is odd.

(2) Median = $\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th}$ term, if n is even.

Median of Continuous Frequency Distributions

Median of grouped data is the data that is continuous and is in the form of frequency distribution. Median is the middle most value of the given data that separates the higher half of the data from the lower half.

Step I: Find the mid-points of the class intervals and denote them by x_r

Step II: Find the cumulative frequency, where cumulative frequency represents the sum of all previous frequencies up to the current value.

Step III: Denote $n = \sum f_i$.

Step IV: Identify median class, i.e., the class containing the cumulative frequency $\frac{n}{2}$.

Step V: Calculate median of the given data using the formula

$$\text{Median} = l + \frac{\frac{n}{2} - C}{f} \times h$$

Where l = lower limit of median class

h = width of median class

f = frequency of median class

C = cumulative frequency just preceding median class.

TOPIC 2

MEAN DEVIATION ABOUT MEAN AND MEDIAN

Mean Deviation

Mean deviation about a number represents the mean of the deviations from that number. In this section, we shall learn the methods of calculator mean deviation about mean and mean deviation about median for various types of data.

Mean Deviation About Mean of Ungrounded Data

Consider the up grouped data $x_1, x_2, \ldots, x_n$ consisting of n values. In order to find the mean of deviation about mean, we have the following working rule:

Step I: Calculate mean of the given data using the formula

Mean $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$.

Step II: Compute the deviation d_i of values of x_i from mean, ignoring the ± signs.

i.e., $|d_i| = |x_i - \bar{x}|$. Then

Mean deviation about mean = $\frac{\sum |d_i|}{n}$

Example 1.2: Find the mean deviation about the mean for the data:

38, 70, 48, 40, 42, 55, 63, 46, 54, 44 [NCERT]

Ans. Mean of the given data = $\frac{\text{Sum of all terms}}{\text{Total number of terms}}$

$$\bar{x} = \frac{38+70+48+40+42+55+63+46+54+44}{10}$$

$$= \frac{500}{10}$$

$$= 50$$

x_i	$x_i - \bar{x}$	$\|x_i - \bar{x}\|$
38	38 – 50 = –12	|–12| = 12
70	70 – 50 = 20	|20| = 20
48	48 – 50 = – 2	|–2| = 2
40	40 – 50 = –10	|–10| = 10
42	42 – 50 = –8	|–8| = 8
55	55 – 50 = 5	|5| = 5
63	63 – 50 = 13	|13| = 13
46	46 – 50 = –4	|–4| = 4
54	54 – 50 = 4	|4| = 4
44	44 – 50 = –6	|–6| = 6
		$\sum_{1}^{10} \|x_i - \bar{x}\| = 84$

Mean deviation about mean

$$= \frac{\Sigma |x_i - \bar{x}|}{10}$$

$$= \frac{84}{10} = 8.4$$

Example 1.3: Find the mean deviation about the mean for the data.

x_i	10	30	50	70	90
f_i	4	24	28	16	8

[NCERT]

Ans. First, we will calculate mean

x_i	f_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i\|x_i - \bar{x}\|$
10	4	4 × 10 = 40	|10 – 50| = |–40| = 40	4 × 40 = 160
30	24	24 × 30 = 720	|30 – 50| = |–20| = 20	24 × 20 = 480
50	28	28 × 50 = 1440	|50 – 50| = |0| = 0	28 × 0 = 0
70	16	16 × 70 = 1120	|70 – 50| = |20| = 20	16 × 20 = 320
90	8	8 × 90 = 720	|90 – 50| = |40| = 40	8 × 40 = 320
$\sum f_i$	= 80	$\sum f_i x_i = 4000$		$\sum f_i \|x_i - \bar{x}\| = 1280$

$$\text{Mean}(\bar{x}) = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$$\bar{x} = \frac{4000}{80}$$

$$\bar{x} = 50$$

Mean deviation about mean = $\frac{\Sigma f_i |x_i - \bar{x}|}{\Sigma f_i}$

Putting $\Sigma f_i |x_i - \bar{x}| = 1280$, $\Sigma f_i = 80$

$\therefore$ M.D. $(\bar{x}) = \frac{1}{80} \times 1280 = 16$

Example 1.4: Find the mean deviation about the mean for the data 4, 7, 8, 9, 10, 12, 13, 17. [NCERT]

Ans. The given data is

4, 7, 8, 9, 10, 12, 13, 17

Mean of the data,

$$\bar{x} = \frac{4+7+8+9+10+12+13+17}{8}$$

$$= \frac{80}{8} = 10$$

The deviation of the respective observations from the mean $\bar{x}$, i.e., $x_i - \bar{x}$, i.e. $x_i - 10$

are –6, –3, –2, –1, 0, 2, 3, 7

The absolute values of the deviations, i.e.,

$|x_i - \bar{x}|$, are

6, 3, 2, 1, 0, 2, 3, 7

The required mean deviation about the mean is

$$\text{M.D. } (\bar{x}) = \frac{\sum_{i=1}^{8} |x_1 - \bar{x}|}{8}$$

$$= \frac{6+3+2+1+0+2+3+7}{8} = \frac{24}{8} = 3$$

Mean Deviation About Mean of Discrete Distribution

Consider a discrete frequency distribution consisting of m distinct values $x_1, x_2, \dots, x_m$ with frequencies $f_1, f_2, \dots, f_m$ respectively. In order to find the mean deviation about mean, we have the following working rule:

Step I: Calculate mean of the given data using the formula

$$\text{Mean, } \bar{x} = \frac{\sum f_i x_i}{\sum fi}$$

Step II: Compute the deviations d_i of values x_i from mean, ignoring the ± signs.

i.e., $|d_i| = |x_i - \bar{x}|$. Then

$$\text{Mean deviation about mean} = \frac{\sum f_i |d_i|}{\sum f_i}$$

Mean Deviation About Mean of Continuous Distribution

Consider a continuous frequency distribution consisting of class intervals with their corresponding frequencies as f_r. In order to find the mean deviation about mean, we have the following working rule.

Step I: Find the mid-points of the class intervals and denote the by x_r

Step II: Calculate mean of the given data using the formula

$$\text{Mean, } \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Step III: Compute the deviations d_i of values x_i from mean, ignoring the ± signs.

i.e., $|d_i| = |x_i - \bar{x}|$.

Then, mean deviation about mean

$$= \frac{\sum f_i |d_i|}{\sum f_i}.$$

Example 1.5: **Find the mean deviation about the mean of the data.**

Height in cms	95-105	105-115	115-125	125-135	135-145	145-155
Number of boys	9	13	26	30	12	10

[NCERT]

Ans.

Height (in cms)	Number of boys f_i	Mid-point x_i	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
95 – 105	9	100	900	$\lvert 100 - 125.3 \rvert = 25.3$	227.7
105 – 115	13	110	1430	$\lvert 110 - 125.3 \rvert = 15.3$	198.9
115 – 125	26	120	3120	$\lvert 120 - 125.3 \rvert = 5.3$	137.8
125 – 135	30	130	3900	$\lvert 130 - 125.3 \rvert = 4.7$	141
135 – 145	12	140	1680	$\lvert 140 - 125.3 \rvert = 14.7$	176.4
145 – 155	10	150	1500	$\lvert 150 - 125.3 \rvert = 24.7$	247
	$\Sigma f_i = 100$		12530		$\sum f_i \lvert x_i - \bar{x} \rvert = 1128.8$

$\Sigma f_i = 100$ and $\Sigma f_i x_i = 12530$

$$\therefore \quad \text{Mean } \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

$$= \frac{12530}{100} = 125.3$$

Now,

$$\Sigma f_i = 100$$

$$\Sigma f_i |x_i - \bar{x}| = 1128.8$$

$$\text{Mean deviation about mean} = \frac{\sum f_i |x_i - \bar{x}|}{f_i}$$

$$= \frac{1128.8}{100}$$

$$= 11.288$$

Mean Deviation about Median of Discrete Distribution

Consider a discrete frequency distribution consisting of m distinct values $x_1, x_2, \dots, x_m$ with frequencies, $f_1, f_2, \dots, f_m$ respectively. In order to find the mean deviation about median. We have the following working rule:

Step I: Arrange the given data (i.e., values of x_i) in ascending order.

Step II: Find the cumulative frequency, when cumulative frequency represents the sum of all previous frequencies up to the current value.

Step III: Calculate median of the given data using the formula given below:

(1) $\text{Median} = \left(\frac{n+1}{2}\right)^{th}$ term, if n is odd.

(2) $\text{Median} = \dfrac{\left(\frac{n}{2}\right)^{th} \text{term} + \left(\frac{n}{2}+1\right)^{th} \text{term}}{2}$ if, n is even.

Step IV: Compute the deviations d_i of values x_i from median (M). Ignoring the ± signs.

i.e., $|d_i| = |x_i - M|$. Then

Mean deviation about median = $\dfrac{\sum f_i |d_i|}{\sum f_i}$

Example 1.6: Find the mean deviation about the median for the data.
13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17 [NCERT]

Ans. The given data is

13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17.

Here, the number of observations are 12, which is even.

Arranging the data in ascending order, we obtain

10, 11, 11, 12, 13, 13, 14, 16, 16, 17, 17, 18

$$\text{Median} = \frac{6^{th}\text{ observation} + 7^{th}\text{ observation}}{2}$$

$$= \frac{13+14}{2} = \frac{27}{2} = 13.5$$

The deviations of the respective observation from the median, i.e., $x_i - M$, i.e. $x_i - 13.5$ are −3.5, −2.5, −2.5, −1.5, −0.5, −0.5, 0.5, 2.5, 2.5, 3.5, 3.5, 4.5.

The absolute values of the deviations, $|x_i - M|$, are 3.5, 2.5, 2.5, 1.5, 0.5, 0.5, 0.5, 2.5, 2.5, 3.5, 3.5, 4.5.

The required mean deviation about the median is

$$M.D.(M) = \frac{\sum_{i=1}^{12} |x_i - M|}{12}$$

$$= \frac{3.5+2.5+2.5+1.5+0.5+0.5+0.5+2.5+2.5+3.5+3.5+4.5}{12}$$

$$\frac{12530}{100} = 125.3$$

$$= \frac{28}{12} = 2.33$$

Mean Deviation About Median of Continuous Distribution

Consider a continuous frequency distribution consisting of class interval with their corresponding frequencies, as f_i. To find the mean deviation about median. We have the following working rule.

Step I: Find the mid-points of the class intervals and denote them by x_r

Step II: Find the cumulative frequency, where cumulative frequency represents the sum of all previous frequencies up to the current value.

Step III: Denote $n = \Sigma f_r$

Step IV: Identify median class, i.e., the class containing the cumulative frequency $\dfrac{n}{2}$.

Step V: Calculate median of the given data using the formula given below:

$$\text{Median} = l + \frac{\frac{n}{2} - C}{f} \times h$$

Where, l = lower limit of median class

H = width of median class

f = frequency of median class

C = Cumulative frequency just preceding median class

Step VI: Compute the deviations d_i of values x_i from median (M), ignoring the ± signs,

i.e., $|d_i| = |x_i - M|$. Then, mean deviation about median

$$\frac{\sum f_i |d_i|}{\sum f_i}$$

Example 1.7: Calculate the mean deviation about median age for the age distribution of 100 persons given below:

Age (in years)	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55
Number	5	6	12	14	26	12	16	9

Converting the given data in continuous frequency by subtracting 0.5 in upper limit. [NCERT]

Ans.

Age	Number	Cumulative Frequency (c.f.)	Mid Point (x_i)	$\|x - M\|$	$f_i \|x_i - M\|$
15.5 − 20.5	5	5	18	$\|18-38\| = 20$	5 × 20 = 100
20.5 − 25.5	6	5 + 6 = 11	23	$\|23-38\| = 20$	6 × 15 = 90
25.5 − 30.5	12	11 + 12 = 23	28	$\|28-38\| = 20$	12 × 10 = 120
30.5 − 35.5	14	23 + 14 = 37	33	$\|33-38\| = 20$	14 × 5 = 70
35.5 − 40.5	26	37 + 26 = 63	38	$\|38-38\| = 20$	26 × 0 = 0
40.5 − 45.5	12	63 + 12 = 75	43	$\|43-38\| = 20$	12 × 5 = 60
45.5 − 50.5	16	75 + 16 = 91	48	$\|48-38\| = 20$	16 × 10 = 160
50.5 − 55.5	9	91 + 9 = 100	53	$\|53-38\| = 20$	9 × 15 = 135
Σf_i = 100					$\Sigma f_i \|x_i - M\|$ = 735

$N = \Sigma f_i = 100$

$$\text{Median class} = \left(\frac{N}{2}\right)^{th} \text{term}$$

$$= \left(\frac{100}{2}\right)^{th} \text{term}$$

$$= 50^{th} \text{term}$$

In the above data, cumulative frequency of class 35.5 – 40.5 is 63 which is greater than 50.

∴ Median class = 35.5 – 40.5

$$\text{Median} = l + \frac{\frac{N}{2} - C}{f} \times h$$

Where,

l = lower limits of median class

N = sum of frequencies

f = frequency of median class

C = cumulative frequency of class before median class.

Here, $l = 35.5$, $N = 100$, $h = 5$, $f = 26$

$$\text{Median} = 35.5 + \frac{\frac{100}{2} - 37}{26} \times 5$$

$$= 35.5 + \frac{50 - 37}{26} \times 5$$

$$= 35.5 + \frac{13}{26} \times 5$$

$$= 35.5 + 2.5 = 38$$

Now,

$$\Sigma f_i = 100$$

$$\Sigma f_i |x_i - M| = 735$$

$$\therefore \text{Mean deviation (M)} = \frac{\sum f_i |x_i - M|}{f_i}$$

$$= \frac{735}{100}$$

$$= 7.35$$

Example 1.8: **Find the mean deviation about the median for the data.**

x_i	15	21	27	30	35
f_i	3	5	6	7	8

[NCERT]

Ans. The given observation is already in ascending order. Adding a column corresponding to cumulative frequencies of the given data, we obtain the following table.

x_i	f_i	c.f.
15	3	3
21	5	8
27	6	14
30	7	21
35	8	29

$N = 29$ (odd)

cf = cumulative frequency

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{term}$$

$$= \left(\frac{29+1}{2}\right)^{th} \text{term}$$

$$= \left(\frac{30}{2}\right)^{th} \text{term}$$

$$= 15^{th} \text{term}$$

As, 15^{th} term lies in cf of 21.

x_i	f_i	$\lvert x_i - M \rvert$	$f_i \lvert x_i - M \rvert$
15	3	$\lvert 15 - 30 \rvert = 15$	$3 \times 15 = 45$
21	5	$\lvert 21 - 30 \rvert = 9$	$5 \times 9 = 45$
27	6	$\lvert 27 - 30 \rvert = 3$	$6 \times 3 = 18$
30	7	$\lvert 30 - 30 \rvert = 0$	$7 \times 0 = 0$
35	8	$\lvert 35 - 30 \rvert = 5$	$8 \times 5 = 40$
$\Sigma f_i = 29$			$\Sigma f_i \lvert x_i - M \rvert = 148$

$$\Sigma f_i = 29$$

$$\Sigma f_i |x_i - M| = 148$$

$$\therefore \text{Mean Deviation} = \frac{\Sigma f_i |x_i - M|}{f_i}$$

$$= \frac{148}{29}$$

$$= 5.1$$

Example 1.9: **Find the mean deviation about median for the following data:**

Class	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60
Frequency	6	8	14	16	4	2

[NCERT]

Ans.

Class	Frequency	Cumulative Frequency	Mid point (x_i)
0 – 10	6	6	5
10 – 20	8	6 + 8 = 14	15
20 – 30	14	14 + 14 = 28	25
30 – 40	16	28 + 16 = 44	35
40 – 50	4	44 + 4 = 48	45
50 – 60	2	48 + 2 = 50	55
	$\Sigma f_i = 50$		

$$N = \Sigma f_i = 50$$

$$\text{Median class} = \left(\frac{N}{2}\right)^{th} \text{ term}$$

$$= \left(\frac{50}{2}\right)^{th} \text{ term}$$

$$= 25^{\text{th}} \text{ term}$$

In above data, cumulative frequency of class 20 – 30 is 28 which is slightly greater than 25.

$\therefore$ Median class = 20 – 30

$$\text{Median} = l + \frac{\frac{N}{2} - c}{f} \times h$$

Where,

l = lower limits of median class

N = sum of frequencies

f = frequency of median class

C = Cumulative frequency of class before median class

Here, $l = 20$, $N = 50$, $C = 14$, $h = 10$, $f = 14$

$$\text{Median} = l + \frac{\frac{N}{2} - c}{f} \times h$$

$$= 20 + \frac{\frac{50}{2} - 14}{14} \times 20 = 20 + \frac{25 - 14}{14} \times 10$$

$$= 20 + \frac{11}{14} \times 10$$

$$= 20 + 7.8 = 27.8$$

Finding mean deviation about median

$$= \frac{\sum f_i |x_i - M|}{\sum f_i}$$

Class	Frequency	Cumulative frequency	Mid point (x_i)	$\lvert x_i - M \rvert$	$f_i \lvert x_i - M \rvert$
0 – 10	6	6	5	$\lvert 5 - 27.8 \rvert = 22.8$	$6 \times 22.8 = 136.8$
10 – 20	8	6 + 8 = 14	15	$\lvert 15 - 27.8 \rvert = 12.8$	$8 \times 12.8 = 102.4$
20 – 30	14	14 + 14 = 28	25	$\lvert 25 - 27.8 \rvert = 2.8$	$14 \times 2.8 = 39.2$
30 – 40	16	28 + 16 = 44	35	$\lvert 35 - 27.8 \rvert = 7.2$	$16 \times 7.2 = 115.2$
40 – 50	4	44 + 4 = 48	45	$\lvert 45 - 27.8 \rvert = 17.2$	$4 \times 17.2 = 68.8$
50 – 60	2	48 + 2 = 50	55	$\lvert 55 - 27.8 \rvert = 27.2$	$2 \times 27.2 = 54.5$
	$\Sigma f_i = 50$				$\Sigma f_i \lvert x_i - M \rvert = 516.8$

$$\Sigma f_i = 50$$

$$\Sigma f_i |x_i - \bar{x}| = 516.8$$

$$\therefore \text{Mean deviation } (M) = \frac{\sum f_i |x_i - M|}{f_i}$$

$$= \frac{516.8}{50}$$

$$= 10.336$$

$$= 10.34$$

Example 1.10: **Case Based:**

Consider the data

Class	*Frequency*
0–10	*6*
10–20	*7*
20–30	*15*
30–40	*16*
40–50	*4*
50–60	*2*

Based on above information, answer the following questions.

(A) Assertion (A): Median is calculated by using the formula, $M = l - \frac{\frac{N}{2} - cf}{f} \times h$

Reason (R): The median is the value in the middle of a data set.

(a) Both (A) and (R) are true and (R) is the correct explanation of (A).
(b) Both (A) and (R) are true but (R) is not the correct explanation of (A).
(c) (A) is true but (R) is false.
(d) (A) is false but (R) is true.

(B) Mean deviation about median is calculated by using the formula:

(a) $MD = \frac{\Sigma f_i |x_i + M|}{N}$ (b) $MD = \frac{\Sigma f_i |x_i - M|}{N}$

(c) $MD = \frac{\Sigma |x_i - M|}{N}$ (d) none of these

(C) Total frequency of the given data is:
(a) 10 (b) 20
(c) 50 (d) 60

(D) Find the median of the given data.
(E) Find the deviation about median.

Ans. Let us make the following table from the given data.

Class	f_i	cf	Mid point (x_i)	$\|x_i - M\|$ $M = 28$	$f_i\|x_i - M\|$
0 – 10	6	6	5	23	138
10 – 20	7	13	15	13	91
20 – 30	15	28	25	3	45
30 – 40	16	44	35	7	112
40 – 50	4	48	45	17	68
50 – 60	2	50	55	27	54
Total	50				508

Here, $N = 50$

$\Rightarrow \frac{N}{2} = \frac{50}{2} = 25$

which item lies in the cumulative frequency 28. Therefore, 20-30 is the median class.

So, we have, $l = 20$, $cf = 13$, $f = 15$, $h = 10$ and $N = 50$

Now, Median, $M = l + \frac{\frac{N}{2} - cf}{f} \times h$

$= 20 + \frac{25 - 13}{15} \times 10$

$= 20 + 8 = 28$

(A) *(d) (A) is false but (R) is true.*

Explanation: We know that, the median is the middle valueof a data set.

Median, $M = l + \frac{\frac{N}{2} - cf}{f} \times h$

(B) *(b)* $MD = \frac{\Sigma f_i |x_i - M|}{N}$

Explanation: Mean deviation.

$MD = \frac{\Sigma f_i |x_i - M|}{N}$

(C) *(c) 50*

Explanation: $N = \Sigma f_i = 50$

(D) Median = 28

(E) The mean deviation about median is given by

$MD(M) = \frac{1}{N}\sum_{i=1}^{6} f_i |x_i - M|$

$= \frac{1}{50} \times 508 = 10.16$

Hence, the mean deviation about median is 10.16.

OBJECTIVE Type Questions

[1 mark]

Multiple Choice Questions

1. The mean of the first three prime number is:
(a) 3 (b) 3.3
(c) 4 (d) none

Ans. *(b) 3.3*

Explanation: Here first 3 number which are prime also 2, 3, 5.

So, mean $\bar{x} = \frac{2+3+5}{3} = \frac{10}{3} = 3.3$

2. The mean deviation from the mean of the set observation –1, 0 and 4 is:
(a) 3 (b) 1
(c) –2 (d) 2

Ans. *(d) 2*

Explanation: $\bar{x} = -\frac{-1+0+4}{3} = 1$

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

$= \frac{|-1-1| + |0-1| + |4-1|}{3}$

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

3. **Calculate mean deviation about the median from the following data.**
21, 30, 24, 32, 31

(a) 5.28 (b) 5.18
(c) 3.38 (d) 3.6 [Diksha]

Ans. (d) 3.6

Explanation: x_i = 21, 24, 30, 31, 32

Number of observation = n = 5 (odd)

Since x is an odd

$$\text{Median} = \left(\frac{5+1}{2}\right)^{th} \text{term}$$
$$= 3^{rd} \text{ term}$$
$$M = 30$$
$$\text{M.D} = \frac{9+6+0+1+2}{5}$$
$$= \frac{18}{5}$$
$$= 3.6$$

4. **Mean deviation about median for 3, 4, 9, 5, 3, 12, 10, 18, 7, 19, 21.**

(a) 4.27 (b) 5.24
(c) 5.27 (d) 4.24

[Delhi Gov. Term-1 SQP 2021]

Ans. (c) 5.27

Explanation: The given data is

3, 3, 4, 5, 7, 9, 10, 12, 18, 19, 21

$$\text{Median} = \left(\frac{11+1}{2}\right)^{th} \text{term} = (6^{th}) \text{ term} = 9$$

Now, $|r_n - M|$ are

6, 6, 5, 4, 2, 0, 1, 3, 9, 10, 12

$$\sum|r_n - M| = 6 + 6 + 5 + 4 + 2 + 0 + 1 + 3 + 9 + 10 + 12$$
$$= 58$$

$$\text{Therefore, } MD(M) = \frac{1}{11}\sum|r_n - M|$$
$$= \frac{1}{11} \times 58$$
$$= 5.27$$

5. **Consider the following data, The mean deviation about median for the data is:**

r_n	15	21	27	30	35
f_i	3	5	6	7	8

(a) 5 (b) 5.3
(c) 5.1 (d) 5.2

Ans. (c) 5.1

Explanation:

N = 29 (odd)

$c.f$ = cumulative frequency

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{term}$$
$$= \left(\frac{29+1}{2}\right)^{th} \text{term}$$
$$= \left(\frac{30}{2}\right)^{th} \text{term}$$
$$= 15^{th} \text{ term}$$

As, 15^{th} term lies in c.f of 21, so median = 30.

x_i	f_i	c.f	$\lvert x_i - M \rvert$	$f_i\lvert x_i - M \rvert$
15	3	3	$\lvert 15 - 30 \rvert = 15$	45
21	5	8	$\lvert 21 - 30 \rvert = 9$	45
27	6	14	$\lvert 27 - 30 \rvert = 3$	18
30	7	21	$\lvert 30 - 30 \rvert = 0$	0
35	8	29	$\lvert 35 - 30 \rvert = 5$	40
	29			148

$$\text{M.D } (M) = \frac{\sum f_i(x_i - M)}{\sum f_i}$$
$$= \frac{148}{29} = 5.1$$

6. **Find the mean deviation about the mean of the distribution is:**

Size	20	21	22	23	24
Frequency	6	4	5	1	4

(a) 1.25 (b) 1
(c) 1.50 (d) 2

Ans. (a) 1.25

Explanation: Given data distribution

Now we have to find the mean deviation about the mean of the distribution construct a table of the given data

Size (x_i)	Frequency (f_i)	f_ix_i
20	6	20 × 6 = 120
21	4	21 × 4 = 84
22	5	22 × 5 = 110
23	1	23 × 1 = 23
24	4	24 × 4 = 96
Total	20	433

We know that mean, $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{433}{20} = 21.65$

To find mean deviation, we have to construct another table

Size (x_i)	Frequency (f_i)	f_ix_i	$d_i = \vert x_i - \text{mean}\vert$	f_id_i
20	6	120	1.65	9.90
21	4	84	0.65	2.60
22	5	110	0.35	1.75
23	1	23	1.35	1.35
24	4	96	2.35	9.40
Total	20	433		25.00

Hence, Mean Deviation becomes,

$$M.D = \frac{\Sigma f_i d_i}{\Sigma f_i} = \frac{25}{20} = 1.25$$

Therefore, the mean deviation about the mean of the distribution is 1.25.

7. The mean deviation from the median of the first three natural number is:

(a) 0.667 (b) 0.253
(c) 0.456 (d) none of these

Ans. *(a) 0.667*

Explanation: First 3 natural numbers are 1, 2, 3.

Then median = $\left(\frac{n+1}{2}\right)^{th}$ term

$= \left(\frac{3+1}{2}\right)^{th}$

$= 2^{nd}$

Hence, median = 2

So $\Sigma|x_i - Med| = 1 + 0 + 1 = 2$

$$M.D. = \frac{2}{3} = 0.667$$

8. Median is calculated by using the formula:

(a) $M = l + \frac{\frac{N}{2} - cf}{f} \times h$

(b) $M = l + \frac{\frac{N}{2} + cf}{f} \times h$

(c) $M = l - \frac{\frac{N}{2} - cf}{f} \times h$

(d) None of these

Ans. *(a)* $M = l + \frac{\frac{N}{2} - cf}{f} \times h$

Explanation: Correct formula for finding median for a given data is $l + \frac{\frac{N}{2} - cf}{f} \times h$.

9. Mean deviation about median is calculated by using the formula:

(a) $M.D = \frac{\Sigma f_i |x_i + M|}{N}$

(b) $M.D = \frac{\sum f_i |x_i - M|}{N}$

(c) $M.D = \frac{\Sigma |x_i - M|}{N}$

(d) None of these

Ans. *(b)* $M.D = \frac{\sum f_i |x_i - M|}{N}$

Explanation: The correct formula to finding mean deviation about median is $\frac{\Sigma f_i |x_i - M|}{N}$.

Assertion Reason Questions

Direction: In the following questions, a statement of Assertion (A) is followed by a statement of Reason (R).

Choose the correct answer out of the following choices.

(a) Both (A) and (R) are true and (R) is the correct explanation of (A).

(b) Both (A) and (R) are true but (R) is not the correct explanation of (A).

(c) (A) is true but (R) is false.

(d) (A) is false but (R) is true.

10. Assertion (A): The mean deviation about the mean to find a measure of dispersion has certain limitations.

Reason (R): The sum of deviations from the mean is more than the sum of deviations from the median. Therefore, the mean deviation about the mean is not very scientific, where degree of variability is very high.

Ans. *(a) Both (A) and (R) are true and (R) is the correct explanation of (A).*

Explanation: The sum of the deviations from the mean (minus signs ignored) is more than the sum of the deviations from median.

Therefore, the mean deviation about the mean is not very scientific. Thus, in many cases, mean deviation may give unsatisfactory results. Also mean deviation is calculated on the basis of absolute values of the deviations and therefore, cannot be subjected to further algebraic treatment. This implied that we must have some other measure of dispersion. Standard deviation is such a measure of dispersion.

11. Assertion (A): The average marks of boys in a class is 52 and that of girls is 42. The average marks of boy and girls combined is 50. The percentage of boys in the class is 80%.

Reason (R): Mean marks scored by the students of a class is 53 and the mean marks of the boys is 50 and the mean marks of the girls is 55. The percentage of girls in the class is 64%.

Ans. *(c) A is true but R is false.*

Explanation: Let the number of boys and girls be x and y.

$\therefore \quad 52x + 42y = 50(x + y)$

$\Rightarrow \quad 2x = 8y$

$\Rightarrow \quad x = 4y$

$\therefore$ Total number of students in the class

$= x + y = 5y$

$\therefore$ Required percentage of boys

$$= \frac{4y}{5y} \times 100\%$$

$= 80\%$

Let the number of boys be x and number of girls be y.

$\therefore \quad 53(x + y) = 55y + 50x$

$\Rightarrow \quad 3x = 2y$

$\Rightarrow \quad x = \frac{2y}{3}$

$\therefore$ Total number of students

$$= x + y = \frac{2y}{3} + y = \frac{5}{3}y$$

Hence, required percentage

$$= \frac{y}{\frac{5y}{3}} \times 100\%$$

$$= \frac{3}{5} \times 100\% = 60\%$$

12. Assertion (A): The weights (in kg) of 5 students are as follows

31, 35, 27, 29, 32, 43, 37, 41, 34, 28, 36, 44, 45, 42, 30.

If the weight 44 kg is replaced by 46 kg and 27 kg is by 25 kg, then the new median is 35.

Reason (R): The mean deviation from the median of the weights (in kg) 54, 50, 40, 42, 51, 45, 455, 57 is 4.78.

Ans. *(b) Both A and R are true but R is not the correct explanation of A.*

Explanation: Since, 44 kg is replaced by 46 kg and 27 kg is replaced by 25 kg, then the given series becomes 31, 35, 25, 29, 32, 43, 37, 41, 34, 28, 36, 46, 45, 42, 30.

On arranging this series in ascending order, we get 25, 28, 29, 30, 31, 32, 34, 35, 36, 37, 41, 42, 43, 45, 46.

Total number of students are 15, therefore the middle term is 8^{th} term whose corresponding value is 35.

On arranging the terms in increasing order of magnitude

40, 42, 45, 47, 50, 51, 54, 55, 57

Number of terms, $N = 9$

$$\therefore \text{Median} = \left(\frac{9+1}{2}\right)^{th} \text{term} = 5^{th} \text{terms} = 50 \text{ kg}$$

Weight (in kg)	Deviation from Median (d)	| d |
40	–10	10
42	–8	8
45	–5	5
47	–3	3
50	0	0
51	1	1
54	4	4
55	5	5
57	7	7
		| d | = 43

$$\text{MD from median} = \frac{43}{9} = 4.78\,\text{kg}$$

13. Assertion (A): The proper measure of dispersion about the mean of a set of observations, i.e. standard deviation, is expressed as the positive square root of the variance.

Reason (R): The units of individual observations x_i and the unit of their mean are different,then that of variance. Since, variance involves sum of square of $(x - \bar{x})$.

Ans. *(a) Both A and R are true and R is the correct explanation of A.*

Explanation: In the calculations of variance, we find that the units of individual observations x_i and the unit of their mean $\bar{x}$ are different from that of variance, since variance involves the sum of squares of $(x_i - \bar{x})$.

For this reason, the proper measure of dispersion about the mean of a set of observations is expressed as the positive square root of the variance and is called standard deviation.

CASE BASED Questions (CBQs)

[4 & 5 marks]

Read the following passages and answer the questions that follow:

14. *A stopwatch was used to find the time that it took a group of students to run 100 m.*

Time (in Seconds)	0–20	20–40	40–60	60–80	80–100
No. of students	8	10	13	6	3

(A) The mean time taken by a student to finish the race is:

(a) 54 (b) 63
(c) 43 (d) 50

(B) The construction of cumulative frequency table is useful in determining:

(a) mean (b) median
(c) mode (d) all of the above

(C) The lower limit of the median class is:

(a) 20 (b) 40
(c) 60 (d) 80

(D) The algebraic sum of absolute values of deviations from the mean is:

(a) 0 (b) 788
(c) 1720 (d) none of these

(E) The value of mean deviation from mean is:

(a) 19.7 (b) 20.7
(c) 18.7 (d) 17.7

Ans. Calculation of Mean, Mean deviation and cumulative

C.I.	Mid-Valcue x_i	Fre-quency fi	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$	$\Sigma f_i \lvert x_i - \bar{x} \rvert$
0–20	10	8	80	33	264
20–40	30	10	300	13	130
40–60	50	13	650	7	91
60–80	70	6	420	27	162
80–100	90	3	270	47	141
		N = 40	$\Sigma f_i x_i$ = 1720		$\Sigma f_i \lvert x_i - \bar{x} \rvert$ = 788

$\therefore$ Mean is given by $\bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{1720}{40} = 43$

And Mean deviation from mean

$$= \frac{\Sigma f_i |x_i - \bar{x}|}{N} = \frac{788}{40} = 19.7$$

And $\frac{N}{2} = \frac{40}{2} = 20$, which lies in the class 40-60

$\therefore$ Median class is 40 – 60.

(A) *(c) 43*

Explanation: mean time = 43

(B) *(b) Median*

Explanation: The construction of the cumulative frequency table is useful in determining the median.

(C) *(b) 40*

Explanation: The median class is 40–60. So, the lower limit = 40.

(D) *(b) 788*

Explanation: The absolute value of algebraic sum of deviation from the mean is 788.

(E) *(a) 19.7*

Explanation: From the calculation above, mean deviation from mean is 19.7.

15. *A seminar on prime numbers was organised in a college to motivate students.*

(A) Find the mean deviation of the first three primes from their median.

(B) Find the median of first three primes and find the algebraic sum of deviations of first three prime numbers from the mean.

(C) Find the mean deviation from the mean for the following data

6.5, 5, 5.25, 5.5, 4.75, 4.5, 6.25, 7.75, 8.5

Ans. (A) The first three prime numbers in ascending order are 2, 3 and 5.

Median of these three prime numbers = Value of 2nd item = 3.

The sum of absolute values of deviations from the median

$= |2-3| + |3-3| + |5-3| = 1 + 0 + 2 = 3$

Mean deviations of first three prime numbers from median = $\frac{3}{3} = 1$

(B) Median is 3.

He algebraic sum of deviations from the mean is always zero.

(C) Given observations are

6.5, 5, 5.25, 5.5, 4.75, 4.5, 6.25, 7.75, 8.5

Here number of observation $n = 9$

Let $\bar{x}$ be the mean of given data.

Then,

$$\bar{x} = \frac{6.5+5+5.25+5.5+4.75+4.5+6.25+7.75+8.5}{9}$$

$$= \frac{54}{9} = 6$$

Let us make the table for deviation and absolute deviation.

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
6.5	0.5	0.50
5.0	–1	1.00
5.25	–0.75	0.75
5.5	–0.5	0.50
4.75	–1.25	1.25
4.5	–1.50	1.50
6.25	0.25	0.25
7.75	1.75	1.75
8.5	2.5	2.50
Total		$\sum_{i=1}^{9} \lvert x_i - \bar{x} \rvert = 10.00$

∴ Mean deviation about mean,

$$MD(\bar{x}) = \frac{\sum_{i=1}^{6} \lvert x_i - \bar{x} \rvert}{9} = \frac{10}{9} = 1.1$$

Hence, the mean deviation about mean is 1.1.

VERY SHORT ANSWER Type Questions (VSA)

[1 mark]

16. Find the range of the following data

x	10–20	20–30	30–40	40–50
f	4	5	6	5

[Diksha]

Ans.

x	10–20	20–30	30–40	40–50
f	4	5	6	5

The highest class is 40–50

And the lowest class is 10–20

Hence, range = upper limit of highest class – lower limit of lowest class

= 50 – 10

= 40

17. Find the mean deviation from the mean for the data: 6, 7, 10, 12, 13, 4, 8, 20.

Ans. Let $\bar{x}$ be the mean of the given data. Then,

$$\bar{x} = \frac{6+7+10+12+13+4+8+20}{8} = 10$$

x_i	$\lvert d_i \rvert = \lvert x_i - \bar{x} \rvert = \lvert x_i - 10 \rvert$
6	4
7	3
10	0
12	2
13	3
4	6
8	2
20	10
Total	$\Sigma d_i = 30$

Now, $\Sigma \lvert d_i \rvert = 30$ and $n = 8$

$$\text{M.D.} = \frac{1}{n}\sum \lvert d_i \rvert = \frac{30}{8} = 3.75$$

Thus, the mean deviation from the mean for the given data is 3.75.

18. Find the mean deviation about median for the following information. [Diksha]

x	15	21	27	30
f	3	5	6	7

Ans.

x_i	f_i	c.f	$d_i = \lvert x_i - 27 \rvert$	$f_i d_i$
15	3	3	12	36
21	5	8	6	30
27	6	14	0	0
30	7	21	3	21
	21		21	87

$$\text{Median} = \left(\frac{21+1}{2}\right)^{th}$$

$= 11^{th}$ term i.e. 27

$$\text{M.D} = \frac{\Sigma f_i d_i}{\Sigma f_i}$$

$$= \frac{87}{21} = 4.14$$

19. Find the mean of the following data:

x_i	2	5	6	8	10	12
f_i	2	8	10	7	8	5

Ans. We have the following table:

x_i	f_i	$f_i x_i$
2	2	4
5	8	40
6	10	60
8	7	56
10	8	80
12	5	60
	$\Sigma f_i = 40$	$\Sigma f_i x_i = 300$

Hence, mean $\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{300}{40} = 7.5$

SHORT ANSWER Type-I Questions (SA-I)

[**2** marks]

20. Find the mean deviation about the median of the data: 36, 46, 42, 45, 55, 46, 51, 49.

Ans. Arranging the given data in ascending order : 36, 42, 45, 46, 46, 49, 51, 55.

Here, number of terms, $n = 8$, which is even.

So, median = $\frac{\left(\frac{n}{2}\right)^{th} term + \left(\frac{n}{2}+1\right)^{th} term}{2}$

$= \frac{4^{th} term + 5^{th} term}{2} = \frac{46+46}{2} = 46$

Calculation of Mean Deviation About Median

x_i	$\lvert d_i \rvert = \lvert x_i - 46 \rvert$
36	10
42	4
45	1
46	0
46	0
49	3
51	5
55	9
	$\Sigma \lvert d_i \rvert = 32$

Hence, mean deviation about median

$= \frac{\sum \lvert d_i \rvert}{n} = \frac{32}{8} = 4.$

21. The following data given the monthly income (in Rs) of 10 families in a city:

4600, 5560, 6440, 4530, 7670, 6850, 6750, 7910, 5490, 6800.

Find the mean using direct method.

Ans. Given data is 4600 5560, 6440, 4530, 7670, 6850, 6750, 7910, 5490, 6800.

Here, number of terms $n = 10$.

Hence, mean

$$\bar{x} = \frac{4600+5560+6440+4530+7670+6850+6750+7910+5490+6800}{10}$$

$= \frac{62600}{10} = 6260.$

22. Calculate mean deviation about mean from the following data:

x_i	3	9	17	23	27
f_i	8	10	12	9	5

[Diksha]

Ans.

x_i	f_i	$f_i x_i$	$\lvert x_i - 15 \rvert$	$f_i \lvert x_i - 15 \rvert$
3	8	24	12	96
9	10	90	6	60
17	12	204	2	24
23	9	207	8	72
27	5	135	12	60
	$N = \Sigma f_i = 44$	$\Sigma f_i x_i = 660$		$\Sigma f_i \lvert x_i - 15 \rvert = 312$

Mean $= X = \frac{1}{N}\left(\sum f_i x_i\right) = \frac{660}{44} = 15$

Mean deviation = M.D.

$= \frac{1}{N}\sum f_i \lvert x_i - 15 \rvert = \frac{312}{44} = 7.09$

23. Calculate the mean deviation about the mean for the following frequency distribution:

Class Interval	0–4	4–8	8–12	12–16	16–20
Frequency	4	6	8	5	2

[NCERT Exemplar]

Ans. Given the frequency distribution

Now we have to find the mean deviation about the mean.

Let us make a table of the given data and append other columns after calculations

Class Interval	Mid-Value (x_i)	Frequency (f_i)	$f_i x_i$
0–4	2	4	8
4–8	6	6	36
8–12	10	8	80
12–16	14	5	70
16–20	18	2	36
		Total = 25	230

Here mean, $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{230}{25} = 9.2$

Now we have to find mean deviation

Class Interval	Mid-Value (x_i)	Frequency (f_i)	$f_i x_i$	$d_i = \lvert x_i - \text{mean}\rvert$	$f_i d_i$
0–4	2	4	8	7.2	28.8
4–8	6	6	36	3.2	19.2
8–12	10	8	80	0.8	6.4
12–16	14	5	70	4.8	24
16–20	18	2	36	8.8	17.6
		Total = 25	230		96

Hence, mean deviation becomes,

$$M.D = \frac{\Sigma f_i d_i}{\Sigma f_i} = \frac{96}{25} = 3.84$$

Therefore, the mean deviation about the mean of the distribution is 3.84.

24. Find the median age of the distribution of 100 persons given below:

Age	16 – 20	21 – 25	26 – 30	31 – 35	36 – 40	41 – 45	46 – 50	51 – 55
Number	5	6	12	14	26	12	16	9

Ans. Given data in inclusive form is

Age	16–20	21–25	26–30	31–35	36–40	41–45	46–50	51–55
Number	5	6	12	14	26	12	16	9

which can be written in exclusive form as

Age	15.5–20.5	20.5–25.5	25.5–30.5	30.5–35.5	35.5–40.5	40.5–45.5	45.5–50.5	50.5–55.5
Number	5	6	12	14	26	12	16	9

We have the following table:

Class	Frequency (f_i)	Cumulative Frequency
15.5-20.5	5	5
20.5-25.5	6	11
25.5-30.5	12	23
30.5-35.5	14	37
35.5-40.5	26	63
40.5-45.5	12	75
45.5-50.5	16	91
50.5-55.5	9	100
	$n = \Sigma f_i = 100$	

Here, $n = 100$. So, $\frac{n}{2} = 50$.

The cumulative frequency just greater than 50 is 63, which belongs to the class 35.5 - 40.5.

So, 35.5-40.5 is the median class.

Lower limit of median class, $l = 35.5$

Width of median class, $h = 5$

Frequency of median class, $f = 26$.

Cumulative frequency of class just preceding median class, $C = 37$.

Hence, median

$$= l + \frac{\frac{n}{2} - C}{f} \times h$$

$$= 35.5 + \frac{50 - 37}{26} \times 5 = 38$$

SHORT ANSWER Type-II Questions (SA-II)

[**3** marks]

25. Find the mean deviation about the median for the following data. **[Diksha]**

Class	0 – 6	6 – 12	12 – 18	18 – 24	24 – 30
Frequency	8	10	12	9	5

Ans. Calculation of mean deviation about the Median

Class	Mid–Values (x)	Fre-quency (f_i)	Cumula-tive Fre-quency (C. f)	$\lvert x_i - 14\rvert$	$f_i \lvert x_i - 14\rvert$
0–6	3	8	8	11	88
6–12	9	10	18	5	50
12–18	15	12	30	1	12
18–24	21	9	39	7	63
24–30	27	5	44	13	65
		$N = 44$			$\Sigma f_i \lvert x_i - 14\rvert = 278$

$N = 44$, so $\frac{N}{2} = 22$ and the cumulative frequency just greater than $\frac{N}{2}$ is 30. Thus, 12– 18 is the median class.

$\therefore$ Median $= l + \frac{N/2 - F}{f} \times h$,

where $l = 12$, $h = 6$, $f = 12$, $F = 18$.

$\Rightarrow$ Median $= 12 + \frac{22-18}{12} \times 6 = 12 + \frac{4 \times 6}{12} = 14$.

Clearly, $\sum f_i \lvert x_i - 14\rvert = 278$

$\therefore$ Mean deviation about median

$= \frac{1}{N}\sum f_i \lvert x_i - 14\rvert = \frac{278}{44} = 6.318$

26. Calculate the mean deviation from the median for the following distribution:

x_i	10	15	20	25	30	35	40	45
f_i	7	3	8	5	6	8	4	9

Ans. We have to calculate mean deviation about median. So, first, we calculate median.

x_i	f_i	Cumu-lative fre-quency	$\lvert d_i\rvert = \lvert x_i - 30\rvert$	$f_i \lvert d_i\rvert$
10	7	7	20	140
15	3	10	15	45
20	8	18	10	80
25	5	23	5	25
30	6	29	0	0
35	8	37	5	40
40	4	41	10	40
45	9	50	15	135
	$N = \Sigma f_i$			$\Sigma f_i \lvert d_i\rvert = 505$

Clearly, $N = 50$

$\Rightarrow \frac{N}{2} = 25$

The cumulative frequency just greater than 25 is 29 and the corresponding value of x is 30.

Therefore, median = 30

Clearly, $\Sigma f_i \lvert x_i - 30\rvert = \Sigma f_i d_i = 505$ and $N = 50$.

$\therefore$ Mean deviation $= \frac{1}{N}\Sigma f_i \lvert d_i\rvert = \frac{505}{50} = 10.1$

27. Calculate the mean deviation from the median of the distribution

Class Interval	0–6	6–12	12–18	18–24	24–30
Frequency	4	5	3	6	2

[NCERT Exemplar]

Ans. Given the frequency distribution

Now we have to find the mean deviation from the median

Let us make a table of the given data and append other columns after calculations.

Class Interval	Mid–Value (x_i)	Fre-quency (f_i)	c.f.	$d_i = \lvert x_i -$ Median $M\rvert$	$f_i d_i$
0–6	3	4	4	11	44
6–12	9	5	9	5	25
12–18	15	3	12	1	3
18–24	21	6	18	7	42
24–30	27	2	20	13	26
	Total	**20**			**140**

Now, here $N = 20$, which is even.

Here, median class $= \frac{N}{2} = 10^{th}$ term.

This observation lie in the class interval 12–18, so median can be written as,

$$M = l + \frac{\frac{N}{2} - cf}{f} \times h$$

Here $l = 12$, $cf = 9$, $f = 3$, $h = 6$ and $N = 20$, substituting these values, the above equation becomes,

$$M = 12 + \frac{\frac{20}{2} - 9}{3} \times 6$$

$$\Rightarrow M = 12 + \frac{10-9}{3} \times 6$$

$$\Rightarrow M = 12 + \frac{1 \times 6}{3}$$

$$\Rightarrow M = 12 + 2 = 14$$

Hence, mean deviation becomes,

$$M.D. = \frac{\sum f_i d_i}{\sum f_i} = \frac{140}{20} = 7$$

Therefore, the mean deviation about the median of the distribution is 7.

28. Find the mean deviation about the median of the following distribution:

Marks Obtained	10	11	12	14	15
No. of students	2	3	8	3	4

[NCERT Exemplar]

Ans. Given data distribution

Now we have to find the mean deviation about the median

Let us make a table of the given data and append other columns after calculations

Now, here $N = 20$, which is even.

Here median,

$$M = \frac{1}{2}\left[\left(\frac{N}{2}\right)^{th} \text{observation} + \left(\frac{N}{2}+1\right)^{th} \text{observation}\right]$$

$$M = \frac{1}{2}\left[\left(\frac{20}{2}\right)^{th} \text{observation} + \left(\frac{20}{2}+1\right)^{th} \text{observation}\right]$$

$$M = \frac{1}{2}[10^{th}\text{observation} + 11^{th}\text{observation}]$$

Both these observations lie in cumulative frequency 13, for which the corresponding observation is 12.

$$M = \frac{1}{2}[12+12] = 12$$

So, the above table with more columns is as shown below,

Marks Obtained x_i	Number of Students (f_i)	Cumulative frequency (cf)	$d_i = \lvert x_i - M \rvert$	$f_i d_i$
10	2	2	2	4
11	3	2 + 3 = 5	1	3
12	8	5 + 8 = 13	0	0
14	3	13 + 3 = 16	2	6
15	4	16 + 4 = 20	3	12
Total	20			25

Hence mean deviation becomes,

$$M.D = \frac{\sum f_i d_i}{\sum f_i} = \frac{25}{20} = 1.25$$

Therefore, the mean deviation about the median of the distribution is 1.25.

LONG ANSWER Type Questions (LA)

[4 & 5 marks]

29. Calculate the mean deviation about the mean of the set of first n natural numbers when n is an even number. [NCERT Exemplar]

Ans. Given set of first n natural numbers when n is an even number.

Now we have to find the mean deviation about the mean.

We know first n natural numbers are 1, 2, 3,...,n and given n is an even number.

So mean is,

$$\bar{x} = \frac{1+2+3+...+n}{n} = \frac{\frac{n(n+1)}{2}}{n} = \frac{(n+1)}{2}$$

The deviation of numbers from the mean are as shown below,

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

Or,

$$\frac{2-(n+1)}{2}, \frac{4-(n+1)}{2}, \frac{6-(n+1)}{2}, ..., \frac{n-2-(n+1)}{2}, \frac{n-(n+1)}{2}, \frac{(n+2)-(n+1)}{2}, ..., \frac{2n-(n+1)}{2}$$

The above equation can be written as

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

Or

$$\frac{1-n}{2}, \frac{3-n}{2}, \frac{5-n}{2}, ..., \frac{-3}{2}, \frac{-1}{2}, \frac{1}{2}, ..., \frac{n-1}{2}.$$

So the absolute values of deviation from the mean is

$$|x_i-\bar{x}| = \frac{(n-1)}{2}, \frac{(n-3)}{2}, \frac{(n-5)}{2}, \ldots, \frac{3}{2}, \frac{1}{2}, \frac{1}{2}, \ldots, \frac{n-1}{2}.$$

The sum of absolute values of deviations from the mean, is

$$\Sigma|x_i-\bar{x}| = \frac{(n-1)}{2}+\frac{(n-3)}{2}+\frac{(n-5)}{2} + \ldots + \frac{3}{2}+\frac{1}{2}+\frac{1}{2}+\ldots+\frac{n-1}{2}$$

We can write as

$$\Sigma|x_i-\bar{x}| = \left(\frac{1}{2}+\frac{3}{2}+\ldots+\frac{(n-1)}{2}\right)\left(\frac{n}{2}\right)$$

Now we know sum of first n odd natural numbers = n^2

Therefore, mean deviation about the mean is

$$M.D = \frac{\Sigma|x_i-\bar{x}|}{n} = \frac{\left(\frac{1}{2}+\frac{3}{2}+\ldots+\frac{(n-1)}{2}\right)\left(\frac{n}{2}\right)}{n}$$

$$= \frac{\left(\frac{n}{2}\right)\left(\frac{n}{2}\right)}{n}$$

$$M.D = \frac{\Sigma|x_i-\bar{x}|}{n} = \frac{\left(\frac{n}{2}\right)^2}{n}$$

$$M.D = \frac{\Sigma|x_i-\bar{x}|}{n} = \frac{n^2}{4n} = \frac{n}{4}$$

Hence, the mean deviation about the mean of the set of first n natural numbers when n is an even number is $\frac{n}{4}$.

30. Calculate mean and mean deviation about median for the following data:

Class	0–10	10–20	20–30	30–40	40–50	50–60
Frequency	6	7	15	16	4	2

Ans. Calculation of mean deviation from median:

C.I.	Mid-values x_i	Frequency f_i	Cumulative Frequency	$\|x_i -$ median$\|$ = $\|x_i - 28\|$	$f_i\|x_i -$ median$\|$
0–10	5	6	6	23	138
10–20	15	7	13	13	91
20–30	25	15	28	3	45
30–40	35	16	44	7	112
40–50	45	4	48	17	68
50–60	55	2	50	27	54
		$N = 50$			$\Sigma f_i\|x_i -$ median$\|$ = 508

Here, $N = 50$

$\therefore \frac{N}{2} = \frac{50}{2} = 25$, so median class is 20 – 30.

From the table: $l = 20$, h = 10, $f = 15$, $cf = 13$

Using the formula:

$$\text{Median} = l + \frac{h}{f}\left(\frac{N}{2} - cf\right)$$

$$= 20 + \frac{10}{15}\left(\frac{50}{2} - 13\right) = 20 + \frac{2}{3}(25-13)$$

$$= 20 + \frac{2}{3} \times 12 = 20 + 8 = 28$$

Thus, mean deviation about median

= M.D. (Median)

$$= \frac{\Sigma f_i|x_i - \text{Median}|}{N} = \frac{508}{50} = 10.16$$

Calculation of Mean:

C.I.	Mid-values x_i	Frequency f_i	$\mu_i = \frac{x_i-a}{h} = \frac{x_i-35}{10}$	$f_i\mu_i$
0–10	5	6	–3	–18
10–20	15	7	–2	–14
20–30	25	15	–1	0
30–40	35	16	0	0
40–50	45	4	1	4
50–60	55	2	2	4
		$N = \Sigma f_i = 50$		$\Sigma f_i\mu_i = -39$

Let the assume mean, $a = 35$

$\therefore$ Mean is given by

$$\bar{x} = a + h\frac{\Sigma f_i\mu_i}{N} = 35 + 10 \times \frac{(-39)}{50} = 35 - 7.8 = 27.2$$

31. Calculate the mean deviation about the mean of the set of first n natural numbers when n is an odd natural number.

[NCERT Exemplar]

Ans. Since, n is an odd natural number. Therefore, $n = 2m + 1$ for some natural number m.

Let $$\bar{x} = \frac{1+2+3+\ldots+(n-1)+n}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$$

$$\Rightarrow \bar{x} = \frac{2m+1+1}{2} = m+1$$

The mean deviation (M. D.) about mean is given by

$$M.D. = \frac{1}{n}\sum_{r=1}^{n}|r-\bar{x}|$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\sum_{r=1}^{2m+1}|r-(m+1)|$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}\sum_{r=1}^{2m+1}|r-(m+1)| \\ +\sum_{r=1}^{2m+1}|r-(m+1)|\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}\sum_{r=1}^{m}|r+(m+1)| \\ +\sum_{r=1}^{m}\{r-(m+1)\}\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}-\sum_{r=1}^{m}r+(m+1)\sum_{r=1}^{m}1+ \\ \sum_{r=m+1}^{2m+1}r-(m+1)\sum_{r=m+1}^{2m+1}1\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}-\frac{m(m+1)}{2}+m(m+1)+\left(\frac{m+1}{2}\right) \\ \{(m+1)+(2m+1)-(m+1)(m+1)\}\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}-\frac{m(m+1)}{2}+m(m+1)+ \\ \frac{1}{2}(m+1)(3m+2)-(m+1)^2\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{1}{2m+1}\left\{\begin{array}{l}\frac{m(m+1)}{2}+\frac{1}{2}(m+1) \\ (3m+2)-(m+1)^2\end{array}\right\}$$

$$\Rightarrow \quad M.D = \frac{m+1}{2(2m+1)}\{m+(3m+2)-2(m+1)\}$$

$$\Rightarrow \quad M.D = \frac{m+1}{2(2m+1)}(2m) = \frac{m(m+1)}{2m+1}$$

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

[$\because$ $n = 2m + 1$]

$$\Rightarrow \quad M.D = \frac{1}{n}\left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right) = \frac{n^2-1}{4n}$$

VARIANCE AND STANDARD DEVIATION 2

TOPIC 1

STANDARD DEVIATION

Standard Deviation of Ungrouped Data

Consider the ungrouped data $x_1, x_2, ... x_n$ consisting of n values. In order to find the variance and standard deviation. We have the following working rule:

Step I: Calculate mean of the given data using the formula

$$\text{Mean } \bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$$

Step II: Compute the variance and standard deviation of the formula

$$\text{Variance, } \sigma^2 = \frac{\sum (x_i - x)^2}{n}$$

$$\text{And standard Deviation, } \sigma = \sqrt{\frac{\sum (x_i - x)^2}{n}}$$

Example 2.1: **Find the mean and variance for the first 10 multiples of 3.** **[NCERT]**

Ans The first 10 multiplies of 3 are

3, 6, 9, 12, 15, 18, 21, 24, 27, 30

$$\text{Mean} = \frac{\text{sum of observations}}{\text{Number of observations}}$$

$$= \frac{3+6+9+12+15+18+21+24+27+30}{10}$$

$$= \frac{165}{10}$$

$$= 16.5$$

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
3	3 – 16.5 = –13.5	$(-13.5)^2 = 182.25$
6	6 – 16.5 = –10.5	$(-10.5)^2 = 110.25$
9	9 – 16.5 = –4.5	$(-7.5)^2 = 56.25$
12	12 – 16.5 = –4.5	$(-4.5)^2 = 20.25$
15	15 – 16.5 = –1.5	$(-1.5)^2 = 2.25$
18	18 – 16.5 = 1.5	$(1.5)^2 = 2.25$
21	12 – 16.5 = 4.5	$(4.5)^2 = 20.25$
24	24 – 16.5 = 7.5	$(7.5)^2 = 56.25$
27	27 – 16.5 = 10.5	$(10.5)^2 = 110.25$
30	30 – 16.5 = 13.5	$(13.5)^2 = 182.25$
		$\Sigma(x_i - \bar{x})^2 = 742.5$

$$\therefore \quad \Sigma(x_i - \bar{x})^2 = 742.5$$

$$\text{Variance, } (\sigma)^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{742.5}{10} = 74.25$$

Standard Deviation of Discrete Frequency Distribution

Consider a discrete frequency distribution consisting of m distinct values $x_1, x_2, ... x_m$ with frequencies, $f_1, f_2, ... f_m$ respectively. In order to find the variance and standard deviation, we have the following working rule:

Step I: Calculate mean of the given data using the formula

$$\text{Mean, } \bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

Step II: Compute the variance and standard deviation using the formula

$$\text{Variance, } \sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

$$\text{And standard deviation, } \sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

Standard Deviation Of Continuous Distribution

Consider a continuous frequency distribution consisting of class intervals with their corresponding frequencies as, f_r. In order to find the variance and standard deviation. We have the following working rule:

Step I: Find the mid-points of the class intervals and denote them by x_r

Step II: Calculate mean of the given data using the formula

$$\text{Mean} = \frac{\sum f_i x_i}{\sum f_i}$$

Step III: Compute the variance and standard deviation using the formula

$$\text{Variance, } \sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

$$\text{And standard deviation, } \sigma = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

Example 2.2: Find the mean and variance for the following frequency distribution.

Class	0–30	30–60	60–90	90–120	120–150	150–180	180–210
Frequency	2	3	5	10	3	5	2

[NCERT]

Ans.

Class	Frequency f_i	Mid-point x_i	$f_i x_i$
0–30	2	15	30
30–60	3	45	135
60–90	5	75	375
90–120	10	105	1050
120–150	3	135	405
150–180	5	165	825
180–210	2	195	390
	$\Sigma f_i = 30$		$\Sigma f_i x_i = 3210$

$\therefore \quad \Sigma f_i x_i = 3210$

$\Sigma f_i = 30$

Mean $\quad (\bar{x}) = \dfrac{\sum f_i x_i}{\sum f_i} = \dfrac{3210}{30} = 107$

Class	(f_i)	(x_i)	$u_i = \dfrac{x_i - A}{h}$	u_i^2	$f_i u_i$	$f_i u_i^2$
0–30	2	15	–3	9	–6	18
30–60	3	45	–2	4	–6	12
60–90	5	75	–1	1	–5	5
90–120	10	105	0	0	0	0
120–150	3	135	1	1	3	3
150–180	5	165	2	4	10	20
180–210	2	195	3	9	6	18
Total	30				2	76

Here, $N = 30$, $h = 30$

$$\text{Variance } (\sigma^2) = \frac{h^2}{N^2}\left[N\sum_{i=1}^{7} f_i u_i^2 - \left(\sum_{i=1}^{7} f_i u_i\right)^2\right]$$

$$= \frac{(30)^2}{(30)^2}\,[30 \times 76 - (2)^2]$$

$$= 2280 - 4$$

$$= 2276$$

Short-Cut Method For S.D. of Continuous Distribution

Consider a continuous frequency distribution consisting of class intervals with their corresponding frequencies as f_r. Let x_i be the mid-points of the class intervals.

Variance, $$\sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}$$

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

Hence, variance, $$\sigma^2 = \frac{\sum f_i x_i^2}{\sum f_i} - \left(\frac{\sum f_i x_i}{\sum f_i}\right)^2$$

And standard deviation,

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

Short-Cut Method for S.D. of Discrete Distribution

Consider a discrete frequency distribution consisting of m distinct values $x_1, x_2, \ldots x_m$ with frequencies, $f_1, f_2, \ldots f_m$ respectively.

Then, variance, $$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{\sum f_i}$$

$$= \frac{\sum \left(x^2_i - \bar{x}^2 - 2x_i x\right)}{\sum f_i}$$

$$= \frac{\sum f_i x^2_i}{\sum f_i} + \frac{\sum \bar{x}^2 f_i}{\sum f_i} - \frac{\sum 2x f_i \bar{x}_i}{\sum f_i}$$

$$= \frac{\sum f_i x_i^2}{\sum f_i} + \frac{x^2 \sum f_i}{\sum f_i} - \frac{2x \sum f_i x_i}{\sum f_i}$$

$$= \frac{\sum f_i x_i^2}{\sum f_i} + x^2 - 2x^2$$

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

Hence, variance, $$\sigma^2 = \frac{\sum f_i x^2_i}{\sum f_i} - \left(\frac{\sum f_i x_i}{\sum f_i}\right)^2$$

And standard deviation,

$$\sigma = \sqrt{\frac{\sum f_i x^2_i}{\sum f_i} - \left(\frac{\sum f_i x_i}{\sum f_i}\right)^2}$$

The coefficient of variation is defined as

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

where σ and $\bar{x}$ are the standard deviation and mean of the data, respectively and $\bar{x} \neq 0$.

For comparing the variability or dispersion of two series, we calculate the coefficient of variation for each series. We make the following conclusions:

(i) The series having greater C.V. is said to be more variable than the other.

(ii) The series having lesser C.V. is said to be more consistent than the other.

Example 2.3: Find the mean and standard deviation using the short-cut method. [NCERT]

Ans.

x_i	f_i	$y_i = x_i - A$	y_i^2	$f_i y_i$	$f_i y_i^2$
60	2	60 − 64= −3	16	−4 × 2 = −8	16 × 2 = 32
61	1	61 − 64= −3	9	−3 × 1 = −3	9 × 1 = 9
62	12	62 − 64= −3	4	−2 × 12 = −24	12 × 4 = 48
63	29	63 − 64= −3	1	−1 × 29 = −29	29 × 1 = 29
64	25	64 − 64= −3	0	0 × 25 = 0	25 × 0 = 0
65	12	65 − 64= −3	1	1 × 12 = 12	12 × 1 = 12
66	10	66 − 64= −3	4	2 × 10 = 20	10 × 4 = 40
67	4	67 − 64= −3	9	3 × 4 = 12	4 × 9 = 36
68	5	68 − 64= −3	16	4 × 5 = 20	5 × 16 = 80
	100	200		0	286

Let A = assumed mean = 64

h = width = 61 − 60 = 1

$$\text{Mean} = a + \frac{\sum f_i y_i}{\sum f_i} \times h$$

$$= 64 + \frac{0}{100} \times 1$$

$$= 64$$

$$\text{Standard Deviation} = \frac{h}{N}\sqrt{N\sum f_i y_i^2 - (\sum f_i y_i)^2}$$

$$= \frac{1}{100}\sqrt{100 \times 286 - (0)^2}$$

$$= \frac{1}{100}\sqrt{28600}$$

$$= \frac{169.1}{100}$$

$$= 1.691$$

```
          169.1
    1  | 28600.00
       | 1
   26  | 186
       | 156
  329  | 3000
       | 2961
 3381  | 3900
       | 3381
       |  519
```

Example 2.4: Case Based:

For a group of 200 candidates, the mean and the standard deviation of scores were found to be 40 and 15, respectively. Later on it was discovered that the scores of 43 and 35 were misread as 34 and 53, respectively.

Student	Eng	Hindi	S.St.	Science	Maths
Ramu	39	59	84	80	41
Rajitha	79	92	68	38	75
Komala	41	60	38	71	82
Patil	77	77	87	75	42
Pursi	72	65	69	83	67
Gayathri	46	96	53	71	39

Answer the following questions on the basis of above information.

(A) Find the sum of correct scores.

(a) 7991 (b) 8000

(c) 8550 (d) 6572

(B) Find the correct mean.

(a) 42.924 (b) 39.955

(c) 38.423 (d) 41.621

(C) The formula of variance is:

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

(c) $\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\Sigma f_i}$ (d) $\sum_{i=1}^{n} f_i(x_i - \bar{x})^2$

(D) Find the correct variance.

(a) 280.3 (b) 235.6

(c) 224.143 (d) 226.521

(E) Find the correct standard deviation.

(a) 14.971 (b) 11.321

(c) 16.441 (d) 12.824

Ans. (A) *(a) 7981*

Explanation: We have, n = 200, incorrect mean = 40

and incorrect standard deviation = 15

Now, incorrect mean = 40

$$\Rightarrow \frac{\text{Incorrect } \Sigma x_i}{200} = 40$$

Incorrect Σx_i = 8000

Correct Σx_i = 8000 − (34 + 53) + (43 + 35)

= 8000 − 87 + 78

= 7991

(B) *(b) 39.955*

Explanation: Correct mean

$$= \frac{7991}{200} = 39.955$$

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

(D) (c) 224.143

Explanation: Incorrect SD = 15

$\Rightarrow$ Incorrect variance = $(15)^2 = 225$

$\Rightarrow \dfrac{\text{Incorrect}\,\Sigma x_i^2}{200} - (\text{Incorrect mean})^2 = 225$

$\dfrac{\text{Incorrect}\,\Sigma x_i^2}{200} - (40)^2 = 225$

$\Rightarrow$ Incorrect $\Sigma x_i^2 = 200(1600 + 225)$

$= 200 \times 1825$

$= 365000$

Now, Correct Σx_i^2 = Incorrect Σx_i^2

$-(34^2 + 53^2) + (43^2 + 35^2)$

$= 365000 - 3965 + 3074$

$= 364109$

So, correct variance

$= \dfrac{1}{200}$ (correct Σx_i^2) – (correct mean)^2

$= \dfrac{1}{200}(364109) - \left(\dfrac{7991}{200}\right)^2$

$= 1820.545 - 1596.402$

$= 224.143$

(E) (a) 14.971

Explanation: Correct standard deviation

$= \sqrt{\text{correct variance}}$

$= \sqrt{224.143}$ [using part (iv)]

$= 14.971$

OBJECTIVE Type Questions

[1 mark]

Multiple Choice Questions

1. Mean and standard deviation of 100 items are 50 and 4, respectively. The sum of the squares of the items is:

(a) 25000 (b) 251600

(c) 26000 (a) none of these

[NCERT Exemplar]

Ans. (b) 251600

Explanation: Given mean and standard deviation of 100 items are 50 and 4, respectively

Now we have to find the sum of the squares of the items.

As per given criteria,

Number of items, $n = 100$

Mean of the given items, $\bar{x} = 50$

But we know,

$$\bar{x} = \frac{\Sigma x_i}{n}$$

Substituting the corresponding values, we get

$$50 = \frac{\Sigma x_i}{100}$$

$\Rightarrow \quad \Sigma x_i = 50 \times 100 = 5000$

Hence the sum of all the 100 items = 5000.

Also, given the standard deviation of the 100 items is 4.

i.e., $\sigma = 4$

But we know

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$$

Substituting the corresponding values, we get

$$4 = \sqrt{\frac{\Sigma x_i^2}{100} - \left(\frac{5000}{100}\right)^2}$$

Now taking square on both sides, we get

$$4^2 = \frac{\Sigma x_i^2}{100} - (50)^2$$

$\Rightarrow \quad 16 = \dfrac{\Sigma x_i^2}{100} - 2500$

$\Rightarrow \quad 16 + 2500 = \dfrac{\Sigma x_i^2}{100}$

On rearranging, we get

$\Rightarrow \quad \dfrac{\Sigma x_i^2}{100} = 2516$

$\Rightarrow \quad \Sigma x_i^2 = 2516 \times 100$

$\Rightarrow \quad \Sigma x_i^2 = 251600$

The sum of the squares of all the 100 items is 251600.

2. If for a distribution $\Sigma(x - 5) = 3$, $\Sigma(x - 5)^2 = 43$ and the total number of item is 18, the standard deviation is:

(a) 5.17 and 1.54 (b) 4.27 and 4.81

(c) 5.23 and 4.81 (d) 4.03 and 4.81

[NCERT Exemplar]

Ans. *(a) 5.17 and 1.54*

Explanation: Given for a distribution $\Sigma(x-5)=3$, $\Sigma(x-5)^2=43$ and the total number of items is 18.

Now, we have to find the mean and standard deviation.

As per given criteria,

Number of items, $n = 18$

And given $\Sigma(x-5)=3$,

And also given, $\Sigma(x-5)^2 = 43$

But, we know mean can be written as,

$$\bar{x} = A + \frac{\Sigma(x-5)}{n}$$

Here assumed mean is 5, so substituting the corresponding values in above equation, we get

$$\bar{x} = 5+\frac{3}{18} = \frac{18\times5+3}{18} = \frac{93}{18} = 5.17$$

And we know the standard deviation can be written as,

$$\sigma = \sqrt{\frac{\Sigma(x-5)^2}{n} - \left(\frac{\Sigma(x-5)}{n}\right)^2}$$

Substituting the corresponding values, we get

$$\sigma = \sqrt{\frac{43}{18} - \left(\frac{3}{18}\right)^2}$$

$$\sigma = \sqrt{2.39-(0.166)^2}$$

$$\sigma = \sqrt{2.39-0.027} = \sqrt{2.363}$$

Hence, $\sigma = 1.54$

So, the mean and standard deviation of given items is 5.17 and 1.54 respectively.

3. Mean and standard deviation of 100 observation were found to be 40 and 10, respectively. If at the time of calculation two observation were wrongly taken as 30 and 70 in place of 3 and 27 respectively, find the correct standard deviation.

(a) 14.24 (b) 10.24

(c) 19.23 (d) 20

[NCERT Exemplar]

Ans. *(b) 10.24*

Explanation: Given mean and standard deviation of 100 observations were found to be 40 and 10, respectively. If at the time of calculation, two observations were wrongly taken as 30 and 70 in place of 3 and 27 respectively.

Now we have to find the correct standard deviation.

As per given criteria,

Number of observations, $n = 100$

Mean of the given observations before correction, $\bar{x} = 40$

But we know,

$$\bar{x} = \frac{\Sigma x_i}{100}$$

$$40 = \frac{\Sigma x_i}{100}$$

$$\Rightarrow \quad \Sigma x_i = 40\times100 = 4000$$

It is said two observations were wrongly taken as 30 and 70 in place of 3 and 27 respectively.

So, $\Sigma x_i = 4000 - 30 - 70 + 3 + 27 = 3930$

So, the correction mean after correction is

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{3930}{100} = 39.3$$

Also given the standard deviation of the 100 observations is 10 before correction i.e., $\sigma = 10$

But we know

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$$

Substituting the corresponding values, we get

$$10 = \sqrt{\frac{\Sigma x_i^2}{100} - \left(\frac{4000}{100}\right)^2}$$

Now taking square on both sides, we get

$$10^2 = \frac{\Sigma x_i^2}{100} - (40)^2$$

$$\Rightarrow \quad 100 = \frac{\Sigma x_i^2}{100} - 1600$$

$$\Rightarrow \quad 100 + 1600 = \frac{\Sigma x_i^2}{100}$$

$$\Rightarrow \quad \frac{\Sigma x_i^2}{100} = 1700$$

$$\Rightarrow \quad \Sigma x_i^2 = 170000$$

It is said two observation were wrongly taken as 30 and 70 in place of 3 and 27 respectively, so correction is

$\Rightarrow \Sigma x_i^2 = 170000 - (30)^2 - (70)^2 + 3^2 + (27)^2$

$\Rightarrow \Sigma x_i^2 = 170000 - 900 - 4900 + 9 + 729$

$= 164938$

$\Rightarrow \Sigma x_i^2 = 164938$

So, the correct standard deviation after correction is

$$\sigma = \sqrt{\frac{164938}{100} - \left(\frac{3930}{100}\right)^2}$$

$$\sigma = \sqrt{1649.38 - (39.3)^2}$$

$$\sigma = \sqrt{1649.38 - 1544.49} = \sqrt{104.89}$$

$$\sigma = 10.24$$

Thus, the correct standard deviation is 10.24.

4. While calculating the mean and variance of 10 readings, a student wrongly used the reading 52 for the correct reading 25. He obtained the mean and variance as 45 and 16 respectively. Find the correct mean and the variance.

(a) 42.3 and 43.81 (b) 43.3 and 43.81

(c) 42.3 and 42.81 (d) 40.3 and 43.81

[NCERT Exemplar]

Ans. *(a) 42.3 and 43.81*

Explanation: Given while calculating the mean and variance of 10 readings, a student wrongly used the reading 52 for the correct reading 25. He obtained the mean and variance as 45 and 16 respectively

Now we have to find the correct mean and the variance.

As per given criteria,

Number of reading, $n = 10$

Mean of the given reading before correction,

$$\bar{x} = 45$$

But we know,

$$\bar{x} = \frac{\Sigma x_i}{n}$$

Substituting the corresponding values, we get

$$45 = \frac{\Sigma x_i}{n}$$

$$\Rightarrow \quad \Sigma x_i = 45 \times 10 = 450$$

It is said one reading 25 was wrongly taken as 52,

So, $\bar{x}_i = 45 \times 10 = 450$

It is said one reading 25 was wrongly taken as 52,

So, $\bar{x}_i = 450 - 52 + 25 = 423$

So the correct mean after correction is

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{423}{10} = 42.3$$

Also given the variance of the 10 reading is 16 before correction,

i.e. $\sigma^2 = 16$

But we know

$$\sigma^2 = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2$$

Substituting the corresponding values, we get

$$16 = \frac{\Sigma x_i^2}{10} - (45)^2$$

$$\Rightarrow \quad 16 = \frac{\Sigma x_i^2}{10} - 2025$$

$$\Rightarrow \quad 16 + 2025 = \frac{\Sigma x_i^2}{10}$$

$$\Rightarrow \quad \frac{\Sigma x_i^2}{10} = 2041$$

$$\Rightarrow \quad \Sigma x_i^2 = 20410$$

It is said one reading 25 wrongly taken as 52, so

$\Rightarrow \Sigma x_i^2 = 20410 - (52)^2 + (25)^2$

$\Rightarrow \Sigma x_i^2 = 20410 - 2704 + 625$

$\Rightarrow \Sigma x_i^2 = 18331$

So the correct variance after correction is

$$\sigma^2 = \frac{18331}{10} - \left(\frac{423}{10}\right)^2$$

$$\sigma^2 = 1833.1 - (42.3)^2 = 1833.1 - 1789.29$$

$$\sigma^2 = 43.81$$

Thus, the corrected mean and variance is 42.3 and 43.81 respectively.

5. The variance of the following data is:

Intervals	4–8	8–12	12–16	16–20
Frequency	3	6	4	7

(a) 13 (b) 18

(c) 19 (d) 20

Ans. *(c) 19*

Explanation:

Class	x_i	f_i
4 – 8	6	3
8 – 12	10	6
12 – 16	14	4
16 – 20	18	7

$$\Sigma f_i x_i = 18 + 60 + 56 + 126 = 260$$

$$\Sigma f_i = 20$$

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = 13$$

$$\text{Variance} = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} = \frac{380}{20}$$

$$= 19$$

6. The standard deviation for the data 10, 20, 30, 40, 50, is:

(a) $5\sqrt{2}$ (b) $8\sqrt{2}$

(c) $10\sqrt{2}$ (d) $15\sqrt{2}$ [Diksha]

Ans. *(c)* $10\sqrt{2}$

Explanation:

$$\bar{x} = \frac{10+20+30+40+50}{5} = \frac{150}{5} = 30$$

x	$d = x - \bar{x}$	d^2
10	$10 - 30 = -20$	400
20	$20 - 30 = -10$	100
30	$30 - 30 = 0$	0.
40	$40 - 30 = 10$	100
50	$50 - 30 = 20$	400
	Σd^2	1000

$$\text{Variance} = \frac{\sum d^2}{n} = \frac{1000}{5}$$

Standard deviation

$$= \sqrt{\text{variance}} = \sqrt{200} = 10\sqrt{2}$$

7. The standard deviation of data 7, 6, 10, 14, 13, 9 and 11 is:

(a) $\sqrt{\frac{52}{7}}$ (b) $\frac{52}{7}$

(c) $\sqrt{6}$ (d) 6

Ans. *(a)* $\sqrt{\frac{52}{7}}$

Explanation:

x_i	x_i^2
7	49
6	36
10	100
14	196
13	169
9	81
11	121
70	752

$$S.D. = \sqrt{\frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2}$$

$$= \sqrt{\frac{752}{7} + \left(\frac{70}{7}\right)^2}$$

$$= \sqrt{\frac{5264 - 4900}{49}}$$

$$= \sqrt{\frac{364}{49}}$$

$$= \sqrt{\frac{52}{7}}$$

8. Find the mean and the variance of the 8, 9, 12, 14, 15, 6, 10, 14.

(a) 8.25 (b) 9.25

(c) 65.2 (d) 55

Ans. *(b)* 9.25

Explanation Given data is 8, 9, 12, 14, 15, 6, 10, 14

$$\text{Mean } \bar{x} = \frac{\sum_{i=1}^{8} x_i}{n}$$

$$= \frac{8+9+12+14+15+6+10+14}{8}$$

$$= \frac{88}{8} = 11$$

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
8	−3	9
9	−2	4
12	1	1
14	3	9
15	4	16
6	−5	25
10	−1	1
14	3	9
		74

$$\therefore \text{Variance } (\sigma^2) = \frac{1}{n}\sum_{i=1}^{8}(x_i - \bar{x})^2 = \frac{1}{8} \times 74 = 9.25$$

9.

x_i	10	15	18	20	25
f_i	3	2	6	7	2

The variance for the data is:

(a) 22.5 (b) 22.159

(c) 21.159 (d) 16.79

Ans. *(d)* 16.79

Explanation:

x_i	f_i	$f_i x_i$	$x_i - x$	$(x_i - x)^2$	$f_i (x_i - x)^2$
10	3	30	–7.9	62.41	187.23
15	2	30	–2.9	8.41	16.82
18	6	108	0.1	0.01	0.06
20	7	140	2.1	4.41	30.87
25	2	50	7.1	50.41	100.82
Total	20	358			335.8

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{358}{20} = 17.9$$

$$\sigma^2 = \frac{1}{N}\sum f_i (x_i - \underline{x})^2$$

$$= \frac{335.8}{20} = 16.79$$

10. The variance of 20 observations is 5. If each observation is multiplied by 2, then the new variance of resulting observation is

(a) 5 (b) 10

(c) 20 (d) none of these

[Delhi Gov. Term-1 SQP 2021]

Ans. *(c)* 20

Explanation:

$\Rightarrow$ Let the observations be $x_1, x_2, x_3, ..., x_{20}$ and x be their mean.

Given, variance = 5 and $n = 20$

$$\Rightarrow \quad \text{Variance} = \frac{1}{n}\sum (x_i - x)^2$$

$$5 = \frac{1}{20}\sum (x_i - x)^2$$

$$\therefore \quad \sum (x_i - x)^2 = 100 \quad ...(i)$$

If each observation is multiplied by 2, we get new observations.

Let new observations by $y_1, y_2, y_3, ...y_{20}$

Where, $y_1 = 2(x_i)$ ...(ii)

We need to find variance of new observations

i.e., $$\text{New variance} = \frac{1}{n}\sum (y_1 - y)^2$$

$$\text{Calculating } y = \frac{1}{n}\sum y_i$$

$$y = \frac{1}{20}\sum 2x_i$$

$$y = 2\left(\frac{1}{20}\sum x_i\right)$$

$$y = 2x \quad ...(iii)$$

$$\Rightarrow \quad \sum (x_1 - x)^2 = 100 \quad \text{[From (i)]}$$

$$\Rightarrow \quad \sum \left(\frac{1}{2}y_1 - \frac{1}{2}y\right)^2 = 100 \quad \text{[From (ii) and (iii)]}$$

$$\Rightarrow \quad \left(\frac{1}{2}\right)^2 \sum (y_1 - y)^2 = 100$$

$$\Rightarrow \quad \sum (y_1 - y)^2 = 400$$

$\Rightarrow$ Now, New variance $= \frac{1}{n}\sum (y_1 - y)^2$

$$= \frac{1}{20} \times 400$$

$$= 20$$

11. The standard deviation of the first ten natural numbers is:

(a) 5.5 (b) 2.97

(c) 3.87 (d) 2.87 [Diksha]

Ans. *(d)* 2.87

Explanation: We know that the standard deviation of first n natural number is $\sqrt{\frac{n^2 - 1}{12}}$.

$\therefore$ Standard deviation of first 10 natural numbers

$$= \sqrt{\frac{10^2 - 1}{12}}$$

$$= \sqrt{\frac{99}{12}}$$

$$= \sqrt{8.25}$$

$$= 2.87$$

12. Let x is the variance and y is the standard deviation then which of the following is correct.

(a) $x = y^2$ (b) $xy = 1$

(c) $x^2 = y$ (d) $x^2 y^2 = 1$

Ans. *(a)* $x = y^2$

Explanation: The standard deviation is the square root of the variance.

$\therefore$ If x is the variance and y is the standard deviation, then $y = \sqrt{x}$, or $y^2 = x$

CASE BASED Questions (CBQs)

[4 & 5 marks]

Read the following passages and answer the questions that follow:

13. *For a group of 200 candidates, the mean and the standard deviation of scores were found to be 40 and 15, respectively. Later on, it was discovered that the scores of 43 and 35 were misread as 34 and 53, respectively:*

(A) The sum of correct scores is:

(a) 7991 (b) 8000

(c) 8550 (d) 6572

(B) The correct mean is:

(a) 42.924 (b) 39.955

(c) 38.423 (d) 41.621

(C) The correct variance is:

(a) 280.3 (b) 235.6

(c) 224.143 (d) 226.521

(D) The correct standard deviation is:

(a) 14.971 (b) 11.321

(c) 16.441 (d) 12.824

(E) Assertion (A): The variance of first n even natural numbers is $\frac{n^2-1}{4}$.

Reason (R): The sum of first n natural numbers is $\frac{n(n+1)}{2}$ and the sum of squares of first n natural number is $\frac{n(n+1)(2n+1)}{6}$.

Ans. (A) *(a) 7991*

Explanation: We have, n = 200, incorrect mean = 40 and incorrect deviation = 15

Now, incorrect mean = 40

$$\Rightarrow \frac{\text{Incorrect} \sum x_i}{200} = 40$$

Incorrect $\sum x_i = 8000$

$\Rightarrow$ correct $\sum x_i = 8000 - (34 + 53) + (43 + 35)$
$= 8000 - 87 + 78 = 7991$

(B) *(b) 39.955*

Explanation: Correct mean $= \frac{7991}{200}$
$= 39.955$

(C) *(c) 224.143*

Explanation: Incorrect *SD* = 15

$\Rightarrow$ incorrect variance $= (15)^2 = 225$

$$\Rightarrow \frac{\text{Incorrect} \sum x_i^2}{200} - (\text{incorrect mean})^2 = 225$$

$$\Rightarrow \frac{\text{Incorrect} \sum x_i^2}{200} - (40)^2 = 225$$

$\Rightarrow$ Incorrect $\sum x_i^2 = 200(1600 + 225)$
$= 200 \times 1825 = 365000$

$\Rightarrow$ correct $\sum x_i^2$ = Incorrect $x_i^2 - (34^2 + 53^2) + (43^2 + 35^2)$
$= 365000 - 3965 + 3074$
$= 364109$

So, correct variance

$$= \frac{1}{200}(\text{correct} \sum x_i^2) - (\text{correct mean})^2$$

$$= \frac{1}{200}(364109) - \left(\frac{7991}{200}\right)^2$$

$= 1820.545 - 1596.402$

$= 224.143$

(D) *(a) 14.971*

Explanation: Correct standard deviation

$= \sqrt{\textit{correct variance}}$

$= \sqrt{224.143}$ (using part (C)

$= 14.971$

(E) **Assertion:** Sum of n even natural numbers

$= n(n + 1)$

$$\text{Mean } (\bar{x}) = \frac{n(n+1)}{n} = n+1$$

$$\text{Variance} = \left[\frac{1}{n}\Sigma(x_i)^2\right] - (\bar{x})^2$$

$$= \frac{1}{n}[2^2 + 4^2 + ... + (2n)^2] - (n+1)^2$$

$$= \frac{1}{n}2^2[1^2 + 2^2 + ... + n^2] - (n+1)^2$$

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

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

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

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

14. *The following values are calculated in respect of length and mass of the students of class XII*

	Length	Mass
Mean	164cm	64kg
Variance	125cm^2	25kg^2

(A) Find the coefficient of variation of length.

(B) Find the coefficient of variation of mass and standard deviation of length.

(C) Find the variance and standard deviation for the following data 6, 7, 10, 12, 13, 4, 8, 12.

Ans. (A) S.D. of length = $\sqrt{25}$ = 11.18

The coefficient of variation of length

$$= \frac{S.D.}{Mean} \times 100$$

$$= \frac{11.18}{164} \times 100$$

= 6.81

(B) S.D. of mass = $\sqrt{25}$ = 5

$$C.V. \text{ of mass} = \frac{S.D}{Mean} \times 100$$

$$= \frac{5}{64} \times 100$$

= 7.81

$$S.D. = \sqrt{\text{variance}} = \sqrt{125} = 11.18$$

(C) Given observations are 6, 7, 10, 12, 13, 4, 8, 12

Number of observations = 8

$$\therefore \text{Mean}(\bar{x}) = \frac{6+7+10+12+13+4+8+12}{8}$$

$$= \frac{72}{8} = 9$$

Now, let us make the following table for deviation.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
6	−3	9	13	4	16
7	−2	4	4	−5	25
10	1	1	8	−1	1
12	3	9	12	3	9
	Total	74		Total	74

$\therefore$ Sum of squares of deviations

$$= \sum_{i=1}^{8}(x_i - \bar{x})^2 = 74$$

$$\text{Hence, variance, } \sigma^2 = \frac{\sum_{i=1}^{8}(x_i - \bar{x})^2}{n} = \frac{74}{8} = 9.25$$

$$\text{and standard deviation} = \sqrt{\sigma} = \sqrt{9.25} = 3.04$$

VERY SHORT ANSWER Type Questions (VSA)

[1 mark]

15. Find the coefficient of variation when particular data will have a variance 4 and mean 5. [Diksha]

Ans. Coefficient of variation = $\frac{S.D}{Mean} \times 100$

Here variance = 4

So

$$S.D = \sqrt{4} = 2$$

Mean = 5

So,

$$CV = \frac{2}{5} \times 100$$

$$CV = 40$$

16. The sum of the squares of deviation of 10 observations taken from their mean 50 is 250. Find Standard Deviation. [Delhi Gov. QB 2022]

Ans. We have,

$$\bar{X} = 50, n = 10$$

$$\sum_{i=1}^{10}(x_i - \bar{X})^2 = 250$$

$$\therefore \quad SD = \sqrt{\text{Variance of } X}$$

$$= \sqrt{\frac{\sum_{i=1}^{10}(x_i - \bar{x})^2}{n}}$$

$$= \sqrt{\frac{250}{10}} = 5$$

17. Find the variance of the following data 5, 4, 8, 11, 7

Ans. Given series is 5, 8, 11, 14, 17

$$\text{Mean} = \frac{5+8+11+14+17}{5} = \frac{55}{5}$$

Mean = μ = 11

Variance = $\Sigma(x_i - \mu)^{2/n}$

$$= \frac{1}{5}[(5-11)^2 + (8-11)^2 + (11-11)^2 + (14-11)^2 + (17-11)^2]$$

$$= \frac{1}{5}[36 + 9 + 0 + 9 + 36]$$

$$= \frac{90}{5}$$

$$= 18$$

18. **Find the variance for the data 7, 6, 10, 12, 13, 4, 8, 12.**

Ans.

x_i	x_i^2
6	36
7	49
10	100
12	144
13	169
4	64
8	64
12	144
Total = 72	722

$$n = 8$$

$$\text{Variance} = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2$$

$$= \frac{722}{8} - \left(\frac{72}{8}\right)^2$$

$$= 9.25$$

19. **Find the standard deviations of first 10 even natural numbers.** **[Delhi Gov. QB 2022]**

Ans. We know standard deviation of first n even natural number is, $\sigma = \sqrt{\frac{n^2-1}{3}}$

Here, $n = 10$

$$\therefore \quad \sigma = \sqrt{\frac{100-1}{3}} = \sqrt{33} = 5.74$$

20. **If the coefficient of variation is 45% and the mean is 12, then find its standard deviation.** **[Delhi Gov. QB 2022]**

Ans. Since, coeffcient of variance

$$= \frac{\text{standard deviation}}{\text{Mean}} \times 100$$

$$45 = \frac{\sigma}{12} \times 100$$

$$\sigma = \frac{45 \times 12}{100}$$

$$\sigma = 5.4$$

21. **Variance of the data 2, 4, 5, 6, 8, 17 is 23.33. Then, find the variance of 4, 8, 10, 12, 16 34.**

Ans. When each observation is multiplied by 2. Then variance is also multiplied by 2.

Variance new series = 2 × variance of given data

= 2 × 23.33

= 46.66

SHORT ANSWER Type-I Questions (SA-I)

[2 marks]

22. **If the mean and standard deviation of 100 observations are 50 and 4 respectively. Find the sum of all observations and the sum of their squares.** **[NCERT Exemplar]**

Ans. Let $x_1, x_2, ... x_{100}$ be 100 observations and their mean and standard deviations be $\bar{X}$ and, σ respectively. Then,

$$\bar{X} = \frac{1}{100}\sum_{i=1}^{100} x_i \text{ and, } \sigma^2 = \frac{1}{100}\sum_{i=1}^{100} x_i^2 - \bar{x}$$

$$\Rightarrow \quad 50 = \frac{1}{100}\sum_{i=1}^{100} x_i \text{ and, } 16 = \frac{1}{100}\sum_{i=1}^{100} x_i^2 - 50^2$$

$$[\because \bar{X} = 50 \text{ and } \sigma = 4]$$

$$\Rightarrow \quad 5000 = \sum_{i=1}^{100} x_i \text{ and, } 1600 = \sum_{i=1}^{100} x_i^2 - 250000$$

$$\Rightarrow \quad \sum_{i=1}^{100} x_i = 5000 \text{ and } \sum_{i=1}^{100} x_i^2 - 251600$$

$$\Rightarrow \quad \sum_{i=1}^{100} x_i = 5000 \text{ and } \sum_{i=1}^{100} x_i^2 = 251600$$

23. **Consider the following data**

Marks obtained	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80
Number of stu-dents	4	6	16	28	16	6	4

Then, find the mean deviation about the mean.

Ans. Take the assumed mean $a = 45$ and $h = 10$, and form the following table:

Marks Obtained	No. of students (f_i)	Mid value (x_i)	$d_i = \frac{x_i - 45}{10}$	$f_i d_i$	$\lvert x_i - \bar{x}\rvert$	$f_i\lvert x_i - \bar{x}\rvert$
10 – 20	4	15	–3	–12	30	120
20 – 30	6	25	–2	–12	20	120
30 – 40	16	35	–1	–16	10	160
40 – 50	28	45	0	0	0	0
50 – 60	16	55	1	16	10	160
60 – 70	6	65	2	12	20	120
70 – 80	4	75	3	12	30	120
	40			0		800

Therefore, $\bar{x} = a + \frac{\sum_{i=1}^{7} f_i d_i}{N} \times h$

$= 45 + \frac{0}{40} \times 10$

$= 45$

And M.D. $(\bar{x}) = \frac{1}{N}\sum_{i=1}^{7} f_i \lvert x_i - \bar{x}\rvert$

$= \frac{800}{40} = 20$

24. Calculate the mean deviation from the median of the below data:
44, 48, 52, 54, 56, 73, 80, 58, 64, 65.

Ans. Arranging the data in ascending order, we have
44, 48, 52, 54, 56, 58, 64, 65, 73, 80
Here, $n = 10$
So, median is the mean of 5th and 6th terms

$\therefore$ Median $(M) = \left(\frac{56+58}{2}\right) = 57$

We make the table from the given data

Scores (x_i)	Deviation form Median ($x_i - M$)	$\lvert x_i - M\rvert$
44	44 – 57 = –13	13
48	78 – 57 = –9	9
52	52 – 57 = –5	5
54	54 – 57 = –3	3
56	56 – 57 = –1	1
58	58 – 57 = 1	1
64	64 – 57 = 7	7
65	65 – 57 = 8	8
73	73 – 57 = 16	16
80	80 – 57 = 23	23
Total		**86**

$\therefore$ Mean deviation $= \frac{\Sigma\lvert x_i - M\rvert}{n} = \frac{86}{10} = 8.6$

Hence, the mean deviation from the median is 8.6.

25. The frequency distribution:

x	A	2A	3A	4A	5A	6A
F	2	1	1	1	1	1

where A is a positive integer, has a variance of 160. Determine the value of A.

[NCERT Exemplar]

Ans. Given frequency distribution table, where variance = 160

Now we have to find the value of A, where A is a positive number.

Now, we have to construct a table of the given data

Size (x_i)	Frequency (f_i)	$f_i x_i$	$f_i x_i^2$
A	2	$2A$	$2A^2$
$2A$	1	$2A$	$4A^2$
$3A$	1	$3A$	$9A^2$
$4A$	1	$4A$	$16A^2$
$5A$	1	$5A$	$25A^2$
$6A$	1	$6A$	$36A^2$
Total	**7**	**$22A$**	**$92A^2$**

And we know variance is

$$\sigma = \frac{\Sigma f_i x_i^2}{n} - \left(\frac{\Sigma f_i x_i}{n}\right)^2$$

Substituting values from above table and also given variance = 160, we get

$$160 = \frac{92A^2}{7} - \left(\frac{22A}{7}\right)^2$$

$$160 = \frac{92A^2}{7} - \frac{484A^2}{49}$$

$$160 = \frac{7 \times 92A^2 - 484A^2}{49}$$

$$160 = \frac{644A^2 - 484A^2}{49}$$

$$160 = \frac{160A^2}{49}$$

$\Rightarrow \quad A^2 = 49$

$\Rightarrow \quad A = 7$

Hence, the value of A is 7.

26. **Find the mean and variance of the frequency distribution given below:**

x	$1 \le x < 3$	$3 \le x < 5$	$5 \le x < 7$	$7 \le x < 10$
f	6	4	5	1

Ans. Given the frequency distribution.

Now we have to find the mean and variance.

Converting the ranges of x to groups, the given table can be rewritten as shown below,

X (class)	f_i	x_i	$f_i x_i$	$f_i x_i^2$
1 – 3	6	2	12	24
3 – 5	4	4	16	64
5 – 7	5	6	30	180
7–10	1	8.5	8.5	72.25
Total	**16**		**66.5**	**340.25**

And we know variance can be written as

$$\sigma = \frac{\Sigma f_i x_i^2}{n} - \left(\frac{\Sigma f_i x_i^2}{n}\right)$$

Substituting values from above table, we get

$$\sigma^2 = \frac{340.25}{16} - \left(\frac{66.5}{16}\right)^2$$

On simplifying, we get

$$\sigma^2 = 21.265 - (4.16)^2$$

$$\sigma^2 = 21.265 - 17.305 = 3.96$$

We also know that mean can be written as

$$\bar{x} = \frac{\Sigma f_i x_i}{n}$$

Substituting values from above table, we get

$$\bar{x} = \frac{66.5}{16} = 4.16$$

Hence, the mean and variance of the given frequency distribution are 4.16 and 3.96 respectively.

SHORT ANSWER Type-II Questions (SA-II)

[3 marks]

27. **For the frequency distribution:**

x	2	3	4	5	6	7
f	4	9	16	14	11	6

Find the standard distribution.

[NCERT Exemplar]

Ans. Now, we have to find the standard deviation.

Let us make a table of the given data and append other columns after calculations

Size (x_i)	Frequency (f_i)	$f_i x_i$	$f_i x_i^2$
2	4	8	16
3	9	27	81
4	16	64	256
5	14	70	350
6	11	66	396
7	6	42	294
Total	**60**	**277**	**1393**

And we know standard deviation is

$$\sigma = \sqrt{\frac{\Sigma f_i x_i^2}{n} - \left(\frac{\Sigma f_i x_i^2}{n}\right)}$$

Substituting values from above table, we get

$$\sigma = \sqrt{\frac{1393}{60} - \left(\frac{277}{60}\right)^2}$$

$$\sigma = \sqrt{23.23 - (4.62)^2}$$

$$\sigma = \sqrt{23.23 - 21.34}$$

$$\Rightarrow \quad \sigma = 1.37$$

Hence, the standard deviation is 1.37.

28. **The mean of 5 observations is 4.4 and their variance is 8.24. If three of the observations are 1, 2, and 6. Find the other two observations.** **[Delhi Gov. QB 2022]**

Ans. Let the other two observations be x and y.

Then,

$$\text{Mean} = 4.4$$

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

$$\Rightarrow \quad x + y = 13$$

$$\text{Variance} = 8.24$$

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

$$\Rightarrow \frac{41 + x^2 + y^2}{5} - 19.36 = 8.24$$

$$\Rightarrow \quad x^2 + y^2 + 41 = 138$$

$$\Rightarrow \quad x^2 + y^2 = 97 \quad \text{...(ii)}$$

Now, $(x - y)^2 + (x + y)^2$

$$\Rightarrow \quad (x - y)^2 + 169 = 2 \times 97 \text{[Using (i) and (ii)]}$$

$$\Rightarrow \quad (x - y)^2 = 25$$

$$\Rightarrow \quad x - y = 5 \quad \text{...(iii)}$$

Solving (i) and (iii), we get $x = 9$ and $y = 4$.

29. **Find the variance and standard deviation for the following distribution:**

X	4.5	14.5	24.5	34.5	44.5	54.5	64.5
f	1	5	12	22	17	9	4

Ans. Calculation of variance and standard Deviation

x_i	f_i	$d_i = xi - 34.5$	$u_i = \frac{x_i - 345}{10}$	$f_i u_i$	u_i^2	$f_i u_i^2$
4.5	1	−30	−3	−3	9	9
14.5	5	−20	−2	−10	4	20
24.5	12	−10	−1	−12	1	12
34.5	22	0	0	0	0	0
44.5	17	10	1	17	1	17
54.5	9	20	2	18	4	36
64.5	4	30	3	12	9	36
	$N = \Sigma f_i = 70$			$\Sigma f_i u_i = 22$		$\Sigma f_i u_i^2 = 130$

Here $N = 70, \Sigma f_i u_i = 22, \Sigma f_i u_i^2 = 130$ and $h = 10$

$\therefore$ $\text{var}(X) = h^2[(\frac{1}{N}\sum f_i u_i^2) - (\frac{1}{N}\sum f_i u_i)^2]$

$\Rightarrow$ $\text{Var}(X) = 100\left[\frac{130}{70} - \left(\frac{22}{70}\right)^2\right]$

$= 100\left[\frac{13}{7} - \left(\frac{11}{35}\right)^2\right]$

$= 100\,[1.857 - 0.098] = 175.822$

Hence, S.D. $= \sqrt{Var(X)} = \sqrt{175.822} = 13.259$

30. **The following table gives the distribution of income of 100 families in a village. Calculate the standard deviation:**

Income (₹)	0–1000	1000–2000	2000–3000	3000–4000	4000–5000	5000–6000
No. of Families	18	26	30	12	10	4

Ans. Calculation of standard deviation

Income (₹)	Mid-values x_i	No. of families (frequencies) f_i	$u_i = \frac{x_i - 2500}{1000}$	$f_i u_i$	u_i^2	$f_i u_i^2$
0–1000	500	18	−2	−36	4	72
1000–2000	1500	26	−1	−26	1	26
2000–3000	2500	30	0	0	0	0
3000–4000	3500	12	1	12	1	12
4000–5000	4500	10	2	20	4	40
5000–6000	5500	4	3	12	9	36
		$\Sigma f_i = 100$		$\Sigma f_i u_i = -18$		$\Sigma f_i u_i^2 = 186$

Here, $N = 100$, $\Sigma f_i u_i = -18$, $\Sigma f_i u_i^2 = 186$ and, $h = 1000$

$\therefore$ $\text{Var}(X) = h^2\left\{\frac{1}{N}(\Sigma f_i ui^2) - \left(\frac{1}{N}\Sigma f_i ui\right)^2\right\}$

$= (1000)^2\left\{\frac{186}{100} - \left(\frac{-18}{100}\right)^2\right\}$

$= 1827600$

Hence, S.D. $= \sqrt{Var(X)} = \sqrt{1827600} = 1351.88$

31. **If for a distribution of 18 observations, $\Sigma(x_i - 5) = 10$ and $\Sigma(x_i - 5)^2 = 50$, find the mean and standard deviation.**

Ans. We have, $\sum_{i=1}^{18}(x_i - 5) = 10$ *and* $\sum_{i=1}^{18}(x_i - 5)^2 = 50$

$\Rightarrow$ $\sum_{i=1}^{18} x_i - \sum_{i=1}^{18} 5 = 10$ and

$\sum_{i=1}^{18} x_i^2 - 10\sum_{i=1}^{18} x_i + \sum_{i=1}^{18} 25 = 50$

$\Rightarrow$ $\sum_{i=1}^{18} x_i - 18 \times 5 = 10$ and,

$\sum_{i=1}^{18} x_i^2 - 10\sum_{i=1}^{18} x_i - 18 \times 25 = 50$

$\Rightarrow$ $\sum_{i=1}^{18} x_i = 100$ and, $\sum_{i=1}^{18} x_i^2 - 10$

$\times 100 - 18 \times 25 = 50$

$\Rightarrow$ $\sum_{i=1}^{18} x_i = 100$ and, $\sum_{i=1}^{18} x_i^2 = 600$

$\therefore$ Mean $= \frac{1}{18}\sum_{i=1}^{18} x_i = \frac{100}{18} = 5.55$

S.D. $= \sqrt{\frac{1}{18}\sum_{i=1}^{18} x_i^2 - (\frac{1}{18}\sum_{i=1}^{18} x_i)^2} = \sqrt{\frac{600}{18} - \left(\frac{100}{18}\right)^2}$

$= \sqrt{\frac{10800 - 10000}{324}} = \sqrt{\frac{800}{324}}$

$= \sqrt{2.46} = 1.56$

32. **Find the variance and standard deviation of the following frequency distribution:**

Variable (x_i)	2	4	6	8	10	12	14	16
Frequency (f_i)	4	4	5	15	8	5	4	5

Ans. Calculation of variance and Standard Deviation

Variable x_i	Frequency f_i	$f_i x_i$	$x_i - \bar{x} = x_i - 9$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
2	4	8	–7	49	196
4	4	16	–5	25	100
6	5	30	–3	9	45
8	15	120	–1	1	15
10	8	80	1	1	8
12	5	60	3	9	45
14	4	56	5	25	100
16	5	80	7	49	245
	$N = \Sigma f_i = 50$	$\Sigma f_i x_i = 450$			$\Sigma f_i (x_i - \bar{x})^2 = 754$

Here, $N = 50$, $\Sigma f_i x_i = 450$ and, $\Sigma f_i (x_i - \bar{X})^2 = 754$

$$\therefore \quad \bar{X} = \frac{1}{N} \Sigma f_i x_i = \frac{450}{50} = 9$$

And, $\quad \text{Var}(X) = \frac{1}{N}\left\{\Sigma f_i (x_i - \bar{X})^2\right\} = \frac{754}{50} = 15.08$

Hence, $\quad S.D. = \sqrt{Var(X)} = \sqrt{15.08} = 3.88$

LONG ANSWER Type Questions (LA)

[**4** & **5** marks]

33. **There are 60 students in a class. The following is the frequency distribution of the marks obtained by the students in a test:**

Marks	0	1	2	3	4	5
Frequency	$x - 2$	x	x^2	$(x + 1)^2$	$2x$	$x + 1$

Where x is a positive integer. Determine the mean and standard deviation of the marks.

[NCERT Exemplar]

Ans. Given there are 60 students in a class. The frequency distribution of the marks obtained by the students in a test is also given.

Now, we have to find the mean and standard deviation of the marks.

It is given there are 60 students in the class, so

$$\Sigma f_i = 60$$

$\Rightarrow \quad (x - 2) + x + x^2 + (x + 1)^2 + 2x + x + 1 = 60$

$\Rightarrow \quad 5x - 1 + x^2 + x^2 + 2x + 1 = 60$

$\Rightarrow \quad 2x^2 + 7x = 60$

$\Rightarrow \quad 2x^2 + 7x - 60 = 0$

Splitting the middle term, we get

$\Rightarrow \quad 2x^2 + 15x - 8x - 60 = 0$

$\Rightarrow \quad x(2x + 15) - 4(2x + 15) = 0$

$\Rightarrow \quad (2x + 15)(x - 4) = 0$

$\Rightarrow \quad 2x + 15 = 0 \text{ or } x - 4 = 0$

$\Rightarrow \quad 2x = -15 \text{ or } x = 4$

Given x is a positive number, so x can take 4 as the only value.

And let assumed mean, $a = 3$

Now put $x = 4$ and $a = 3$ in the frequency distribution table and add other columns after calculation, we get

Marks (x_i)	Frequency (f_i)	$d_i = x_i - a$	$f_i d_i$	$f_i d_i^2$
0	$x - 2 = 4 - 2$	–3	–6	18
1	$x = 4$	–2	–8	16
2	$x^2 = 4^2 = 16$	–1	–16	16
3	$(x + 1)^2 = 25$	0	0	0
4	$2x = 8$	1	8	8
5	$x + 1 = 5$	2	10	20
Total	60		–16	78

And we know standard deviation is

$$\sigma = \sqrt{\frac{\Sigma f_i d_i^2}{n} - \left(\frac{\Sigma f_i d_i}{n}\right)^2}$$

Substituting values from the above table, we get

$$\sigma = \sqrt{\frac{78}{60} - \left(\frac{-16}{60}\right)^2}$$

$$\sigma = \sqrt{1.3 - 0.07}$$

$\Rightarrow \quad \sigma = 1.10$

Hence, the standard deviation is 1.10.

Now, mean is

$$\bar{x} = A + \frac{\Sigma f_i d_i}{N} = 3 + \left(-\frac{12}{60}\right)$$

$$= 3 - \frac{1}{5} = \frac{14}{5} = 2.8$$

Hence, the mean and standard deviation of the marks are 2.8 and 1.10 respectively.

34. **Determine mean and standard deviation of first n terms of an A.P. whose first term is a and common difference is d.**

[NCERT Exemplar]

Ans. Given first n term of an A.P. whose first term is a and common difference is d.

Now, we have to find mean and standard deviation.

The given A.P. in tabular form is as shown below,

x_i	$d_i = x_i - a$	d_i^2
a	0	0
$a+b$	d	d^2
$a+2d$	$2d$	$4d^2$
$a+3b$	$3d$	$9d^2$
—	—	—
$a+(n-1)$	$(n-1)d$	$(n-1)^2d^2$

Here we have assumed a as mean.

Given the A.P. have n terms, and we know the sum of all the terms of AP can be written as

$$\Sigma x_i = \frac{n}{2}[2a+(n-1)d]$$

Now we will calculate the actual mean,

$$\bar{x} = \frac{\Sigma x_i}{n}$$

Substituting the corresponding values, we get

$$\bar{x} = \frac{\frac{n}{2}[2a+(n-1)d]}{n}$$

The above equation can be written as

$$\bar{x} = \frac{[2a+(n-1)d]}{2}$$

$$\bar{x} = a + \frac{[(n-1)d]}{2}$$

$$\bar{x} = a + \frac{(n-1)}{2}d$$

We also have,

$$\Sigma d_i = \Sigma(x_i - a) = d[1 + 2 + 3 + ... + (n-1)]$$

$$= d\left(\frac{n(n-1)}{2}\right)$$

$$\Sigma d_i^2 = \Sigma(x_i - a)^2 = d^2[1^2 + 2^2 + 3^2 + ... + (n-1)^2]$$

$$= d^2\left(\frac{n(n-1)(2n-1)}{6}\right)$$

Now we know standard deviation is given by

$$\sigma = \sqrt{\frac{\Sigma(x_i-a)^2}{n} - \left(\frac{\Sigma(x_i-a)^2}{n}\right)}$$

Substituting the corresponding values, we get

$$\sigma = \sqrt{\frac{d^2\left(\frac{n(n-1)(2n-1)}{6}\right)}{n} - \left(\frac{d\left(\frac{n(n-1)}{2}\right)}{n}\right)^2}$$

$$\sigma = \sqrt{d^2\left(\frac{n(n-1)(2n-1)}{6n}\right) - d^2\left(\frac{n^2(n-1)^2}{4n^2}\right)}$$

Cancelling the like terms, we get

$$\sigma = \sqrt{d^2\left(\frac{(n-1)(2n-1)}{6}\right) - d^2\left(\frac{(n-1)^2}{4}\right)}$$

Taking out common terms we get

$$\sigma = \sqrt{\left(\frac{d^2(n-1)}{2}\right)\left(\frac{(2n-1)}{3} - \frac{n-1}{2}\right)}$$

By taking the LCM, we get

$$\sigma = \sqrt{\left(\frac{d^2(n-1)}{2}\right)\left(\frac{2(2n-1)-3(n-1)}{6}\right)}$$

$$\sigma = \sqrt{\left(\frac{d^2(n-1)}{2}\right)\left(\frac{2(2n-1)-3(n-1)}{6}\right)}$$

$$\sigma = \sqrt{\left(\frac{d^2(n-1)}{2}\right)\left(\frac{n+1}{6}\right)}$$

$$\sigma = d\sqrt{\frac{(n^2-1)}{12}}$$

Hence, the mean and standard deviation of the given A.P. is

$$a + \frac{(n-1)}{2}d \text{ and } d\sqrt{\frac{(n^2-1)}{12}} \text{ respectively.}$$

35. Find the standard deviation of the first n natural numbers.

[NCERT Exemplar, Delhi Gov. QB 2022]

Ans. Given set of first n natural numbers:

Now, we have to find the standard deviation.

Given first n natural numbers, we can write in table as shown below

So, the sums of these are

$$\Sigma x_i = 1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}$$

And

$$\Sigma x_i^2 = 1^2 + 2^2 + 3^3 + ... + n^2 = \frac{n(n+1)(2n+1)}{6}$$

Therefore, the standard deviation can be written as,

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$$

Substituting the values we get

$$\sigma = \sqrt{\frac{\frac{n(n+1)(2n+1)}{6}}{n} - \left(\frac{\frac{n(n+1)}{2}}{n}\right)^2}$$

On simplifying

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

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

Multiplying the numerator we get

$$\sigma = \sqrt{\frac{2n^2+n+2n+1}{6} - \frac{n^2+2n+1}{4}}$$

Taking *LCM* and simplifying we get

$$\sigma = \sqrt{\frac{2(2n^2+3n+1)-3(n^2+2n+1)}{12}}$$

$$\sigma = \sqrt{\frac{4n^2+6n+2-3n^2-6n-3}{12}}$$

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

Hence, the standard deviation of the first n natural numbers is $\sqrt{\frac{n^2-1}{12}}$.

36. The mean and standard deviation of some data for the time taken to complete a test are calculated with the following results:

Number of observation = 25, mean = 18.2 seconds, standard deviation = 3.25 seconds.

Further, another set of 15 observation $x_1, x_2, ..., x_{15}$ also in seconds, is now available and we have

$$\sum_{i=1}^{15} x_i = 279 \quad \text{and} \quad \sum_{i=1}^{15} x_i^2 = 5524$$

Calculate the standard derivation based on all 40 observations. [NCERT Exemplar]

Ans. Given, number of observation = 25, mean = 18.2 seconds, standard deviation = 3.25 seconds. Another set of 15 observation $x_1, x_2, ..., x_{15}$ also in seconds, is $\sum_{i=1}^{15} x_i = 279$ and $\sum_{i=1}^{15} x_i^2 = 5524$

Now, we have to find the standard derivation based on all 40 observation.

As per the given criteria,

In first set,

Number of observation, $n_1 = 25$

Mean, $\bar{x}_1 = 18.2$

And standard deviation, $\sigma_1 = 3.25$

And

In second set,

Number of observation, $n_2 = 15$

$$\sum_{i=1}^{15} x_i = 279 \ \alpha\nu\delta \ \sum_{i=1}^{15} x_i^2 = 5524$$

For the first set we have

$$\bar{x}_1 = 18.2 = \frac{\Sigma x_i}{25}$$

$$\Sigma x_i = 25 \times 18.2 = 455$$

Therefore the standard deviation becomes,

$$\sigma_1^2 = \frac{\Sigma x_i^2}{25} - (18.2)^2$$

Substituting the values, we get

$$(3.25)^2 = \frac{\Sigma x_1^2}{25} - 331.24$$

$$\Rightarrow \quad 10.5625 + 331.24 = \frac{\Sigma x_i^2}{25}$$

On rearranging, we get

$$\Rightarrow \quad \frac{\Sigma x_i^2}{25} = 341.8025$$

On cross multiplication, we get

$\Rightarrow \quad \Sigma x_1^2 = 25 \times 341.8025 = 8545.06$

For the combined standard deviation of the 40 observation, $n = 40$

And

$\Rightarrow \quad \Sigma x_i^2 = 8545.06 + 5524 = 14069.06$

$\Rightarrow \quad \Sigma x_i = 455 + 279 = 734$

Therefore the standard deviation can be written as,

$$\sigma = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$$

Substituting the values, we get

Therefore the standard deviation can be written as,

$$\sigma = \sqrt{\frac{14069.06}{40} - \left(\frac{734}{40}\right)^2}$$

On simplifying we get

$$\sigma = \sqrt{351.7265 - (18.35)^2}$$

$$\sigma = \sqrt{351.7265 - 336.7225}$$

$$\sigma = \sqrt{15.004}$$

$$\sigma = 3.87$$

Hence, the mean standard deviation based on all 40 observations is 3.87.

37. The mean and standard deviation of a set of n_1 observation are $\bar{x}_1$ and s_1, respectively while the mean and standard deviation of another set of n_2 observations are x_2 and

s_2, respectively. Show that the standard deviation of the combined set of $(n_1 + n_2)$ observation is given by

$$\sqrt{\frac{n_1(s_1)^2 + n_2(s_2)^2}{n_1 + n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)}}$$

Ans. Given the mean and standard deviation of a set of n_1 observations are $\bar{x}_1$ and s_1, respectively while the mean and standard deviation of another set of n_2 observations are $\bar{x}_2$ and s_2, respectively.

To show that the standard deviation of the combined set of $(n_1 + n_2)$ observations is given by

$$\sqrt{\frac{n_1(s_1)^2 + n_2(s_2)^2}{n_1 + n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}}$$

As per given criteria,

For first set

Let x_i where $i = 1, 2, 3, 4, ..., n_1$

For second set

And y_j where $j = 1, 2, 3, 4, ..., n_2$

And the means are

$$\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n} x_i, \bar{x}_2 = \frac{1}{n_2}\sum_{j=1}^{n} y_j$$

Now mean of the combined series is given by

$$\bar{x} = \frac{1}{n_1 + n_2} = \left[\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} y_j\right] = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \quad ...(i)$$

And the corresponding square of standard deviation is

$$\sigma_1^2 = \frac{1}{n_1}\sum_{i=1}^{n}(x_i - \bar{x}), \sigma_2^2 = \frac{1}{n_2}\sum_{j=1}^{n}(y_j - \bar{x})^2$$

Therefore, square of standard deviation becomes,

$$\sigma^2 = \sigma_1^2 + \sigma_2^2$$

$$= \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{j=1}^{n}(y_j - \bar{x})^2\right] \quad ...(ii)$$

Now,

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

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \bar{x}_j)^2 + n_1(\bar{x}_j - \bar{x})^2 + 2(\bar{x}_j - \bar{x})^2 \sum_{i=1}^{n}(x_i - \bar{x}_j)^2$$

But the algebraic sum of the deviation of values of first series from their mean is zero.

$$\sum_{i=1}^{n}(x_i - x_j)^2 = 0$$

Also,

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = n_1 s_1^2 + n_1\left(\bar{x}_1 - \bar{x}\right)^2 \quad ...(iii)$$

But

$$d_1 = \bar{x}_1 - \bar{x}$$

Substituting value from equation (i), we get

$$\bar{x}_1 - \bar{x} = \bar{x}_1 - \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$\bar{x}_1 - \bar{x} = \frac{(\bar{x}_1)(n_1 + n_2) - (n_1\bar{x}_1 + n_2\bar{x}_2)}{n_1 + n_2}$$

$$\bar{x}_1 - \bar{x} = \frac{(n_1\bar{x}_1 + n_2\bar{x}_1) - (n_1\bar{x}_1 + n_2\bar{x}_2)}{n_1 + n_2}$$

$$\bar{x}_1 - \bar{x} = \frac{(n_2\bar{x}_1) - (n_2\bar{x}_2)}{n_1 + n_2}$$

Substituting this value in equation (iii), we get

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = n_1 s_1^2 + n_1\left(\frac{n_2(\bar{x}_1 - \bar{x}_2)}{n_1 + n_2}\right)^2$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = n_1 s_1^2 + \frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2} \quad ...(iv)$$

Similarly, we have

$$\sum_{j=1}^{n}(y_i - \bar{x})^2 = \sum_{j=1}^{n}(y_j - \bar{x}i + \bar{x}_i - \bar{x})^2$$

$$\sum_{j=1}^{n}(y_j - \bar{x})^2 = \sum_{j=1}^{n}(y_j - \bar{x}_j)^2 + n_2(\bar{x}_j - \bar{x})^2 + 2(\bar{x}_j - \bar{x})\sum_{j=1}^{n}(y_j - \bar{x}_j)^2$$

But the algebraic sum of the deviation of values of the second series from their mean is zero.

$$\sum_{j=1}^{n}(y_j - \bar{x}_i)^2 = 0$$

Also,

$$\sum_{j=1}^{n}(y_j - \bar{x})^2 = n_2 s_2^2 + n_2\left(\bar{x}_2 - \bar{x}\right)^2 \quad ...(v)$$

But $d_2 = \bar{x}_2 - \bar{x}$

Substituting value from equation (i), we get

$$x_2 - \bar{x} = \bar{x}_2 - \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$x_2 - \bar{x} = \frac{(\bar{x}_2)(n_1 + n_2) - (n_1\bar{x}_1 + n_2\bar{x}_2)}{n_1 + n_2}$$

$$\bar{x}_2 - \bar{x} = \frac{(n_1\bar{x}_2 + n_2\bar{x}_2) - (n_1\bar{x}_1 + n_2\bar{x}_2)}{n_1 + n_2}$$

$$\bar{x}_2 - \bar{x} = \frac{(n_1\bar{x}_2) - (n_1\bar{x}_2)}{n_1 + n_2}$$

$$\bar{x}_2 - \bar{x} = \frac{n_1(\bar{x}_2 - \bar{x}_1)}{n_1 + n_2}$$

Substituting this value in equation (v), we get

$$\sum_{j=1}^{n}(y_j - \bar{x})^2 = n_2 s_2^2 + n_2\left(\frac{n_1(\bar{x}_2 - \bar{x}_1)}{n_1 + n_2}\right)^2$$

$$\sum_{j=1}^{n}(y_j - \bar{x})^2 = n_2 s_2^2 + \frac{n_2 n_1^2(\bar{x}_2 - \bar{x}_1)^2}{(n_1 + n_2)^2} \quad \text{...(vi)}$$

Substituting equation (iv) and (vi) i nequation (ii), we get

$$\sigma^2 = \sigma_1^2 + \sigma_2^2$$

$$= \frac{1}{n_1 + n_2}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_j - \bar{x})^2\right]$$

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1 s_1^2 + \frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2} + n_2 s_2^2 + \frac{n_2 n_1^2(\bar{x}_2 - \bar{x}_1)^2}{(n_1 + n_2)^2}\right]$$

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2} + \frac{n_2 n_1^2(-(\bar{x}_1 - \bar{x}_2))^2}{(n_1 + n_2)^2}\right]$$

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2} + \frac{n_2 n_1^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}\right]$$

$$\sigma^2 = \frac{1}{n_1 + n_2}\left[n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}(n_2 + n_1)\right]$$

$$\sigma^2 = \left[\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \frac{\frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}(n_2 + n_1)}{n_1 + n_2}\right]$$

$$\sigma^2 = \left[\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}\right]$$

So the combined standard deviation

$$S.D.\ (\sigma) = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1 + n_2)^2}}$$

Hence, proved.