STATISTICS

CHAPTER 13

Data and its types

A group of information that represents the qualitative or quantitative attributes of a variable or set of variables is called data.

There are two types of data. These are :

(i) Ungrouped data : In an ungrouped data, data is listed in series e.g., 1, 4, 9, 16, 25, etc., this is also called in divided data.

(ii) Grouped data-It is of two types:

(a) Discrete data : In this type, data is presented in such a way that exact measurements of items are clearly shown.

(b) Continuous data : In this type, data is arranged in groups or classes but they are not exactly measurable, they form a continuous series.

➢ **Measures of Central tendency**

A certain value that represent the whole data and signifying its characteristics is called measure of central tendency mean, median and mode are the measures of central tendency.

- **Mean**

Mean of ungrouped data : The mean of n observations $x_1, x_2, x_3, ..., x_n$ is given by

$$\text{Mean } (\bar{x}) = \frac{x_1 + x_2 + x_3 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Mean of grouped data : Let $x_1, x_2, ..., x_n$ be in observations with respective frequencies $f_1, f_2, ..., f_n$.

Then, Mean

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

➢ **Limit of the class**

The starting and ending values of each class are called lower and upper limits.

➢ **Class Interval**

The difference between upper and lower boundary of a class is called class interval or size of the class.

➢ **Primary and secondary data**

The data collected by the investigator himself is known as the primary data, while the data collected by a person, other than the investigator, is known as secondary data.

➢ **Frequency**

The number of times an observation occurs in the given data, is called the frequency of the observation.

Measure of Dispersion

The measures of central tendency are not sufficient to give complete information about given data. Variability is another factor which is required to be studied under statistics. The single number that describes variability is called measure of dispersion. It is the measure of scattering of the data about some central tendency.

There are following measures of dispersion

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

Range

Range is the difference of maximum and minimum value of data

Range = maximum value – minimum value.

for eg given marks of sameer and Suresh as follows–

Sameer	=	79, 62, 40, 5
Suresh	=	60, 45, 52, 42
For Sameer, Range	=	79 – 5 = 74
For Suresh, Range	=	60 – 42 = 18

Thus, range of Sameer > range of Suresh.

So, the scores are scattered or dispersed in case of Sameer while for Suresh, these are close to each other. The range of data gives us a rough idea of Variability or Scatter but does not tell about the dispersion of the data from the measure of central tendency.

Mean Deviation

Mean deviation is an important measure of deviation, which depend upon the deviations of the observations from a central tendency. It is defined as the arithmetic mean of the absolute deviations of all the values taken from a central value or a fixed number a. The mean divination from 'a' is denoted by MD (a) and is defined by–

Statistics

Trace the Mind Map ‣ First Level

Measures of Dispersion

The degree to which numerical data tend to spread about an average value is called the dispersion of the data. The four dispersions are: Range, Mean deviation, Standard deviation and Variance.

Range

The difference between the highest and the lowest element of a data called its range.

$range = x_{max} - x_{min}$

e.g: Find the range of the given data:

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

Here, $x_{max} = 17$ and $x_{min} = 4$

Range = 17–4 = 13

Mean Deviation

M.D.

Mean Deviation (*M.D.*) for ungrouped data

- *M.D.* about mean i.e., $M.D.(\bar{x}) = \frac{\sum |x_i - \bar{x}|}{n}$, where $\bar{x}$ is mean and n = no. of terms
- *M.D.* about median i.e., $M.D.(M) = \frac{\sum |x_i - M|}{n}$, where M is median and n = no. of terms

Mean Deviation (*M.D*) for grouped data

M.D. about mean i.e., $M.D.(\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{N}$

M.D. about median i.e., $M.D.(M) = \frac{1}{N} \sum f_i |x_i - M|$

where $N = \sum f_i$

Median class

Median class is the class interval whose *C.F.* is just greater than or equal to $\frac{N}{2}$, *f*, ***h*** **and** ***C*** **are** respectively the lower limit, the frequency, width of the median class and *C* the cumulative frequency (*C.F.*) of class just preceding to median class.

Variance & Standard Deviation of a Continuous frequency distribution

Variance of a continuous frequency distribution

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

Standard deviation of a continuous frequency distribution

$S.D.\ (\sigma) = \frac{1}{N} \sqrt{N \sum f_i x_i^2 - (\sum f_i x_i)^2}$

Step Deviation or Shortcut Method

- $M.D.(\bar{x}) = \frac{1}{N} \sum f_i |x_i - \bar{x}|$

where $\bar{x} = a + \frac{\sum_{i=1}^{n} f_i d_i}{N} \times h$

and $d_i = \frac{x_i - a}{h}$

Here, a = assumed mean

h = common factor

N = sum of frequencies

- $M.D.(M) = \frac{1}{N} \sum f_i |x_i - M|$

where, M(median) $= l + \frac{\frac{N}{2} - C}{f} \times h$

Shortcut method to find variance and standard deviation

Variance

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

Standard Deviation

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

where $y_i = \frac{x_i - A}{h}$

$N = \sum_{i=1}^{n} f_i$

A = assumed mean

h = width of class interval

Variance & Standard Deviation of a discrete frequency distribution

Variance of a discrete frequency distribution

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

Standard Deviation (*S.D.*) of a discrete frequency distribution

$S.D.\ (\sigma) = \sqrt{\frac{1}{N} \sum f_i (x_i - \bar{x})^2}$

where, $N = \sum_{i=1}^{n} f_i$ and $\bar{x}$ = mean

Variance and Standard Deviation for ungrouped data

Variance of ungrouped data

Var $(\sigma^2) = \frac{1}{n} \sum (x_i - \bar{x})^2$

Standard Deviation (*S.D.*) for un grouped data:

$S.D.\ (\sigma) = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$

$$M.D.\ (a) = \frac{\text{Sum of absolute values of deviations from 'a'}}{\text{Number of observations}}$$

➢ **Mean deviation for ungrouped data**

Let $x, x_2, ..., x_n$ be n observations. Then, mean deviation about means $(\bar{x})$ or median (M) can be found by the formula

$$\frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n} \text{ or } \frac{\sum_{i=1}^{n}|x_i - M|}{n}$$, where n is the number of observations

➢ **Mean deviation for grouped data**

(i) For discrete frequency distribution : Let the data have 'n' district values $x_1, x_2, ..., x_n$ and their corresponding frequencies are $f_1, f_2 ..., x_n$ respectively. Then, this data can be represented in the tabular form as

x_i	x_1	x_2	x_3	... x_n
f_i	f_1	f_2	f_3	... f_n

and is called discrete frequency distribution. There, mean deviation about mean or median is given by

$$\frac{\sum_{i=1}^{n} f_i |x_i - A|}{N}, \text{where } N = \sum_{i=1}^{n} f_i = \text{total frequency},$$

and A = mean or median.

(ii) For continuous frequency distribution : A continuous frequency distribution is a series in which the data is classified into different class intervals without gaps along with their respective frequencies.

Mean deviation about mean $(\bar{x})$, i.e.,

$$MD\ (\bar{x}) = \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - \bar{x}|$$, where 'x_i' s' are the mid-point of the intervals and

Also, mean $(\bar{x}) = \frac{1}{N}\sum_{i=1}^{n} f_i x_i$

Mean deviation about median (M), i.e.,

$$MD\ (M) = \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - M|$$, where x_i' s are the mid-points of the intervals and $\sum_{i=1}^{n} f_i = N$. Also,

median $M = l + \frac{\frac{N}{2} - C_f}{f} \times h$, where l, f, h and C_f are lower limit, the frequency, the width of median class and cumulative frequency of class just preceding the median class.

➢ **Shortcut (Step-deviation) Method for calculating the Mean deviation about Mean :** This method is used to manage large data. In this method, we take an assumed mean, which is in the middle or just close to it, in the data. we denote the new variable by and is defined by

$u_i = \frac{x_i - a}{h}$, where a is the assumed mean and h is the common factor or length of class interval the mean $\bar{x}$ by step deviation method is given by

$$\bar{x} = a + \frac{\sum_{i=1}^{n} f_i u_i}{N} \times h$$

➢ **Limitations of Mean Deviation**

(i) If the data is more scattered or the degree of variability is very high, then the median is not a valid representative. Thus, the mean deviation about the median is not fully relied.

(ii) The sum of the deviations from the mean is more than the sum of the deviations from the median. Therefore, the mean deviation about mean is not very scientific.

(iii) The mean deviation is calculated on the basis of absolute values of the deviations and so cannot be subjected to further algebraic treatment. Sometimes, it gives unsatisfactory results.

➢ **Extra Information**

- The cumulative frequency of a class is the frequency obtained by adding the frequencies of all the classes preceding the given class and the frequency of given class.
- Mean deviation may be obtained from any measure of central tendency. However, we study deviations from mean and median in this chapter.

Variance

The mean of squares of deviations from mean is called the variance and it is denoted by the symbol 'σ^2'.

The variance of 'n' observations $x_1, x_2, ..., x_n$ is given by :

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

➢ **Significance of deviation**

(i) If the deviation is zero, it means there is no deviation at all and all observations are equal to mean.

(ii) If deviation is small, this indicates that the observations are close to the mean.

(iii) If the deviation is large, there is a high degree of dispersion of the observation from the mean.

Standard Deviation

Standard deviation is the square root of the arithmetic mean of the squares of deviations from mean and it is denoted by the symbol σ.

or

The square root of variance, is called standard deviation i.e., $\sqrt{\sigma^2} = \sigma$. It is also known as root mean square deviation.

➢ **Variance and Standard deviation of ungrouped data**

Variance of n observations $x_1, x_2, \ldots, x_n$ is given by

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

and,Standard Deviation, $\sigma = \sqrt{\text{Variance}} = \sqrt{\sigma^2}$

$$\therefore \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \text{ or } \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}$$

$$= \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}$$

➢ **Variance and Standard deviation of Grouped data**

(i) For discrete frequency distribution

Let the discrete frequency distribution be $x_1, x_2, x_3, \ldots, x_n$ and $f_1, f_2, f_3, \ldots, f_n$. Then by direct method :

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

$$= \frac{1}{N}\sum f_i x_i^2 - \left(\frac{\sum f_i x_i}{N}\right)^2$$

andStandard deviation (σ)

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

$$= \frac{1}{N}\sqrt{N\sum f_i x_i^2 - \left(\sum f_i x_i\right)^2}$$

where $N = \sum_{i=1}^{n} fi$

By short cut method, variance (σ^2)

$$= \frac{1}{N}\sum_{i=1}^{n} f_i d_i^2 - \left(\frac{\sum f_i d_i}{N}\right)^2$$

andstandard deviation (σ)

$$= \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i d_i^2 - \left(\frac{\sum f_i d_i}{N}\right)^2}$$

where $d_i = x_i - a$, deviation from assumed mean and a = assumed mean.

(ii) For Continuous frequency distribution

Direct Method : If there is a frequency distribution of n classes and each class defined by its mid-point x_i with corresponding frequency f_i, then the variance and standard deviation are :

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

$$\text{andStandard deviation } (\sigma) = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}$$

$$\text{or,} \quad \sigma^2 = \frac{1}{N^2}\left[N\sum f_i x_i^2 - \left(\sum f_i x_i\right)^2\right]$$

$$\text{and} \quad \sigma = \frac{1}{N^2}\sqrt{N\sum f_i x_i^2 - \left(\sum f_i x_i\right)^2}$$

Step-Deviation (short-cut method) : Sometimes the values of mid points x_i of different classes in a continuous distribution are large and so the calculation of mean and variance becomes tedious and time consuming. For this, we use the step-deviation method. Here,

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

and Standard deviation (σ)

$$= h\sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i u_i^2 - \left(\frac{\sum_{i=1}^{n} f_i u_i}{N}\right)^2}$$

where, $u_i = \frac{x_i - a}{h}$, a = assumed mean and

h = width of class interval.

➢ **Extra Information**

- A characteristics that varies in magnitude from observation to observation e.g., weight, height, income, age, etc. are variables.
- Due to the limitations of mean deviation, some other method is required for measure of dispersion. Standard deviation is such a measure of dispersion.
- The ratio *S.D.* (σ) and the *A.M.* $(\bar{x})$ is called the coefficient of standard deviation $\left(\frac{\sigma}{\bar{x}}\right)$.
- The percentage form of coefficient of *S.D.* i.e., $\left(\frac{\sigma}{\bar{x}}\right)$ is called coefficient of variation.
- The distribution for which the coefficient of variation is less is called more consistent.

- Standard deviation of first n natural numbers is $\sqrt{\frac{n^2-1}{12}}$.
- Standard deviation is independent of change of origin, but it depends of change of scale.

NCERT CORNER

EXERCISE - 13.1

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

4, 7, 8, 9, 10, 12, 13, 17 **[NCERT Ex. 13.1. Q. 1. Page 270]**

Sol. Given data is : 4, 7, 8, 9, 10, 12, 13, 17.

$$\therefore \quad \text{Arithmetic mean, } \bar{x} = \frac{4+7+8+9+10+12+13+17}{8} = \frac{80}{8} = 10$$

$$\sum_{i=1}^{8}|x_i - \bar{x}| = 6+3+2+1+0+2+3+7 = 24$$

$$\therefore \quad \text{Mean deviation about mean, } M.D.\left(\bar{x}\right) = \frac{\sum|x_i-\bar{x}|}{n} = \frac{24}{8} = 3$$

Q. 2. Find the mean deviation about the mean for the data :

38, 70, 48, 40, 42, 55, 63, 46, 54, 44. **[NCERT Ex. 13.1. Q. 2. Page 270]**

Sol. Here,

$$\text{mean, } \bar{x} = \frac{38+70+48+40+42+55+63+46+54+44}{10} = \frac{500}{10} = 50$$

$$\sum_{i=1}^{10}|x_i - \bar{x}| = 12+20+2+10+8+5+13+4+4+6 = 84$$

$\therefore$ Mean deviation about mean $\left(\bar{x}\right)$ i.e.,

$$M.D.\left(\bar{x}\right) = \frac{\sum_{i=1}^{10}|x_i-\bar{x}|}{n} = \frac{84}{10} = 8.4.$$

Q. 3. Find the mean deviation about the median for the following data :

13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17. **[NCERT Ex. 13.1. Q. 3. Page 270]**

Sol. Let us first arrange the data in ascending order as :

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

Here, $n = 12$, which is even.

Median observations are $\frac{12}{2}$ and $\frac{12}{2}+1$ i.e., 6th and 7th observations.

$$\textbf{6}^{\textbf{th}} \textbf{ observation} = 13, 7^{th} \text{ observation} = 14$$

$$\therefore \quad \text{Median, } M = \frac{13+14}{2} = 13.5.$$

$$\Sigma |x_i - M| = 3.5 + 2.5 + 2.5 + 1.5 + 0.5 + 0.5 + 0.5 + 2.5 + 2.5 + 3.5 + 3.5 + 4.5$$
$$= 28$$

∴ **Mean deviation about the median** M**,**

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

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

Q. 4. Find the mean deviation about the median for the data :

36, 72, 46, 42, 60, 45, 53, 46, 51, 49. **[NCERT Ex. 13.1. Q. 4. Page 270]**

Sol. Let us arrange the data in ascending order as :

36, 42, 45, 46, 46, 49, 51, 53, 60, 72

∴ No. of observations, $n = 10$, which is even.

$$\therefore \frac{10^{th}}{2} \text{ observation} = 5^{th} \text{ observation} = 46$$

$$\text{and } \left(\frac{10}{2}+1\right) \text{th observation} = 6^{th} \text{ observation} = 49.$$

$$\therefore \quad \text{Median} = \frac{1}{2}(46+49) = 47.5$$

$$\text{and} \quad \sum_{i=1}^{10} |x_i - M| = 11.5 + 5.5 + 2.5 + 1.5 + 1.5 + 1.5 + 3.5 + 5.5 + 12.5 + 24.5$$

$$= 70$$

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

$$= \frac{70}{10} = 7$$

Q. 5. Find the mean deviation about the mean for the data : **[NCERT Ex. 13.1. Q. 5. Page 270]**

x_i	**5**	**10**	**15**	**20**	**25**
f_i	**7**	**4**	**6**	**3**	**5**

Sol.

x_i	f_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
5	7	35	9	63
10	4	40	4	16
15	6	90	1	6
20	3	60	6	18
25	5	125	11	55
Total	25	350		158

$$\therefore \quad \bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{350}{25} = 14$$

$$\text{and} \quad \textbf{M.D. } (\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i} = \frac{158}{25} = 6.32$$

Q. 6. Find the mean deviation about mean for the data : **[NCERT Ex. 13.1. Q. 6. Page 270]**

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

Sol.

x_i	f_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
10	4	40	40	160
30	24	720	20	480
50	28	1400	0	0
70	16	1120	20	320
90	8	720	40	320
Total	80	4000		1280

$\therefore$ Mean $\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{4,000}{80} = 50$

$\therefore$ Mean deviation about mean $(\bar{x})$ i.e.,

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

$$= \frac{1,280}{80}$$

$$= 16$$

Q. 7. Find the mean deviation about the median for the following data: **[NCERT Ex. 13.1. Q. 7. Page 270]**

x_i	5	7	9	10	12	15
f_i	8	6	2	2	2	6

Sol.

x_i	f_i	*c.f.*	$\|x_i - 7\|$	$f_i\|x_i - 7\|$
5	8	8	2	16
7	6	14	0	0
9	2	16	2	4
10	2	18	3	6
12	2	20	5	10
15	6	26	8	48
	$\Sigma f = 26$		20	$\Sigma f_i \|x_i - 7\| = 84$

Here $\frac{N}{2} = \frac{26}{2} = 13$

The *c. f.* just greater than 13 is 14 and the corresponding value of *x* is 7.

$\therefore$ Median = 7

$\therefore$ Mean deviation about median $= \frac{\Sigma f_i |x_i - M|}{\Sigma f}$

$$= \frac{84}{26}$$

$$= 3.23$$

Q. 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 Ex. 13.1. Q. 8. Page 270]

Sol. Let us find cumulative frequency of the given distribution.

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

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

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

$$= 15\text{th observation}$$

Since, 15 occurs in the cumulative frequency 21, so the corresponding observation is 30.

Hence, $M = 30$

$\therefore$ Mean deviation about median $= \dfrac{\Sigma f_i \lvert x_i - M \rvert}{\Sigma f_i}$

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

$$= 5.1$$

Q. 9. Find the mean deviation about the mean for the following data:

Income per day	**0 – 100**	**100 – 200**	**200 – 300**	**300 – 400**	**400 – 500**	**500 – 600**	**600 – 700**	**700 – 800**
Number of Persons	**4**	**8**	**9**	**10**	**7**	**5**	**4**	**3**

[NCERT Ex. 13.1. Q. 9. Page 271]

Sol. Let us calculate Mean Deviation.

Classes	f_i	x_i	$d_i = \dfrac{x_i - 350}{100}$	$f_i d_i$	$\lvert x_i - 358 \rvert$	$f_i \lvert x_i - 358 \rvert$
0 – 100	4	50	– 3	– 12	308	1232
100 – 200	8	150	– 2	– 16	208	1664
200 – 300	9	250	– 1	– 9	108	972
300 – 400	10	350	0	0	8	80
400 – 500	7	450	1	7	92	644
500 – 600	5	550	2	10	192	960
600 – 700	4	650	3	12	292	1168
700 – 800	3	750	4	12	392	1176
	$\Sigma f_i = 50$		**4**	**$\Sigma f_i d_i = 4$**		**$\Sigma f_i \lvert x_i - 358 \rvert = 7896$**

Here, $\text{Mean } (\bar{x}) = \text{assumed mean} + \dfrac{\sum f_i d_i}{\sum f_i} \times h$

$$= 350 + \frac{4}{50} \times 100$$

$$= 358$$

$\therefore$ Mean deviation about mean $(\bar{x})$,

$$M.D.\ (\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i} = \frac{7896}{50} = 157.92$$

Q. 10. Find the mean deviation about mean for the data :

Height (cm)	**90 – 105**	**105 – 115**	**115 – 125**	**125 – 135**	**135 – 145**	**145 – 155**
No. of boys	**9**	**13**	**26**	**30**	**12**	**10**

[NCERT Ex. 13.1. Q. 10. Page 271]

Sol.

Class Interval	**Mid Value (x_i)**	$d_i = \frac{x_i - 130}{10}$	f_i	$f_i d_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
95 – 105	100	– 3	9	– 27	25.3	227.7
105 – 115	110	– 2	13	– 26	15.3	198.9
115 – 125	120	– 1	26	– 26	5.3	137.8
125 – 135	130	0	30	0	4.7	141
135 – 145	140	1	12	12	14.7	176.4
145 – 155	150	2	10	20	24.7	247
			100	**– 47**		**1128.8**

Let the assumed mean be 130.

Then,

$$\bar{x} = 130 + \frac{\sum f_i d_i}{\sum f_i} \times 10 = 130 - \frac{47}{100} \times 10 = 125.3$$

$\therefore$ Mean deviation about mean $(\bar{x})$ i.e.,

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

$$= \frac{1,128.8}{100}$$

$$= 11.288$$

Q. 11. Find the mean deviation about median for the following data :

Marks	**0 – 10**	**10 – 20**	**20 – 30**	**30 – 40**	**40 – 50**	**50 – 60**
Number of girls	**6**	**8**	**14**	**16**	**4**	**2**

[NCERT Ex. 13.1. Q. 11. Page 271]

Sol. Let us find cumulative frequencies and $\Sigma f_i |x_i - M|$ as follows :

Marks	**Mid Value (x_i)**	f_i	*c.f.*	$\|x_i - M\|$	$f_i \|x_i - M\|$
0 – 10	5	6	6	22.86	137.16
10 – 20	15	8	14	12.86	102.88
20 – 30	25	14	28	2.86	40.04
30 – 40	35	16	44	7.14	114.24
40 – 50	45	4	48	17.14	68.56
50 – 60	55	2	50	27.14	54.28
		N = 50			**517.16**

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

Here, the *c. f.* just greater than 25 is 28, so median class is (20–30).

$$\text{Median} = 20 + \frac{25-14}{14}\times 10$$

$$= = 20 + \frac{110}{14}$$

$$= 20 + 7.86 = 27.86$$

$\therefore$ Mean deviation about median $= \dfrac{\sum_{i=1}^{n} f_i\,|x_i - M|}{N}$

$$= \frac{517.16}{50}$$

$$= 10.34$$

Q. 12. Calculate the mean deviation about median age for the age distribution of 100 persons given below :
[NCERT Ex. 13.1. Q. 12. Page 271]

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

Sol.

Age	Class	Mid Value (x_i)	f_i	*c.f.*	$\|x_i - M\|$	$f_i\|x_i - M\|$
16 – 20	15.5 – 20.5	18	5	5	20	100
21 – 25	20.5 – 25.5	23	6	11	15	90
26 – 30	25.5 – 30.5	28	12	23	10	120
31 – 35	30.5 – 35.5	33	14	37	5	70
36 – 40	35.5 – 40.5	38	26	63	0	0
41 – 45	40.5 – 45.5	43	12	75	5	60
46 – 50	45.5 – 50.5	48	16	91	10	160
51 – 55	50.5 – 55.5	53	9	100	15	135
			N = 100			**735**

Here $\dfrac{N}{2} = \dfrac{100}{2} = 50,$

It occurs in the cumulative frequency 63. Here, the median class is 35.5–40.5.

$\therefore$ Median, $M = l + \dfrac{\frac{N}{2} - c}{f_m}\times h$

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

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

$$= 35.5 + 2.5$$

$\Rightarrow$ $M = 38$

Now, mean deviation about the median

$$\text{M. S. }(M) = \frac{\sum_{i=1}^{8} f_i\,|x_i - M|}{N}$$

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

$$= 7.35$$

EXERCISE - 13.2

Q. 1. Find the mean and variance for the for the following data :

6, 7, 10, 12, 13, 4, 8, 12.

[NCERT Ex. 13.2. Q. 1. Page 281]

Sol. Here,

$$\text{Mean, } \bar{x} = \frac{6+7+10+12+13+4+8+12}{8} = \frac{72}{8} = 9$$

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

$$\sum (x_i - \bar{x})^2 = 9 + 4 + 1 + 9 + 16 + 25 + 1 + 9 = 74$$

$$\therefore \text{ Variance} = \frac{\sum (x_i - \bar{x})^2}{\sum f_i} = \frac{74}{8} = 9.25$$

Q. 2. Find the mean and variance for first n natural numbers. [NCERT Ex. 13.2. Q. 2. Page 281]

Sol. The first n natural numbers are 1, 2, 3, ..., n.

$$\therefore \quad \text{Mean, } \bar{x} = \frac{1+2+3+...+n}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$$

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

$$= \frac{1}{n}\left[\sum x_i^2 - 2\bar{x}x_i + \bar{x}^2\right]$$

$$= \frac{\sum x_i^2}{n} - \bar{x}\frac{\sum x_i}{n} + \frac{\bar{x}\cdot n}{n}$$

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

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

[Since, frequency of each variable is one]

$$\Sigma x_i^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\therefore \quad \text{Variance} = \frac{n(n+1)(2n+1)}{6n} - \left(\frac{n+1}{2}\right)^2$$

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

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

$$= \frac{(n+1)(n-1)}{12}$$

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

Q. 3. Find the mean and variance for the first 10 multiples of 3. [NCERT Ex. 13.2. Q. 3. Page 281]

Sol. First 10 multiples of 3 are :

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

x_i	$y_i = \frac{x_i - 15}{3}$	y_i^2
3	– 4	16
6	– 3	9
9	– 2	4
12	– 1	1
15	0	0
18	1	1
21	2	4
24	3	9
27	4	10
30	5	25
Total	**5**	**85**

$$\textbf{Mean, } \bar{x} = a + \frac{\sum y_i}{n} \times h$$

$$= 15 + \frac{5}{10} \times 3$$

$$= 15 + 1.5$$

$$= 16.5$$

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

$$= \frac{9}{100}(850 - 25)$$

$$= 74.25$$

$\therefore$ Variance = 74.25 and mean = 16.5

Q. 4. Find the mean and standard deviation for the following data : [NCERT Ex. 13.2. Q. 4. Page 281]

x_i	6	10	14	18	24	28	30
f_i	2	4	7	12	8	4	3

Sol.

x_i	f_i	$f_i x_i$	$f_i x_i^2$
6	2	12	72
10	4	40	400
14	7	98	1372
18	12	216	3888
24	8	192	4608
28	4	112	3136
30	3	90	2700

130	$\Sigma f_i = 40$	$\Sigma f_i x_i = 760$	$\Sigma f_i x_i^2 = 16176$

$\therefore$ Mean $= \dfrac{\sum f_i x_i}{\sum f_i} = \dfrac{760}{40} = 19$

and Standard deviation,

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

$$= \sqrt{\frac{16,176}{40} - \left(\frac{760}{40}\right)^2}$$

$$= \sqrt{404.4 - 361}$$

$$= \sqrt{43.4}$$

$$= 6.59$$

$\therefore$ Mean $= 19$, $S.D. = 6.59$.

Q. 5. Find the mean and standard deviation for the following data : **[NCERT Ex. 13.2. Q. 5. Page 281]**

x_i	**92**	**93**	**97**	**98**	**102**	**104**	**109**
f_i	**3**	**2**	**3**	**2**	**6**	**3**	**3**

Sol.

x_i	f_i	$d = x_i - 98$	$f_i d_i$	$f_i d_i^2$
92	3	– 6	– 18	108
93	2	– 5	– 10	50
97	3	– 1	– 3	3
98	2	0	0	0
102	6	4	24	96
104	3	6	18	108
109	3	11	33	363
N = 22			$\Sigma f_i d_i = 44$	$\Sigma f_i d_i^2 = 728$

$\therefore$ Mean $= a + \dfrac{\sum f_i d_i}{N} = 98 + \dfrac{44}{22} = 100$

and Standard deviation, $S.D. = \sqrt{\dfrac{\sum f_i d_i^2}{N} - \left(\dfrac{\sum f_i d_i}{N}\right)^2}$

$$= \sqrt{\frac{728}{22} - \left(\frac{44}{22}\right)^2}$$

$$= \sqrt{33.0909 - 4}$$

$$= \sqrt{29.0909}$$

$$= 5.09$$

$\therefore$ Mean $= 100$, $S.D. = 5.09$

Q. 6. Find the mean and standard deviation using short-cut method. **[NCERT Ex. 13.2. Q. 6. Page 281]**

x_i	**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**
f_i	**2**	**1**	**12**	**29**	**25**	**12**	**10**	**4**	**5**

Sol. Let $y_i = \dfrac{x_i - a}{i} = x_i - 64$ [$\because$ $i = 1$ and 'a' is assumed to be 64]

x_i	f_i	$y_i = x_i - 64$	$f_i y_i$	y_i^2	$f_i y_i^2$
60	2	– 4	– 8	16	32
61	1	– 3	– 3	9	9
62	12	– 2	– 24	4	48
63	29	– 1	– 29	1	29
64	25	0	0	0	0
65	12	1	12	1	12
66	10	2	20	4	40

67	4	3	12	9	36
68	5	4	20	16	80
Total	**100**		**0**	**60**	**286**

Here, Mean, $\bar{x} = 64 + \frac{\sum f_i y_i}{\sum f_i}$

$= 64 + 0$

$= 64$

and Variance, $\sigma^2 = \frac{1}{\sum f_i}\left[\sum f_i y_i^2 - n\bar{y}^2\right]$

$= \frac{1}{100}[286] = 2.86$

$\therefore$ $\sigma = \sqrt{2.86} = 1.69$

Q. 7. Find the mean and variance for the following frequency distribution. **[NCERT Ex. 13.2. Q. 7. Page 281]**

Classes	**0 – 30**	**30 – 60**	**60 – 90**	**90 – 120**	**120 – 150**	**150 – 180**	**180 – 210**
Frequencies	**2**	**3**	**5**	**10**	**3**	**5**	**2**

Sol.

Classes	**Mid Value** $x_i x_i$	**Frequency** (f_i)	$y_i = \frac{x_i - 105}{30}$	$f_i y_i$	y_i^2	$f_i y_i^2$
0 – 30	15	2	– 3	– 6	9	18
30 – 60	45	3	– 2	– 6	4	12
60 – 90	75	5	– 1	– 5	1	5
90 – 120	105	10	0	0	0	0
120 – 150	135	3	1	3	1	3
150 – 180	165	5	2	10	4	20
180 – 210	195	2	3	6	9	18
Total		**30**		**2**		**76**

Here, $y_i = \frac{x_i - 9}{h} = \frac{x_i - 105}{30}$

$\bar{x} = a + i\bar{y}$

$= 105 + 30 \times \frac{\sum f_i y_i}{\sum f_i}$

$= 105 + 30 \times \frac{2}{30}$

$= 107$

and, Variance, $\sigma^2 = \frac{h^2}{\sum f_i}\left[\sum f_i y_i^2 - n\bar{y}^2\right]$ $\quad h = (l_2 - l_1)$, $h = (30 - 0) = 30$

$= \frac{30^2}{30}\left[76 - 30 \times \frac{1}{225}\right]$

$= 30\,[76 - 0.13]$

$= 2276$

$\therefore$ Mean = 107 and variance = 2276.

Q. 8. Find the mean and variance for the following frequency distribution. **[NCERT Ex. 13.2. Q. 8. Page 282]**

Classes	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50
Frequencies	**5**	**8**	**15**	**16**	**6**

Sol.

Classes	Mid Value f_i	Frequency (f_i)	$y_i = \frac{x_i - 25}{10}$	$f_i y_i$	$f_i y_i^2$	$f_i y_i^2$
0 – 10	5	5	– 2	– 10	4	20
10 – 20	15	8	– 1	– 8	1	8
20 – 30	25	15	0	0	0	0
30 – 40	35	16	1	16	1	16
40 – 50	45	6	2	12	4	27
Total		**50**		**10**		**68**

Here, $$y_i = \frac{x_i - a}{i} = \frac{x_i - 25}{10}$$ [$\because a = 25, i = 10$]

$\therefore$
$$\bar{x} = a + i\bar{y}$$
$$= 25 + 10 \times \frac{\sum f_i y_i}{\sum f_i}$$
$$= 25 + 10 \times \frac{10}{50}$$
$$= 27$$

Variance, $$\sigma^2 = \frac{h^2}{\sum f_i}\left[\sum f_i y_i^2 - n\bar{y}^2\right]$$
$$= \frac{10^2}{50}\left[68 - 50 \times (0.2)^2\right]$$
$$= \frac{100}{50}[68 - 50 \times 0.04]$$
$$= 2\,(68 - 2)$$
$$= 132$$

$\therefore$ Mean = 27 and variance = 132.

Q. 9. Find the mean, variance and standard deviation using short cut method. **[NCERT Ex. 13.2. Q. 9. Page 282]**

Height (in cm)	70 – 75	75 – 80	80 – 85	85 – 90	90 – 95	95 – 100	100 – 105	105 – 110	110 – 115
No. of Children	**3**	**4**	**7**	**7**	**15**	**9**	**6**	**6**	**3**

Sol.

Classes	Mid Value x_i	Frequency (f_i)	$y_i = \frac{x_i - 92.5}{5}$	$f_i y_i$	y_i^2	$f_i y_i^2$
70 – 75	72.5	3	– 4	– 12	16	48
75 – 80	77.5	4	– 3	– 12	9	36
80 – 85	82.5	7	– 2	– 14	4	28
85 – 90	87.5	7	– 1	– 7	1	– 7
90 – 95	92.5	15	0	0	0	0
95 – 100	97.5	9	1	9	1	9
100 – 105	102.5	6	2	12	4	24
105 – 110	107.5	6	3	18	9	54
110 – 115	112.5	3	4	12	16	48
Total		**60**		**6**		**254**

Here, $y_i = \frac{x_i - 92.5}{5}$ $\because$ $h = 5$ and $a = 92.5$

∴ $$\bar{x} = a+\frac{\sum f_i y_i}{\sum f_i}\times h = 92.5+\frac{6}{60}\times 5 = 93$$

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

$$= \frac{25}{3,600}\left[60\times 254-6^2\right]=\frac{1}{44}(15240-36)=\frac{15204}{144}$$

$$= 105.58$$

∴ Mean = 93, Variance = 105.58

and Standard deviation $= \sqrt{105.58} = 10.27$

Q. 10. The diameter of circles (in mm) drawn in a design are given below :

Diameter (in mm)	33 – 36	37 – 40	41 – 44	45 – 48	49 – 52
No. of Circles	15	17	21	22	25

Calculate the standard deviation and mean diameter of the circles.

[Hint : First make the data continuous by making the classes as 32.5 – 36.5, 36.5 – 40.5, 40.5 – 44.5, 44.5 – 48.5, 48.5 – 52.5 and then proceed]. **[NCERT Ex. 13.2. Q. 10. Page 282]**

Sol. After making the classes continuous, we have :

Diameter	f_i	x_i	$d_i = \frac{x_i - a}{h}$	$f_i d_i$	$f_i d_i^2$
32.5 – 36.5	15	34.5	– 2	– 30	60
36.5 – 40.5	17	38.5	– 1	– 17	17
40.5 – 44.5	21	42.5	0	0	0
44.5 – 48.5	22	46.5	1	22	22
48.5 – 52.5	25	50.5	2	50	100
Total	**100**			**25**	**199**

Mean, $$\bar{x} = a+\frac{\sum f_i d_i}{N}\times h$$

$$= 42.5+\frac{25\times 4}{100}$$

$$= 43.5$$

and Standard deviation, $$\sigma = h\sqrt{\frac{\sum f_i d_i^2}{N}-\left(\frac{\sum f_i d_i}{N}\right)^2}$$

$$= 4\sqrt{\frac{199}{100}-\left(\frac{25}{100}\right)^2}$$

$$= 4\sqrt{1.99-0.0625}$$

$$= 4\sqrt{1.9275}$$

$$= 5.55$$

∴ Mean, $\bar{x}$ = 43.5 and S.D. = 5.55.

MISCELLANEOUS EXERCISE

Q. 1. The mean and variance of eight observations are 9 and 9.25, respectively. If six of the observations are 6, 7, 10, 12, 12 and 13, find the remaining two observations. **[NCERT Misc. Ex. Q. 1. Page 286]**

Sol. Let the remaining two observations be x and y.

Therefore, the observations are 6, 7, 10, 12, 12, 13, x, y.

Mean, $\bar{x} = \dfrac{6+7+10+12+12+13+x+y}{8} = 9$

$\Rightarrow \quad 60 + x + y = 72$

$\Rightarrow \quad x + y = 12 \quad$ ...(1)

$$\text{variance} = 9.25 = \frac{1}{n}\sum_{i=1}^{8}(x_1 - \bar{x})^2$$

$$9.25 = \frac{1}{8}\left[(-3)^2 + (-2)^2 + (1)^2 + (3)^2 + (4)^2 + x^2 + y^2 - 2\times 9(x+y) + 2\times(9)^2\right]$$

$$9.25 = \frac{1}{8}\left[9+4+1+9+9+16+x^2+y^2-18(12)+162\right] \quad \text{...[using (1)]}$$

$$9.25 = \frac{1}{8}\left[48+x^2+y^2-216+162\right]$$

$$9.25 = \frac{1}{8}\left[x^2+y^2-6\right]$$

$\Rightarrow \quad x^2 + y^2 = 80 \quad$ ...(2)

From (1), we obtain

$x^2 + y^2 + 2xy = 144 \quad$ ...(3)

From (2) and (3), we obtain

$2xy = 64 \quad$ ...(4)

Subtracting (4) from (2), we obtain

$x^2 + y^2 - 2xy = 80 - 64 = 16$

$\Rightarrow \quad x - y = \pm 4 \quad$ ...(5)

Therefore, from (1) and (5), we obtain

$x = 8$ and $y = 4$, when $x - y = 4$

$x = 4$ and $y = 8$, when $x - y = -4$

Thus, the remaining observations are 4 and 8.

Q. 2. The mean and variance of 7 observations are 8 and 16, respectively. If five of the observations are 2, 4, 10, 12 and 14. Find the remaining two observations. [NCERT Misc. Ex. Q. 2. Page 286]

Sol. Let the remaining two observations be x and y.

The observations are 2, 4, 10, 12, 14, x, y.

Mean, $\bar{x} = \dfrac{2+4+10+12+14+x+y}{7}$

$\Rightarrow \quad 56 = 42 + x + y$

$\Rightarrow \quad x + y = 14 \quad$ ...(1)

$$\text{Variance} = 16 = \frac{1}{n}\sum_{i=1}^{7}(x_i - \bar{x})^2$$

$$16 = \frac{1}{7}\left[(2-8)^2 + (4-8)^2 + (10-8)^2 + (12-8)^2 + (14-8)^2 + (x-8)^2 + (y-8)^2\right]$$

$$16 = \frac{1}{7}\left[36+16+4+16+36+x^2+y^2-16(14)+2(64)\right] \quad \text{...[using (1)]}$$

$$16 = \frac{1}{7}\left[108+x^2+y^2-224+128\right]$$

$$16 = \frac{1}{7}\left[12+x^2+y^2\right]$$

$\Rightarrow \quad x^2 + y^2 = 112 - 12 = 100$

$\Rightarrow \quad x^2 + y^2 = 100 \quad$ ...(2)

From (1), we obtain

$x^2 + y^2 + 2xy = 196 \quad$ ...(3)

From (2) and (3), we obtain

$$2xy = 196 - 100$$

$$\Rightarrow \quad 2xy = 96 \qquad ...(4)$$

Subtracting (4) from (2), we obtain

$$x^2 + y^2 - 2xy = 100 - 96$$

$$\Rightarrow \quad (x - y)^2 = 4$$

$$\Rightarrow \quad x - y = \pm 2 \qquad ...(5)$$

Therefore, from (1) and (5), we obtain

$x = 8$ and $y = 6$ when $x - y = 2$

$x = 6$ and $y = 8$ when $x - y = -2$

Thus, the remaining observations are 6 and 8

Q. 3. The mean and standard deviation of six observations are 8 and 4, respectively. If each observation is multiplied by 3, find the new mean and new standard deviation of the resulting observations.

[NCERT Misc. Ex. Q. 3. Page 286]

Sol. Let the observations be x_1, x_2, x_3, x_4, x_5 and x_6.

It is given that mean is 8 and standard deviation is 4.

$$\text{Mean } \bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6}{6} \qquad ...(1)$$

If each observation is multiplied by 3 and the resulting observations and y_1, then

$$y_1 = 3x_1 \text{ i.e., } x_1 = \frac{1}{3} y_1, \text{ for } i = 1 \text{ to } 6$$

$$\therefore \quad \text{New Mean, } \bar{y} = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{6}$$

$$= \frac{3(x_1 + x_2 + x_3 + x_4 + x_5 + x_6)}{6}$$

$$= 3 \times 8 \qquad ...[\text{Using (1)}]$$

$$= 24$$

$$\text{Standard deviation, } \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{6}(x_1 - \bar{x})^2}$$

$$\therefore \quad (4)^2 = \frac{1}{6}\sum_{i=1}^{6}(x_1 - \bar{x})^2$$

$$\sum_{i=1}^{6}(x_1 - \bar{x})^2 = 96 \qquad ...(2)$$

From (1) and (2), it can be observed that,

$$\bar{y} = 3\bar{x}$$

$$\bar{x} = \frac{1}{3}\bar{y}$$

Substituting the values of x_1 and $\bar{x}$ in (2), we obtain

$$\sum_{i=1}^{6}\left(\frac{1}{3}y_1 - \frac{1}{3}\bar{y}\right)^2 = 96$$

$$\Rightarrow \quad \sum_{i=1}^{6}(y_1 - \bar{y})^2 = 864$$

Therefore, variance of new observations $= \left(\frac{1}{6} \times 864\right) = 144$

Hence, the standard deviation of new observations is $\sqrt{144} = 12 \ = 12$

Q. 4. Given that $\bar{x}$ is the mean and σ^2 is the variance of n observations $x_1, x_2, ..., x_n$. Prove that the mean and variance of the observations $ax_1, ax_2, ax_3, ..., ax_n$ are $a\bar{x}$ and $a^2\sigma^2$, respectively ($a \neq 0$). [NCERT Misc. Ex. Q. 4. Page 286]

Sol. The given n observations are $x_1, x_2 ... x_n$.

$$\text{Mean} = \bar{x}$$

$$\text{Variance} = \sigma^2$$

$$\therefore \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} y_1(x_i - \bar{x})^2 \quad ...(1)$$

If each observation is multiplied by a and the new observations are y_1, then

$$y_1 = ax_i \text{ i.e., } x_i = \frac{1}{a}y_1$$

$$\therefore \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_1 = \frac{1}{n}\sum_{i=1}^{n} ax_1 = \frac{a}{n}\sum_{i=1}^{n} x_1 = a\bar{x} \qquad \left(\because \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_1\right)$$

Therefore, mean of the observations $ax_1, ax_2 ... ax_n$, is $a\bar{x}$

Substituting the values of x_1 and $\bar{x}$ in (1), we obtain

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{a}y_1 - \frac{1}{a}\bar{y}\right)^2$$

$$\Rightarrow \quad a^2\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_1 - \bar{y}\right)^2$$

Thus, the variance of the observations, $ax_1, ax_2 ... ax_n$, is $a^2\sigma^2$.

Q. 5. The mean and standard deviation of 20 observations are found to be 10 and 2 respectively. On rechecking, it was found that an observation 8 was incorrect. Calculate the correct mean and standard deviation in each of the following cases :

(i) if wrong item is omitted

(ii) if it is replaced by 12. [NCERT Misc. Ex., Q. 5, Page 286]

Sol. (i) Incorrect Mean of 20 observations = 10

Incorrect Standard Deviation = 2

$$X^{-} = \frac{\sum_{i=1}^{20} x_i}{n}$$

$$\mathbf{10} = \frac{1}{2}\sum_{i=1}^{20} x_i$$

$$\Rightarrow \quad \frac{1}{2}\sum_{i=1}^{20} x_i = 20$$

$\therefore$ The incorrect sum of observations = 200

Correct sum of observations = 200 – 8 = 192

$$\Rightarrow \quad \text{Correct mean} = \frac{\text{Correct Sum}}{19}$$

$$= \frac{192}{9} = 10.1$$

$$\text{Standard deviation } \sigma = \sqrt{\frac{1}{n}\Sigma x_i^2 - \left(\frac{\Sigma x_i}{n}\right)^2}$$

$$2 = \sqrt{\frac{1}{20}\text{Incorrect} - \Sigma x_i^2 - 10^2}$$

$$\Rightarrow \quad 4 = \frac{1}{20} \quad \text{Incorrect } \Sigma x_i^2 - 100$$

$$104 = \frac{\text{Incorrect}}{20}\Sigma x_i^2$$

$$\Rightarrow \quad \text{Incorrect } \Sigma x_i^2 = 2080$$

$$\therefore \quad \text{Correct } \sum_{i=1}^{20} x_i^2 = \text{Incorrect } \sum_{i=1}^{20} x_i^2 - 8^2$$

$$= 2080 - 64$$

$$= 2016$$

$$\therefore \quad \text{Correct Standard Deviation} = \sqrt{\frac{\text{Correct } \Sigma x_i^2}{n} - (\text{Correct Mean})^2}$$

$$= \sqrt{\frac{\text{Correct } \Sigma x_i^2}{n} - (\text{Correct Mean})^2}$$

$$= \sqrt{106.1 - 102.01}$$

$$= \sqrt{4.09}$$

$$= 2.02$$

(ii) When 8 is replaced by = 12

Incorrect sum of observation = 200

$\therefore$ Correct Sum of Observations = 200 – 8 + 12 = 204

$$\therefore \quad \text{Correct Mean} = \frac{\text{Correct Sum}}{20} = \frac{204}{20} = 10.2$$

$$\text{Standard Deviation } \sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}$$

$$2 = \sqrt{\frac{\text{Incorrect } \Sigma x_i^2}{20} - 10^2}$$

$$\Rightarrow \quad 4 = \frac{1}{20} \quad \text{Incorrect } \Sigma x_i^2 - 100$$

$$\text{Incorrect } \Sigma x_i^2 = 2080$$

$$\therefore \quad \textbf{Correct } \Sigma x_i^2 = 2080 - 8^2 + 12^2 = 2160$$

$$\text{Correct Standard Deviation} = \sqrt{\frac{\text{Correct } \Sigma x_i^2}{n} - (\text{Correct Mean})^2}$$

$$= \sqrt{\frac{2160}{20} - (10.2)^2}$$

$$= \sqrt{108 - 104.04}$$

$$= \sqrt{3.96}$$

$$\sigma = 1.98$$

Q. 6. The mean and standard deviation of a group of 100 observations were found to be 20 and 3, respectively. Later on it was found that three observations were incorrect, which were recorded as 21, 21 and 18. Find the mean and standard deviation if the incorrect observations are omitted. [NCERT Mise Ex. Q. 6. Page 286]

Sol.

$$\text{Number of observations } (n) = 100$$

$$\text{Incorrect mean } (\bar{x}) = 20$$

$$\text{Incorrect standard deviation } (\sigma) = 3$$

$$\Rightarrow \quad 20 = \frac{1}{100}\sum_{i=1}^{100} x_i$$

$$\Rightarrow \quad \sum_{i=1}^{100} x_i = 20 \times 100 = 2{,}000$$

$$\therefore \quad \text{Incorrect sum of observations} = 2{,}000$$

$$\Rightarrow \quad \text{Correct sum of observations} = 2{,}000 - 21 - 21 - 18 = 2{,}000 - 60 = 1940$$

❒❒❒