Example 1

Find the mean deviation about the mean for the following data: $6, 7, 10, 12, 13, 4, 8, 12$

Solution

Given, $6, 7, 10, 12, 13, 4, 8, 12$

Mean of the given data is

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

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

The deviations of the given observations from the mean $\overline{x}$,

That is $x_i - \overline{x}$ are

$$x_1 - \overline{x} = 6 - 9 = -3$$
$$x_2 - \overline{x} = 7 - 9 = -2$$
$$x_3 - \overline{x} = 10 - 9 = 1$$
$$x_4 - \overline{x} = 12 - 9 = 3$$
$$x_5 - \overline{x} = 13 - 9 = 4$$
$$x_6 - \overline{x} = 4 - 9 = -5$$
$$x_7 - \overline{x} = 8 - 9 = -1$$
$$x_8 - \overline{x} = 12 - 9 = 3$$

The absolute values of the deviation

That is $|x_i - \overline{x}|$ are $3, 2, 1, 3, 4, 5, 1, 3$

The mean deviation of the required data is

Mean deviation $(\overline{x}) = \dfrac{\sum_{i=1}^{8} |x_i - \overline{x}|}{8}$

$$= \frac{3+2+1+3+4+5+1+3}{8}$$

$$= \frac{22}{8} = 2.75$$

The mean deviation of given data is 2.75.

Example 2

Find the mean deviation about the mean for the following data:

$12, 3, 18, 17, 4, 9, 17, 19, 20, 15, 8, 17, 2, 3, 16, 11, 3, 1, 0, 5$

Solution

Given data $12, 3, 18, 17, 4, 9, 17, 19, 20, 15, 8, 17, 2, 3, 16, 11, 3, 1, 0, 5$

To find the mean $(\overline{x})$ of the given data

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

Sum of all term is

$$12+3+18+17+4+9+17+19+20+15+8+17+2+3+16+11+3+1+0+5=200$$

Total number of terms is 20

So, $\overline{x}=\dfrac{200}{20}=10$

The absolute values for the deviation from the definition

Then, $|x_i-\overline{x}|$ are $2, 7, 8, 7, 6, 1, 7, 9, 10, 5, 2, 7, 8, 7, 6, 1, 7, 9, 10, 5$

The sum of the absolute value of deviation is

$$\sum_{i=1}^{20}|x_i-\overline{x}|=124$$

Mean deviation of the given data

$$\text{M.D.}(\overline{x})=\frac{124}{20}=6.2$$

The mean deviation of the given data is 6.2.

Example 3

Find the mean deviation about the median for the following data: $3, 9, 5, 3, 12, 10, 18, 4, 7, 19, 21.$

Solution

Given, $3, 9, 5, 3, 12, 10, 18, 4, 7, 19, 21.$

Arrange the data into decreasing order

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

The total number of observations is 11 which is odd

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

$$M = \left(\frac{12}{2}\right)^{\text{th}} \text{ observation}$$

6^{th} observation $= 9$

Median of the given data is 9

The correct values of the deviations from the median,

So, $|x_i - \text{M}|$ are $6, 6, 5, 4, 2, 0, 1, 3, 9, 10, 12$

Therefore, the sum of the total values

$$\sum_{i=1}^{11} |x_i - \text{M}| = 58$$

The mean deviation of the data is

$$\text{M.D.(M)} = \frac{1}{11}\sum_{i=1}^{11}|x_i - \text{M}|$$

$$= \frac{1}{11} \times 58 = 5.27$$

The mean deviation of the given data is 5.27

Example 4

Find mean deviation about the mean for the following data:

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

Solution

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

Given,

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

First to determine the $f_i x_i$

Multiply each the term

$f_i x_i$ of the given data $4,40,60,56,80,60$

The mean $(\bar{x})$ of the given data

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{6} f_i x_i$$

$$= \frac{1}{40} \times 300$$

$$= 7.5$$

Table shows

x_i	f_i	$f_i x_i$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
2	2	4	5.5	11
5	8	40	2.5	20
6	10	60	1.5	15
8	7	56	0.5	3.5
10	8	80	2.5	20
12	12	60	4.5	22.5
	40	300		92

Mean Deviation about the mean

$$\bar{x} = \frac{\sum_{i=1}^{6} f_i \lvert x_i - \bar{x}\rvert}{\sum_{i=1}^{6} f_i}$$

$$= \frac{92}{40}$$

$$= 2.3$$

The mean deviation following data is 2.3.

Example 5

Find the mean deviation about the median for the following data:

x_i	**3**	6	9	12	13	15	21 14	22
f_i	3	4	5	2	4	5	4	3

Solution

The given observations are already in ascending order. Adding a row corresponding to cumulative frequencies to the given data

x_i	**3**	6	9	12	13	15	21	22
f_i	3	4	5	2	4	5	4	3
$c.f$	**3**	7	12	14	18	23	27	30

Now, $N = 30$ which is even.

Median is the mean of the 15^{th} and 16^{th} observations.

Both of these observations lie in the cumulative frequency 18, for which the corresponding observation is 13 .

Therefore,

$$\text{Median } M = \frac{15^{th} \text{ observation } + 16^{th} \text{ observation}}{2}$$

$$= \frac{13+13}{2} = 13$$

Now, absolute values of the deviations from median,

That $|x_i - M|$ are shown in

$\lvert x_i - \mathbf{M}\rvert$	10	7	4	1	0	2	8	9
f_i	3	4	5	2	4	5	4	3
$f_i \lvert x_i - \mathbf{M}\rvert$	30	28	20	2	0	10	32	27

Sum of the terms $\sum_{i=1}^{8} f_i = 30$

Sum of the terms $\sum_{i=1}^{8} f_i |x_i - M| = 149$

Therefore, mean deviation of data

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

$$= \frac{1}{30} \times 149 = 4.97$$

The mean deviation of given data is 4.97.

Example 6

Find the mean deviation about the mean for the following data.

Marks obtained	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80
Number of students	2	3	8	14	8	3	2

Solution

We make the following from the given data:

Marks obtained	Number of students f_i	Mid-points x_i	$f_i x_i$	$\lvert x_i - \bar{\chi} \rvert$.	$f_i \lvert x_i - \bar{x} \rvert$
10 – 20	2	15	30	30	60
20 – 30	3	25	75	20	60
30 – 40	8	35	280	10	80
40 – 50	14	45	630	0	0
50 – 60	8	55	440	10	80
60 – 70	3	65	195	20	60
70 – 80	2	75	150	30	60
	40		1800		400

Total number of students is $N = \sum_{i=1}^{7} f_i = 40$

Sum of the $f_i x_i$ is $\sum_{i=1}^{7} f_i x_i = 1800$

Sum of the $f_i \lvert x_i - \bar{x} \rvert$ is $\sum_{i=1}^{7} f_i \lvert x_i - \bar{x} \rvert = 400$

Find the mean deviation about the mean for the following data.

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

$$= \frac{1800}{40}$$

$$= 45$$

Mean deviation about mean M.D.

$$(\bar{x}) = \frac{1}{N}\sum_{i=1}^{7} f_i\left|x_i - \bar{x}\right|$$

$$= \frac{1}{40} \times 400 = 10$$

The mean deviation about mean is 10.

Example 7

Calculate the 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

Solution

Form the following from the given data

Class	Frequency f_i	Cumulative frequency (c.f.)	Mid-points x_i	$\lvert x_i - \text{Med}\rvert$	$f_i\lvert x_i - \text{Med}\rvert$
$0-10$	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
$50-60$	5	50	55	27	54
	50				508

Sum of the f_i $N = \sum f_i = 50$

The class interval contains $\frac{N^{th}}{2}$ or 25^{th} item is $20-30$.

Therefore, $20-30$ is the median class.

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

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 $l = 20, C = 13, f = 15, h = 10$ and $N = 50$

Therefore, Median $= 20 + \frac{25-13}{15} \times 10 = 20 + 8 = 28$

Thus, mean deviation about median is given by

$$\text{M.D. (M)} = \frac{1}{N} \sum_{i=1}^{6} f_i \left| x_i - M \right|$$

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

Example 8

Find the variance of the following data:

$6, 8, 10, 12, 14, 16, 18, 20, 22, 24$

Solution

From the given data

The mean does calculate by the step-deviation method using 14 since the estimated mean. The number of observations is $n = 10$

x_i	$d_i = \frac{x_i - 14}{2}$	Deviations from mean $(x_i - \bar{x})$	$(x_i - \bar{x})^2$
6	–4	–9	81
8	–3	–7	49
10	–2	–5	25
12	–1	–3	9
14	0	–1	1
16	1	1	1
18	2	3	9
20	3	5	25
22	4	7	49
24	5	9	81
	5		330

So, the mean of the given data

Mean $\overline{x} =$ assumed mean $+\frac{\sum_{i=1}^{n} d_i}{n} \times h$

$= 14 + \frac{5}{10} \times 2$

$= 15$

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

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

The variance of the given data is

$$\left(\sigma^2\right) = \frac{1}{n} \sum_{i=1}^{10} \left(x_i - \overline{x}\right)^2$$

$= \frac{1}{10} \times 330$

$= 33$

The standard deviation, usually denoted by C, is given by

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

So, Standard deviation $(\sigma) = \sqrt{33} = 5.74$

Example 9

Find the variance and standard deviation for the following data:

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

Solution

Presenting the data in tabular form

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

	30	420			1374

Sum of the f_i is $\sum f_i = 30$

Sum of the $x_i f_i$ is $\sum x_i f_i = 420$

Sum of the $f_i (x_i - \bar{x})^2$ is $\sum_{i=1}^{7} f_i (x_i - \bar{x})^2 = 1374$

So, the mean of the given data is

$$\bar{x} = \frac{\sum_{i=1}^{7} f_i x_i}{\mathrm{N}}$$

$$= \frac{1}{30} \times 420$$

$$= 14$$

Therefore,

$$\text{variance } (\sigma^2) = \frac{1}{\mathrm{N}} \sum_{i=1}^{7} f_i (x_i - \bar{x})^2$$

$$= \frac{1}{30} \times 1374 = 45.8$$

and

Standard deviation $(\sigma) = \sqrt{45.8} = 6.77$

Example 10

Calculate the mean, variance and standard deviation for the following distribution:

Class	$30-40$	$40-50$	$50-60$	$60-70$	$70-80$	$80-90$	$90-100$
Frequenc-y	3	7	12	15	8	3	2

Solution

From the given data, create a table for following data

Class	Frequency (f_i)	Midpoint (x_i)	$f_i x_i$	$(x_i - \overline{x})^2$	$f_i (x_i - \overline{x})^2$
$30-40$	3	35	105	729	2187
$40-50$	7	45	315	289	2023
$50-60$	12	55	660	49	588
$60-70$	15	65	975	9	135
$70-80$	8	75	600	169	1352
$80-90$	3	85	255	529	1587
$90-100$	2	95	190	1089	2178
	50		3100		10050

Sum of the f_i is $\sum f_i = 50$

Sum of the $x_i f_i$ is $\sum x_i f_i = 3100$

Thus, the mean of the given data

$$\text{Mean } \overline{x} = \frac{1}{\text{N}} \sum_{i=1}^{7} f_i x_i$$

$$= \frac{3100}{50} = 62$$

Sum of the $f_i (x_i - \overline{x})^2$ is $\sum_{i=1}^{7} f_i (x_i - \overline{x})^2 = 10050$

So, the Variance

$$(\sigma^2) = \frac{1}{\text{N}} \sum_{i=1}^{7} f_i (x_i - \overline{x})^2$$

$$= \frac{1}{50} \times 10050 = 201$$

and

Standard deviation $(\sigma) = \sqrt{201} = 14.18$

Example 11

Find the standard deviation for the following data:

x_i	3	8	13	18	23
f_i	7	10	15	10	6

Solution

Let us form the following

x_i	f_i	$f_i x_i$	x_i^2	$f_i x_i^2$
3	7	21	9	63
8	10	80	64	640
13	15	195	169	2535
18	10	180	324	3240
23	6	138	529	3174
	48	614		9652

Now, by formula of standard deviation

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

Standard deviation of following data

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

Substitute the values in the table

$$= \frac{1}{48}\sqrt{48 \times 9652 - (614)^2}$$

$$= \frac{1}{48}\sqrt{463296 - 376996}$$

$$= \frac{1}{48} \times 293.77 = 6.12$$

So, Standard deviation $(c) = 6.12$

Examples 12

Calculate mean, variance and standard deviation for the following distribution.

Classes	$30-40$	$40-50$	$50-60$	$60-70$	$70-80$	$80-90$	$90-100$
Frequency	3	7	12	15	8	3	2

Solution

Let the assumed mean A $=65$.

Here $h=10$

Obtain the following data into table.

Class	Frequency f_i	Mid-point x_i	$y_i=\frac{x_i-65}{10}$	y_i^2	f_iy_i	$f_iy_i^2$
$30-40$	3	35	-3	9	-9	27
$40-50$	7	45	-2	4	-14	28
$50-60$	12	55	-1	1	-12	12
$60-70$	15	65	0	0	0	0
$70-80$	8	75	1	1	8	8
$80-90$	3	85	2	4	6	12
$90-100$	2	95	3	9	6	18
	$N=50$				-15	105

From the table,

$$\sum f_iy_i=-15$$

$$\sum f_iy_i^2=105$$

$$N=\sum f_i=50$$

So, the mean of the give data is

$$\overline{x}=A+\frac{\sum f_iy_i}{50}\times h$$

$$=65-\frac{15}{50}\times 10=62$$

So, the variance is $\sigma^2=\frac{h^2}{N^2}\left[N\Sigma f_iy_i^2-\left(\sum f_iy_i\right)^2\right]$

$$=\frac{(10)^2}{(50)^2}\left[50\times 105-(-15)^2\right]$$

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

and standard deviation $(\sigma) = \sqrt{201} = 14.18$

Example 13

Two plants A and B of a factory show following results about the number of workers and the wages paid to them.

	A	B
No. of workers	5000	6000
Average monthly wages	$Rs\ 2500$	$Rs\ 2500$
Variance of distribution of wages	81	100

In which plant, A or B is there greater variability in individual wages?

Solution

To compare the variation, we have to calculate coefficient of variation.

Coefficient of variation $(c \cdot v) = \frac{\text{stardard Deviation}}{\text{mean}} \times 100$.

since, Average monthly wages are some, we compare standard deviation of the distribution.

variance in plant $A = 81$

Therefore, standard deviation of the distribution of wages in plant $\mathrm{A}(\sigma_1) = 9$

Variance in plant $B = 100$

Standard deviation of the distribution of wages in plant $\mathrm{B}(\sigma_2) = 10$

Since the average monthly wages in both plants are same.

Therefore, the plant with greater standard deviation will have more variability.

Thus, the plant B has greater variability in the individual wages.

Example 14

Coefficient of variation of two distributions are 60 and 70, and their standard deviations are 21 and 16, respectively. What are their arithmetic means.

Solution

Given

The coefficient of variation (CV) for the first distribution is 60.

The coefficient of variation (CV) for the second distribution is 70.

The standard deviation for the first distribution is $\sigma_1 = 21$

The standard deviation for the second distribution is $\sigma_2 = 16$

Let $\bar{x}_1$ and $\bar{x}_2$ be the means of 1st and 2nd distribution, respectively.

So,

$$\text{C.V. (1 st distribution)} = \frac{\sigma_1}{\bar{x}_1} \times 100$$

Substitute the values in equation

$$60 = \frac{21}{\bar{x}_1} \times 100$$

Hence,

$$\bar{x}_1 = \frac{21}{60} \times 100 = 35$$

Similarly,

$$\text{C.V. (2 nd distribution)} = \frac{\sigma_2}{\bar{x}_2} \times 100$$

Substitute the values in equation

$$70 = \frac{16}{\bar{x}_2} \times 100$$

Hence,

$$\bar{x}_2 = \frac{16}{70} \times 100 = 22.85$$

Example 15

The following values are calculated in respect of heights and weights of the students of a section of Class XI

	Height	Weight
Mean	162.6 cm	52.36 kg
Variance	127.69 cm^2	23.1361 kg^2

Can we say that the weights show greater variation than the heights?

Solution

To compare the variability of the given data, To calculate their coefficients of variation.

Given

Variance of height $=127.69\,\text{cm}^2$

Therefore, Standard deviation of height $=\sqrt{127.69}\,\text{cm}=11.3\,\text{cm}$

Also

Variance of weight $=23.1361\,\text{kg}^2$

Therefore Standard deviation of weight $=\sqrt{23.1361}\,\text{kg}=4.81\,\text{kg}$

Now, the coefficient of variations (C.V.) are given by

$$\text{(C.V.) in heights }=\frac{\text{Standard Deviation}}{\text{Mean}}\times 100$$

$$=\frac{11.3}{162.6}\times 100=6.95$$

$$\text{and (C.V.) in weights }=\frac{4.81}{52.36}\times 100=9.18$$

Clearly C.V. in weights is greater than the C.V. in heights

Hence, we can tell that weights explain more variability than heights

Example 16

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

Solution

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

Given variance $=5$ and $n=20$. We know that formula for variance

$$\text{Variance }\left(\sigma^2\right)=\frac{1}{n}\sum_{i=1}^{20}\left(x_i-\bar{x}\right)^2,$$

$$\text{That is }\ 5=\frac{1}{20}\sum_{i=1}^{20}\left(x_i-\bar{x}\right)^2$$

$$\sum_{i=1}^{20}\left(x_i-\bar{x}\right)^2=100.......(1)$$

If each observation is multiplied by 2, and the new resulting observations are y_i, then

$y_i = 2x_i$

That is $y_i = 2x_i$ i.e., $x_i = \frac{1}{2} y_i$

Therefore

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{20} y_i$$

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

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

$$\bar{y} = 2\bar{x} \text{ or } \bar{x} = \frac{1}{2} \bar{y}$$

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

$$\sum_{i=1}^{20} \left(\frac{1}{2} y_i - \frac{1}{2} \bar{y} \right)^2 = 100$$

$$\sum_{i=1}^{20} (y_i - \bar{y})^2 = 400$$

Thus, the variance of new observations $= \frac{1}{20} \times 400 = 20$

Example 17

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.

Solution

Given, Mean of 5 observations is 4.4

variance is 8.24.

Let the next two observations be x and y.

Therefore, the series is 1,2,6, x, y.

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

$22 = 9 + x + y$

Therefore

$x+y=13$ ....(1)

variance $=8.24=\frac{1}{n}\sum_{i=1}^{5}\left(x_i-\bar{x}\right)^2$

That is $8.24=\frac{1}{5}\left[(3.4)^2+(2.4)^2+(1.6)^2+x^2+y^2-2\times 4.4(x+y)+2\times(4.4)^2\right]$

or $41.20=11.56+5.76+2.56+x^2+y^2-8.8\times 13+38.72$

Therefore $x^2+y^2=97$.......(2)

But from (1), we have

$ $x^2+y^2+2xy=169$

From (2) and (3),we have

$2xy=72$

Subtracting (4) from (2), we get

$x^2+y^2-2xy=97-72$

That is $(x-y)^2=25$

or $x-y=\pm 5 \ldots$ (5)

So, from (1) and (5), we get

$x=9, y=4$ when $x-y=5$

or $x=4, y=9$ when $x-y=-5$

Thus, the remaining observations are 4 and 9.

Example 18

If each of the observation $x_1, x_2, \ldots, x_n$ is increased by ' a ', where a is a negative or positive number, show that the variance remains unchanged.

Solution

Let $\bar{x}$ be the mean of $x_1, x_2, \ldots, x_n$.

Then the variance is given by

$$\sigma_1^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$$

If a is calculated to every observation, these new observations will be

$$y_i = x_i + a$$

Let the mean of the new observations be $\bar{y}$.

Then $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(x_i + a)$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} a\right] = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{na}{n} = \bar{x} + a$$

That is $\bar{y} = \bar{x} + a$

Thus, the variance of the new observations

Thus, the variance of the new observations

$$\sigma_2^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i + a - \bar{x} - a)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sigma_1^2$$

Hence, the variance of these new observations is this equal while that of the original observations.

Example 19

The mean and standard deviation of 100 observations were calculated as 40 and 5.1, respectively by a student who took by mistake 50 instead of 40 for one observation. What are the correct mean and standard deviation?

Solution

Given that number of observations $(n) = 100$

Incorrect $\text{mean}(\bar{x}) = 40$

Incorrect standard deviation $(\sigma) = 5.1$

We know that the mean of a data is

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

Substitute the values

$$40 = \frac{1}{100}\sum_{i=1}^{100} x_i$$

or $\sum_{i=1}^{100} x_i = 4000$

Incorrect sum of observations $= 4000$

Thus, the correct sum of observations $=$ Incorrect sum $-50+40$

$= 4000-50+40 = 3990$

Hence

$$\text{Correct mean } = \frac{\text{correct sum}}{100} = \frac{3990}{100} = 39.9$$

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

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

So, $5.1 = \sqrt{\frac{1}{100} \times \text{ Incorrect } \sum_{i=1}^{n} x_i^2 - (40)^2}$

And $26.01 = \frac{1}{100} \times \text{ Incorrect } \sum_{i=1}^{n} x_i^2 - 1600$

Hence, Incorrect $\sum_{i=1}^{n} x_i^2 = 100(26.01+1600) = 162601$

Now, Correct $\sum_{i=1}^{n} x_i^2 =$ Incorrect $\sum_{i=1}^{n} x_i^2 - (50)^2 + (40)^2$

$= 162601 - 2500 + 1600 = 161701$

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

$$= \sqrt{\frac{161701}{100} - (39.9)^2}$$

$= \sqrt{1617.01 - 1592.01} = \sqrt{25} = 5$

Hence, the correct mean and standard deviation of observation is 39.9 and 5.

EXERCISE 15.1

Find the mean deviation about the mean for the data in Exercises 1 and 2.

1. $4, 7, 8, 9, 10, 12, 13, 17$

Solution

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 deviations of the respective observations from the mean $\bar{x}$,

So, $x_i - \bar{x}$ are $-6, -3, -2, -1, 0, 2, 3, 7$

The absolute values of the deviations, $|x_i - \bar{x}|$ are $6, 3, 2, 1, 0, 2, 3, 7$

The necessary mean deviation regarding the mean is

$$M \cdot D \cdot (\bar{x}) = \frac{\sum_{i=1}^{8} |x_i - \bar{x}|}{8}$$
$$= \frac{6+3+2+1+0+2+3+7}{8}$$
$$= \frac{24}{8}$$
$$= 3$$

2. $\mathbf{38, 70, 48, 40, 42, 55, 63, 46, 54, 44}$

The given data is $38, 70, 48, 40, 42, 55, 63, 46, 54, 44$

Mean of the data

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

The deviations of the respective observations from the mean $\bar{x}$,

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

$-12, 20, -2, -10, -8, 5, 13, -4, 4, -6$

The absolute values of the deviations, $|x_i - \bar{x}|$ are

$12, 20, 2, 10, 8, 5, 13, 4, 4, 6$

The necessary mean deviation regarding the mean is

$$M \cdot D \cdot (\bar{x}) = \frac{\sum_{i=1}^{10} |x_i - \bar{x}|}{10}$$

$$= \frac{12+20+2+10+8+5+13+4+4+6}{10}$$

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

$$= 8.4$$

Find the mean deviation about the median for the data in Exercises 3 and 4.

3. $13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17$

Solution

The given data is \$13,17,16,14,11,13,10,16,11,18,12,17\$

Here, these numbers from observations do 12, that is, even.

Arranging the above data in ascending order, we obtain

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

Median of the data

$$M = \frac{\left(\frac{12}{2}\right)^{\text{th}} \text{ observation } + \left(\frac{12}{2}+1\right)^{\text{th}} \text{ observation}}{2}$$

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

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

$$= \frac{27}{2}$$

$$= 13.5$$

The deviations of the respective observations from the median, i.e., $x_i - M$ 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 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 necessary mean deviation regarding the mean is

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

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

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

$$= 2.33$$

4. $36, 72, 46, 42, 60, 45, 53, 46, 51, 49$

Solution

The given data is 36,72,46,42,60,45,53,46,51,49

Here, these numbers from observations do 10, that is, even

Arranging the above data in ascending order,

we obtain

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

Median of the data

$$M = \frac{\left(\frac{10}{2}\right)^{\text{th}} \text{ observation } + \left(\frac{10}{2}+1\right)^{\text{th}} \text{ observation}}{2}$$

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

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

$$= \frac{95}{2}$$

$$= 47.5$$

The deviations of the respective observations from the median

That is $x_i - M$ are

$-11.5, -5.5, -2.5, -1.5, -1.5, 1.5, 3.5, 5.5, 12.5, 24.5$

The absolute values of the deviations, $|x_i - M|$ are

11.5,5.5,2.5,1.5,1.5,1.5,3.5,5.5,12.5,24.5

The necessary mean deviation regarding the mean is

$$M \cdot D \cdot (M) = \frac{\sum_{i=1}^{10} |x_i - M|}{10}$$

$$= \frac{11.5+5.5+2.5+1.5+1.5+1.5+3.5+5.5+12.5+24.5}{10}$$

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

$$= 7$$

Find the mean deviation about the mean for the data in Exercises 5 and 6

5.

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

Solution

Form the data in to a table

x_i	f_i	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
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
	25	350		158

Sum of the f_i is $N = \sum_{i=1}^{5} f_i = 25$

Sum of the $f_i x_i$ is $\sum_{i=1}^{5} f_i x_i = 350$

So, the mean of the data is

$\bar{x} = \frac{1}{N}\sum_{i=1}^{5} f_i x_i$

$= \frac{1}{25} \times 350$

$= 14$

Mean deviation about the mean

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

$= \frac{1}{25} \times 158$

$= 6.32$

The mean deviation of the following data is 6.32

Find the mean deviation about the mean for the data in Exercises 5 and 6

6.

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

Solution

Form a table of given data

x_i	f_i	$f_i x_i$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
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
	80	4000		1280

Sum of the f_i is $\sum_{i=1}^{5} f_i = 80$

Sum of the $f_i x_i$ is $\sum_{i=1}^{5} f_i x_i = 4000$

Hence, the mean of the given data

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{5} f_i x_i$$

$$= \frac{1}{80} \times 4000$$

$$= 50$$

Mean deviation about the mean

$$M \cdot D \cdot (\overline{x}) = \frac{1}{N}\sum_{i=1}^{5} f_i \left|x_i - \overline{x}\right|$$

$$= \frac{1}{80} \times 1280$$

$$= 16$$

So, mean deviation of the given data is 16.

Find the mean deviation about the median for the data in Exercises 7 and 8.

7.

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

Solution

The given observations are in ascending order.

Adding one column corresponding at cumulative frequencies from the given data, we collect the following table.

x_i	f_i	C.F
5	8	8
7	6	14
9	2	16
10	2	18
12	2	20

15	6	26

Here, $N = 26$, which is even.

Consequently, Median means the average of 13th and 14th observations.

Both these observations extend into the cumulative frequency 14, to which this corresponding observation is 7."

$$\text{Median} = \frac{13^{\text{th}} \text{ observation } + 14^{\text{th}} \text{ observation}}{2} = \frac{7+7}{2} = 7$$

The absolute values of the deviations. $|x_i - M|$ are

$\lvert x_i - M \rvert$	2	0	2	3	5	8
f_i	8	6	2	2	2	6
$f_i \lvert x_i - M \rvert$	16	0	4	6	10	48

$$\sum_{i=1}^{6} f_i = 26 \text{ and } \sum_{i=1}^{6} f_i |x_i - M| = 84$$

Hence, the mean deviation about the median

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

$$= \frac{1}{26} \times 84$$

$$= 3.23$$

Find the mean deviation about the median for the data

8.

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

Solution

The given observations are in ascending order.

Adding one column corresponding at cumulative frequencies from the given data, we collect the following table.

x_i	f_i	$C.F$
15	3	3
21	5	8
27	6	14
30	7	21
35	8	29

Here, $N = 29$, which is odd.

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

The observation extends within the cumulative frequency 21, during which the similar observation is 30.

The absolute values of the deviations, $|x_i - M|$ are

$\lvert x_i - M\rvert$	15	9	3	0	5
f_i	3	5	6	7	8
$f_i\lvert x_i - M\rvert$	45	45	18	0	40

From above table,

Here, $\sum_{i=1}^{5} f_i = 29$ and $\sum_{i=1}^{5} f_i |x_i - M| = 148$

Mean deviation of data is

$$M \cdot D \cdot (M) = \frac{1}{N}\sum_{i=1}^{5} f_i |x_i - M|$$

$$= \frac{1}{29} \times 148$$

$$= 5.1$$

Infinity Learn by Sri Chaitanya Educational Institutions

Find the mean deviation about the mean for the data in Exercises 9 and 10.

9.

Income per day in ₹	$0-100$	$100-200$	$200-300$	$300-400$	$400-500$	$500-600$	$600-700$	$700-800$
Number of persons	4	8	9	10	7	5	4	3

Solution

Form the following table

Income per day	Number of persons f_i	Mid-point x_i	$f_i x_i$	$\lvert x_i - \bar{x}\rvert$	$f_i \lvert x_i - \bar{x}\rvert$
$0-100$	4	50	200	308	1232
$100-200$	8	150	1200	208	1664
$200-300$	9	250	2250	108	972
$300-400$	10	350	3500	8	80
$400-500$	7	450	3150	92	644
$500-600$	5	550	2750	192	960
$600-700$	4	650	2600	292	1168
$700-800$	3	750	2250	392	1176
	50		17900		7896

From table

Here, $N=\sum_{i=1}^{8} f_i = 50$ and $\sum_{i=1}^{8} f_i x_i = 17900$

So, Mean

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{8} f_i x_i$$

$$= \frac{1}{50}\times 17900$$

$$= 358$$

The mean deviation about the mean

$$M \cdot D \cdot (\bar{x}) = \frac{1}{N} \sum_{i=1}^{8} f_i |x_i - \bar{x}|$$

$$= \frac{1}{50} \times 7896$$

$$= 157.92$$

10. Find the mean deviation about the mean for 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

Solution

The following table is formed

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

From above table

$N = \sum_{i=1}^{6} f_i = 100$ **and** $\sum_{i=1}^{6} f_i x_i = 12530$

Therefore, the mean is

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{6} f_i x_i$$
$$= \frac{1}{100} \times 12530$$
$$= 125.3$$

So, the mean deviation of the mean

$$M.D.(\bar{x}) = \frac{1}{N}\sum_{i=1}^{6} f_i |x_i - \bar{x}|$$
$$= \frac{1}{100} \times 1128.8$$
$$= 11.28$$

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

Solution

Form of the following tables

Marks	No. of girls	c.f.	Mid-point (x_i)	$\|x_i - M\|$	$f_i \|x_i - M\|$
$0-10$	6	6	5	22.85	137.1
$10-20$	8	14	15	12.85	102.8
$20-30$	14	28	25	2.85	39.9
$30-40$	16	44	35	7.15	114.4
$40-50$	4	48	45	17.15	68.6

$50-60$	2	50	55	27.15	54.3
	50				517.1

Here, $\sum_{i=1}^{6} f_i = 50$ and $\sum_{i=1}^{6} f_i \left| x_i - M \right| = 517.1$

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

Median

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

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

$$= 20 + 7.85$$

$$= 27.85$$

Hence, the mean deviation about median

$$M \cdot D \cdot (M) = \frac{1}{N} \sum_{i=1}^{6} f_i \left| x_i - M \right|$$

$$= \frac{1}{50} \times 517.1$$

$$= 10.34$$

12. 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

Solution

The given data is not continuous.

Hence, this holds to do converted within continuous frequency distribution by subtracting 0.5 from this lower limit also adding 0.5 to this upper limit from each class interval.

Age (in years)	Number of persons f_i	c. f	Mid-point x_i	$\lvert x_i - M\rvert$	$f_i\lvert x_i - M\rvert$
$15.5-20.5$	5	5	18	20	100
$20.5-25.5$	6	11	23	15	90
$25.5-30.5$	12	23	28	10	120
$30.5-35.5$	14	37	33	5	70
$35.5-40.5$	26	63	38	0	0
$40.5-45.5$	12	75	43	5	60
$45.5-50.5$	16	91	48	10	160
$50.5-55.5$	9	100	53	15	135
	100				735

The table is formed as follows

The class interval including this $\left(\frac{N}{2}\right)^{th}$ about 50^{th} item is $35.5-40.5$

Therefore, is the $35.5-40.5$ is the median class

It is known that,

$$=l+\left(\frac{\frac{N}{2}-C}{f}\right)\times h$$

Median

Here, $l=35.5, C=37, f=26, h=5, N=100$

Median,

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

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

$$=35.5+2.5$$

$$=38$$

So, mean deviation about the median is given by,

$$M.D.(M)=\frac{1}{N}\sum_{i=1}^{8} f_i\lvert x_i - M\rvert$$

$= \frac{1}{100} \times 735$

$= 7.35$

Exercise 15.2

Find the mean and variance for each of the data in Exercise 1 to 5.

1. $6, 7, 10, 12, 13, 4, 8, 12$

Solution

The given data is $6, 7, 10, 12, 13, 4, 8, 12$

Mean of the data

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

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

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

$$= 9$$

The following table is obtained from the given above data

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

			74

Variance of the data

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

$$=\frac{1}{8}\times 74$$

$$=9.25$$

2. First n natural numbers

Solution

The mean of first n natural numbers is calculated

$$\frac{\text{Sum of all observations}}{\text{Number of observations}}=\frac{\frac{n(n+1)}{2}}{n}=\frac{n+1}{2}$$

Variance,

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

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

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

Rearrange the equation

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

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

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

Simplify,

$$= (n+1)\left[\frac{4n+2-3n-3}{12}\right]$$

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

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

3. First 10 multiples of 3

Solution

The first ten multiples of 3 are $3, 6, 9, 12, 15, 18, 21, 24, 27, 30$

The number of observations, $n = 10$

Mean of the data

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10}$$

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

$$= 16.5$$

The following table is obtained for the first 10 multiples of 3

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

24	7.5	56.25
27	10.5	110.25
30	13.5	182.25

Variance

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

$$=\frac{1}{10}\times 742.5$$

$$=74.25$$

4.

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

Solution

The table shows

x_i	f_i	$f_i x_i$	$x_i-\bar{x}$	$\left(x_i-\bar{x}\right)^2$	$f_i\left(x_i-\bar{x}\right)^2$
6	2	12	−13	169	338
10	4	40	−9	81	324
14	7	98	−5	25	175
18	12	216	−1	1	12
24	8	192	5	25	200
28	4	112	9	81	324
30	3	90	11	121	363
	40	760			1736

Here,

$$N=\sum_{i=1}^{7} f_i=40,\ \text{and}\ \sum_{i=1}^{7} f_1 x_1=760$$

So,

$$\bar{x}=\frac{\sum_{i=1}^{7} f_1 x_1}{N}$$

$$=\frac{760}{40}$$

$$=19$$

Variance,

$$\left(\sigma^2\right)=\frac{1}{n}\sum_{i=1}^{7}\left(x_i-\bar{x}\right)^2$$

$$=\frac{1}{40}\times 1736$$

$$=43.4$$

5.

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

Solution

Following the table

x_i	f_i	$f_i x_i$	$x_i-\bar{x}$	$\left(x_i-\bar{x}\right)^2$	$f_i\left(x_i-\bar{x}\right)^2$
92	3	276	-8	64	192
93	2	186	-7	49	98
97	3	291	-3	9	27
98	2	196	-2	4	8
102	6	612	2	4	24
104	3	312	4	16	48

109	3	327	9	81	243
	22	2200			640

Here, $N=\sum_{i=1}^{7} f_i = 22$ and $\sum_{i=1}^{7} f_1 x_1 = 2200$

So, the mean

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{7} f_i x_i$$

$$=\frac{1}{22}\times 2200$$

$$=100$$

Variance

$$\left(\sigma^2\right)=\frac{1}{N}\sum_{i=1}^{7}\left(x_i-\bar{x}\right)^2$$

$$=\frac{1}{22}\times 640$$

$$=29.09$$

6. Find the mean and standard deviation using short-cut method.

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

Solution

Form of the following

x_i	f_i	$f_i=\frac{x_i-64}{1}$	y_1^2	$f_1 y_1$	$f_1 y_1^2$
60	2	-4	16	-8	32
61	1	-3	9	-3	9

62	12	-2	4	-24	48
63	29	-1	1	-29	29
64	25	0	0	0	0
65	12	1	1	12	12
66	10	2	4	20	40
67	4	3	9	12	36
68	5	4	16	20	80
	100	220		0	286

Mean,

$$\bar{x} = A + \frac{\sum_{i=1}^{9} f_i y_i}{N} \times h$$

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

$$= 64$$

Variance,

$$\sigma^2 = \frac{h^2}{N^2}\left[N\sum_{i=1}^{9} f_i y_i^2 - \left(\sum_{i=1}^{9} f_i y_i\right)^2\right]$$

$$= \frac{1}{100^2}[100 \times 286 - 0]$$

$$= 2.86$$

Standard deviation,

$$(\sigma) = \sqrt{2.86} = 1.69$$

Find the mean and variance for the following frequency distributions in

Exercises 7

Classes	$0-30$	$30-60$	$60-90$	$90-120$	$120-150$	$150-180$	$180-210$
Frequencies	2	3	5	10	3	5	2

Solution

The following tables shows

Class	Frequency f_i	Mid-point x_i	$y_i = \dfrac{x_i - 105}{30}$	y_i^2	$f_i x_i$	$f_i y_i^2$
$0-30$	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
					2	76

Mean of given data

$$\overline{x} = A + \frac{\sum_{i=1}^{7} f_i y_i}{N} \times h$$

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

$$= 105 + 2$$

$$= 107$$

Variance of the given data

$$\sigma^2 = \frac{h^2}{N^2}\left[N\sum_{i=1}^{9} f_i y_i^2 - \left(\sum_{i=1}^{9} f_i y_i\right)^2\right]$$

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

$$= 2280-4$$

$$= 2276$$

Find the mean and variance for the following frequency distributions in Exercises 8

Classes	$0-10$	$10-20$	$20-30$	$30-40$	$40-50$
Frequencies	5	8	15	16	6

Solution

The following data is

Class	Frequency f_i	Mid-point (x_i)	$y_i = \frac{x_i - 25}{10}$	y_i^2	$f_i y_i$	$f_i y_i^2$
$0-10$	5	5	-2	4	-10	20
$10-20$	8	15	-1	1	-8	8
$20-30$	15	25	0	0	0	0
$30-40$	16	35	1	1	16	16
$40-50$	6	45	2	4	12	24
	50				10	68

Mean,

$$\bar{x} = A + \frac{\sum_{i=1}^{5} f_i y_i}{N} \times h$$

$$= 25 + \frac{10}{50} \times 10$$

$$= 25 + 2$$

$$= 27$$

Variance,

$$\left(\sigma^2\right) = \frac{h^2}{N^2}\left[N\sum_{i=1}^{5} f_i y_i^2 - \left(\sum_{i=1}^{5} f_i y_i\right)^2\right]$$

Substitute the values

$$= \frac{(10)^2}{(50)^2}\left[50 \times 68 - (10)^2\right]$$

$$= \frac{1}{25}[3400 - 100]$$

$$= \frac{3300}{25}$$

$$= 132$$

9. Find the mean, variance and standard deviation using short-cut method

Height in cms	$70-75$	$75-80$	$80-85$	$85-90$	$90-95$	$95-100$	$100-105$	$105-110$	$110-115$
No. of children	3	4	7	7	15	9	6	6	3

Solution

The following table shows

After taking the mid-values of class-intervals, let us take the assumed mean $A = 92.5$. Here $h = 5$

We receive the following table of the given data.

Height (in cms)	No. of children f_i	Mid point x_i	$y_i=\frac{x_i-92.5}{5}$	y_i^2	f_iy_i	$f_iy_i^2$
$70-75$	3	72.5	-4	16	-12	48
$75-80$	4	77.5	-3	9	-12	36
$80-85$	7	82.5	-2	4	-14	28
$85-90$	7	87.5	-1	1	-7	7
$90-95$	15	92.5	0	0	0	0
$95-100$	9	97.5	1	1	9	9
$100-105$	6	102.5	2	4	12	24
$105-110$	6	107.5	3	9	18	54
$110-115$	3	112.5	4	16	12	48
	60				6	254

Here,

$$\mathrm{N}=\sum_{i=1}^{9} f_i=60,\text{ and }\sum_{i=1}^{9} f_iy_i=6$$

So, the mean

$$\bar{x}=\mathrm{A}+\frac{\sum_{i=1}^{9} f_iy_i}{\mathrm{N}}\times h$$

$$=92.5+\frac{6}{60}\times 5$$

$$=92.5+0.5$$

$$=93$$

and Variance $\left(\sigma^2\right)=h^2\left[\frac{1}{\mathrm{N}}\sum f_iy_i^2-\left(\frac{\sum f_iy_i}{N}\right)^2\right]$

$$=\frac{(5)^2}{(60)^2}\left[60\times 254-(6)^2\right]$$

$$=\frac{25}{3600}\times 15204$$

$$=105.58$$

Standard deviation,

$$(\sigma)=\sqrt{105.58}=10.27$$

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

Diameters	$33-36$	$37-40$	$41-44$	$45-48$	$49-52$
No. of circles	15	17	21	22	25

Calculate the standard deviation and mean diameter of the circles

Solution

First of all, let us make the data continuous.

Let us make the data continuous by taking classes as $32.5-36.5, 36.5-40.5, 40.5-44.5, 44.5-48.5, 48.5-52.5$.

Diameter	Mid value x_i	No. of Circles f_i	$y_i=\frac{x_i-42.5}{4}$	$f_i y_i$	$f_i y_i^2$
$32.5-36.5$	34.5	15	-2	-30	60
$36.5-40.5$	38.5	17	-1	-17	17
$40.5-44.5$	42.5	21	0	0	0
$44.5-48.5$	46.5	22	1	22	22
$48.5-52.5$	50.5	25	2	50	100
		$\Sigma f_i=100$		$\Sigma f_i y_i=25$	$\Sigma f_i y_i^2=199$

$N = \Sigma f_i = 100,$

$\Sigma f_i y_i = 25,$

$\Sigma f_i y_i^2 = 199$

$A = 42.5, h = 4$

We know that Variance

$$\sigma^2 = h^2\left[\frac{1}{N}\sum f_i y_i^2 - \left(\frac{\sum f_i y_i}{N}\right)^2\right]$$

$$= 16\left[\frac{1(199)}{100} - \left(\frac{25}{100}\right)^2\right]$$

$$= 16[1.99 - 0.0625]$$

$$\sigma^2 = 16 \times 1.9275 = 30.84$$

Standard deviation

$$(\sigma) = \sqrt{30.84} = 5.55$$

EXERCISE 15.3

1. From the data given below state which group is more variable, A or B ?

Marks	$10-20$	$20-30$	$30-40$	$40-50$	$50-60$	$60-70$	$70-80$
Group A	9	17	32	33	40	10	9
Group B	10	20	30	25	43	15	7

Solution

For group A

Marks	Group A f_i	Mid-point x_i	$y_i = \frac{x_i - 45}{10}$	y_i^2	$f_i y_i$	$f_i y_i^2$
$10-20$	9	15	-3	9	-27	81
$20-30$	17	25	-2	4	-34	68
$30-40$	32	35	-1	1	-32	32
$40-50$	33	45	0	0	0	0
$50-60$	40	55	1	1	40	40
$60-70$	10	65	2	4	20	40
$70-80$	9	75	3	9	27	81

	150				–6	342

Here, $N=150, h=10, A=45$

Mean,

$$\bar{x}=A+\frac{\sum_{i=1}^{7} f_i y_i}{N}\times h$$

$$=45+\frac{(-6)}{150}\times 10$$

$$=45-0.4$$

$$=44.6$$

Variance,

$$\left(\sigma_1^2\right)=\frac{h^2}{N^2}\left[N\sum_{i=1}^{7} f_i y_i-\left(\sum_{i=1}^{7} f_i y_i\right)^2\right]$$

$$=\frac{100}{22500}\left[150\times 342-(-6)^2\right]$$

$$=\frac{1}{225}\times 51264$$

$$=227.84$$

Standard deviation,

$$\left(\sigma_1\right)=\sqrt{227.84}=15.09$$

Group B

Marks	Group B f_i	Mid-point x_i	$y_i=\frac{x_i-45}{10}$	y_i^2	$f_i y_i$	$f_i y_i^2$
$10-20$	10	15	–3	9	–30	90
$20-30$	20	25	–2	4	–40	80
$30-40$	30	35	–1	1	–30	30
$40-50$	25	45	0	0	0	0
$50-60$	43	55	1	1	43	43
$60-70$	15	65	2	4	30	60
$70-80$	7	75	3	9	21	63
	150				–6	366

Mean,

$$\bar{x} = A + \frac{\sum_{i=1}^{7} f_i y_i}{N} \times h$$

$$= 45 + \frac{(-6)}{150} \times 10$$

$$= 45 - 0.4 = 44.6$$

Variance,

$$\sigma_2^2 = \frac{h^2}{N^2}\left[N\sum_{i=1}^{7} f_i y_i - \left(\sum_{i=1}^{7} f_i y_i\right)^2\right]$$

$$= \frac{100}{22500}\left[150 \times 366 - (-6)^2\right]$$

$$= \frac{1}{225} \times 54864$$

$$= 243.84$$

Standard deviation,

$$(\sigma_2) = \sqrt{243.84} = 15.61$$

For this mean of both these groups remains same, the group by greater standard deviation order be and variable.

Group B has more variability in the marks.

2. From the prices of shares X and Y below, find out which is more stable in value:

X	35	54	52	53	56	58	52	50	51	49
Y	108	107	105	105	106	107	104	103	104	101

Solution

The prices of the shares X are $35, 54, 52, 53, 56, 58, 52, 50, 51, 49$

Here, the number of observations, $N = 10$

Mean,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{10} x_i$$

$$= \frac{1}{10} \times 510$$

$$= 51$$

The following data is given

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
35	-16	256
54	3	9
52	1	1
53	2	4
56	5	25
58	7	49
52	1	1
50	-1	1
51	0	0
49	-2	4
		350

Variance

$$(\sigma_1^2) = \frac{1}{N}\sum_{i=1}^{10}(x_i - \bar{x})^2$$

$$= \frac{1}{10} \times 350$$

$$= 35$$

Standard deviation

$$(\sigma_1) = \sqrt{35} = 5.91$$

C.V.(Shares X) $=\frac{\sigma_1}{X}\times 100$

$=\frac{5.91}{51}\times 100$

$=11.58$

The prices of the shares Y are 108,107,105,105,106,107,104,103,104,101

The number of observations, $N=10$

Mean,

$\bar{y}=\frac{1}{N}\sum_{i=1}^{10} y_i$

$=\frac{1}{10}\times 1050=105$

The following table is corresponding to shares Y.

y_i	$(y_i-\bar{y})$	$(y_i-\bar{y})^2$
108	3	9
107	2	4
105	0	0
105	0	0
106	1	1
107	2	4
104	-1	1
103	-2	4
104	-1	1
101	-4	16
		60

Variance,

$$\left(\sigma_2^2\right)=\frac{1}{N}\sum_{i=1}^{10}\left(y_i-\bar{y}\right)^2$$

$$=\frac{1}{10}\times 40$$

$$=4$$

Standard deviation, $\left(\sigma_2\right)=\sqrt{4}=2$

C.V. (Shares Y) $=\frac{\sigma_2}{X}\times 100$

$$=\frac{2}{105}\times 100$$

$$=1.9$$

C.V of prices of shares X is greater than the C V of prices of shares Y.

Thus, the prices of shares Y are more stable than the prices of shares X .

3. An analysis of monthly wages paid to workers in two firms A and B, belonging to the same industry, gives the following results:

	Firm A	**Firm B**
No. of wage earners	586	648
Mean of monthly wages	Rs 5253	Rs 5253
Variance of the distribution of wages	100	100

(i) Which firm A or B pays larger amount as monthly wages?

(ii) Which firm, A or B, shows greater variability in individual wages?

Solution

(i) Monthly wages of firm $\mathrm{A}=\mathrm{Rs}5253$

Number of wage earners in firm $\mathrm{A}=586$

Total amount paid $=$ Rs. 5253×586

Monthly wages of firm $\mathrm{B}=\mathrm{Rs}5253$

Number of wage earners in firm $\mathrm{B}=648$

Total amount paid= Rs. 5253×648

So, firm B pays the larger amount as monthly wages as the number of wage earners in firm B are more than the number of wage earners in firm A .

(ii) Variance from this distribution of wages of firm $A\left(\sigma_1^2\right)=100$

Standard deviation from this distribution of wages within firm $A\left(\sigma_1\right)=\sqrt{100}=10$

Variance from this distribution of wages of firm $B\left(\sigma_1^2\right)=121$

$\therefore$ Standard deviation of the distribution of wages in firm $A\left(\sigma_1\right)=\sqrt{121}=11$

The mean of monthly wages of both the firms is same.

Therefore, this firm among greater standard deviation order have more variability.

So, firm B has greater variability within each individual wages

4. The following is the record of goals scored by team A in a football session:

No. of goals scored	**0**	**1**	**2**	**3**	**4**
No. of matches	**1**	**9**	**7**	**5**	**3**

For the team B, mean number of goals scored per match was 2 with a standard deviation 1.25 **goals. Find which team may be considered more consistent?**

Solution

The mean including standard deviation from goals scored on team A

No. of goals scored	No. of matches	$f_i x_i$	x_i^2	$f_i x_i^2$
0	1	0	0	0
1	9	9	1	9
2	7	14	4	28
3	5	15	9	45

4	3	12	16	48
	25	50		130

Mean

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{5} f_i x_i$$

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

$$= 2$$

Therefore, this mean of both these teams remains same

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

$$= \frac{1}{25}\sqrt{25\times 130 - (50)^2}$$

$$= \frac{1}{25}\sqrt{750}$$

$$= \frac{1}{25}\times 27.38$$

$$= 1.09$$

The standard deviation of team B is 1.25 goals.

The average number of goals scored on both these teams is the same

That is 2.

Accordingly, the team by lower standard deviation determination remain and consistent."

Thus, team A is more consistent than team B.

5. The sum and sum of squares corresponding to length x (in cm) and weight y (in gm) of 50 plant products are given below:

$$\sum_{i=1}^{50} x_i = 212, \sum_{i=1}^{50} x_i^2 = 902.8, \sum_{i=1}^{50} y_i = 261, \sum_{i=1}^{50} y_i^2 = 1457.6$$

Which is more varying, the length or weight?

Solution

Given,

$$\sum_{i=1}^{50} x_i = 212, \sum_{i=1}^{50} x_i^2 = 902.8$$

Here, $N = 50$

Mean,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{50} x_i$$

$$= \frac{212}{50}$$

$$= 4.24$$

Variance,

$$\left(\sigma_1^2\right) = \frac{1}{N}\sum_{i=1}^{50}\left(x_i - \bar{x}\right)^2$$

$$= \frac{1}{50}\sum_{i=1}^{50}\left(x_i - 4.24\right)^2$$

$$= \frac{1}{50}\sum_{i=1}^{50}\left[x_i^2 - 8.48x_i + 17.97\right]$$

$$= \frac{1}{50}\left[\sum_{i=1}^{50} x_i^2 - 8.48\sum_{i=1}^{50} x_i + 17.97 \times 50\right]$$

$$= \frac{1}{50}[902.8 - 8.48 \times (212) + 898.5]$$

$$= \frac{1}{50}[1801.3 - 1797.76]$$

$$= \frac{1}{50} \times 3.54$$

$$= 0.07$$

Standard variation $\sigma_2(\text{ length }) = \sqrt{0.07} = 0.26$

$$\text{C.V (length)} = \frac{\text{standard deviation}}{\text{mean}} \times 100$$

$$= \frac{0.26}{4.24} \times 100$$

$$= 6.13$$

$$\sum_{i=1}^{50} y_i = 261, \sum_{i=1}^{50} y_i^2 = 1457.6$$

Here, $N = 50$

Mean, $\bar{y} = \frac{1}{N}\sum_{i=1}^{50} y_i$

$$= \frac{261}{50}$$

$$= 5.22$$

Variance,

$$\left(\sigma_2^2\right) = \frac{1}{N}\sum_{i=1}^{50}\left(y_i - \bar{y}\right)^2$$

$$= \frac{1}{50}\sum_{i=1}^{50}\left(y_i - 5.22\right)^2$$

$$= \frac{1}{50}\sum_{i=1}^{50}\left[y_i^2 - 10.44 y_i + 27.24\right]$$

$$= \frac{1}{50}\left[\sum_{i=1}^{50} y_i^2 - 10.44\sum_{i=1}^{50} y_i + 27.24 \times 50\right]$$

$$= \frac{1}{50}[1457.6 - 10.44 \times (261) + 1362]$$

$$= \frac{1}{50}[2819.6 - 2724.84]$$

$$= \frac{1}{50} \times 94.76$$

$$= 1.89$$

Standard variation $\sigma_2(\text{ weight }) = \sqrt{1.89} = 1.37$

C.V (weight) $= \frac{\text{standard deviation}}{\text{mean}} \times 100$

$$= \frac{1.37}{5.22} \times 100$$

$$= 26.24$$

So, C.V from weights comprises greater than C.V from lengths,

Therefore, weights vary more than the lengths.

Miscellaneous Exercise On Chapter 15

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.

Solution

Let the last two observations be x and y.

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

Mean,

$$\overline{x} = \frac{6+7+10+12+12+13+x+y}{8}$$

$$9 = \frac{60+x+y}{8}$$

$$60+x+y = 72$$

$$x+y = 12 \text{..........}(1)$$

Variance,

$$9.25 = \frac{1}{n}\sum_{i=1}^{8}\left(x_i - \overline{x}\right)^2$$

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

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

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

$$9.25 = \frac{1}{8}\left[x^2+y^2-6\right] \quad x^2+y^2 = 80 \quad \text{...........}(2)$$

From (1), we obtain

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

From (2) and (3), we obtain

$$2xy = 64 \quad \text{..............}(4)$$

Subtracting (4) from (2), we obtain

$$x^2+y^2-2xy=16$$
$$x-y=\pm 4 \ldots\ldots\ldots\ldots\ldots(5)$$

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

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

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

So, the two observations are 4 and 8.

2. The mean and variance of 7 observations are 8 and 16, respectively. If five of the observations are $2,4,10,12,14$. Find the remaining two observations.

Solution

Let the last two observations be x and y.

Therefore, the observations are $2,4,10,12,14,x,y$

Mean,

$$\bar{x}=\frac{2+4+10+12+14+x+y}{7}$$

$$8=\frac{42+x+y}{7}$$
$$42+x+y=56$$
$$x+y=14 \ \ldots\ldots\ldots\ldots(1)$$

Variance,

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

$$16=\frac{1}{7}\left[(-6)^2+(-4)^2+(2)^2+(4)^2+(6)^2+x^2+y^2-2\times 8(x+y)+2\times(8)^2\right]$$

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

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

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

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

From (1), we obtain

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

From (2) and (3), we obtain

$2xy = 196 - 100$

$2xy = 96$ ....(4)

Subtracting (4) from (2), we obtain

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

$(x-y)^2 = 4$

$x - y = \pm 2$

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

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

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

So, the last observations are 6 and 8.

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.

Solution

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

It implies that the mean equals 8 and standard deviation equals 4.

Mean,

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6}{6} = 8$$

If every observation involves multiplied with 3 and this resulting observations do y_i,

then $y_i = 3x_i$

That is $x_i = \frac{1}{3} y_i$, for $i = 1$ to 6

So, new mean,

$$\bar{y} = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{6}$$

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

$$= 3 \times 8 = 24$$

Standard deviation,

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

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

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

From both equations

it can be observed that, $\bar{y} = 3\bar{x}$ and $\bar{x} = \frac{1}{3}\bar{y}$

Substituting the values of x_1 and $\bar{x}$ in $\sum_{i=1}^{6}(x_1 - \bar{x})^2 = 96$,

we obtain

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

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

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

So, the standard deviation of new observations equals $\sqrt{144} = 12$

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

Solution

The given n observations are $x_1, x_2, \ldots, x_n$

Mean $= \bar{x}$

Variance $== \sigma^2$

Therefore,

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

If every observation is involved multiplied with a and this resulting observations will

y_i, then $y_i = ax_i$ i.e., $x_i = \frac{1}{a} y_i$

Hence,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$= \frac{1}{n}\sum_{i=1}^{n} ax_i$$

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

$$= a\bar{x}$$

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

Add the values of x_i and $\bar{x}$ in $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} y_i\left(x_i - \bar{x}\right)^2$, we obtained

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

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

So, the variance of the observations, $ax_1, ax_2, \ldots\ldots\ldots\ldots..x_n$, is $a^2\sigma^2$

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.

Solution

(i) Number of observations $(n)=20$

Incorrect mean $=10$

Incorrect standard deviation $=2$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{20} x_i$$

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

$$\sum_{i=1}^{20} x_i=200$$

That is, incorrect sum of observations $=200$

Correct sum of observations $=200-8=192$

Therefore, correct mean $=\dfrac{\text{correct sum}}{19}=\dfrac{192}{19}=10.1$

Standard deviation,

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

$$2=\sqrt{\frac{1}{20}\text{ Incorrect }\sum_{i=1}^{n} x_i^2-(10)^2}$$

$$4=\frac{1}{20}\text{ Incorrect }\sum_{i=1}^{n} x_i^2-100$$

$$\text{Incorrect }\sum_{i=1}^{n} x_i^2=2080$$

$$\text{Correct }\sum_{i=1}^{n} x_i^2=\text{ Incorrect }\sum_{i=1}^{n} x_i^2-(8)^2$$

$$=2080-64$$

$$=2016$$

Correct standard deviation $=\sqrt{\dfrac{\text{Correct }\sum_{i=1}^{n} x_i^2}{n}-(\text{ correct mean })^2}$

$$= \sqrt{\frac{2016}{19} - \left(\frac{192}{19}\right)^2}$$

$$= \sqrt{\frac{1440}{361}}$$

$$= \sqrt{3.988}$$

$$= 1.99$$

(ii) When 8 is replaced by 12,

Incorrect sum of observations $= 200$

Correct sum of observations $= 200 - 8 + 12 = 204$

Hence, Correct mean $= \frac{\text{correct sum}}{20} = \frac{204}{20} = 10.2$

Standard deviation,

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

$$2 = \sqrt{\frac{1}{20}\ \text{Incorrect} \sum_{i=1}^{n} x_i^2 - (10)^2}$$

$$4 = \frac{1}{20}\ \text{Incorrect} \sum_{i=1}^{n} x_i^2 - 100$$

$$\text{Incorrect} \sum_{i=1}^{n} x_i^2 = 2080$$

So, Correct $\sum_{i=1}^{n} x_i^2 =$ Incorrect $\sum_{i=1}^{n} x_i^2 - (8)^2 + (12)^2$

$$= 2080 - 64 + 144$$

$$= 2160$$

Correct standard deviation $= \sqrt{\frac{\text{Correct} \sum x_i^2}{n} - (\text{ correct mean })^2}$

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

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

$= \sqrt{3.96}$

$= 1.98$

6. The mean and standard deviation of marks obtained by 50 students of a class in three subjects, Mathematics, Physics and Chemistry are given below:

Subject	Mathematics	Physics	Chemistry
Mean	42	32	40.9
Standard deviation	12	15	20

Which of the three subjects shows the highest variability in marks and which shows the lowest?

Solution

Given

Standard deviation of mathematics $= 12$

Standard deviation of Physics $= 15$

Standard deviation of Chemistry $= 20$

The coefficient of variation (C.V) is given by $\frac{\text{standard deviation}}{\text{mean}} \times 100$

$$C.V(\text{ Mathematics }) = \frac{12}{42} \times 100 = 28.57$$

$$C.V(\text{ Physics }) = \frac{15}{32} \times 100 = 46.87$$

$$C.V(\text{ Chemistry }) = \frac{20}{40.9} \times 100 = 48.89$$

The subject with greater C.V is more variable than others.

Since, this highest variability into marks remains in Chemistry and the lowest variability in marks is in Mathematics.

7. 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.

Solution

Number of observations $= 100$

Incorrect mean $(\overline{x}) = 20$

Incorrect standard deviation $(\sigma) = 3$

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

$$\sum_{i=1}^{100} x_i = 20 \times 100$$

$$= 2000$$

Incorrect sum of observations $= 2000$

Correct sum of observations $= 2000 - 21 - 21 - 18 = 2000 - 60 = 1940$

Therefore, Correct mean $= \frac{\text{correct sum}}{100 - 3} = \frac{1940}{97} = 20$

Standard deviation

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

$$\Rightarrow 3 = \sqrt{\frac{1}{100} \times \text{ Incorrect } \sum x_i^2 - (20)^2}$$

$$\Rightarrow \text{ Incorrect } \sum x_i^2 = 100(9 + 400) = 40900$$

$$\text{Correct } \sum_{i=1}^{n} x_i^2 = \text{ Incorrect } \sum_{i=1}^{n} x_i^2 - (21)^2 - (21)^2 - (18)^2$$

$$= 40900 - 441 - 441 - 324$$

$$= 39694$$

Correct standard deviation

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

$$= \sqrt{\frac{39694}{97} - (20)^2}$$

$$= \sqrt{409.216 - 400}$$

$$= \sqrt{9.216}$$

$$= 3.036$$