statistics

13 Chapter

Exercise 13.1

1. Find the mean deviation about the mean for the data $4, 7, 8, 9, 10, 12, 13, 17$.

Ans: Consider the given data, which is, $4, 7, 8, 9, 10, 12, 13, 17$.

Therefore, the mean of the data is,

$$\bar{x} = \frac{4+7+8+9+10+12+13+17}{8} = \frac{80}{8} = 10$$

Observe that the deviations of the respective observations from the mean $(\bar{x})$, that is, $x_i - \bar{x}$ can be calculated as, $-6, -3, -2, -1, 0, 2, 3, 7$ and therefore, the absolute value of the deviations calculated by $|x_i - \bar{x}|$ are $6, 3, 2, 1, 0, 2, 3, 7$.

Now, the mean deviation about the mean is,

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

2. Find the mean deviation about the mean for the data $38, 70, 48, 40, 42, 55, 63, 46, 54, 44$.

Ans: Consider the given data, which is, $38, 70, 48, 40, 42, 55, 63, 46, 54, 44$.

Therefore, the mean of the data is,

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

Observe that the deviations of the respective observations from the mean $(\bar{x})$, that is, $x_i - \bar{x}$ can be calculated as, $-12, 20, -2, -10, -8, 5, 13, -4, 4, -6$ and therefore, the absolute value of the deviations calculated by $|x_i - \bar{x}|$ are $12, 20, 2, 10, 8, 5, 13, 4, 4, 6$.

Now, the mean deviation about the mean is,

$$\therefore \text{M.D.}(\bar{x}) = \frac{\sum_{i=1}^{10}|x_i - \bar{x}|}{10} = \frac{12+20+2+10+8+5+13+4+4+6}{10} = \frac{84}{10} = 8.4$$

3. **Find the mean deviation about the median for the data $13,17,16,14,11,13,10,16,11,18,12,17$.**

Ans: Consider the given data, which is, $13,17,16,14,11,13,10,16,11,18,12,17$.

Observe that the number of observations in this case is 12, that is, even and on arranging the data in ascending order it can be obtained as, $10,11,11,12,13,13,14,16,16,17,17,18$ Therefore, the median of the data is the average of the 6^{th} and the 7^{th} observations**,**

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

Observe that the deviations of the respective observations from the median (M) , that is, $x_i - M$ can be calculated as, $-3.5,-2.5,-2.5,-1.5,-0.5,-0.5,0.5,2.5,2.5,3.5,3.5,4.5$ and therefore, the absolute value of the deviations calculated by $|x_i - M|$ are $3.5,2.5,2.5,1.5,0.5,0.5,0.5,2.5,2.5,3.5,3.5,4.5$.

Now, the mean deviation about the median is,

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

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

$$\Rightarrow \text{M.D.}(M) = \frac{28}{12} = 2.33$$

4. **Find the mean deviation about the median for the data $36,72,46,42,60,45,53,46,51,49$.**

Ans: Consider the given data, which is, $36,72,46,42,60,45,53,46,51,49$**.**

Observe that the number of observations in this case is 10, that is, even and on arranging the data in ascending order it can be obtained as, $36,42,45,46,46,49,51,53,60,72$ Therefore, the median of the data is the average of the 5^{th} and the 6^{th} observations**,**

$$\therefore M = \frac{46+49}{2} = \frac{95}{2} = 47.5$$

Observe that the deviations of the respective observations from the median (M), that is, $x_i - M$ can be calculated as,

$-11.5, -5.5, -2.5, -1.5, -1.5, 1.5, 3.5, 5.5, 12.5, 24.5$ and therefore, the absolute value of the deviations calculated by $|x_i - M|$ are

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

Now, the mean deviation about the median is,

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

$$= \frac{11.5+5.5+2.5+1.5+1.5+1.5+3.5+5.5+12.5+24.5}{10} \Rightarrow \text{M.D.}(M) = \frac{70}{10} = 7$$

5. **Find the mean deviation about the mean for the data.**

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

Ans: Consider the given data and observe the table as shown below, which is,

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

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = \frac{1}{N}\sum_{i=1}^{5} f_i x_i = \frac{1}{25} \times 350 = 14$$

Therefore, the mean deviation about the mean is,

$$\text{M.D.}(\bar{x}) = \frac{\sum_{i=1}^{5} f_i |x_i - \bar{x}|}{N} = \frac{158}{25} = 6.32$$

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

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

Ans: Consider the given data and observe the table as shown below, which is,

x_i	f_i	$f_i x_i$	$\left\|x_i - \bar{x}\right\|$	$f_i \left\|x_i - \bar{x}\right\|$
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

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = \frac{1}{N}\sum_{i=1}^{5} f_i x_i = \frac{1}{80} \times 4000 = 50$$

Therefore, the mean deviation about the mean is,

$$\text{M.D.}(\bar{x}) = \frac{\sum_{i=1}^{5} f_i \left|x_i - \bar{x}\right|}{N} = \frac{1280}{80} = 16$$

7. Find the mean deviation about the median for the data.

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

Ans: It can be clearly observed that the given observations are already in ascending order and hence on adding a column corresponding to the cumulative frequencies of the given data, the following table can be obtained as shown below,

x_i	f_i	c.f.
5	8	8
7	6	14

9	2	16
10	2	18
12	2	20
15	6	26

Observe that the number of observations in this case is 26, that is, even and therefore, the median is the mean of the 13^{th} and the 14^{th} observations. Observe that both of these observations lie in the cumulative frequency 14, for which the corresponding observation is obtained as 7,

$$\therefore M = \frac{7+7}{2} = \frac{14}{2} = 7$$

Now, the absolute values of the deviations from the median can be calculated using $|x_i - M|$ and therefore observe the table as shown below,

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

Therefore, the mean deviation about the mean is,

$$\text{M.D.}(M) = \frac{\sum_{i=1}^{6} f_i |x_i - M|}{N} = \frac{16+4+6+10+48}{26} = \frac{84}{26} = 3.23$$

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**

Ans: It can be clearly observed that the given observations are already in ascending order and hence on adding a column corresponding to the cumulative frequencies of the given data, the following table can be obtained as shown below,

x_i	f_i	c.f.
15	3	3
21	5	8

27	6	14
30	7	21
35	8	29

Observe that the number of observations in this case is 29, that is, odd and therefore, the median is the 15^{th} observation. Observe that this observations lie in the cumulative frequency 21, for which the corresponding observation is obtained as 30,

$\therefore M = 30$

Now, the absolute values of the deviations from the median can be calculated using $|x_i - M|$ and therefore observe the table as shown below,

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

Therefore, the mean deviation about the median is,

$$M.D.(M) = \frac{\sum_{i=1}^{5} f_i |x_i - M|}{N} = \frac{45+45+18+0+40}{29} = \frac{148}{29} = 5.1$$

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

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

Ans: Consider the given data and observe the table as shown below, which is,

Income per day	Number of persons (f_i)	Midpoint (x_i)	$f_i x_i$	$\left\|x_i - \bar{x}\right\|$	$f_i \left\|x_i - \bar{x}\right\|$
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	4	550	2750	192	960
600 - 700	5	650	2600	292	1168
700 - 800	3	750	2250	392	1176
	50		17900		7896

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = \frac{1}{N}\sum_{i=1}^{8} f_i x_i = \frac{1}{50} \times 17900 = 358$$

Therefore, the mean deviation about the mean is,

$$\text{M.D.}\left(\bar{x}\right) = \frac{\sum_{i=1}^{8} f_i \left|x_i - \bar{x}\right|}{N} = \frac{7896}{50} = 157.92$$

10. **Find the mean deviation about the mean for the data.**

Height in cms	**Number of boys**
95-105	**9**
105-115	**13**
115-125	**26**

125-135	**30**
135-145	**12**
145-155	**10**

Ans: Consider the given data and observe the table as shown below, which is,

Height in cms	Number of boys (f_i)	Midpoint (x_i)	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
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

The mean of the data can be calculated as shown below,

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

Therefore, the mean deviation about the mean is,

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

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

Ans:

Marks	Mid values x_i	f_i	e. f.	$[x_i-27.86]$	$F_i[x_i-27.86]$
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
		50			517.16

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

$\therefore$ Median class is $20 - 30$ $\therefore$

Median = $20 + \frac{25-14}{14} \times 10 = 20 + 7.86 = 27.86$

M.D. about median = $\frac{1}{N}\sum_{i=1}^{n} fi\ [X_i - M] = \frac{1}{50} \times 517.16 = 10.34$

12. Calculate the mean deviation about the median age for the age distribution of 100 persons.

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

Ans: It can be clearly observed that the given data is not continuous and therefore it needs to be converted into continuous frequency distribution, which can be done by subtracting 0.5 from the lower limit and adding 0.5 to the upper limit of each class interval. Now, observe the table as shown below, which is,

Age	Number (f_i)	Cumulative Frequency (c.f.)	Midpoint (x_i)	$\|x_i - M\|$	$f_i\|x_i - M\|$
$15.5-20.5$	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

Observe that the class interval containing the $\left(\frac{N}{2}\right)^{th}$ item or the 50^{th} item is $35.5-40.5$. Thus, $35.5-40.5$ is the median class.

Therefore, median of the data can be calculated as shown below,

$$\therefore M = l + \frac{\frac{N}{2}-C}{f}\times h \quad [\text{ where } l = 35.5,\ C = 37,\ f = 26,\ h = 5 \text{ and } N = 100]$$

$$\Rightarrow M = 35.5 + \frac{50-37}{26}\times 5 = 35.5 + \frac{13}{26}\times 5 = 35.5 + 2.5 = 38$$

Therefore, the mean deviation about the median is,

$$\text{M.D.}(M) = \frac{\sum_{i=1}^{8} f_i\left|x_i - M\right|}{N} = \frac{735}{100} = 7.35$$

Exercise 13.2

1. Find the mean and variance for the data $6, 7, 10, 12, 13, 4, 8, 12$.

Ans: Consider the given data which is, $6, 7, 10, 12, 13, 4, 8, 12$.

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = \frac{1}{n}\sum_{i=1}^{8} x_i = \frac{6+7+10+12+13+4+8+12}{8} = \frac{72}{8} = 9$$

Now, observe the table as shown below, which is,

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

Therefore, the variance is,

$$\mathrm{Var}\left(\sigma^2\right)=\frac{\sum_{i=1}^{8}\left(x_i-\bar{x}\right)^2}{n}=\frac{74}{8}=9.25$$

2. Find the mean and variance for the first n natural numbers.

Ans: The mean of the first n natural numbers can be calculated as shown below,

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

Now, the variance can be calculated as,

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

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

$$\Rightarrow \mathrm{Var}\left(\sigma^2\right)=\frac{1}{n}\sum_{i=1}^{n}x_i^{\,2}-\frac{1}{n}\sum_{i=1}^{n}2\left(\frac{n+1}{n}\right)x_i+\frac{1}{n}\sum_{i=1}^{n}\left(\frac{n+1}{2}\right)^2$$

$$\Rightarrow \mathrm{Var}\left(\sigma^2\right)=\frac{(n+1)(2n+1)}{6}-\frac{(n+1)^2}{2}+\frac{(n+1)^2}{4}$$

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

$$\Rightarrow \mathrm{Var}\left(\sigma^2\right)=\frac{n^2-1}{12}$$

3. Find the mean and variance for the first 10 multiples of 3.

Ans: Observe that the first 10 multiples of 3 are, $3,6,9,12,15,18,21,24,27,30$.

The mean of the data can be calculated as shown below,

$$\therefore \bar{x}=\frac{1}{n}\sum_{i=1}^{10}x_i=\frac{3+6+9+12+15+18+21+24+27+30}{10}=\frac{165}{10}=16.5$$

Now, observe the table as shown below, which is,

x_i	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$
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
		742.5

Therefore, the variance is,

$$\mathrm{Var}\left(\sigma^2\right) = \frac{\sum_{i=1}^{10}\left(x_i - \bar{x}\right)^2}{n} = \frac{742.5}{10} = 74.25$$

4. **Find the mean and variance for the data.**

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

Ans: Consider the given data and observe the table as shown below, which is,

x_i	f_i	$f_i x_i$	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$	$f_i\left(x_i - \bar{x}\right)^2$
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
	30	760			1736

The mean of the data can be calculated as shown below**,**

$$\therefore \bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{1}{40} \times 760 = 19$$

Therefore, the variance is,

$$\mathrm{Var}\left(\sigma^2\right) = \frac{\sum_{i=1}^{7} f_i\left(x_i - \bar{x}\right)^2}{N} = \frac{1736}{40} = 43.4$$

5. Find the mean and variance for the data.

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

Ans: Consider the given data and observe the table as shown below, which is,

x_i	f_i	$f_i x_i$	$\left(x_i - \bar{x}\right)$	$\left(x_i - \bar{x}\right)^2$	$f_i\left(x_i - \bar{x}\right)^2$
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

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{1}{22} \times 2200 = 100$$

Therefore, the variance is,

$$\operatorname{Var}\left(\sigma^2\right) = \frac{\sum_{i=1}^{7} f_i \left(x_i - \bar{x}\right)^2}{N} = \frac{640}{22} = 29.09$$

6. Find the mean, variance and standard deviation using the shortcut method.

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

Ans: Consider the given data and observe the table as shown below, which is,

x_i	f_i	$y_i = \frac{x_i - 64}{1}$	y_i^2	$f_i y_i$	$f_i y_i^2$
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	5	3	9	12	36
68	4	4	16	20	80
	100			0	286

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = A + \frac{\sum_{i=1}^{9} f_i y_i}{N} \times h = 64 + \frac{0}{100} \times 1 = 64$$

Therefore, the variance is,

$$\operatorname{Var}\left(\sigma^2\right) = \frac{h^2}{N^2}\left[N\sum_{i=1}^{9} f_i y_i^{\,2} - \left(\sum_{i=1}^{9} f_i y_i\right)^2\right] = \frac{1}{(100)^2}\left[100 \times 286 - 0\right] = 2.86$$

Now the standard deviation can be calculated as shown below,

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

7. Find the mean and variance for the following frequency distribution.

Classes	**0-30**	**30-60**	**60-90**	**90-120**	**120-150**	**150-180**	**180-210**
Freque-ncies	**2**	**3**	**5**	**10**	**3**	**5**	**2**

Ans: Consider the given data and observe the table as shown below, which is,

Class	Frequency f_i	Midpoint (x_i)	$y_i = \frac{x_i - 105}{30}$	$y_i^{\,2}$	$f_i y_i$	$f_i y_i^{\,2}$
$0-30$	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
	30				2	76

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = A + \frac{\sum_{i=1}^{7} f_i y_i}{N} \times h = 105 + \frac{2}{30} \times 30 = 107$$

Therefore, the variance is,

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

$$= \frac{(30)^2}{(30)^2}\left[30 \times 76 - (2)^2\right] = 2280 - 4 = 2276$$

8. Find the mean and variance for the following frequency distribution.

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

Ans: Consider the given data and observe the table as shown below, which is,

Class	Frequency f_i	Midpoint (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

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = A + \frac{\sum_{i=1}^{5} f_i y_i}{N} \times h = 25 + \frac{10}{50} \times 10 = 27$$

Therefore, the variance is,

$$\mathrm{Var}\left(\sigma^2\right)=\frac{h^2}{N^2}\left[N\sum_{i=1}^{5}f_iy_i^{\,2}-\left(\sum_{i=1}^{5}f_iy_i\right)^2\right]=\frac{(10)^2}{(50)^2}\left[50\times 68-(10)^2\right]$$

$$\Rightarrow \mathrm{Var}\left(\sigma^2\right)=\frac{1}{25}[3400-100]=132$$

9. Find the mean, variance and standard deviation using the shortcut method.

Height in cms	**Number of children**
70-75	**3**
75-80	**4**
80-85	**7**
85-90	**7**
90-95	**15**
95-100	**9**
100-105	**6**
105-110	**6**
110-115	**3**

Ans: Consider the given data and observe the table as shown below, which is,

Class interval	Frequency f_i	Midpoint (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

The mean of the data can be calculated as shown below,

$$\therefore \bar{x} = A + \frac{\sum_{i=1}^{9} f_i y_i}{N} \times h = 92.5 + \frac{6}{60} \times 5 = 92.5 + 0.5 = 93$$

Therefore, the variance is,

$$Var\left(\sigma^2\right) = \frac{h^2}{N^2}\left[N\sum_{i=1}^{9} f_i y_i^2 - \left(\sum_{i=1}^{9} f_i y_i\right)^2\right] = \frac{(5)^2}{(100)^2}\left[60 \times 254 - (6)^2\right]$$

$$\Rightarrow Var\left(\sigma^2\right) = \frac{25}{3600}(15204) = 105.58$$

Now the standard deviation can be calculated as shown below,

$$\therefore \sigma = \sqrt{105.58} = 10.27$$

10. The diameters of circles (in mm) drawn in a design are given below. Find the mean, variance and standard deviation using the shortcut method.

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

Ans: Consider the given data and observe the table as shown below, which is,

Class interval	Frequency f_i	Midpoint (x_i)	$y_i = \frac{x_i - 92.5}{5}$	${y_i}^2$	f_iy_i	$f_i{y_i}^2$
$32.5-36.5$	15	34.5	–2	4	–30	60
$36.5-40.5$	17	38.5	–1	1	–17	17
$40.5-44.5$	21	42.5	0	0	0	0
$44.5-48.5$	22	46.5	1	1	22	22
$48.5-52.5$	25	50.5	2	4	50	100
	100				25	199

The mean of the data can be calculated as shown below,

$$\therefore \overline{x} = A + \frac{\sum_{i=1}^{5} f_iy_i}{N} \times h = 42.5 + \frac{25}{100} \times 4 = 43.5$$

Therefore, the variance is,

$$\text{Var}\left(\sigma^2\right) = \frac{h^2}{N^2}\left[N\sum_{i=1}^{5} f_i{y_i}^2 - \left(\sum_{i=1}^{5} f_iy_i\right)^2\right] = \frac{(4)^2}{(100)^2}\left[100 \times 199 - (25)^2\right]$$

$$\Rightarrow \text{Var}\left(\sigma^2\right) = \frac{16}{10000}(19900 - 625) = \frac{16}{10000}(19275) = 30.84$$

Now the standard deviation can be calculated as shown below,

$$\therefore \sigma = \sqrt{30.84} = 5.55$$

Miscellaneous Exercise

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.

Ans: Consider the remaining two observations to be x and y. Thus the eight observations are $6,7,10,12,12,13,x,y$.

Therefore, from the mean of the data it can be obtained that,

$$\overline{x} = \frac{6+7+10+12+12+13+x+y}{8} = 9$$

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

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

Again, from the variance of the data it can be obtained that,

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

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

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

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

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

Observe that form equation (i), it can be obtained that,

$x^2+y^2+2xy=144$ ......(iii)

Also, from equations (ii) and (iii), it can be obtained that,

$2xy=64$ ......(iv)

Now, on subtracting equation (iv) from equation (ii), it can be obtained that,

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

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

It can be calculated from equations (i) and (v) that, $x=8$ and $y=4$, when $x-y=4$ and $x=4$ and $y=8$, when $x-y=-4$. Henceforth, the remaining 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 $2,4,10,12$ and 14. Find the remaining two observations.

Ans: Consider the remaining two observations to be x and y. Thus the eight observations are $2,4,10,12,14,x,y$.

Therefore, from the mean of the data it can be obtained that**,**

$$\bar{x}=\frac{2+4+10+12+14+x+y}{7}=8$$

$\Rightarrow 42+x+y=56$

$\Rightarrow x+y=14$ ......(i)

Again, from the variance of the data it can be obtained that**,**

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

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

$$\Rightarrow \frac{1}{7}\left[36+16+4+16+x^2+y^2-(16)(14)+128\right]=16 \quad \text{[By, using (i)]}$$

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

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

Observe that form equation (i), it can be obtained that,

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

Also, from equations (ii) and (iii), it can be obtained that,

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

Now, on subtracting equation (iv) from equation (ii), it can be obtained that,

$x^2 + y^2 - 2xy = 4$

$\Rightarrow x - y = \pm 2$ ......(v)

It can be calculated from equations (i) and (v) that, $x = 8$ and $y = 6$, when $x - y = 2$ and $x = 6$ and $y = 8$, when $x - y = -2$. Henceforth, the remaining two 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.

Ans: Assume the observations to be $x_1, x_2, x_3, x_4, x_5, x_6$.

Therefore, the mean of the data is,

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

$$\Rightarrow \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6}{6} = 8 \text{ ......(i)}$$

Observe that when each of the observation is multiplied with 3 and if we consider the resulting observations as y_i, then observe as shown below,

$$\therefore y_1 = 3x_1$$

$$\Rightarrow x_1 = \frac{1}{3} y_1, \forall i = 1,2,3,....,6 \;\&\; i \in \mathbb{Z}^+$$

Now, the mean of the new data is,

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

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

$$\Rightarrow \bar{y} = 3 \times 8 \quad \text{[By using (i)]}$$

$$\Rightarrow \bar{y} = 24$$

Therefore, the standard deviation of the data is,

$$\sigma = \sqrt{\frac{\sum_{i=1}^{6} \left(x_i - \bar{x}\right)^2}{n}} = 4$$

$$\Rightarrow (4)^2 = \frac{\sum_{i=1}^{6}\left(x_i - \overline{x}\right)^2}{6}$$

$$\Rightarrow \sum_{i=1}^{6}\left(x_i - \overline{x}\right)^2 = 96 \text{ ......(ii)}$$

Again, it can be observed from equations (i) and (ii) that $\overline{y} = 3\overline{x}$ and $\overline{x} = \frac{\overline{y}}{3}$ and hence on substituting the values of x_i and $\overline{x}$ in equation (ii) it can be clearly obtained as shown below,

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

$$\Rightarrow \sum_{i=1}^{6}\left(y_i - \overline{y}\right)^2 = 864$$

Henceforth, the standard deviation of the new data can be calculated as shown below**,**

$$\therefore \sigma = \sqrt{\frac{864}{6}} = \sqrt{144} = 12$$

4. Given that $\overline{x}$ 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 observations $ax_1, ax_2, ax_3, ..., ax_n$ are $a\overline{x}$ and $a^2\sigma^2$, respectively $(a \neq 0)$.

Ans: Observe that the given observations are $x_1, x_2, ..., x_n$ and the mean and variance of data is $\overline{x}$ and σ^2 respectively.

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

Observe that when each of the observation is multiplied with a and if we consider the resulting observations as y_i, then observe as shown below,

$$\therefore y_i = ax_i$$

$$\Rightarrow x_i = \frac{1}{a} y_i, \forall i = 1, 2, 3, ..., n \ \& \ i \in \mathbb{Z}^+$$

Now, the mean of the new data is**,**

$$\therefore \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} ax_i = \frac{a}{n}\sum_{i=1}^{n} x_i = a\overline{x} \quad \left[\because \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\right]$$

Again on substituting the values of x_i and $\bar{x}$ in equation (i) it can be clearly obtained as shown below,

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

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

Henceforth, it can be clearly proved that the mean and variance of the new data is $a\bar{x}$ and $a^2\sigma^2$, respectively.

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

Ans: Observe that the incorrect number of observations, incorrect mean and the incorrect standard deviation are 20, 10 and 2, respectively. $x_1, x_2, ..., x_n$ and the mean and variance of data is $\bar{x}$ and σ^2 respectively.

$$\therefore \bar{x} = \frac{\sum_{i=1}^{20} x_i}{n} = 10$$

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

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

Now, the correct mean is**,**

$$\therefore \bar{x} = \frac{\sum_{i=1}^{19} x_i}{n}$$

$$\Rightarrow \bar{x} = \frac{192}{19} = 10.1 \quad \left[\because \sum_{i=1}^{19} x_i = 192\right]$$

Again observe as shown below,

$$\therefore \sigma = \sqrt{\frac{1}{20}\sum_{i=1}^{20} x_i^{\ 2} - \left(\overline{x}\right)^2} = 2$$

$$\Rightarrow 4 = \frac{1}{20}\sum_{i=1}^{20} x_i^{\ 2} - \left(\overline{x}\right)^2$$

$$\Rightarrow 2080 = \sum_{i=1}^{20} x_i^{\ 2}$$

Therefore, the correct standard deviation of the data is**,**

$$\sum_{i=1}^{20} x_i^{\ 2} - (8)^2$$

$$\Rightarrow 2080 - 64$$

$$\Rightarrow 2016$$

$$\therefore \sigma = \sqrt{\frac{\sum_{i=1}^{19} x_i^{\ 2}}{n} - \left(\overline{x}\right)^2}$$

$$\Rightarrow \sigma = \sqrt{\frac{2016}{19} - (10.1)^2}$$

$$\Rightarrow \sigma = \sqrt{1061.1 - 102.1} = \sqrt{4.09} = 2.02$$

(ii) If it is replaced by $\mathbf{12}$.

Ans: Observe that the incorrect sum of observations is 200.

$$\therefore \sum_{i=1}^{20} x_i = 200 - 8 + 12 = 204$$

Now, the correct mean is**,**

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

$$\Rightarrow \overline{x} = \frac{204}{20} = 10.2 \quad \left[\because \sum_{i=1}^{20} x_i = 204\right]$$

Again observe as shown below,

$$\therefore \sigma = \sqrt{\frac{1}{20}\sum_{i=1}^{20} x_i^{\ 2} - \left(\overline{x}\right)^2} = 2$$

$$\Rightarrow 4 = \frac{1}{20}\sum_{i=1}^{20} x_i^{\ 2} - \left(\overline{x}\right)^2$$

$$\Rightarrow 2080 = \sum_{i=1}^{20} x_i^2$$

Therefore, the correct standard deviation of the data is**,**

$$\sum_{i=1}^{20} x_i^2 - (8)^2 + (12)^2$$

$$\Rightarrow 2080 - 64 + 144$$

$$\Rightarrow 2160$$

$$\therefore \sigma = \sqrt{\frac{\sum_{i=1}^{20} x_i^2}{n} - \left(\overline{x}\right)^2}$$

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

$$\Rightarrow \sigma = \sqrt{108 - 104.04} = \sqrt{3.96} = 1.98$$

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 3 observations are incorrect which were recorded as $21, 21$ and 18. Find the mean and standard deviation if the incorrect observations are omitted.

Ans: Observe that the number of observations, incorrect mean and the incorrect standard deviation are 100, 20 and 3, respectively. $x_1, x_2, ..., x_n$ and the mean and variance of data is $\overline{x}$ and σ^2 respectively.

$$\therefore \overline{x} = \frac{\sum_{i=1}^{100} x_i}{n} = 20$$

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

$$\Rightarrow \sum_{i=1}^{100} x_i = 2000$$

Now, the correct mean is**,**

$$\therefore \bar{x} = \frac{\sum_{i=1}^{97} x_i}{n}$$

$$\Rightarrow \bar{x} = \frac{2000 - 60}{97} = 20 \quad \left[\because \sum_{i=1}^{97} x_i = 1940\right]$$

Again observe as shown below,

$$\therefore \sigma = \sqrt{\frac{1}{100}\sum_{i=1}^{100} x_i^{\,2} - \left(\bar{x}\right)^2} = 3$$

$$\Rightarrow 9 = \frac{1}{100}\sum_{i=1}^{100} x_i^{\,2} - \left(\bar{x}\right)^2$$

$$\Rightarrow 40900 = \sum_{i=1}^{100} x_i^{\,2}$$

Therefore, the correct standard deviation of the data is**,**

$$\sum_{i=1}^{100} x_i^{\,2} - (2)\left(21\right)^2 - (18)^2$$

$$\Rightarrow 40900 - 1206$$

$$\Rightarrow 39694$$

$$\therefore \sigma = \sqrt{\frac{\sum_{i=1}^{97} x_i^{\,2}}{n} - \left(\bar{x}\right)^2}$$

$$\Rightarrow \sigma = \sqrt{\frac{39694}{97} - \left(20\right)^2}$$

$$\Rightarrow \sigma = \sqrt{409.22 - 400} = \sqrt{9.22} = 3.04$$