MEASURES OF DISPERSION, MOMENTS AND SKEWNESS

4.4 (b) (i) To find the quartile deviation graphically, we locate the two quartiles graphically and then apply the relation

$$Q.D. = \frac{Q_3 - Q_1}{2}.$$

(ii) Calculation of Quartile Deviation by using an appropriate formula:

Income per week (Rs.)	Class Boundaries	No. of Earners, f	F
41–50	40.5–50.5	30	30
51–60	50.5–60.5	36	66
61–70	60.5–70.5	43	109
71–80	70.5–80.5	104	213
81–90	80.5–90.5	73	286
91–100	90.5–100.5	14	300
Total	---	300	---

$Q.D. = \dfrac{Q_3 - Q_1}{2}$, where

Q_1 = Income of $\left(\dfrac{n}{4}\right)$*th* earner

= Income of $\left(\dfrac{300}{4}\right)$*th*, *i.e.* 75*th* earner which lies in the group 60.5–70.5, therefore

$$Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - C\right)$$

$$= 60.5 + \frac{10}{43}(75 - 66) = 60.5 + 2.09 = \text{Rs. } 62.59; \text{ and}$$

$$Q_3 = \text{Income of } \left(\frac{3n}{4}\right)th \text{ earner}$$

$$= \text{Income of } \frac{3(300)}{4} th, \text{ i.e. } 225th \text{ earner which lies in}$$

the group 80.5–90.5; therefore

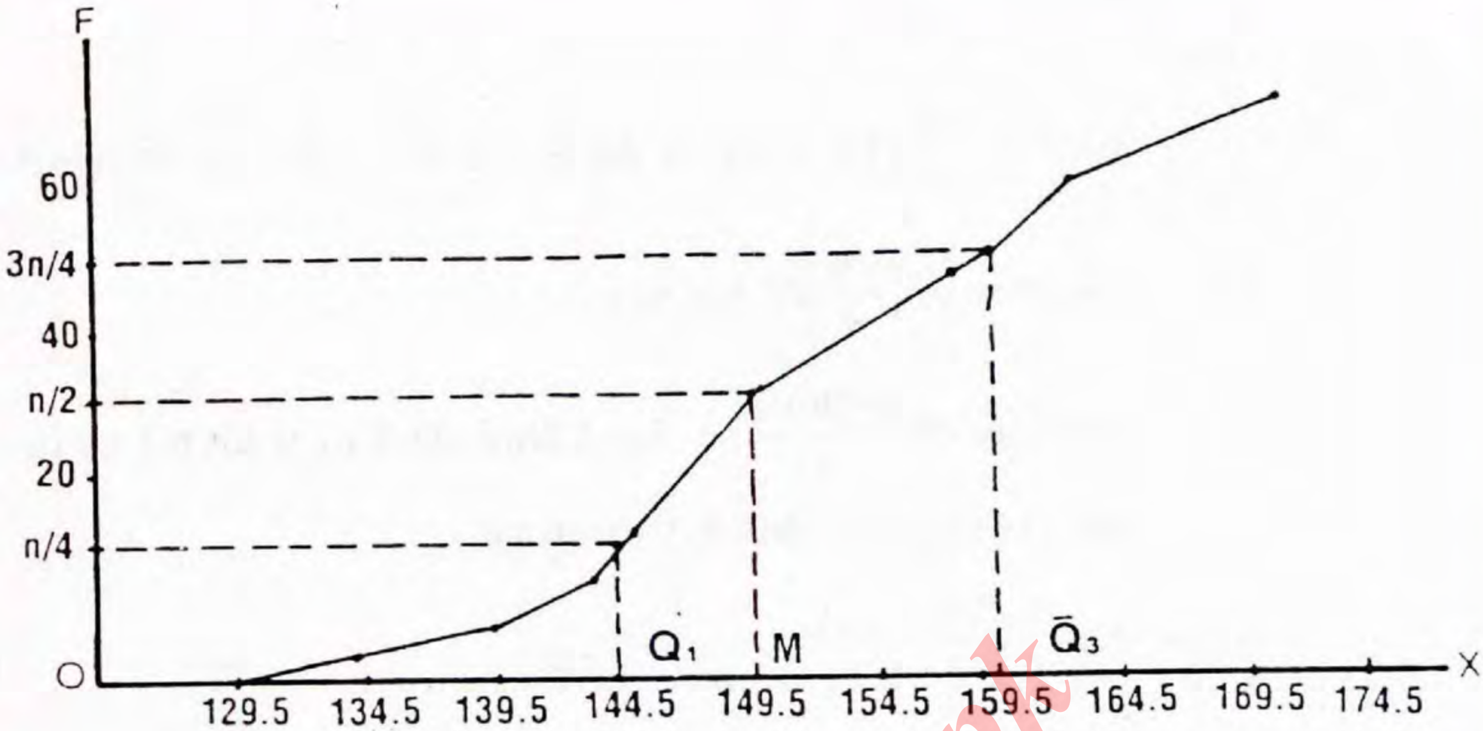
$$Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - C\right)$$

$$= 80.5 + \frac{10}{73}(225 - 213) = 80.5 + 1.64 = \text{Rs. } 82.14$$

$$\text{Hence } Q.D. = \frac{Q_3 - Q_1}{2} = \frac{82.14 - 62.59}{2} = \text{Rs. } 9.775 = \text{Rs. } 9.78$$

4.5 Construction of a Cumulative Frequency Table

Weight (lbs)	Tally	f	Weight	Cum f
129.5 – 134.5	III	3	less than 129.5	0
			less than 134.5	3
134.5 – 139.5	III	3	less than 139.5	6
139.5 – 144.5	~~IIII~~ I	6	less than 144.5	12
144.5 – 149.5	~~IIII~~ ~~IIII~~ IIII	14	less than 149.5	26
149.5 – 154.5	~~IIII~~ IIII	9	less than 154.5	35
154.5 – 159.5	~~IIII~~ III	8	less than 159.5	43
159.5 – 164.5	~~IIII~~ II	7	less than 164.5	50
164.5 – 169.5	~~IIII~~ I	6	less than 169.5	54
169.5 – 174.5	IIII	4	less than 174.5	60
Total	--	60	--	--



From graph, we estimate the approximate values of

Median = 152 lb, $Q_1 \simeq 146$ lb and Q3 $\simeq$ 161 lbs.

$\therefore$ S.I.Q. Range $= \dfrac{Q_3 - Q_1}{2} = \dfrac{161 - 146}{2} = 7.5$ lbs.

Calculation of Mean and Standard Deviation, using the grouped data:

Weight	x_i	f	$u(=x-147)/5$	fu	fu^2
129.5–134.5	132	3	−3	−9	27
134.5–139.5	137	3	−2	−6	12
139.5–144.5	142	6	−1	−6	6
144.5–149.5	147	14	0	0	0
149.5–154.5	152	9	1	9	9
154.5–159.5	157	8	2	16	32
159.5–164.5	162	7	3	21	63
164.5–169.5	167	6	4	24	96
169.5–174.5	172	4	5	20	100
Total	--	60	--	69	345

Now mean, $\bar{x} = a + \frac{\sum fu}{n} \times h$

$$= 147 + \frac{69 \times 5}{60} = 147 + 5.75 = 152.75 \text{ lbs, and}$$

$$S = h\sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= 5\sqrt{\frac{345}{60} - \left(\frac{69}{60}\right)^2} = 5\sqrt{5.75 - 1.3225}$$

$$= 5\sqrt{4.4275} = 5\,(2.104) = 10.52 \text{ lbs.}$$

Mean and Standard Deviation of Original Observations

$$\text{Mean} = \frac{\sum X}{n} = \frac{171 + 160 + \ldots + 144}{60} = \frac{9175}{60} = 152.917 \text{ lbs.}$$

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

$$= \sqrt{\frac{1409399}{60} - \left(\frac{9175}{60}\right)^2} = \sqrt{23489.98333 - 23383.50694}$$

$$= \sqrt{106.47639} = 10.32 \text{ lbs.}$$

Comparing the means and standard deviations computed from the grouped data and the original observations, we find that they are almost the same.

4.6. (b) Calculation of the Mean Deviation from the mean.

Marks	No. of Students, f	x	fx	$x-\bar{x}$	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$
0–9	2	4.5	9.0	–41.67	41.67	83.34
10–19	3	14.5	43.5	–31.67	31.67	95.01
20–29	8	24.5	196.0	–21.67	21.67	173.36
30–39	24	34.5	828.0	–11.67	11.67	280.08
40–49	27	44.5	1201.5	–1.67	1.67	45.09
50–59	40	54.5	2180.0	8.33	8.33	333.20
60–69	11	64.5	709.5	18.33	18.33	201.63
70–79	5	74.5	372.5	28.33	28.33	141.65
Total	120	--	5540	--	--	1353.36

Now $\bar{x} = \frac{\sum fx}{n} = \frac{5540}{120} = 46.17$; and therefore

Mean Deviation, $(M.D.) = \frac{\sum f\,|x-\bar{x}|}{n}$

$= \frac{1353.36}{120} = 11.28$ marks

4.8. Calculation of the Quartile Deviation, Mean Deviation and their Co-efficients.

Height x	Class Boundaries	Group A				Group B			
		f	F	x-Med	$f\|x-$Med$\|$	f	F	$x-$Med	$f\|x-$Med$\|$
58	57.5–58.5	10	10	−3.35	33.50	15	15	−3.26	48.90
59	58.5–59.5	18	28	−2.35	42.30	20	35	−2.26	45.20
60	59.5–60.5	30	58	−1.35	40.50	32	67	−1.26	40.32
61	60.5–61.5	42	100	−0.35	14.70	35	102	−0.26	9.10
62	61.5–62.5	35	135	0.65	22.75	33	135	0.74	24.42
63	62.5–63.5	28	163	1.65	46.20	22	157	1.74	38.28
64	63.5–64.5	16	179	2.65	42.40	20	177	2.74	54.80
65	64.5–65.5	8	187	3.65	29.20	10	187	3.74	37.40
Total	---	187	--	--	271.55	187	--	--	298.42

Group A.

Q_1 = Height of $\left(\frac{n}{4}\right)th$ person

= Height of $\left(\frac{187}{4}\right)th$, *i.e.* $46.75th$ person which lies in the group 59.5 – 60.5,

$Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - C\right)$

$= 59.5 + \frac{1}{30}(46.75-28) = 59.5 + 0.62 = 60.12$ inches

Q_3 = Height of $\left(\frac{3n}{4}\right)th$ person

= Height of $\frac{3\,(187)}{4}th$, *i.e.* $140.25th$ person which lies in the group 62.5 – 63.5

$$\therefore \quad Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - C\right) = 62.5 + \frac{1}{28}(140.25 - 135)$$

$$= 62.5 + 0.19 = 62.69 \text{ inches}$$

$$\text{Median} = \text{Height of } \left(\frac{n}{2}\right)th \text{ person}$$

$$= \text{Height of } \left(\frac{187}{2}\right)th, \text{ i.e. } 93.5th \text{ person which lies}$$

in the group 60.5 – 61.5.

$$\therefore \quad \text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - C\right)$$

$$= 60.5 + \frac{1}{42}(93.5 - 58) = 60.5 + 0.85 = 61.35 \text{ inches}$$

$$\text{Thus } Q.D. = \frac{Q_3 - Q_1}{2} = \frac{62.69 - 60.12}{2} = \frac{2.57}{2} = 1.285 \text{ inches,}$$

$$\text{and Coeff. of } Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{62.69 - 60.12}{62.69 + 60.12} = \frac{2.57}{122.81} = 0.02$$

$$M.D. = \frac{\sum f\,|x - \text{Med}|}{n} = \frac{271.55}{187} = 1.45 \text{ inches;}$$

$$\text{Coeff of } M.D. = \frac{M.D.}{\text{Median}} = \frac{1.45}{61.35} = 0.024$$

Group B.

$$Q_1 = \text{Height of } \left(\frac{n}{4}\right)th \text{ person}$$

$$= \text{Height of } \left(\frac{187}{4}\right)th, \text{ i.e. } 46.75th \text{ person which lies}$$

in the group 59.5 – 60.5

$$\therefore \quad Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - C\right)$$

$$= 59.5 + \frac{1}{32}(46.75 - 35) = 59.5 + 0.37 = 59.87 \text{ inches;}$$

$$Q_3 = \text{Height of } \left(\frac{3n}{4}\right)th \text{ person}$$

$$= \text{Height of } \frac{3\,(187)}{4}th, \text{ i.e. } 140.25th \text{ person which lies}$$

in the group 62.5 – 63.5

$$\therefore \quad Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - C\right) = 62.5 + \frac{1}{22}(140.25 - 135)$$

$$= 62.5 + 0.24 = 62.74 \text{ inches}$$

Similarly, we estimate the median as

$$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - C\right) = 60.5 + \frac{1}{35}(93.5 - 67)$$

$$= 60.5 + 0.76 = 61.26 \text{ inches}$$

$$\text{Now } Q.D. = \frac{Q_3 - Q_1}{2} = \frac{62.74 - 59.87}{2} = 1.435 \text{ inches;}$$

$$M.D. = \frac{\sum f\,|x - \text{Med}|}{n} = \frac{298.42}{187} = 1.60 \text{ inches;}$$

$$\text{Coeff of } Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{62.74 - 59.87}{62.74 + 59.87} = \frac{1.435}{122.61} = 0.012 \text{ and}$$

$$\text{Coeff of } M.D. = \frac{M.D.}{\text{Median}} = \frac{1.60}{61.26} = 0.026$$

4.9. (b) Calculation of population variance and standard deviation

X_i	$X_i - \mu$	$(X_i - \mu)^2$
10	+2.5	6.25
8	+0.5	0.25
7	−0.5	0.25
9	+1.5	2.25
5	−2.5	6.25
12	+4.5	20.25
8	+0.5	0.25
6	−1.5	2.25
8	+0.5	0.25
2	−5.5	30.25
75	--	68.5

$$\therefore \qquad \mu = \frac{\sum X_i}{N} = \frac{75}{10} = 7.5;$$

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{N} = \frac{68.5}{10} = 6.85; \text{ and}$$

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} = \sqrt{\frac{68.5}{10}} = \sqrt{6.85} = 2.62$$

4.10. (b) Calculation of sample mean and standard deviation.

x_i	$x - \bar{x}$	$(x - \bar{x})^2$
70	10	100
50	−10	100
60	0	0
70	10	100
50	−10	100
300	--	400

$$\therefore \qquad \bar{x} = \frac{\sum x}{n} = \frac{300}{5} = 60, \text{ and}$$

$$S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{\frac{400}{5}} = \sqrt{80} = 8.944$$

(i) Now, we calculate the sample mean and the standard deviation of the scores obtained by adding 10 points to them:

	$y(=x+10)$	$(y-\bar{y})$	$(y-\bar{y})^2$
	80	10	100
	60	−10	100
	70	0	0
	80	10	100
	60	−10	100
Σ	350	---	400

$$\therefore \quad \bar{y} = \frac{\sum y}{n} = \frac{350}{5} = 70 = \bar{x} + 10; \text{ and}$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{400}{5}} = \sqrt{80} = 8.944 = S_x$$

(ii) Increasing all scores by 10% implies that each score is to be multiplied by $\frac{110}{100}$. The calculations then become:

	$y(=\frac{110}{100}x)$	$y-\bar{y}$	$(y-\bar{y})^2$
	77	11	121
	55	−11	121
	66	0	0
	77	11	121
	55	−11	121
$\sum$	330	---	484

$$\therefore \quad \bar{y} = \frac{\sum y}{n} = \frac{330}{5} = 66 = \frac{110}{100}(60) = a\bar{x}, \text{ where } a = \frac{110}{100};$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{484}{5}} = \sqrt{96.8} = 9.838$$

$$= \frac{110}{100}(8.944) = a\,S_x$$

Hence we observe that

(i) when 10 is added to all scores, the mean is increased by 10 but standard deviation remains unchanged; and

(ii) when the scores are increased by 10%, *i.e.* multiplied by $\frac{110}{100}$, both the mean and standard deviation are multiplied by $\frac{110}{100}$.
It is based on the following properites. If $y = ax + b$, then $\bar{y} = a\bar{x} + b$, and $S_y = |a|\,S_x$.

4.11. (b) Given $n = 15$, $\sum x = 480$ & $\sum x^2 = 15{,}735$. Therefore

$$\bar{x} = \frac{\sum x}{n} = \frac{480}{15} = 32; \text{ and}$$

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} = \sqrt{\frac{15735}{15} - \left(\frac{480}{15}\right)^2}$$

$$= \sqrt{1049 - 1024} = \sqrt{25} = 5$$

4.12. (b) Calculation of means and standard deviations

	x	$x-\bar{x}$	$(x-\bar{x})^2$
	3	−1	1
	6	2	4
	2	−2	4
	1	−3	9
	7	3	9
	5	1	1
$\sum$	24	---	28

$$\therefore \quad \bar{x} = \frac{\sum x}{n} = \frac{24}{6} = 4, \text{ and}$$

$$S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{28}{6}} = \sqrt{4.6667} = 2.16$$

Now $y = 2x + 5$. Then

	y	$y-\bar{y}$	$(y-\bar{y})^2$
	11	−2	4
	17	4	16
	9	−4	16
	7	−6	36
	19	6	36
	15	2	4
$\sum$	78	---	112

$$\bar{y} = \frac{\sum y}{n} = \frac{78}{6} = 13 = 2\,(4) + 5 = 2\bar{x} + 5, \text{ and}$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{112}{6}} = \sqrt{18.6667} = 4.32$$

$$= 2(2.16) = (2)\,S_x.$$

(c) (i) When the age of the youngest child is 1 year, the ages of children are 1, 2, 3, 4, 5, 6 and 7. Therefore

$$\text{mean} = \frac{\sum x}{n} = \frac{1 + 2 + \ldots + 7}{7} = \frac{28}{7} = 4 \text{ year; and}$$

$$s.d. = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} = \sqrt{\left(\frac{1^2 + 2^2 + \ldots + 7^2}{7}\right) - \left(\frac{28}{7}\right)^2}$$

$$= \sqrt{20 - 16} = \sqrt{4} = 2 \text{ years.}$$

(ii) When the youngest child is 8 years old, the ages of the children are 8, 9, 10, 11, 12, 13 and 14 years. Therefore

$$\text{Mean} = \frac{8 + 9 + \ldots + 14}{7} = \frac{77}{7} = 11 \text{ years; and}$$

$$S.D. = \sqrt{\frac{8^2 + 9^2 + \ldots + 14^2}{7} - \left(\frac{77}{7}\right)^2}$$

$$= \sqrt{125 - 121} = \sqrt{4} = 2 \text{ years.}$$

The standard deviations of (i) and (ii) coincide because they remain unaffected if a constant is added to the values of a variable. Here 7 is added to all values of (i).

4.13. Let the sample of size n consists of $x_1, x_2, \ldots, x_n$ values. Then, assuming x_1 as the smallest observation and x_n, the largest observation, we compute the Range and the Standard deviation by the relations

$$\text{Range} = x_n - x_1, \text{ and}$$

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

When $n = 2$, then assuming $x_2 > x_1$, we have

Range $= x_2 - x_1$; and

$$S.D. = \sqrt{\frac{x_1^2 + x_2^2}{2} - \left(\frac{x_1 + x_2}{2}\right)^2}$$

$$= \sqrt{\frac{1}{4}[2x_1^2 + 2x_2^2 - x_1^2 - x_2^2 - 2x_1x_2]}$$

$$= \sqrt{\frac{1}{4}(x_1^2 + x_2^2 - 2x_1x_2)}$$

$$= \sqrt{\left(\frac{x_2 - x_1}{2}\right)^2} = \frac{x_2 - x_1}{2} = \frac{1}{2}\text{(Range)}$$

Hence we observe that Range and Standard Deviation are related when $n=2$.

4.14. Here $M.D. = \frac{\sum f_i |x_i - \bar{x}|}{n} = \frac{1696.0}{60} = 28.27$ gms

(see Example 4.4)

and $S.D.$ = 34.87 gms (see example 4.7 in the text)

$\therefore \frac{\text{Mean Deviation}}{\text{Standard Deviation}} = \frac{28.27}{34.87} = 0.81$

The statement is almost correct.

4.15. Calculation of the mean and the standard deviation.

x	f	D	fD	fD^2
30	4	−4	−16	64
31	8	−3	−24	72
32	23	−2	−46	92
33	35	−1	−35	35
34	62	0	−121	0
35	44	1	44	44
36	18	2	36	72
37	4	3	12	36
38	1	4	4	16
39	1	5	5	25
Total	200	--	+101 −20	456

$$\therefore \qquad \bar{x} = a + \frac{\sum fD}{n} = 34 + \frac{(-20)}{200} = 33.9, \text{ and}$$

$$s = \sqrt{\frac{\sum fD^2}{n} - \left(\frac{\sum fD}{n}\right)^2} = \sqrt{\frac{456}{200} - \left(\frac{-20}{200}\right)^2}$$

$$= \sqrt{2.28 - 0.01} = \sqrt{2.27} = 1.507$$

4.16. The mean and standard deviation calculated for question 4.15 are

$$\bar{x} = 33.9, \text{ and } S = 1.507,$$

The interval "mean $\pm 2S$" is obtained as below:

$$\bar{x} \pm 2S = 33.9 \pm 2(1.507) = 33.9 \pm 3.014$$

$$= 30.886 \text{ and } 36.914$$

This interval 30.886 – 36.914 will according to Chebyshev's rule, contain *at least* $\left(1 - \frac{1}{2^2}\right)$, *i.e.* $\frac{3}{4}$ of 200 = 150 metal bars. The number of metal bars lying within this interval is 190 which is greater than 150.

4.17. Calculation of means and standard deviations of the expenditure.

Expenditure Rs.	x_i	u_i	Place A			Place B		
			f_1	f_1u	f_1u^2	f_2	f_2u	f_2u^2
30–60	45	−2	28	−56	112	39	−78	156
60–90	75	−1	292	−292	292	284	−284	284
90–120	105	0	389	−348	0	401	−362	0
120–150	135	1	212	212	212	202	202	202
150–180	165	2	59	118	236	48	96	192
180–210	195	3	18	54	162	31	93	279
210–240	225	4	2	8	32	5	20	80
Total	--	--	1000	+392 44	1046	1010	+411 49	1193

Place A:

$$\bar{x} = a + \frac{\sum f_1 u}{n} \times h, \text{ where } u = \frac{x - 105}{30}.$$

$$= \text{Rs. } 105 + \frac{44}{1000} \times 30 = \text{Rs. } 105 + 1.32 = \text{Rs. } 106.32$$

$$s = h \times \sqrt{\frac{\sum f_1 u^2}{n} - \left(\frac{\sum f_1 u}{n}\right)^2}$$

$$= 30 \times \sqrt{\frac{1046}{1000} - \left(\frac{44}{1000}\right)^2} = 30 \times \sqrt{1.046 - 0.0019}$$

$$= 30 \times 1.02 = \text{Rs. } 30.6$$

Place B:

$$\bar{x} = a + \frac{\sum f_2 u}{n} \times h$$

$$= \text{Rs. } 105 + \frac{49}{1010} \times 30 = \text{Rs. } 105 + 1.46 = \text{Rs. } 106.46$$

$$s = h \times \sqrt{\frac{\sum f_2 u^2}{n} - \left(\frac{\sum f_2 u}{n}\right)^2}$$

$$= 30 \times \sqrt{\frac{1193}{1010} - \left(\frac{49}{1010}\right)^2} = 30 \times \sqrt{1.1812 - 0.0025}$$

$$= 30 \times \sqrt{1.1787} = 30 \times 1.09 = \text{Rs. } 32.7$$

4.18. Calculation of the mean wage and the standard deviation.

Classes (Rs.)	Midvalues (x)	u	f	fu	fu^2
4.50–5.50	5.00	−7	6	−42	294
5.50–6.50	6.00	−6	17	−102	612
6.50–7.50	7.00	−5	35	−175	875
7.50–8.50	8.00	−4	48	−192	768
8.50–9.50	9.00	−3	65	−195	585
9.50–10.50	10.00	−2	90	−180	360
10.50–11.50	11.00	−1	131	−131	131
11.50–12.50	12.00	0	173	−1017	0
12.50–13.50	13.00	1	155	155	155
13.50–14.50	14.00	2	117	234	468
14.50–15.50	15.00	3	75	225	675
15.50–16.50	16.00	4	52	208	832
16.50–17.50	17.00	5	21	105	525
17.50–18.50	18.00	6	9	54	324
18.50–19.50	19.00	7	6	42	294
Total	----	---	1000	+ 1023 6	6898

Mean wage or $\bar{x} = a + \frac{\sum fu}{n} \times h$, where $u = x - 12.00$ and $h = 1$.

$$= \text{Rs. } 12.00 + \frac{6}{1000} = \text{Rs. } 12.006, \text{ and}$$

$$s = \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= \sqrt{\frac{6898}{1000} - \left(\frac{6}{1000}\right)^2} = \sqrt{6.898 - 0.00036}$$

$$= \sqrt{6.897964} = \text{Rs. } 2.626$$

4.19. Calculation of the mean and the standard deviation.

	x	$u(=x-90)$	u^2
	95	5	25
	73	−17	289
	82	−8	64
	108	18	324
	103	13	169
	78	−12	144
	79	−11	121
	94	4	16
	97	7	49
	95	5	25
	69	−21	441
	87	−3	9
	130	40	1600
	89	−1	1
	67	−23	529
	93	3	9
	96	6	36
	68	−22	484
	83	−7	49
	117	27	729
Total	1803	3	5113

Now $\bar{x} = a + \frac{\sum u}{n} = 90 + \frac{3}{20} = 90.15$; and

$$S = \sqrt{\frac{\sum u^2}{n} - \left(\frac{\sum u}{n}\right)^2}$$

$$= \sqrt{\frac{5113}{20} - \left(\frac{3}{20}\right)^2} = \sqrt{255.65 - 0.0225}$$

$$= \sqrt{255.6275} = 15.99$$

(i) Mean $\pm S = 90.15 \pm 15.99 = 106.14, 74.16$

Observations lying within these limits = 13

$\therefore$ Percentage of observations $= \frac{13}{20} \times 100 = 65\%$

(ii) Mean $\pm 2S = 90.15 \pm 2(15.99)$

$= 90.15 \pm 31.98 = 122.13, 58.17$

Observations lying within these limits = 19

$\therefore$ Percentage of observations $= \frac{19}{20} \times 100 = 95\%$

(iii) Mean $\pm 3S = 90.15 \pm 3(15.99)$

$= 90.15 \pm 47.97 = 138.12, 42.18$

Observations lying within these limits = 20

Hence %age of observations $= \frac{20}{20} \times 100 = 100\%$.

4.20. Computation of the mean, the standard deviation, etc.

Midvalues (inches) x	No. of Students (f)	u $\left(=\frac{x-14.5}{0.5}\right)$	fu	fu^2
12.5	4	−4	−16	64
13.0	19	−3	−57	171
13.5	30	−2	−60	120
14.0	63	−1	−63	63
14.5	66	0	−196	0
15.0	29	1	29	29
15.5	18	2	36	72
16.0	1	3	3	9
16.5	1	4	4	16
Total	231	--	+72 −124	544

$$\therefore \quad \bar{x} = a + \frac{\sum fu}{n} \times h = 14.5 + \frac{(-124)}{231} \times 0.5$$

$$= 14.5 - 0.27 = 14.23 \text{ inches}$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= 0.5 \times \sqrt{\frac{544}{231} - \left(\frac{-124}{231}\right)^2} = 0.5 \times \sqrt{2.35 - 0.28}$$

$$= 0.5 \times \sqrt{2.07} = 0.72 \text{ inches.}$$

Using the criterion $\bar{x} \pm 3s$, we get

(i) the largest size of the collars $= \bar{x} + 3s + \frac{3}{4}$ inches

$$= 14.23 + 3(0.72) + 0.75$$

$$= 17.14 \text{ inches}$$

(ii) the smallest size of the collars $= \bar{x} - 3s + \frac{3}{4}$ inches

$= 14.23 - 3(0.72) + 0.75$

$= 12.82$ inches.

4.21. Determination of the actual classes.

u	f	fu	fu^2	Actual Classes
−4	2	−8	32	109.5−115.5
−3	5	−15	45	115.5−121.5
−2	8	−16	32	121.5−127.5
−1	18	−18	18	127.5−133.5
0	22	−57	0	133.5−139.5
1	13	13	13	139.5−145.5
2	8	16	32	145.5−151.5
3	4	12	36	151.5−157.5
Σ	80	+41 −16	208	---

Substituting the values in the formula

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}, \text{ we get}$$

$$9.6 = h \times \sqrt{\frac{208}{80} - \left(\frac{-16}{80}\right)^2} = h \times \sqrt{2.60 - 0.04} = 1.6h$$

$\therefore \quad h = 9.6 \div 1.6 = 6$, and

$$\text{mean} = a + \frac{\sum fu}{n} \times h$$

i.e. $135.3 = a + \frac{(-16)}{80} \times 6 = a - 1.2$

$\therefore \quad a = 135.3 + 1.2 = 136.5$

Thus the midpoint of the actual class corresponding to the frequency 22 is 136.5. As the lenght of the class-interval is 6, therefore this class is (136.5−3.0) to (136.5+3.0), *i.e.* 133.5–139.5.

The other classes are then determined by adding to and subtracting from these class limits, the width of the class interval repeatedly. The classes thus determined are shown in the last column of the table on page 65.

4.22. Calculation of the mean, standard deviation, etc.

Source A					Source B				
Life (hrs)	No. of components (f)	u	fu	fu^2	Life (hours)	(f)	u	fu	fu^2
1000–1020	40	−3	−120	360	1030–1040	339	−2	−678	1356
1020–1040	96	−2	−192	384	1040–1050	136	−1	−136	136
1040–1060	364	−1	−364	364	1050–1060	25	0	−814	0
1060–1080	372	0	−676	0	1060–1070	20	1	20	20
1080–1100	85	1	85	85	1070–1080	130	2	260	520
1100–1120	43	2	86	172	1080–1090	350	3	1050	3150
Total	1000	--	+171 −505	1365	---	1000	--	+1330 +516	5182

Source A:

$$\text{Mean life, } i.e.\ \bar{x} = a + \frac{\sum fu}{n} \times h = 1070 + \frac{(-505)}{1000} \times 20$$

$$= 1070 - 10 = 1060 \text{ hours.}$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 20 \times \sqrt{\frac{1365}{1000} - \left(\frac{-505}{1000}\right)^2}$$

$$= 20 \times \sqrt{1.365 - 0.255} = 20 \times \sqrt{1.110} = 21.1 \text{ hours}$$

Source B:

$$\bar{x} = a + \frac{\sum fu}{n} \times h$$

$$= 1055 + \frac{516}{1000} \times 10 = 1055 + 5.16 = 1060 \text{ hours}$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 10 \times \sqrt{\frac{5182}{1000} - \left(\frac{516}{1000}\right)^2}$$

$$= 10 \times \sqrt{5.182 - 0.2663} = 22.2 \text{ hours.}$$

We observe that the mean lives in hours of the components of the two sources are equal and the two sources have nearly the same dispersion, but these data give a false impression as the distribution of source B is U-shaped.

4.24. (b) By definition, $C.V. = \frac{s}{\bar{x}} \times 100$

Now $\bar{x} = a + \frac{\sum fu}{n} \times h = 62 + \frac{140}{210} \times 5 \quad (\because \; u = \frac{x-62}{5})$

$= 62 + 5.83 = 67.83$, and

$s^2 = h^2 \left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right]$ where $n = \sum f = 120$

$= 25\left[\frac{598}{120} - \left(\frac{140}{120}\right)^2\right] = 25(4.9833 - 1.3611) = 90.555$

And s^2(corrected) $= s^2$ (uncorrected) $- \frac{h^2}{12}$

$= 90.555 - \frac{25}{12} = 90.555 - 2.0833 = 88.4717$, so that

s (corrected) $= \sqrt{88.4717} = 9.41$

Hence C.V. $= \frac{9.41}{67.83} \times 100 = 13.87\%$

4.25. (b) Computation of standard deviation for the data in locality A.

Income (Rs.)	f	x	u	fu	fu^2
35–39	13	37	–3	–39	117
40–44	15	42	–2	–30	60
45–49	17	47	–1	–17	17
50–54	28	52	0	–86	0
55–59	12	57	1	12	12
60–64	10	62	2	20	40
65–69	5	67	3	15	45
Total	100	--	--	+47 –39	291

Now $\quad \bar{x} = a + \frac{\sum fu}{n} \times h$, where $u = \frac{x_i - 52}{5}$

$$= 52 + \frac{(-39)}{100} \times 5 = 52 - 1.95 = \text{Rs. } 50.05; \text{ and}$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 5 \times \sqrt{\frac{291}{100} - \left(\frac{-39}{100}\right)^2}$$

$$= 5 \times \sqrt{2.91 - 0.1521} = 5 \times \sqrt{2.7579} = \text{Rs. } 8.3.$$

$\therefore \quad C.V.$ for locality $A = \frac{s}{\bar{x}} \times 100 = \frac{8.3}{50.05} \times 100 = 16.58\%$

Again $C.V.$ for locality $B = \frac{s}{\bar{x}} \times 100 = \frac{4.96}{52.28} \times 100 = 9.50\%$

Locality A has greater variability as the coefficient of variation for A is larger than that for B.

4.26. (b) Computation of the co-efficients of variation for the candidates X and Y.

Paper	X	$X-\bar{X}$	$(X-\bar{X})^2$	Y	$Y-\bar{Y}$	$(Y-\bar{Y})^2$
I	58	−4.7	22.09	39	−24.4	595.36
II	49	−13.7	187.69	38	−25.4	645.16
III	76	13.3	176.89	86	22.6	510.76
IV	80	17.3	299.29	72	8.6	73.96
V	47	−15.7	246.49	75	11.6	134.56
VI	72	9.3	86.49	69	5.6	31.36
VII	61	−1.7	2.89	57	−6.4	40.96
VIII	59	−3.7	13.69	49	−14.4	207.36
IX	77	14.3	204.49	83	19.6	384.16
X	48	−14.7	216.09	66	2.6	6.76
Total	627	--	1456.10	634	--	2630.40

Candidate X:

$$\bar{X} = \frac{\sum X_i}{n} = \frac{627}{10} = 62.7 \text{ marks}$$

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}} = \sqrt{\frac{1456.10}{10}} = \sqrt{145.61} = 12.07 \text{marks}$$

$\therefore$ $C.V.$ for candidate $X = \frac{S_x}{\bar{X}} \times 100 = \frac{12.07}{62.7} \times 100 = 19.25\%$

Candidate Y:

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{634}{10} = 63.4 \text{ marks}$$

$$S_y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n}} = \sqrt{\frac{2630.40}{10}} = \sqrt{263.04} = 16.22 \text{marks}$$

$\therefore$ $C.V.$ for candidate $Y = \frac{S_y}{\bar{Y}} \times 100 = \frac{16.22}{63.4} \times 100 = 25.58\%$

The performance of the candidate X is more consistent than that of the candidate Y as the $C.V.$ for X is smaller than that for Y.

4.27 (a) Here $\bar{x} = 67.45$, s^2 (uncorrected) $= 8.5275$, and $h = 3$

$$\therefore \quad s^2 \text{ (corrected)} = s^2 \text{ (uncorrected)} - \frac{h^2}{12}$$

$$= 8.5275 - \frac{9}{12} = 7.7775, \text{ and}$$

$$s \text{ (corrected)} = \sqrt{7.7775} = 2.79.$$

$$\text{Hence } C.V. \text{ (corrected)} = \frac{s \text{ (corrected)}}{\text{mean}} \times 100$$

$$= \frac{2.79}{67.45} \times 100 = 4.14\%$$

(b) Calculation of the means, standard deviations and the coefficients of variation for Batsmen A and B.

	Batsman A		Batsman B	
	Scores x	x^2	Scores y	y^2
	12	144	47	2209
	15	225	12	144
	6	36	76	5776
	73	5329	48	2304
	7	49	4	16
	19	361	51	2601
	199	39601	37	1369
	36	1296	48	2304
	84	7056	13	169
	29	841	0	0
Σ	480	54938	336	16892

Bastman A:

$$\text{Mean scores} = \frac{\sum x_i}{n} = \frac{480}{10} = 48 \text{ scores,}$$

$$S_x = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = \sqrt{\frac{54938}{10} - \left(\frac{480}{10}\right)^2}$$

$$= \sqrt{5493.80 - 2304} = \sqrt{3189.8} = 56.48 \text{ scores, and}$$

$$C.V. = \frac{S_x}{\bar{x}} \times 100 = \frac{56.48}{478} \times 100 = 117.67\%$$

Bastman B:

$$\text{Mean scores} = \frac{\sum y_i}{n} = \frac{336}{10} = 33.6 \text{ scores,}$$

$$S_y = \sqrt{\frac{\sum y_i^2}{n} - \left(\frac{\sum y_i}{n}\right)^2} = \sqrt{\frac{16892}{10} - \left(\frac{336}{10}\right)^2}$$

$$= \sqrt{1689.20 - 1128.96} = \sqrt{560.24} = 23.67 \text{ scores, and}$$

$$C.V. = \frac{S_y}{\bar{y}} \times 100 = \frac{23.67}{33.6} \times 100 = 70.45\%$$

Batsman A is better as a run getter as A's average score is 48 and B's average score is 33.6.

Batsman B is a more consistent player as the coefficient of variation for batsman B is smaller than that for A.

4.28. (b) Calculation of coefficient of variation.

Weight (lb)	x	f	u	fu	fu^2
118–126	122	3	−3	−9	27
127–135	131	5	−2	−10	20
136–144	140	9	−1	−9	9
145–153	149	12	0	−28	0
154–162	158	5	1	5	5
163–171	167	4	2	8	16
172–180	176	2	3	6	18
Total	---	40	--	+19 −9	95

$$\bar{x} = a + \frac{\sum fu}{n} \times h, \text{ where } h = 9, \text{ and } u = \frac{x - 149}{9}$$

$$= 149 + \frac{(-9)}{40} \times 9 = 149 - 2.025 = 146.975 \text{ lbs.}$$

$$s^2 \text{ (uncorrected)} = h^2 \left[\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2\right]$$

$$= 81 \left[\frac{95}{40} - \left(\frac{-9}{40}\right)^2\right]$$

$$= 81 [2.375 - 0.0506] = 81 \times (2.3244)$$

$\therefore$ $$s \text{ (uncorrected)} = 9 \times \sqrt{2.3244} = 13.716 \text{ lbs; and}$$

$$C.V. = \frac{s}{\text{mean}} \times 100 = \frac{13.716}{146.975} \times 100 = 9.33\%$$

$$\text{Now, } s^2 \text{ (corrected)} = s^2 \text{ (uncorrected)} - \frac{h^2}{12}$$

$$= 81(2.3244) - \frac{81}{12}$$

$$= 81[2.3244 - 0.0833] = 81 \times (2.2411)$$

$\therefore$ $$s \text{ (corrected)} = 9 \times \sqrt{2.2411} = 9 \times 1.497 = 13.473 \text{ lbs; and}$$

$$C.V. \text{(corrected)} = \frac{s \text{ (corrected)}}{\text{mean}} \times 100 = \frac{13.473}{146.975} \times 100 = 9.17\%$$

(c) $$C.V. \text{ for tube } A = \frac{280}{1495} \times 100 = 18.8\%$$

$$C.V. \text{ for tube } B = \frac{310}{1895} \times 100 = 16.4\%$$

(i) Tube B has a greater absolute dispersion as $S_B > S_A$

(ii) Tube A has a greater relative dispersion as the co-efficient of variation for A is larger than that for B.

4.29. Computation of co-efficients of variation.

Expenditure (Rupees)	x	u	Town A			Town B		
			f	fu	fu^2	f	fu	fu^2
21–30	25.5	−3	3	−9	27	2	−6	18
31–40	35.5	−2	61	−122	244	14	−28	56
41–50	45.5	−1	132	−132	132	20	−20	20
51–60	55.5	0	153	−263	0	27	−54	0
61–70	65.5	1	140	140	140	28	28	28
71–80	75.5	2	51	102	204	7	14	28
81–90	85.5	3	2	6	18	2	6	18
Total	--	--	542	+248 / −1.5	765	100	+48 / −6	168

Town A:

$$\bar{x} = a + \frac{\sum fu}{n} \times h = 55.5 + \frac{(-15)}{542} \times 10$$

$$= 55.5 - 0.28 = \text{Rs. } 55.22;$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= 10 \times \sqrt{\frac{765}{542} - \left(\frac{-15}{542}\right)^2} = 10 \times \sqrt{1.4114 - 0.0009}$$

$$= 10 \times \sqrt{1.4105} = 10 \times (1.1876) = \text{Rs. } 11.88; \text{ and}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{\text{Rs. } 11.88}{\text{Rs. } 55.22} \times 100 = 21.51\%$$

Town B:

$$\bar{x} = a + \frac{\sum fu}{n} \times h = 55.5 + \frac{(-6)}{100} \times 10$$

$$= 55.5 - 0.6 = \text{Rs. } 54.9;$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2}$$

$$= 10 \times \sqrt{\frac{168}{100} - \left(\frac{-6}{100}\right)^2} = 10 \times \sqrt{1.68 - 0.0036}$$

$$= 10 \times \sqrt{1.6764} = 10 \times (1.2948) = \text{Rs. } 12.95; \text{ and}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{12.95}{54.9} \times 100 = 23.59\%$$

The co-efficient of variation for town *B* is larger than that for town *A*. Thus there is greater variability in expenditures of families in town *B* than that of town *A*.

4.30. Computation of co-efficients of variation.

Weight (kilograms)	u	Class A			Class B			Class C		
		f	fu	fu^2	f	fu	fu^2	f	fu	fu^2
25	−2	7	−14	28	5	−10	20	6	−12	24
35	−1	10	−10	10	9	−9	9	25	−25	25
45	0	20	−24	0	21	−19	0	24	−37	0
55	1	18	18	18	15	15	15	4	4	4
65	2	7	14	28	6	12	24	3	6	12
Total	--	62	$\frac{+32}{8}$	84	56	$\frac{+27}{8}$	68	62	$\frac{+10}{-27}$	65

Class A:

$$\bar{x} = a + \frac{\sum fu}{n} \times h, \text{ where } h = 10 \text{ and } u = \frac{x - 45}{10}$$

$$= 45 + \frac{8}{62} \times 10 = 45 + 1.29 = 46.29 \text{ kilograms;}$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 10 \times \sqrt{\frac{84}{62} - \left(\frac{8}{62}\right)^2}$$

$$= 10 \times \sqrt{1.3548 - 0.0166} = 10 \times \sqrt{1.3382} = 11.57 \text{ kg; and}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{11.57}{46.29} \times 100 = 24.99\%$$

Class B:

$$\bar{x} = a + \frac{\sum fu}{n} \times h$$

$$= 45 + \frac{8}{56} \times 10 = 45 + 1.43 = 46.43 \text{ kg};$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 10 \times \sqrt{\frac{68}{56} - \left(\frac{8}{56}\right)^2}$$

$$= 10 \times \sqrt{1.2143 - 0.0204} = 10 \times 1.093 = 10.93 \text{ kg; and}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{10.93}{46.43} \times 100 = 23.54\%$$

Class C:

$$\bar{x} = a + \frac{\sum fu}{n} \times h$$

$$= 45 + \frac{(-27)}{62} \times 10 = 45 - 4.35 = 40.65 \text{ kg};$$

$$s = h \times \sqrt{\frac{\sum fu^2}{n} - \left(\frac{\sum fu}{n}\right)^2} = 10 \times \sqrt{\frac{65}{62} - \left(\frac{-27}{62}\right)^2}$$

$$= 10 \times \sqrt{1.0484 - 0.1896} = 10 \times 0.927 = 9.27 \text{ kg; and}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{9.27}{40.65} \times 100 = 22.80\%$$

4.31. By definition, the variance, s^2, is

$$s^2 = \frac{\sum n_r (x_r - \bar{x})^2}{\sum n_r}, \text{ where } n_r \text{ denotes the frequency and } \bar{x} = \frac{\sum n_r x_r}{\sum n_r}$$

$$= \frac{\sum n_r (x_r - k + k - \bar{x})^2}{\sum n_r}, \text{ where } k \text{ is any arbitrary number.}$$

$$= \frac{1}{\sum n_r}\left[\sum n_r \{(x_r - k)^2 + (\bar{x} - k)^2 - 2(x_r - k)(\bar{x} - k)\}\right]$$

$$= \frac{\sum n_r (x_r - k)^2}{\sum n_r} + \frac{(\bar{x}_r - k)^2 \sum n_r}{\sum n_r} - 2(\bar{x} - k)\frac{\sum n_r (x_r - k)}{\sum n_r}$$

$$= \frac{\sum n_r (x_r - k)^2}{\sum n_r} + (\bar{x} - k)^2 - 2(\bar{x} - k)^2$$

$$= \frac{\sum n_r (x_r - k)^2}{\sum n_r} - (\bar{x} - k)^2$$

Hence $s = \sqrt{\frac{\sum n_r (x_r - k)^2}{\sum n_r} - \delta^2}$, where $\bar{x} = k + \delta$.

4.32. (a) Here $n_1 = 50, \bar{x}_1 = 59.5, S_1 = 8.38,$

$n_2 = 40, \bar{x}_2 = 54.0, S_2 = 8.23.$

Let $\bar{x}$ and S denote the mean and standard deviation respectively of the combined group of children. Then

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$= \frac{(50)(59.5) + (40)(54.0)}{50 + 40} = \frac{5135}{90} = 57.06; \text{ and}$$

$$S^2 = \frac{1}{n_1+n_2}\left[n_1S_1^2 + n_1(\bar{x}_1 - \bar{x})^2 + n_2S_2^2 + n_2(\bar{x}_2 - \bar{x})^2\right]$$

$$= \frac{1}{90}[(50)(8.38)^2 + 50(59.5 - 57.06)^2 + (40)(8.23)^2 + 40(54.0 - 57.06)^2]$$

$$= \frac{1}{90}[3511.22 + 297.68 + 2709.32 + 374.54]$$

$$= \frac{6892.76}{90} = 76.5862$$

Hence $S = \sqrt{76.5862} = 8.75$

(b) Here $n_1 = 200, \bar{x}_1 = 25, s_1 = 3,$

$n_2 = 250, \bar{x}_2 = 10, s_2 = 4,$

$n_3 = 300, \bar{x}_3 = 15, s_3 = 5.$

Let $\bar{x}$ and s denote the combined mean and the standard deviation respectively of the combined distribution. Then

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3}$$

$$= \frac{200(25) + 250(10) + 300(15)}{200 + 250 + 300} = \frac{12000}{750} = 16; \text{ and}$$

$$s^2 = \frac{1}{\sum n_i}\left[\sum n_i \ \{s_i^2 + (\bar{x}_i - \bar{x})^2\}\right]$$

$$= \frac{1}{750}[200\{9 + (25 - 16)^2 + 250\{16 + (10 - 16)^2\} + 300\{25 + (15 - 16)^2\}]$$

$$= \frac{1}{750}[18000 + 13000 + 7800] = 51.73, \text{ so that}$$

$$s = \sqrt{51.73} = 7.2; \text{ and hence}$$

$$C.V. = \frac{s}{\bar{x}} \times 100 = \frac{7.2}{16} \times 100 = 45\%$$

4.33 (c) Z-score for the top student $= 50 + 10\left(\frac{x - \bar{x}}{S}\right)$

$$= 50 + 10\left(\frac{98 - 63.7}{12.3}\right)$$

$$= 50 + 28 = 78, \text{ and}$$

Z-score for the bottom student $= 50 + 10\left(\frac{x - \bar{x}}{S}\right)$

$$= 50 + 10\left(\frac{21 - 63.7}{12.3}\right)$$

$$= 50 - 35 = 15.$$

4.34. By definition, a standard Z-score $= 50 + 10\left(\frac{x - \bar{x}}{S}\right)$

Student A:

Z-score on the 1st test $= 50 + 10\left(\frac{70 - 70}{5}\right) = 50$

Z-score on the 2nd test $= 50 + 10\left(\frac{90 - 75}{8}\right) = 69$

$$Z\text{-score on the 3rd test} = 50 + 10\left(\frac{70-60}{12}\right) = 58$$

$$\therefore \text{ Average score of student } A = \frac{50 + 69 + 58}{3} = 59$$

Student B:

$$Z\text{-score on the 1st test} = 50 + 10\left(\frac{90-70}{5}\right) = 90$$

$$Z\text{-score on the 2nd test} = 50 + 10\left(\frac{70-75}{8}\right) = 44$$

$$Z\text{-score on the 3rd test} = 50 + 10\left(\frac{70-60}{12}\right) = 58$$

$$\therefore \text{ Average score of student } B = \frac{90 + 44 + 58}{3} = 64$$

4.35 (b) The data ordered from smallest to largest and the two quartiles are found to be

$$42, 43, 58, \quad \underset{\substack{\uparrow \\ Q_1}}{63}, 65, 67, 68, 72, 75, 75, 78, \underset{\substack{\uparrow \\ Q_3}}{79}, 80, 82, 96.$$

To find the trimmed mean and the trimmed standard deviation, we remove the three observations 42, 43 and 58 below the first quartile and the three observations 80, 82 and 96 above the third quartile. Thus we have nine observations 63, 65, 67, 68, 72, 75, 75, 78, 79 as trimmed data set.

$$\therefore \text{ Trimmed mean} = \frac{63 + 65 + \ldots + 79}{9} = \frac{642}{9} = 71.33, \text{ and}$$

$$\text{Trimmed } S.D. = \sqrt{\frac{(63)^2 + (65)^2 + \ldots + (79)^2}{9} - \left(\frac{642}{9}\right)^2}$$

$$= \sqrt{\frac{46066}{9} - \left(\frac{642}{9}\right)^2}$$

$$= \sqrt{5118.4444 - 5088.4444} = \sqrt{30} = 5.48$$

To find the Winsorized mean and standard deviation, we replace the three observations 42, 43, 58 below Q_1 with 63 and the three observations 80, 82, 96 above Q_3 with 79; and get the Winsorized data set as 63, 63, 63, 63, 65, 67, 68, 72, 75, 75, 78, 79, 79, 79, 79.

$$\text{The Winsorized mean} = \frac{\sum X_i}{n} = \frac{1068}{15} = 71.2, \text{ and}$$

$$\text{the Winsorized } S.D = \sqrt{\frac{76696}{15} - \left(\frac{1068}{15}\right)^2}$$

$$= \sqrt{5113.0667 - 5069.44} = \sqrt{43.6267} = 6.61$$

4.39 (a) The moments about the mean are obtained as below:

$$m_1 = 0$$

$$m_2 = m_2' - (m_1')^2 = 2.5 - (1)^2 = 1.5$$

$$m_3 = m_3' - 3m_2' m_1' + 2(m_1')^3$$

$$= 5.5 - 3(1)(2.5) + 2(1)^3 = 0, \text{ and}$$

$$m_4 = m_4' - 3m_1' m_3' + 6(m_1')^2 m_2' - 3(m_1')^4$$

$$= 16 - 4(1)(5.5) + 6(1)^2(2.5) - 3(1)^4 = 6$$

Again, we have

$$m_1' = \frac{1}{n}\sum f(x-2) \quad \text{.... (1)} \qquad m_2' = \frac{1}{n}\sum f(x-2)^2 \quad \text{.... (2)}$$

$$m_3' = \frac{1}{n}\sum f(x-2)^3 \quad \text{.... (3)} \qquad m_4' = \frac{1}{n}\sum f(x-2)^4 \quad \text{.... (4)}$$

To find the moments about $x=0$, we need the values of $\frac{1}{n}\sum fx$, $\frac{1}{n}\sum fx^2$, $\frac{1}{n}\sum fx^3$ and $\frac{1}{n}\sum fx^4$, which are obtained from the relations (1), (2), (3) and (4).

Thus from (1), we get

$$1 = \frac{\sum fx}{n} - 2 \text{ or } \frac{\sum fx}{n} = 3$$

From (2), we get

$$2.5 = \frac{1}{n}\sum f(x^2 - 4x + 4)$$

or $$2.5 = \frac{1}{n}\sum fx^2 - 4\frac{\sum fx}{n} + 4, \quad \therefore \frac{1}{n}\sum fx^2 = 10.5$$

From (3), we have

$$5.5 = \frac{1}{n}\sum f(x^3 - 6x^2 + 12x - 8)$$

i.e. $$\frac{1}{n}\sum fx^3 - 6\frac{1}{n}\sum fx^2 + 12\frac{\sum fx}{n} - 8 = 5.5$$

or $$\frac{1}{n}\sum fx^3 = 5.5 + 6(10.5) - 12(3) + 8 = 40.5$$

Similarly, from (4) we get $\frac{1}{n}\sum fx^4 = 168.$

Hence the moments about $x=0$ are

$$m_1' = \frac{\sum fx}{n} = 3, \qquad m_2' = \frac{\sum fx^2}{n} = 10.5,$$

$$m_3' = \frac{\sum fx^3}{n} = 40.5, \text{ and} \qquad m_4' = \frac{\sum fx^4}{n} = 168.$$

(b) $$m_1' = \frac{\sum fu}{\sum f} = \frac{-46}{125} = -0.368$$

$$m_2' = \frac{\sum fu^2}{n} = \frac{306}{125} = 2.448$$

$$m_3' = \frac{\sum fu^3}{n} = \frac{-242}{125} = -1.936$$

$$m_4' = \frac{\sum fu^4}{n} = \frac{1962}{125} = 15.696$$

Mean or $\bar{x} = a + \frac{\sum fu}{n} \times h$

$$= 10 + \left(\frac{-46}{125}\right) \times 5 = 10 - 1.84 = 8.16$$

The moments about mean are

$$m_1 = 0$$

$$m_2 = h^2 [m'_2 - (m'_1)^2] = (5)^2 [2.448 - (-0.368)^2]$$

$$= 25(2.313) = 57.825$$

$$m_3 = h^3 [m'_3 - 3m'_1 m'_2 + 2(m'_1)^3]$$

$$= (5)^3 [-1.936 - 3(-0.368)(2.448) + 2(-0.368)^3]$$

$$= 125 [-1.936 + 2.703 - 0.100] = 125(0.667) = 83.375$$

$$m_4 = h^4 [m'_4 - 4m'_1 m'_3 + 6(m'_1)^2 m'_2 - 3(m'_1)^4]$$

$$= (5)^4 [15.696 - 4(-0.368)(-1.936) + 6(-0.368)^2 \times (2.448) - 3(-0.368)^4]$$

$$= 625[15.696 - 2.850 + 1.989 - 0.055] = 625(14.78) = 9237.5$$

Variance $= m_2 = 57.825$

$$b_1 = \frac{m_3^2}{m_2^3} = \frac{(83.375)^2}{(57.825)^3} = \frac{6951.3906}{193351.22} = 0.04, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{9237.5}{(57.825)^2} = \frac{9237.5}{3343.7306} = 2.76.$$

As the value of $b_1 \neq 0$ and $b_2 < 3$, *i.e.* the distribution is platy-kurtic, therefore the distribution would not be considered as normal.

4.40. Calculation of the moments, etc.

x	f	D	fD	fD^2	fD^3	fD^4
1	1	−4	−4	16	−64	256
2	6	−3	−18	54	−162	486
3	13	−2	−26	52	−104	208
4	25	−1	−25	25	−25	25
5	30	0	−73	0	−355	0
6	22	1	22	22	22	22
7	9	2	18	36	72	144
8	5	3	15	45	135	405
9	2	4	8	32	128	512
Σ	113	--	+63 −10	282	357 +2	2058
$\frac{\text{Sum}}{n}$	1	---	−0.089 $= m_1'$	2.496 $= m_2'$	0.018 $= m_3'$	18.212 $= m_4'$

$\therefore$ Moments about the mean are

$$m_1 = 0$$

$$m_2 = m_2' - (m_1')^2 = 2.496 - (0.089)^2 = 2.49$$

$$m_3 = m_3' - 3m_1'm_2' + 2\,(m_1')^3$$

$$= (0.018) - 3(-0.089)\,(2.496) + 2(-0.089)^2 = 0.7$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2\,m_2' + 3(m_1')^4$$

$$= 18.212-4(-0.089)(0.018)+6(-0.089)^2(2.496)-3(-0.089)^4$$

$$= 18.33$$

4.41 Calculation of the first four moments.

Weekly wages x	No. of Labourers (f)	D $(x-20)$	fD	fD^2	fD^3	fD^4
15	6	−5	−30	150	−750	3750
16	19	−4	−76	304	−1216	4864
17	13	−3	−39	117	−351	1053
18	18	−2	−36	72	−144	288
19	20	−1	−20	20	−20	20
20	25	0	−201	0	−2481	0
21	28	1	28	28	28	28
22	34	2	68	136	272	544
23	22	3	66	198	594	1782
24	15	4	60	240	960	3840
Total	200	--	+222 +21	1265	+1854 −627	14169
$\frac{\text{Sum}}{n}$	1	---	0.105 $= m_1'$	6.325 $= m_2'$	−3.135 $= m_3'$	70.845 $= m_4'$

∴ Moments about the mean are:

$$m_1 = 0$$

$$m_2 = m_2' - (m_1')^2 = 6.325 - (0.105)^2 = 6.314$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$= -3.135 - 3(0.105)(6.325) + 2(0.105)^3 = -5.125$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2 m_2' - 3(m_1')^4$$

$$= 70.845 - 4(0.105)(-3.135) + 6(0.105)^2(6.325) - 3(0.105)^4$$

$$= 82.58$$

$$b_1 = \frac{m_3^2}{m_2^3} = \frac{(-5.125)^2}{(6.314)^3} = 0.104, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{82.58}{(6.314)^2} = 2.071.$$

As $b_1 \neq 0$ and $b_2 < 3$, therefore the distribution is not normal.

4.42. Calculation of the first four moments.

Groups	f	x	$u(=\frac{x-11}{2})$	fu	fu^2	fu^3	fu^4
2–4	18	3	−4	−72	288	−1152	4608
4–6	24	5	−3	−72	216	−648	1944
6–8	47	7	−2	−94	188	−376	752
8–10	80	9	−1	−80	80	−80	80
10–12	102	11	0	−318	0	−2256	0
12–14	66	13	1	66	66	66	66
14–16	40	15	2	80	160	320	640
16–18	21	17	3	63	189	567	1701
18–20	15	19	4	60	240	960	3840
Total	413	--	--	+269 −49	1427	+1913 −343	13637
Sum÷n	1	--	--	−0.1186 $= m'_1$	3.4552 $= m'_2$	−0.8305 $= m'_3$	33.0194 $= m'_4$

Hence the moments about the mean and in ordinary units are obtained as below:

$$m_1 = 0$$

$$m_2 = h^2 [m'_2 - (m'_1)^2]$$

$$= 4 [3.4552 - (-0.1186)^2] = 4(3.4411) = 13.76$$

$$m_3 = h^3 [m_3' - 3m_1'm_2' + 2(m_1')^3]$$

$$= 8 [-0.8305 - 3(-0.1186)(3.4552) + 2(-0.1186)^3]$$

$$= 8 (0.3956) = 3.16$$

$$m_4 = h^4 [m_4' - 4m_1'm_3' + 6(m_1')^2 m_2' - 3 (m_1')^4]$$

$$= 16 [33.0194 - 4(-0.1186)(-0.8305) + 6(-0.1186)^2 \times (3.4552) - 3(-0.1186)^4]$$

$$= 528.06$$

$$b_1 = \frac{m_3^2}{m_2^3} = \frac{(3.16)^2}{(13.76)^3} = \frac{9.9856}{2605.2853} = 0.004, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{528.06}{(13.76)^2} = \frac{528.06}{189.3376} = 2.79.$$

4.43. (b) Calculation of b_1 and b_2.

No. of heads (x)	f	D ($x-4$)	fD	fD^2	fD^3	fD^4
0	1	−4	−4	16	−64	256
1	7	−3	−21	63	−189	567
2	26	−2	−52	104	−208	416
3	54	−1	−54	54	−54	54
4	74	0	−131	0	−515	0
5	52	1	52	52	52	52
6	32	2	64	128	256	512
7	9	3	27	81	243	729
8	1	4	4	16	64	256
Total	256	--	147 +16	514	+615 +100	2842
Sum÷n	1	--	−0.0625 $= m_1'$	2.0078 $= m_2'$	0.3906 $= m_3'$	11.1016 $= m_4'$

The moments about the mean are obtained as below:

$m_1 = 0$

$$m_2 = m_2' - (m_1')^2 = 2.0078 - (0.0625)^2 = 2.0039$$

$$m_3 = m_3' - 3m_1'm_2' + 2\,(m_1)^3$$

$$= 0.3906 - 3(0.0625)\,(2.0078) + 2(0.0625)^3 = 0.0146$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2\, m_2' - 3(m_1')^4$$

$$= 11.1016 - 4(0.0625)\,(0.3906) + 6(0.0625)^2\,(2.0078) - 3(0.0625)^4 = 11.0510$$

Hence $b_1 = \dfrac{m_3^2}{m_2^3} = \dfrac{(0.0146)^2}{(2.0039)^3} = 0.0003$, and

$$b_2 = \frac{m_4}{m_2^2} = \frac{11.0510}{(2.0039)^2} = 2.75.$$

4.44. Calculation of b_1 and b_2.

Classes	x	f	$u\left(=\frac{x-19}{2}\right)$	fu	fu^2	fu^3	fu^4
10–12	11	3	−4	−12	48	−192	768
12–14	13	30	−3	−90	270	−810	2430
14–16	15	110	−2	−220	440	−880	1760
16–18	17	218	−1	−218	218	−218	218
18–20	19	275	0	−540	0	−2100	0
20–22	21	222	1	222	222	222	222
22–24	23	108	2	216	432	864	1728
24–26	25	32	3	96	288	864	2592
26–28	27	2	4	8	32	128	512
Total	--	1000	--	+542 +2	1950	2078 −22	10230
Sum÷n	--	1	--	0.002 $= m_1'$	1.950 $= m_2'$	−0.022 $= m_3'$	10.234 $= m_4'$

Now the moments about the mean and in terms of the class interval units are obtained as below:

$$m_1 = 0$$

$$m_2 = m_2' - (m_1')^2 = 1.950 - (0.002)^2 = 1.95$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$= -0.022 - 3(0.002)(1.950) + 2(0.002)^3 = -0.034$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2(m_2') - 3(m_1')^4$$

$$= 10.230 - 4(.002)(-0.022) + 6(.002)^2(1.95) - 3(.002)^4$$

$$= 10.230$$

Hence $b_1 = \dfrac{m_3^2}{m_2^3} = \dfrac{(-0.034)^2}{(1.95)^3} = 0.0002$, and

$$b_2 = \frac{m_4}{m_2^2} = \frac{10.230}{(1.95)^2} = 2.79.$$

Applying Sheppard's corrections, we have

$m_2(\text{corrected}) = m_2 - \dfrac{h^2}{12}$, where m_2 is in ordinary units.

$$= 2^2(1.95) - \frac{2^2}{12} = 7.47$$

$$m_3(\text{corrected}) = m_3$$

$$= (2)^3(-0.034) = -0.27, \text{ and}$$

$$m_4(\text{corrected}) = m_4 - \frac{h^2}{2}m_2 + \frac{7}{240}h^4$$

$$= 2^4\left(10.230 - \tfrac{1}{2} \times 1.95 + \frac{7}{240}\right)$$

$$= 16(10.230 - 0.975 + 0.021) = 148.216$$

Hence $b_1 = \dfrac{m_3^2}{m_2^3} = \dfrac{(-0.27)^2}{(7.47)^3} = 0.002$, and

$$b_2 = \frac{m_4}{m_2^2} = \frac{148.216}{(7.47)^2} = 2.66.$$

The following table is constructed to apply Charlier check.

u	f	$u+1$	$f(u+1)$	$f(u+1)^2$	$f(u+1)^3$	$f(u+1)^4$
−4	3	−3	−9	27	−81	243
−3	30	−2	−60	120	−240	480
−2	110	−1	−110	110	−110	110
−1	218	0	−179	0	−431	0
0	275	1	275	275	275	275
1	222	2	444	888	1776	3552
2	108	3	324	972	2916	8748
3	32	4	12 8	512	2048	8192
4	2	5	10	50	250	1250
Σ	1000	---	+1181 1002	2954	+7265 6834	22850

Now $\sum f(u+1) = \sum fu + n$

$= 2 + 1000 = 1002;$

$\sum f(u+1)^2 = \sum fu^2 + 2\sum fu + n$

$= 1950 + 2(2) + 1000 = 2954;$

$\sum f(u+1)^3 = \sum fu^3 + 3\sum fu^2 + 3\sum fu + n$

$= -22 + 3(1950) + 3(2) + 1000 = 6834$; and

$\sum f(u+1)^4 = \sum fu^4 + 4\sum fu^3 + 6\sum fu^2 + 4\sum fu + n$

$= 10230 + 4(-22) + 6(1950) + 4(2) + 1000 = 22,850$

Hence the result.

4.46 (b) Calculation for skewness.

(i) $$Sk = \frac{Q_1 + Q_3 - 2\text{ Median}}{Q_3 - Q_1} = \frac{37.15 + 61.27 - 2(49.21)}{61.27 - 37.15}$$

$$= \frac{98.42 - 98.42}{24.12} = 0$$

Thus the distribution is symmetrical.

(ii) Since mode (=1487) is greater than mean (1403), therefore the distribution is negatively skewed.

(iii) Given the first three moments about arbirary origin $(x = 16)$ as $m'_1 = -0.35$, $m'_2 = 2.09$, $m'_3 = -1.93$.

We find the third moment about mean to determine the presence of skewness.

$$\therefore \quad m_3 = m'_3 - 3\, m'_1\, m'_2 + 2(m'_1)^3$$

$$= -1.93 - 3\,(-0.35)\,(2.09) + 2(-0.35)^3 = 0.18$$

As m_3 is not equal to zero and is positive, therefore the distribution is positively skewed.

4.47. Calculation of the coefficient of skewness.

Age (years)	No. of Men (f)	x	fx	fx^2	Class Boundries	F
15–19	29	17	493	8381	14.5–19.5	29
20–24	176	22	3872	85184	19.5–24.5	205
25–29	208	27	5616	151632	24.5–29.5	413
30–34	173	32	5536	177152	29.5–34.5	586
35–39	82	37	3034	112258	34.5–39.5	668
40–44	40	42	1680	70560	39.5–44.5	708
45–49	15	47	705	33135	44.5–49.5	723
50–54	3	52	156	8112	49.5–54.5	726
Total	726	--	21092	646414	--	--

(i) Now, $\bar{x} = \frac{\sum fx}{n} = \frac{21092}{726} = 29.05$ years;

$$\text{Mode} = l + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 24.5 + \frac{(208 - 176)}{(208 - 176) + (208 - 173)} \times 5$$

$$= 24.5 + \frac{32}{67} \times 5 = 24.5 + 2.39 = 26.89 \text{ years; and}$$

$$s = \sqrt{\frac{\sum fx^2}{n} - \left(\frac{\sum fx}{n}\right)^2} = \sqrt{\frac{646414}{726} - \left(\frac{21092}{726}\right)^2}$$

$$= \sqrt{890.3774 - 843.9025} = 6.82 \text{ years.}$$

Applying the Pearsonian measure of skewness, we find

$$Sk = \frac{\text{Mean} - \text{Mode}}{s} = \frac{29.05 - 26.89}{6.82} = 0.32$$

(ii) Median = Age of $\left(\frac{n}{2}\right)th$ person

= Age of $\left(\frac{726}{2}\right)th$, *i.e.* $363rd$ man, which lies in the age-group 24.5 – 29.5.

$\therefore$ $$\text{Median} = l + \frac{h}{f}\left(\frac{n}{2} - C\right)$$

$$= 24.5 + \frac{5}{208}(363 - 205) = 24.5 + 3.80$$

$$= 28.30 \text{ years}$$

Similarly, we find

$$Q_1 = l + \frac{h}{f}\left(\frac{n}{4} - C\right)$$

$$= 19.5 + \frac{5}{176}(181.5 - 29) = 19.5 + 4.33$$

$$= 23.83 \text{ years, and}$$

$$Q_3 = l + \frac{h}{f}\left(\frac{3n}{4} - C\right)$$

$$= 29.5 + \frac{5}{173}(544.5 - 413) = 29.5 + 3.80 = 33.30 \text{ yrs}$$

Hence using the Bowley's co-efficient of skewness, we get

$$Sk = \frac{Q_1 + Q_3 - 2\text{ Median}}{Q_3 - Q_1} = \frac{33.30 + 23.83 - 2(28.30)}{33.30 - 23.83}$$

$$= \frac{0.53}{9.47} = 0.06.$$

4.48. Calculation of the first four moments about the mean.

Age (x)	f	$u(=\frac{x-40}{5})$	fu	fu^2	fu^3	fu^4
25	2	−3	−6	18	−54	162
30	8	−2	−16	32	−64	128
35	18	−1	−18	18	−18	18
40	27	0	−40	0	−136	0
45	25	1	25	25	25	25
50	16	2	32	64	128	256
55	7	3	21	63	189	567
60	2	4	8	32	128	512
Total	105	--	+86 +46	252	+470 334	1668
Sum÷n	1	--	0.4381 $= m'_1$	2.4000 $= m'_2$	3.1810 $= m'_3$	15.8857 $= m'_4$

Hence the moments about the mean in terms of class-interval units are obtained below:

$$m_1 = 0;$$

$$m_2 = m'_2 - (m'_1)^2 = 2.4000 - (0.4381)^2 = 2.2081$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$= 3.1810 - 3(0.4381)(2.4000) + 2(0.4381)^3 = 0.1949;$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2(m_2') - 3(m_1')^4$$

$$= 15.8857 - 4(0.4381)(3.1810) + 6(0.4381)^2(2.4000) - 3(0.4381)^4 = 12.9646;$$

$$b_1 = \frac{m_3^2}{m_2^3} = \frac{(0.1949)^2}{(2.2081)^3} = 0.0035, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{12.9646}{(2.2081)^2} = 2.66.$$

Hence the distribution is slightly positively skewed and platykurtic.

4.49. The necessary computations are given below:

(a) Mean, $\bar{X} = 10 + 5\bar{u} = 10 + 5\frac{(-46)}{125} = 8.16$

$$m_1' = \frac{\sum fu}{n} = \frac{-46}{125} = -0.368;$$

$$m_2' = \frac{\sum fu^2}{n} = \frac{306}{125} = 2.448;$$

$$m_3' = \frac{\sum fu^3}{n} = \frac{-242}{125} = -1.936;$$

$$m_4' = \frac{\sum fu^4}{n} = \frac{1962}{125} = 15.696.$$

$$m_2(\text{variance}, s^2) = h^2[m_2' - (m_1')^2] = 25[2.448 - (-0.368)^2]$$

$$= 25(2.313) = 57.825, \text{ so that } s = 7.60$$

$$m_3 = 125[-1.936 - 3(-0.368)(2.448) + 2(-0.368)^3]$$

$$= 125[-1.936 + 2.703 - 0.100] = 125(0.667) = 83.375$$

$$m_4 = (5)^4\,[15.696-4(-0.368)(-1.936)+6(-0.368)^2\,(2.448) - 3(-0.368)^4]$$

$$= 625\,[15.696 - 2.850 + 1.989 - 0.055]$$

$$= 625\,[14.78] = 9237.5$$

$$b_1 = \frac{m_3^2}{m_2^3} = \frac{(83.375)^2}{(57.825)^3} = 0.04, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{9237.5}{(57.825)^2} = 2.76.$$

(b) Mean, $\bar{X} = 20 + \bar{v} = 20 + \dfrac{(21)}{200} = 20.105;$

$$m_1' = \frac{21}{200} = 0.105;\quad m_2' = \frac{1265}{200} = 6.325;$$

$$m_3' = \frac{-627}{200} = -3.135;\quad m_4' = \frac{14169}{200} = 70.845$$

$$m_2(\text{variance, } s^2) = m_2' - (m_1')^2 = 6.325 - (0.105)^2$$

$$= 6.314, \text{ so that } s = \sqrt{6.314} = 2.51.$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$= -3.135 - 3(0.105)\,(6.325) + 2(0.105)^3 = -5.125;$$

Similarly, $m_4 = 82.58;$

$$b_1 = \frac{(-5.125)^2}{(6.314)^3} = 0.104 \text{ and } b_2 = \frac{82.58}{(6.314)^2} = 2.07$$

(i) A distribution with smaller C.V. will be more consistent.

$$\text{C.V. for (a)} = \frac{s}{\bar{x}} \times 100 = \frac{7.60}{8.16} \times 100 = 93.14\%$$

$$\text{C.V. for (b)} = \frac{2.51}{20.105} \times 100 = 12.48\%$$

Hence distribution (b) is more consistent.

(ii) A distribution having m_3 negative, will be negatively skewed. The distribution (b) has $m_3 = -5.125$, so it is negatively skewed.

(iii) A distribution with $b_1 = 0$ and $b_2 = 3$ will be mesokurtic. None of the distributions is mesokurtic.

4.50. (a) For a distribution to be mesokurtic, b_2 equals 3.

We are given that $m_4 = 24.3$. Therefore

$$\frac{m_4}{m_2^2} = b_3 = 3 \quad i.e. \quad \frac{243}{m_2^2} = 3$$

or $$m_2^2 = \frac{243}{3} = 81 \quad \text{or} \quad m_2 = 9$$

Hence the desired value of the standard deviation $= \sqrt{9} = 3$.

(b) First we calculate the moments about the mean. Thus

$$m_1 = 0;$$

$$m_2 = m_2' - (m_1')^2 = 17 - (-1.5)^2 = 14.75;$$

$$m_3 = m_3' - 3m_1'm_2' + 2(m_1')^3$$

$$= -30 - 3(-1.5)(17) + 2(-1.5)^3$$

$$= 76.5 - 36.75 = 39.75; \text{ and}$$

$$m_4 = m_4' - 4m_1'm_3' + 6(m_1')^2(m_2') - 3(m_1')^4$$

$$= 108 - 4(-1.5)(-30) + 6(-1.5)^2(17) - 3(-1.5)^4$$

$$= 337.50 - 195.19 = 142.31$$

Hence $$b_1 = \frac{m_3^2}{m_2^3} = \frac{(39.75)^2}{(14.75)^3} = 0.49, \text{ and}$$

$$b_2 = \frac{m_4}{m_2^2} = \frac{142.31}{(14.75)^2} = 0.65.$$

The distribution is platy-kurtic as $b_2 < 3$.

4.51 **(a)** b_2 in the first case $= \dfrac{m_4}{m_2^2} = \dfrac{230}{(9)^2} = 2.84 < 3$

b_2 in the second case $= \dfrac{m_4}{m_2^2} = \dfrac{780}{(16)^2} = 3.05 > 3$

Hence we may conclude that

(i) the second distribution is lepto-kurtic,

(ii) neither distribution is mesokurtic, and

(iii) the first distribution is platy-kurtic.

(b) We have $b_2 = \dfrac{m_4}{m_2^2} = \dfrac{m_4}{(25)^2}$

(i) For leptokurtic, b_2 must be greater than 3,

$$i.e.\ \frac{m_4}{625} > 3 \text{ or } m_4 > 1875.$$

(ii) For meso-kurtic, b_2 is equal to 3,

$$i.e.\ \frac{m_4}{625} = 3 \text{ or } m_4 = 1875$$

(iii) For platy-kurtic, b_2 must be less than 3,

$$i.e.\ \frac{m_4}{625} < 3 \text{ or } m_4 < 1875$$

●●●●●●●●●