First we form the following table for necessary calculations.

Table 5.5

Calculation for standard deviation

Income	class-mark (x_i)	$y_i = \frac{x_i - 375}{25}$	No. of person (f_i)	$y_i f_i$	$y_i^2 f_i$
150-300	225	-6	232	-1392	8352
300-450	375	0	128	0	0
450-600	525	6	60	360	2160
600-750	675	12	40	480	5760
750-900	825	18	28	504	9072
900-1100	1000	25	12	300	7500
1100-1500	1300	37	6	222	8214
Total	—	—	506	474	41058

Now $s_y^2 = \frac{1}{n}\sum f_i y_i^2 - \left(\frac{1}{n}\sum f_i y_i\right)^2 = \frac{41058}{506} - \left(\frac{474}{506}\right)^2$.

$= 81.14 - 0.8774 = 80.2626$

Here $x_i = 375 + 25 y_i$ $\therefore s_x^2 = (25)^2 \times s_y^2 = 625 \times 80.2626$
$= 50164.125$

Hence $s_x = \sqrt{50164.125} = 224$ units. (Approx)

(d) Quartile deviation : Quartile deviation is a measure of dispersion based on the quartiles. One can take the average of the difference between Q_2 and Q_1 and that between Q_3 and Q_2 as a measure of dispersion, called Quartile deviation (denoted by Q). So,

$$Q = \frac{(Q_2 - Q_1) + (Q_3 - Q_2)}{2} = \frac{Q_3 - Q_1}{2} \quad \text{--- (5k)}$$

If two variables x and y are related as $y = a + bx$, then their quartile deviations are related as $Q(y) = |b| Q(x)$

Example 5.9 The following table gives the distribution of monthly expenditure for 430 families in a particular region:

Monthly expenditure (Rs.)	frequency
Less than 1000	30
1000 – 1250	45
1250 – 1500	70
1500 – 1750	82
1750 – 2000	66
2000 – 2250	57
2250 – 2500	28
2500 – 2750	22
2750 – 3000	18
more than 3000	12

Can you compute the standard deviation from the above data? Find the dispersion using a suitable measure

Solution: Here the terminal classes of the frequency table are open and therefore standard deviation can not be obtained. The quartile deviation is a suitable measure of dispersion here. Now we form the following table for cumulative frequencies.

Table 5.6
Cumulative frequency table

Monthly expenditure (Rs.)	Frequency	Cumulative frequency
Less than 1000	30	30
1000 – 1250	45	75
1250 – 1500	70	145
1500 – 1750	82	227

(contd)

Scanned with CamScanner

(contd)

1750 - 2000	66	293
2000 - 2250	57	350
2250 - 2500	28	378
2500 - 2750	22	400
2750 - 3000	18	418
more than 3000	12	430
Total	430	--

From the cumulative frequency column of Table 5.6 we find that the class interval 1250-1500 and 2000-2250 respectively contain Q_1 and Q_3

Here $Q_1 = 1250 + \frac{107.5 - 75}{70} \times 250 = 1250 + \frac{8125}{70}$

$= 1250 + 116.07 = 1366.07$ (Rs.)

and $Q_3 = 2000 + \frac{322.5 - 293}{57} \times 250 = 2000 + \frac{7375}{57}$

$= 2000 + 129.38 = 2129.38$ (Rs.)

$\therefore Q = \frac{Q_3 - Q_1}{2} = \frac{2129.38 - 1366.07}{2} = \frac{763.31}{2} = 381.65$ (Rs)

6. Moments, skewness and Kurtosis

6.1 Moments we define moment of order r about an arbitrary origin A, denoted by $m_r'(A)$, as

$$m_r'(A) = \frac{1}{n} \sum_{i=1}^{n} (x_i - A)^r, \quad r = 0, 1, 2, \ldots$$

for n values x_i, $i = 1, 2, \ldots, n$ of a variable x.

Scanned with CamScanner

If the given data are arranged in a frequency table, the formula becomes $m'_r(A) = \frac{1}{n} \sum_{i=1}^{k} f_i (x_i - A)^r$

where x_i and f_i are respectively the class mark and frequency of the ith class, $i = 1, 2, \ldots, k$ and $n = \sum_{i=1}^{k} f_i$

Moments about zero are usually termed as raw moments. When the arithmetic mean is taken as the origin, moments are called central moments. So, for n values x_i, $i = 1, 2, \ldots, n$, the central moment of order r (or rth central moment), denoted by m_r, is given by $m_r = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^r$ while for a frequency distribution

$$m_r = \frac{1}{n} \sum_{i=1}^{k} f_i (x_i - \bar{x})^r$$

It is evident that the mean of a variable is its first raw moment and variance is the second central moment.

Note: When $y = a + bx$, then $m_r(y) = b^r m_r(x)$

This shows that the central moments are independent of origin but depend on change of scale

6.2 Relationship between central and non-central moments

(a) Central moments expressed in terms of moments about an arbitrary origin

Let A be an arbitrary origin. Then we have

$$\bar{x} - A = \frac{1}{n} \sum_{i=1}^{n} x_i - A = \frac{1}{n} \sum_{i=1}^{n} (x_i - A) = m'_1(A)$$

Now central moment of order r is given by

$$m_r = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^r = \frac{1}{n} \sum_{i=1}^{n} \{(x_i - A) - (\bar{x} - A)\}^r$$

Scanned with CamScanner