Example 29. The regression coefficient of Y on X, *i.e.*,

$$byx = 1.2, \text{ If } u = \frac{X-100}{2} \text{ and } v = \frac{Y-100}{3}, \text{ Find } bvu.$$

Solution:

$$byx = \frac{\Sigma(X-\overline{X})(Y-\overline{Y})}{\Sigma(X-\overline{X})^2}$$

$$u = \frac{X-100}{2} \qquad v = \frac{Y-100}{3}$$

$$\Rightarrow \quad X = 2u+100 \qquad \Rightarrow \quad Y = 3v+100$$

$$\overline{X} = 2\overline{u}+100 \qquad \overline{Y} = 3\overline{v}+100$$

$$\therefore \quad (X-\overline{X}) = 2(u-\overline{u}) \qquad \therefore \quad (Y-\overline{Y}) = 3(v-\overline{v})$$

Now,
$$byx = \frac{2\times 3\,\Sigma(u-\overline{u})(v-\overline{v})}{4\cdot\Sigma(u-\overline{u})^2} = 1.5\,bvu$$

So,
$$bvu = \frac{byx}{1.5}$$

$$bvu = \frac{1.2}{1.5} = 0.8$$

EXERCISE 2.5

1. The following data relate to marks in English and Maths:

Mean marks in English	39.5
Mean marks in Maths	47.6
S.D. of marks in English	10.8
S.D. of marks in Maths	16.9
Coefficient of correlation between English and Maths = + 0.42	

(*i*) Obtain the two regression equations.

(*ii*) Calculate the expected average marks in Maths of candidates who received 50 marks in English.

[**Ans.** $X = 0.268Y + 26.73$, $Y = 0.657X + 21.64$, 54.5]

2. There are two series of index numbers, P for price index and S for stock commodities. The mean and standard deviation of P are 100 and 8 respectively and S are 103 and 4. The correlation coefficient between the two series is 0.4. With this data work out a linear equation to read off values of P for various values of S. Can the same equation be used to read off the values of S for different values of P? If not, give the appropriate equation.

[**Ans.** $P = 0.8S + 17.6$; No; $S = 0.2P + 83$]

3. The following results were worked out from marks in Statistics and Mathematics in a certain examination:

	Marks in Statistics (X)	Marks in Maths (Y)
A.M.	39.5	47.5
S.D.	10.8	17.8
Coefficient of correlation = 0.42		

(*i*) Find the two regression equations.

(*ii*) Estimate the value of Y when $X = 50$ and X, when $Y = 30$.

[**Ans.** (*i*) $Y = 0.6922X + 20.1581$, $X = 0.2548Y + 27.397$; (*ii*) 54.76, 35.04]

4. You are given the following data about sales and advertisement expenditure of a firm:

	Sales (Rs. crore)	Adv. Expenditure (Rs. crore)
Arithmetic Mean:	50	10
Standard Deviation:	10	2
Coefficient of correlation = +0.9		

(*i*) Calculate the two regression equations.

(*ii*) Estimate the likely sales for a proposed advertisement expenditure of Rs. 13.5 crore.

(*iii*) What should be the advertisement budget if the company wants to achieve a sales target of Rs. 70 crore?

[**Ans.** (*i*) $X = 4.5Y + 5$, $Y = 0.18X + 1$, (*ii*) 65.75 crores, (*iii*) 13.6 crores]

5. Given the following data, what would be the possible yield when rainfall is 29"?

	Rainfall	Yield per acre
Mean	25"	40
Variance	9"	36
Coefficient of correlation between rainfall and production = 0.8		

[**Ans.** 46.4]

6. For a given set of bivariate data, the following results were obtained:
$\overline{X} = 53.2$, $\overline{Y} = 27.9$, regression coefficient of Y on $X = -1.5$
Regression coefficient of X on $Y = -0.2$. Find the most probable value of Y when $X = 60$. Also find the coefficient of correlation.

[**Ans.** $Y_{60} = 17.7$, $r = -0.548$]

7. If $\overline{X} = 45$, $\sigma_x = 2.5$, $\overline{Y} = 60$, $\sigma_y = 2.2$, $r = 0.75$
Estimate (*i*) Value of Y when $X = 35$
(*ii*) Value of X when $Y = 20$.

[**Ans.** (*i*) 53.4, (*ii*) 10.92]

8. Following information are given:

	X Series	Y Series
Mean:	10	8
Standard Deviation:	6	5

The covariance between X and Y is 15. Estimate the value of X when $Y = 9$.

[Ans. $X = 0.6Y + 5.2, X_9 = 10.6$**]**

9. If $\overline{Y} = 15, \overline{X} = 3.5, b_{yx} = 2.5$
Obtain estimate of Y when $X = 5$. **[Ans. 18.75]**

10. If $\sigma_x^2 = 0.75, \sigma_y^2 = 1.2, r_{xy} = 0.65$, find b_{xy}. **[Ans. 0.52]**

11. If $\sigma_x^2 = 25, \sigma_y^2 = 625, bxy = 0.16$, find '$r$'. **[Ans.** $r = 0.8$**]**

12. If $byx = 0.50, bxy = 1.5$, find r. **[Ans.** $r = 0.87$**]**

13. A group of 20 students was observed for weight (X) and height (Y). The variance of height was found to be 9 cm and that of weight 1600 gm. If the correlation coefficient between height and weight was 0.5, obtain an average absolute increase in weight in response to height. **[Ans. 6.67 gm]**
[Hint: Find *bxy*]

14. In a regression analysis, the following two regression coefficients were obtained
$bxy = 3.5$ and $byx = 0.5$
Comment on these values. **[Ans.** $r^2 = 1.75 \Rightarrow r = 1.32 \ngtr 1$**, Inconsistent values]**

◘ TO OBTAIN REGRESSION EQUATIONS IN CASE OF GROUPED DATA

For obtaining regression equations from grouped data, first of all we have to construct a correlation table. After that, we find out $\overline{X}, \overline{Y}$ and the regression coefficients *byx* and *bxy*. Special adjustment must be made while calculating the value of regression coefficients because regression coefficients are independent of change of origin but not of scale. In grouped data, the regression coefficients (*byx* and *bxy*) are computed by using the following formulae:

(i) $$bxy = \frac{N \times \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \times \Sigma fdy^2 - (\Sigma fdy)^2} \times \frac{i_x}{i_y}$$

(ii) $$byx = \frac{N \,.\, \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \,.\, \Sigma fdx^2 - (\Sigma fdx)^2} \times \frac{i_y}{i_x}$$

Where, i_x = Common factor of X–variable

i_y = Common factor of Y–variable.

The following examples makes the computation of regression equations more clear:

Example 30. Obtain the two regression equations from the following bivariate frequency distribution:

Y	X: 0—20	X: 20—40	X: 40—60
10—25	10	5	3
25—40	4	40	8
40—55	6	9	15

Also compute Karl Pearsons' coefficient correlation from two regression coefficients.

Solution: **(Table Given at Page 120)**

$$\overline{X} = A + \frac{\Sigma fdx}{N} \times i_x$$

$$= 30 + \frac{6}{100} \times 20$$

$$= 30 + \frac{120}{100}$$

$$= 30 + 1.2 = 31.2$$

$$\overline{Y} = A + \frac{\Sigma fdy}{N} \times i_y$$

$$= 32.5 + \frac{12}{100} \times 15$$

$$= 32.5 + 1.8 = 34.3$$

$$bxy = \frac{N \times \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \times \Sigma fdy^2 - (\Sigma fdy)^2} \times \frac{i_x}{i_y}$$

$$= \frac{(100)(16) - (6)(12)}{(100)(48) - (12)^2} \times \frac{20}{15}$$

$$= \frac{1600 - 72}{4800 - 144} \times \frac{20}{15} = \frac{1528}{4656} \times \frac{20}{15}$$

$$= \frac{30560}{69840} = +0.43$$

$$byx = \frac{N \,.\, \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \,.\, \Sigma fdx^2 - (\Sigma fdx)^2} \times \frac{i_y}{i_x}$$

$$= \frac{(100)(16) - (6)(12)}{(100)(46) - (6)^2} \times \frac{15}{20}$$

$$= \frac{1600 - 72}{4600 - 36} \times \frac{15}{20} = \frac{1528}{4564} \times \frac{15}{20}$$

$$= \frac{22920}{91280} = +0.25$$

Solution 30.

Let $dx = \frac{X-30}{20}$, $dy = \frac{Y-32.5}{15}$

Y ↓	M.V.		dy \ dx	X → 0—20	20—40	40—60	f	fdy	fdy²	fdxdy
			M.V.	10	30	50				
				−20	0	+20				
			dx	−1	0	+1				
10–25	17.5	−15	−1	1 / 10 / 10	0 / 5 / 0	−1 / 3 / −3	18	−18	18	7
25–40	32.5	0	0	0 / 4 / 0	0 / 40 / 0	0 / 8 / 0	52	0	0	0
40–55	47.5	+15	+1	−1 / 6 / −6	0 / 9 / 0	1 / 15 / 15	30	+30	30	9
			f	20	54	26	N = 100	Σ fdy = +12	Σ fdy² = 48	Σ fdxdy = 16
			fdx	−20	0	+26	Σ fdx = 6			
			fdx²	20	0	+26	Σ fdx² = 46			
			fdxdy	4	0	12	Σ fdxdy = 16			

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$
$$Y - 34.3 = 0.25(X - 31.2)$$
$$Y - 34.3 = 0.25X - 7.8$$
$$Y = 0.25X + 26.5$$

Regression Equation of X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$
$$X - 31.2 = 0.43(Y - 34.3)$$
$$X - 31.2 = 0.43Y - 14.749$$
$$X = 0.43Y + 16.451$$

Coefficient of Correlation (r)

$$r = \sqrt{bxy\,.\,byx}$$
$$= \sqrt{(+0.43)(+0.25)}$$

$\therefore\ r = +0.33$ approx.

Aliter: r can also be computed by using the formula:

$$r = \frac{N\,.\Sigma fdxdy - \Sigma fdx\,.\Sigma fdy}{\sqrt{N\,.\Sigma fdx^2 - (\Sigma fdx)^2}\ \sqrt{N\,.\Sigma fdy^2 - (\Sigma fdy)^2}}$$
$$= \frac{100 \times 16 - (6)(+12)}{\sqrt{100 \times 46 - (6)^2}\ \sqrt{100 \times 48 - (12)^2}}$$
$$= \frac{1600 - 72}{\sqrt{4600 - 36}\ \sqrt{4800 - 144}} = \frac{1528}{\sqrt{4564}\ \sqrt{4656}}$$
$$= \frac{1528}{4609.77} = 0.33$$

Example 31. Following is the distribution of students according to their height and weight:

Height (in inches)	Weight (in lbs.) 90—100	100—110	110—120	120—130
50—55	4	7	5	2
55—60	6	10	7	4
60—65	6	12	10	7
65—70	3	8	6	3

Calculate: (*i*) the two regression coefficients.

(*ii*) the two regression equations.

(*iii*) coefficient of determination.

Solution: **(Table given at page 122)** Let X denote height and Y denote weight.

Solution 31.

Let $dx = \dfrac{X-57.5}{5}$, $dy = \dfrac{Y-105}{10}$

X ↓ M.V. / Y → M.V.		dx \ dy		90—100	100—110	110—120	120—130	f	fdx	fdx²	fdxdy
			M.V.	95	105	115	125				
				−10	0	+10	+20				
			dy	−1	0	+1	+2				
50—55	52.5	−5	−1	^{1} 4 $_{4}$	^{0} 7 $_{0}$	$^{-1}$ 5 $_{-5}$	$^{-2}$ 2 $_{-4}$	18	−18	18	−5
55—60	57.5	0	0	^{0} 6 $_{0}$	^{0} 10 $_{0}$	^{0} 7 $_{0}$	^{0} 4 $_{0}$	27	0	0	0
60—65	62.5	+5	+1	$^{-1}$ 6 $_{-6}$	^{0} 12 $_{0}$	^{1} 10 $_{10}$	^{2} 7 $_{14}$	35	35	35	18
65—70	67.5	+10	+2	$^{-2}$ 3 $_{-6}$	^{0} 8 $_{0}$	^{2} 6 $_{12}$	^{4} 3 $_{12}$	20	40	80	18
			f	19	37	28	16	N=100	57	133	31
			fdy	−19	0	+28	32	41			
			fdy²	19	0	28	64	111			
			fdxdy	−8	0	17	22	31			

(i) Regression Coefficient of Y on X

$$byx = \frac{N.\Sigma fdxdy - \Sigma fdx.\Sigma fdy}{N.\Sigma fdx^2 - (\Sigma fdx)^2} \times \frac{i_y}{i_x}$$

$$byx = \frac{100\times 31-(57)(41)}{100\times 133-(57)^2} \times \frac{10}{5}$$

$$= \frac{3100-2337}{13300-3249} \times \frac{2}{1} = \frac{763}{10051} \times 2 = \frac{1526}{10051} = 0.151 \approx 0.15$$

(ii) Regression Coefficient of X on Y

$$bxy = \frac{N\times \Sigma fdxdy - \Sigma fdx.\Sigma fdy}{N\times \Sigma fdy^2 - (\Sigma fdy)^2} \times \frac{i_x}{i_y}$$

$$= \frac{100\times 31-(57)(41)}{100\times 111-(41)^2} \times \frac{5}{10} = \frac{3100-2337}{11100-1681} \times \frac{1}{2}$$

$$= \frac{763}{9419} \times \frac{1}{2} = \frac{763}{18838} = 0.040 \approx 0.04$$

Calculation of $\overline{X}$ and $\overline{Y}$

$$\overline{X} = A + \frac{\Sigma fdx}{N} \times i_x$$

$$= 57.5 + \frac{57}{100} \times 5 = 57.5 + \frac{285}{100}$$

$$= 57.5 + 2.85 = 60.35$$

$$\overline{Y} = A + \frac{\Sigma fdy}{N} \times i_y$$

$$= 105 + \frac{41}{100} \times 10 = 105 + \frac{410}{100}$$

$$= 105 + 4.10 = 109.10$$

Regression Equation of Y on X

$$Y - \overline{Y} = byx(X - \overline{X})$$

$$Y - 109.10 = 0.15(X - 60.35)$$

$$Y = 0.15X - 9.0525 + 109.10$$

$$Y = 0.15X + 100.04$$

Regression Equation of X on Y

$$X - \overline{X} = bxy(Y - \overline{Y})$$

$$X - 60.35 = 0.04(Y - 109.10)$$

$$X - 60.35 = 0.04Y - 4.364$$

$$X = 0.04Y + 55.98$$

(iii) **Correlation Coefficient (*r*)**

$$r = \sqrt{byx \,.\, bxy}$$

$$= \sqrt{0.15 \times 0.04} = \sqrt{0.006} = 0.077$$

Thus, Coefficient of Determination $= r^2 = (0.077)^2 = 0.006$

Example 32. From the following grouped data, find two regression equations and correlation co-efficient.

Wife's Age	Husband's Age			
	20—25	25—30	30—35	35—40
15—20	20	10	3	2
20—25	4	28	6	4
25—30	—	5	11	—
30—35	—	—	2	—
35—40	—	—	—	5

Solution: **(Table Given at Page 125)**

Regression Coefficient of X on Y

$$bxy = \frac{N \times \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \times \Sigma fdy^2 - (\Sigma fdy)^2} \times \frac{i_x}{i_y}$$

$$= \frac{100(138) - (-80)(-100)}{100(204) - (-100)^2} \times \frac{5}{5}$$

$$= \frac{13800 - 8000}{20400 - 10000} = \frac{5800}{10400} = 0.558$$

$\therefore$ $bxy = 0.558$

Regression Coefficient of Y on X

$$byx = \frac{N \,.\, \Sigma fdxdy - \Sigma fdx \,.\, \Sigma fdy}{N \,.\, \Sigma fdx^2 - (\Sigma fdx)^2} \times \frac{i_y}{i_x}$$

$$= \frac{100(138) - (-80)(-100)}{100(150) - (-80)^2} \times \frac{5}{5}$$

$$= \frac{13800 - 8000}{15000 - 6400} = \frac{5800}{8600} = 0.674$$

$\therefore$ $byx = 0.674$

$$\bar{X} = A + \frac{\Sigma fdx}{N} \times i_x = 32.5 + \left(\frac{-80}{100}\right)(5) = 28.5$$

$$\bar{Y} = A + \frac{\Sigma fdy}{N} \times i_y = 27.5 + \left(\frac{-100}{100}\right)(5) = 22.5$$

Solution 32.

Let $dx = \frac{X - 32.5}{5}$, $dy = \frac{Y - 27.5}{5}$

Y ↓	M.V.	dy	dx	20—25	25—30	30—35	35—40	f	fdy	fdy²	fdxdy
X →			M.V.	22.5	27.5	32.5	37.5				
			dx	–10	–5	0	5				
			dy	–2	–1	0	1				
15—20	17.5	–10	–2	4 20 80	2 10 20	0 3 0	–2 2 –4	35	–70	140	96
20—25	22.5	–5	–1	2 4 8	1 28 28	0 6 0	–1 4 –4	42	–42	42	32
25—30	27.5	0	0	—	0 5 0	0 11 0	—	16	0	0	0
30—35	32.5	+5	1	—	—	0 2 0	—	2	2	2	0
35—40	37.5	+10	2	—	—	—	2 5 10	5	10	20	10
			f	24	43	22	11	N = 100	–100	204	138
			fdx	–48	–43	0	11	= –80			
			fdx²	96	43	0	11	= +150			
			fdxdy	88	48	0	2	= +138			

Regression Equation of X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$
$$X - 28.5 = 0.558(Y - 22.5)$$
$$X - 28.5 = 0.558Y - 12.555$$
$$X = 15.945 + 0.558Y$$

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$
$$Y - 22.5 = 0.674(X - 28.5)$$
$$Y - 22.5 = 0.674X - 19.209$$
$$Y = 0.674X - 19.209 + 22.5$$
$$Y = 3.291 + 0.674X$$

Now $r = \sqrt{bxy \times byx}$

$= \sqrt{0.558 \times 0.674} = +0.613$

EXERCISE 2.6

1. Obtain two regression equations for the following grouped data:

Sales (Rs. '000) (Y)	Advertisement Exp. (Rs. '000) (X)			
	5—15	15—25	25—35	35—45
75—125	4	1	—	—
125—175	7	6	2	1
175—225	1	3	4	2
225—275	1	1	3	4

Also find coefficient of correlation. **[Ans.** $X = 0.134Y - 1.45$, where X denotes Adv. Exp. $Y = 2.65X + 119.13$ where Y denotes sales, $r = +0.595$]

2. Obtain the regression equations from the following data:

Marks in English (X)	Marks in Statistics (Y)					Total
	10—20	20—30	30—40	40—50	50—60	
10—20	6	3	—	—	—	9
20—30	3	16	10	—	—	29
30—40	—	10	15	7	—	32
40—50	—	—	7	10	4	21
50—60	—	—	—	4	5	9
Total	9	29	32	21	9	100

Also find coefficient of determination. **[Ans.** $X = 6.77 + 0.802Y$; $Y = 6.77 + 0.802X$, $r^2 = 0.802$]

3. The following table shows the frequency distribution of 50 couples classified according to their ages.

Age of Wives (X)	Age of Husband (Y)			Total
	20—25	25—30	30—35	
16—20	9	14	—	23
20—24	6	11	3	20
24—28	—	—	7	7
Total	15	25	10	50

Estimate (*i*) the age of husband when wife's age is 20 years and (*ii*) the age of wife when husband's age is 30 years.

[Ans. $X = 0.47Y + 8.03$; $Y = 0.72X + 12.02$; $Y = 26.42$ years; $X = 22.13$ years]

4. Find regression equations from the following data:

Marks in Statistics	Marks in Economics				Total
	4—8	8—12	12—16	16—20	
8—14	11	6	2	1	20
14—20	5	12	15	8	40
20—26	—	2	3	15	20
Total	16	20	20	24	80

[Ans. $X = 0.67Y + 1.27$; $Y = 0.611X + 9.3$]

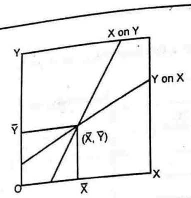
5. Find coefficient of correlation and regression equations from the following data:

Y \ X	5—15	15—25	25—35	35—45
0—10	1	1	—	—
10—20	3	6	5	1
20—30	1	8	9	2
30—40	—	3	9	3
40—50	—	—	4	4

[Ans. $r = +0.53$; $X = 15.58 + 0.42Y$; $Y = 8.91 + 0.67X$]

TO OBTAIN THE MEAN VALUES AND CORRELATION COEFFICIENT FROM THE REGRESSION EQUATIONS

(1) To Find the Mean Values from the Regression Equations: Two regression lines intersect each other at mean values ($\bar{X}$ and $\bar{Y}$) points. In other words, the point of intersection of the two lines of regression give the mean values of both X and Y variables, *i.e.*, $\bar{X}$ and $\bar{Y}$. This is clear from the following diagram:



The above diagram makes it clear that two regression lines Y on X and X on Y intersect each other at $\bar{X}$ and $\bar{Y}$ points. Therefore, it is clear that if both regression lines/regression equations are known, then solving them out gives the mean values X and Y series, *i.e.*, $\bar{X}$ and $\bar{Y}$.

(2) To Find the Coefficient of Correlation from two Regression Equations: Correlation coefficient can be worked out from the regression coefficients *bxy* and *byx*. From the regression equation of X on Y, we can find out *bxy* and from the regression equation of Y on X, we can find out *byx* and then correlation coefficient *r* can be derived as $r = \sqrt{byx \cdot bxy}$.

Note: But sometimes the regression equations are given in such way that by inspection, it is difficult to make out which one is the regression equation of X on Y and which one is Y on X. In such a case, any of them, taken as a regression equation of X on Y, is used to compute *bxy*. Similarly, with the help of other equation *byx* is computed. If the product of *bxy* and *byx* yields more than unity (1), then this follows that our supposition is wrong, because r^2 cannot exceed unity. That means we have to change our supposition. This time we have to make the supposition other way round, *i.e.*, previously taken to be X on Y should now be taken for Y on X. Thus, the product of regression coefficients shall not exceed unity and our results will be correct.

Alternative Method: To find out which regression equation is X on Y and which is Y on X, the following alternative method can also be applied:

Suppose two regression equations are as follows:

(1) $a_1x + b_1y + c_1 = 0$

(2) $a_2x + b_2y + c_2 = 0$

(*i*) If $a_1\ b_2 \le a_2\ b_1$ (in magnitude, *i.e.*, ignoring signs), then

$a_1x + b_1y + c_1 = 0$ is the regression of Y on X and

$a_2x + b_2y + c_2 = 0$ is the regression of X on Y.

(*ii*) If $a_1\ b_2 > a_2\ b_1$ (in magnitude), then

$a_1x + b_1y + c_1 = 0$ is the regression of X on Y and

$a_2x + b_2y + c_2 = 0$ is the regression of Y on X.

Example 33. From the following two regression equations, identify which one is of X on Y and which one is of Y on X:

$$2X + 3Y = 42$$
$$X + 2Y = 26$$

Solution: In the absence of any clear cut indication, let us assume that equation first to be Y on X and equation second to be of X on Y.

Let equation first be regression equation of Y on X

$$2X + 3Y = 42$$
$$3Y = 42 - 2X$$
$$\Rightarrow \quad Y = \frac{42}{3} - \frac{2X}{3} \quad \text{...}(i)$$

From this it follows that

$$byx = \text{Coefficient of } X \text{ in } (i) = \frac{-2}{3}$$

Now, equation (*ii*) be regression equation of X on Y

$$X + 2Y = 26$$
$$X = 26 - 2Y \quad \text{...}(ii)$$

From this it follows that

$$bxy = \text{Coefficient of } Y \text{ in } (ii) = -2$$

Now, we calculate '*r*' on the basis of the above values of two regression coefficient, we get

$$r^2 = byx \cdot bxy = \frac{-2}{3} \times -2 = \frac{4}{3} > 1$$

Here, $r^2 > 1$ which is impossible as $r^2 \le 1$. So, our assumption is wrong. We now choose equation (*i*) as regression of *X* on *Y* and (*ii*) as regression equation of *Y* on *X*:

Assuming the first equation as of *X* on *Y*, we have

$$2X + 3Y = 42$$

or

$$2X = 42 - 3Y$$
$$X = \frac{42}{2} - \frac{3Y}{2} \quad \text{...}(iii)$$

From this, it follows that

$$bxy = \text{Coefficient of Y in } (iii) = \frac{-3}{2}$$

Now, assuming the second equation as Y on X, we have

$$X + 2Y = 26$$

or

$$2Y = 26 - X$$
$$\Rightarrow \quad Y = \frac{26}{2} - \frac{1}{2}X \quad \text{...}(iv)$$

From this it follows that

$$byx = \text{Coefficient of } X \text{ in } (iv) = -\frac{1}{2}$$

Now, $r^2 = bxy \cdot byx = -\frac{3}{2} \times \frac{-1}{2} = \frac{3}{4} = 0.75$

Here, $r^2 < 1$ which is possible. r^2 is within the limit, *i.e.*, $r^2 \leq 1$.
Hence, it is proved that the first equation is of X on Y and the second equation is of Y on X.

Aliter:
$$2X + 3Y = 42 \quad \text{...(i)}$$
$$X + 2Y = 26 \quad \text{...(ii)}$$

As $2 \times 2 > 3 \times 1$

∴ $2X + 3Y = 42$ is the regression of X on Y and
$X + 2Y = 26$ is the regression of Y on X.

Example 34. Given the regression equations:
$$3X + 4Y = 44$$
$$5X + 8Y = 80$$
Variance of $X = 30$
Find $\overline{X}, \overline{Y}, r$ and σ_y

Solution: **Calculation of $\overline{X}$ and $\overline{Y}$**

The regression equations are:
$$3X + 4Y = 44 \quad \text{...(i)}$$
$$5X + 8Y = 80 \quad \text{...(ii)}$$

Multiply (*i*) by 2 and substracting (*ii*) from it
$$6X + 8Y = 88$$
$$5X + 8Y = 80$$
$$- \quad - \quad -$$
$$X = 8 \quad \text{or} \quad \overline{X} = 8$$

Substituting the value of $X = 8$ in (*i*), we get
$$3(8) + 4Y = 44$$
$$24 + 4Y = 44$$
$$4Y = 20$$
$$Y = 5 \quad \text{or} \quad \overline{Y} = 5$$

∴ $\overline{X} = 8, \overline{Y} = 5$

Calculation of Correlation Coefficient

Suppose (*i*) be regression of Y on X
$$3X + 4Y = 44$$
$$4Y = 44 - 3X$$
$$Y = \frac{44}{4} - \frac{3}{4}X \quad \text{...(iii)}$$

∴ byx = Coefficient of X in (*iii*) $= -\frac{3}{4}$

Let equation (*ii*) be regression of X on Y
$$5X + 8Y = 80$$
$$5X = 80 - 8Y$$
$$X = \frac{80}{5} - \frac{8}{5}Y \quad \text{...(iv)}$$

∴ bxy = Coefficient of Y in (*iv*) $= -\frac{8}{5}$

Now $$r^2 = byx \cdot bxy = -\frac{3}{4} \times \frac{-8}{5} = \frac{24}{20} > 1$$

This is impossible as $r^2 \leq 1$. So our assumption is wrong.
We now choose equation (*i*) as regression of X on Y and (*ii*) as the regression of Y on X.

Let equation (*i*) be regression equation of X on Y
$$3X + 4Y = 44$$
$$3X = 44 - 4Y$$
$$X = \frac{44}{3} - \frac{4}{3}Y$$

∴ $$bxy = -\frac{4}{3} \quad \text{...(iii)}$$

Equation (*ii*) be regression equation of Y on X
$$5X + 8Y = 80$$
$$8Y = 80 - 5X$$
$$Y = \frac{80}{8} - \frac{5}{8}X$$

∴ $$byx = -\frac{5}{8} \quad \text{...(iv)}$$

∴ $$r = -\sqrt{bxy \cdot byx} = -\sqrt{\left(\frac{-4}{3}\right) \times \left(\frac{-5}{8}\right)}$$
$$= -\sqrt{\frac{5}{6}} = -\sqrt{0.8333} = -0.912 = -0.91$$

Calculation of σ_y

We know, $bxy = r \cdot \frac{\sigma_x}{\sigma_y}$

Given: $bxy = \frac{-4}{3}, r = -0.91, \sigma_x^2 = 30 \Rightarrow \sigma_x = \sqrt{30} = 5.47$

Substituting the given values in the formula of bxy, we get

$$-1.33 = -0.91 \times \frac{5.47}{\sigma_y}$$

$$\sigma_y = \frac{0.91 \times 5.47}{1.33} = 3.74$$

Example 35. In a partially destroyed laboratory record of an analysis of correlation data, the following results are legible:

Variance of X = 9

Regression equations

$$8X - 10Y + 66 = 0$$
$$40X - 18Y = 214$$

Find (*i*) $\overline{X}$ and $\overline{Y}$ (*ii*) r_{xy} (*iii*) S.D. of Y.

Solution: (*i*) **Calculation of $\overline{X}$ and $\overline{Y}$**

$$8X - 10Y + 66 = 0$$

$$\Rightarrow \quad 8X - 10Y = -66 \quad ...(i)$$

$$40X - 18Y = 214 \quad ...(ii)$$

Multiplying equation (*i*) by 5 and subtracting (*ii*) from it

$$40X - 50Y = -330$$
$$40X - 18Y = 214$$
$$- \quad + \quad -$$
$$-32Y = -544$$

$$\therefore \quad Y = \frac{544}{32} = 17 \text{ or } \overline{Y} = 17$$

By putting the value of Y = 17 in equation (*i*)

$$8X - 10(17) = -66$$
$$8X = -66 + 170$$
$$8X = 104$$
$$\therefore \quad X = 13 \text{ or } \overline{X} = 13$$

$$\therefore \quad \overline{X} = 13, \; \overline{Y} = 17$$

(*ii*) **Calculation of Coefficient of Correlation (r_{xy})**

Let us assume that the first equation be regression equation of Y on X.

$$8X - 10Y + 66 = 0$$

or

$$-10Y = -66 - 8X$$
$$10Y = 66 + 8X$$

$$Y = \frac{66}{10} + \frac{8}{10}X$$

$$Y = 6.6 + 0.8X$$

$$\therefore \quad byx = 0.8$$

Assuming second equation as regression equation of X on Y.

$$40X - 18Y = 214$$
$$40X = 214 + 18Y$$
$$X = \frac{214}{40} + \frac{18}{40}Y$$
$$X = 5.35 + 0.45Y$$
$$\therefore \quad bxy = +0.45$$

Now, $r = \sqrt{bxy \cdot byx}$

$$= \sqrt{0.45 \times 0.8} = +0.6$$

(*iii*) **Calculation of σ_y**

Given, variance of $X = \sigma_x^2 = 9 \Rightarrow \sigma_x = 3$

$$bxy = 0.45, \; r = +0.6$$

We know, $bxy = r \cdot \frac{\sigma_x}{\sigma_y}$

Substituting the values, we get

$$0.45 = 0.6 \times \frac{3}{\sigma_y}$$

$$\Rightarrow \quad \sigma_y = \frac{1.8}{0.45} = 4$$

Hence, $\overline{X} = 13, \overline{Y} = 17, r = +0.6, \sigma_y = 4$.

Example 36. The two lines of regression are given as follows:

$$Y = -4 + \frac{2}{3}X \quad X = -5 + \frac{5}{3}Y$$

Find (*i*) Y when $X = 3$ and X when $Y = 3$ (*ii*) r_{xy}

Solution: Given,

$$Y = -4 + \frac{2}{3}X \quad ...(i)$$

$$X = -5 + \frac{5}{3}Y \quad ...(ii)$$

Let equation (*i*) as regression of Y on X.

$$Y = -4 + \frac{2}{3}X$$

$\therefore \quad byx = \frac{2}{3}$

Taking equation (*ii*) as regression of X on Y

$$X = -5 + \frac{5}{3}Y$$

$\therefore \quad bxy = \frac{5}{3}$

$\therefore \; r^2_{xy} = \frac{2}{3} \times \frac{5}{3} = \frac{10}{9} = 1.1 > 1.$

Here, $r^2 > 1$ which is impossible as $r^2 \leq 1$. So our assumption is wrong.

Reversing the assumption and taking equation (*i*) as regression of X on Y

$$Y = -4 + \frac{2}{3}X \;\Rightarrow\; 3Y = -12 + 2X$$

$$\Rightarrow \quad 2X = 3Y + 12$$

$$\Rightarrow \quad X = \frac{3}{2}Y + 6 \qquad \text{...(iii)}$$

$\therefore \quad bxy = \frac{3}{2}$

Taking equation (*ii*) as regression of Y on X

$$X = -5 + \frac{5}{3}Y \quad or \quad 3X = -15 + 5Y$$

$$\Rightarrow \quad 5Y = 3X + 15$$

$$\Rightarrow \quad Y = \frac{3}{5}X + 3 \qquad \text{...(iv)}$$

$\therefore \quad byx = \frac{3}{5}$

$\therefore \quad r^2 = bxy \,.\, byx = \frac{3}{2} \times \frac{3}{5} = \frac{9}{10} = 0.9$

Since, $r^2 < 1$, our supposition that equation (*i*) is the line of regression of X on Y and equation (*ii*) is the line of regression of Y on X is true.

(*i*) To obtain an estimate of Y when X = 3, we use the line of regression of Y on X viz., (*ii*) or (*iv*).

Thus, from (*iv*), $Y = \frac{3}{5}X + 3$

Put $X = 3$, $Y = \frac{3}{5}(3) + 3 = \frac{9}{5} + 3 = \frac{24}{5} = 4.8$

To obtain an estimate of X when Y = 3, we use the line of regression of X on Y viz. (*i*) or (*iii*).

Thus, from (*iii*), $X = \frac{3}{2}Y + 6$

Put $Y = 3$, $X = \frac{3}{2} \times 3 + 6 = \frac{9}{2} + 6 = \frac{21}{2} = 10.5$

(*ii*) $r_{xy} = \sqrt{bxy \,.\, byx} = \sqrt{\frac{3}{2} \times \frac{3}{5}} = \sqrt{\frac{9}{10}} = \sqrt{0.9} = 0.948$

Example 37. A student obtained the following regression equations. Do you agree with him?

$$6X = 15Y + 21$$
$$21X + 14Y = 56$$

Solution: Here we have two possibilities:

Case I: Treating equation (*i*) as regression equation of X on Y:

$$6X = 15Y + 21 \qquad \text{...(i)}$$

or $\quad X = \frac{15}{6}Y + \frac{21}{6}$

Clearly, $bxy = \frac{15}{6}$

Equation (*ii*) as regression equation of Y on X:

$$21X + 14Y = 56 \qquad \text{...(ii)}$$

or $\quad 14Y = 56 - 21X$

or $\quad Y = \frac{56}{14} - \frac{21}{14}X$

$\therefore \quad byx = -\frac{21}{14}$

Now, $r^2 = bxy.\, byx = \frac{15}{6} \times \frac{-21}{14} < 0.$ Here $r^2 < 0$ which is impossible as $r^2 \geq 0$

Case II: Treating equation (*i*) as regression equation of Y on X:

$$6X = 15Y + 21$$

or $\quad 15Y = 6X - 21$

or $\quad Y = \frac{6X}{15} - \frac{21}{15}$

$\therefore \quad byx = \frac{6}{15}$

Equation (*ii*) as regression equation of X on Y

$$21X + 14Y = 56$$
$$21X = 56 - 14Y$$
$$X = \frac{56}{21} - \frac{14}{21}Y$$

$$bxy = -\frac{14}{21}$$

$\therefore$ Now, $r^2 = byx.bxy = \frac{6}{15} \times \frac{-14}{21} < 0$. Here $r^2 < 0$ which is impossible as $r^2 \geq 0$.

Here, $r^2 < 0$ which is impossible as $r^2 \geq 0$. Hence, calculations done by the students are wrong.

Example 38. If the regression coefficient of X on Y is –1/6 and that of Y on X is –3/2. What is the value of correlation coefficient between X and Y?

Solution: Given, $bxy = \frac{-1}{6},\ byx = \frac{-3}{2}$

$$r = -\sqrt{bxy\,.\,byx}$$

$$= -\sqrt{\left(\frac{-1}{6}\right)\left(\frac{-3}{2}\right)} = -\sqrt{\frac{1}{4}} = -\frac{1}{2} = -0.5$$

Hence, $r = -0.5$

IMPORTANT TYPICAL EXAMPLE

Example 39. The regression equation of profits (X) on sales (Y) of a certain firm is $6Y - 10X + 210 = 0$. The average sales of the firm were Rs. 88,000 and the variance of profits is $\frac{16}{25}$th of the variance of sales. Find the average profits and the coefficient of correlation between sales and profits.

Solution: Regression equation of profits (X) on sales (Y) is:

$$6Y - 10X + 210 = 0$$

The average profits can be obtained by putting Y = 88,000 in the regression equation of X on Y as follows:

$$6 \times 88{,}000 - 10X + 210 = 0 \rightarrow 10X = 5{,}28{,}210$$

$\therefore$ $X = 52{,}821$ or $\overline{X} = 52{,}821$

Also we are given:

Variance of profits $(\sigma_x^2) = \frac{16}{25}$ variance of sales (σ_y^2)

$$\Rightarrow \quad \frac{\sigma_x}{\sigma_y} = \frac{4}{5}$$

$6Y - 10X + 210 = 0$ is the regression of X on Y

$$-10X = -210 - 6Y$$

or

$$10X = 210 + 6Y$$

$$X = \frac{210}{10} + \frac{6}{10}Y$$

$$X = 21 + \frac{3}{5}Y.$$

$\therefore$ bxy = regression coefficient of X on $Y = \frac{3}{5}$

Since $bxy = r\,.\,\frac{\sigma_x}{\sigma_y}$

Putting the values, we get

$$\frac{3}{5} = r \times \frac{4}{5}$$

Thus, $r = \frac{3}{4} = 0.75$

Hence, $\overline{X}$=52, 821, $r = 0.75$

EXERCISE 2.7

1. From the following regression equations:

   $20X - 9Y = 107$

   $4X - 5Y = -33$

   Calculate $\overline{X}, \overline{Y}$ and r. **[Ans.** $\overline{X} = 13, \overline{Y} = 17, r = 0.6$**]**

2. Regression equations of two variables X and Y are as follows:

   $2Y - X - 50 = 0$

   $3Y - 2X - 10 = 0$

   (*i*) Identify which of the two can be called regression of Y on X and X on Y.

   (*ii*) Find the means as well as coefficient of correlation between X and Y.

   [Ans. (*i*) (1) Y on X and (2) X on Y; (*ii*) $\overline{X} = 130, \overline{Y} = 90, r = 0.866$**]**

3. The two regression lines are given by:

   $$Y = \frac{40}{18}X - \frac{214}{18} \text{ and } X = \frac{10}{8}Y - \frac{66}{8}$$

   Find (*i*) Correlation coefficient between X and Y

   (*ii*) Y, when $X = 10$

   (*iii*) X, when $Y = 10$

   (*iv*) σ_y, if $\sigma_x^2 = 9$ **[Ans.** (*i*) $r = 0.6$, (*ii*) 14.6, (*iii*) 9.85, (*iv*) $\sigma_y = 4$**]**

4. The lines of regression of a bivariate population are:

   $12X - 15Y + 99 = 0$

   $64X - 27Y = 373$

   The variance of X is 9.

   Find: (*i*) the mean value of X and Y

   (*ii*) Correlation coefficient between X and Y, and

   (*iii*) the Standard deviation of Y.

   [Ans. (*i*) $\overline{X} = 13, \overline{Y} = 17$, (*ii*) $r = +0.58$, (*iii*) $\sigma_y = 4.138$**]**

5. For certain data, the following regression equations were obtained :

$$4X - 5Y + 33 = 0$$
$$20X - 9Y - 107 = 0$$

Estimate Y when X = 20 and X when Y = 20.

[Ans. $Y_{20} = 22.6, X_{20} = 14.35$]

[**Hint:** See Example 64]

6. Given the regression lines as: $3X + 2Y = 26$ and $6X + Y = 31$, find their point of intersection and interpret it. Also find the correlation coefficient between X and Y.

[Ans. Point of intersection of the lines of regression gives the mean values : ($\overline{X} = 4$ and $\overline{Y} = 7$); $r_{xy} = -0.25$]

7. The two regression lines obtained from certain data were : $Y = X + 5$ and $16X = 9Y - 94$. Find the variance of X, if the variance of Y is 16. Also find the covariance between X and Y..

[Ans. $\sigma_x = 3$; Cov $(X, Y) = r\sigma_x\sigma_y = 9$]

8. The line of regression of marks in Statistics (X) on marks in Economics (Y) for a class of 50 boys is $3Y - 5X + 180 = 0$. Average marks in Economics = 44 and variance of marks in Statistics is $\frac{9}{16}$th of variance of marks in Economics. Find (*i*) average marks in statistics (*ii*) Coefficient of correlation between X and Y.

[Ans. (*i*) $\overline{X} = 62.4$, (*ii*) $r = 0.8$]

9. Equations of two regression lines in a regression analysis are as follows:

$3X + 2Y = 26$ and $6X + Y = 31$

A student obtained the mean values $\overline{X} = 7, \overline{Y} = 4$ and the value of the correlation coefficient $r = +05$. Do you agree with him? If not, suggest your results.

[Ans. (*i*) $\overline{X} = 4, \overline{Y} = 7$, (*ii*) $r = -0.5$]

10. For a set of 10 pairs of values of X and Y, the regression line of X on Y is $X - 2Y + 12 = 0$, mean and standard deviation of Y being 8 and 2 respectively. Later it is known that a pair $(X = 3, Y = 8)$ was wrongly recorded and the correct pair detected is $(X = 8, Y = 3)$. Find the correct regression line of X on Y.

[**Hint:** See Example 62] **[Ans.** $X = Y - 3$]

11. A student obtained the two regression equations as

$2X - 5Y - 7 = 0$ and $3X + 2Y - 8 = 0$

Do you agree with him?

[Ans. (*i*) $byx = \frac{2}{5}, byx = -\frac{2}{3}$ (*ii*) $byx = \frac{5}{2}, bxy = -\frac{2}{3}$

$\therefore$ Equations obtained are wrong]

12. The two lines of regression are given as follows:

$5X - 6Y + 90 = 0$, $15X - 8Y - 130 = 0$

(*i*) Find $\overline{X}$ and $\overline{Y}$ (*ii*) Find r_{xy} (*iii*) Estimate Y when $X = 10$ (*iv*) Estimate X when $Y = 20$.

[Ans. $\overline{X} = 30$, $\overline{Y} = 40$, $r = 0.67$, $Y = 23.33$, $X = 19.33$]

STANDARD ERROR OF ESTIMATE

In regression, given the value of independent variable, we estimate the value of dependent variable by using/applying regression equation. To find out an estimate that is 100% accurate is almost impossible. If we want to make sure that to what extent the estimates made by us are accurate or reliable, then this can be done with the help of standard error of estimate. By using standard error of estimate, we can check the reliability of our estimates. Standard error of estimate shows that to what extent the estimated values by regression line are closer to actual values.

For two regression lines (Regression of X on Y and Regression of Y on X), there are two standard error of estimates:

(*i*) **Standard Error of Estimate of Y on X (S_{yx})**

(*ii*) **Standard Error of Estimate of X on Y (S_{xy})**

(*i*) **Standard Error of Estimate of Y on X:** It is denoted by S_{yx}. Its computation is made by the following formulae:

First formula:

$$S_{yx} = \sqrt{\frac{\Sigma(Y - Y_c)^2}{N}}$$

Here, Y = Actual values, Y_c = Estimated values.

Second formula:

$$S_{yx} = \sqrt{\frac{\Sigma Y^2 - a\Sigma Y - b\Sigma XY}{N}}$$

Where, a and b are to be obtained from normal equations and a = intercept, b = slope of line.

Third formula:

$$S_{yx} = \sigma_y\sqrt{1 - r^2}$$

Where, σ_y = SD of Y; r = coefficient of correlation between X and Y.

The third formula is suitable for use when we are given the values of correlation coefficient (r) and standard deviations (σ_x and σ_y).

(*ii*) **Standard Error of Estimate of X on Y:** It is denoted by S_{xy}. Its computation is done by the following formulae:

First formula:

$$S_{xy} = \sqrt{\frac{\Sigma(X - X_c)^2}{N}}$$

Here, X = Actual values, X_c = Estimated values

Second formula:

$$S_{xy} = \sqrt{\frac{\Sigma X^2 - a\Sigma X - b\Sigma XY}{N}}$$

Where, a and b are to be obtained from normal equations and a = intercept, b = slope of line.

Third formula:

$$S_{xy} = \sigma_x \sqrt{1-r^2}$$

Where, σ_x = SD of Y; r = coefficient of correlation between X and Y.

The third formula is suitable for use when we are given the values of correlation coefficient (r) and standard deviations (σ_x and σ_y).

The following examples make the computation of standard error of estimate more clear:

Example 40. Find the 'standard error of the estimates'.

$$\sigma_x = 4.4, \sigma_y = 2.2, r = 0.8$$

Solution: Given, $r = 0.8, \sigma_x = 4.4, \sigma_y = 2.2$

As 'r', σ_x and σ_y are known, the following formulae are used to find the 'standard error of estimates':

$$S_{yx} = \sigma_y \sqrt{1-r^2} \qquad \text{...}(i)$$

$$S_{xy} = \sigma_x \sqrt{1-r^2} \qquad \text{...}(ii)$$

Putting the given values in (*i*) and (*ii*), we get

$$S_{yx} = 2.2 \times \sqrt{1-(0.8)^2} = 2.2 \times \sqrt{1-0.64}$$

$$= 2.2 \times \sqrt{0.36} = 2.2 \times 0.6 = 1.32$$

$$S_{xy} = 4.4 \times \sqrt{1-(0.8)^2} = 4.4 \times \sqrt{1-0.64}$$

$$= 4.4 \times \sqrt{0.36} = 4.4 \times 0.6 = 2.64$$

Example 41. For a set of 10 pairs of reading on X and Y, the coefficient of correlation is 0.856 and the standard deviation of Y is 5.54. Find the standard error of estimate of Y on X.

Solution: We are given:

$r = 0.856,$ $\quad \sigma_y = 5.54$

The standard error of estimate (S_{yx}) is given by

$$S_{yx} = \sigma_y \sqrt{1-r^2}$$

$$= 5.54\sqrt{1-(0.856)^2} = 5.54\sqrt{1-0.7327}$$

$$= 5.54\sqrt{0.2673} = 5.54 \times 0.5170 = 2.864$$

Example 42. From the data given below:

X:	6	2	10	4	8
Y:	9	11	5	8	7

Compute two regression equations and calculate the standard error of the estimates (S_{yx} and S_{xy}).

Solution:

Calculation of Regression Equations

X	$\bar{X}=6$ $(X-\bar{X})$ x	x^2	Y	$\bar{Y}=8$ $(Y-\bar{Y})$ y	y^2	xy
6	0	0	9	1	1	0
2	–4	16	11	3	9	–12
10	4	16	5	–3	9	–12
4	–2	4	8	0	0	0
8	2	4	7	–1	1	–2
$N=5$ $\Sigma X = 30$	$\Sigma x = 0$	$\Sigma x^2 = 40$	$\Sigma Y = 40$	$\Sigma y = 0$	$\Sigma y^2 = 20$	$\Sigma xy = -26$

We have, $\bar{X} = \frac{30}{5} = 6 \quad \bar{Y} = \frac{40}{5} = 8$

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{-26}{20} = -1.3$$

$$byx = \frac{\Sigma xy}{\Sigma x^2} = \frac{-26}{40} = -0.65$$

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$

$$Y - 8 = -0.65(X - 6)$$

$$Y - 8 = -0.65X + 3.9$$

$$Y = -0.65X + 11.9$$

$$Y = 11.9 - 0.65X$$

Regression Equation of X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$

$$X - 6 = -1.3(Y - 8)$$

$$X - 6 = -1.3Y + 10.4$$

$$X = -1.3Y + 10.4 + 6$$

$$X = -1.3Y + 16.4$$

$$X = 16.4 - 1.3Y$$

Thus, the two regression equations are:

$$Y = 11.9 - 0.65X$$

$$X = 16.4 - 1.3Y$$

Calculation of Standard Error of Estimates

From the regression equation of Y on X ($Y_c = 11.9 - 0.65X$) for various values of X, we can find out the corresponding value of Y_c values and from the equation of X on Y ($X_c = 16.4 - 1.3Y$), we can find X_c. These values are:

Computation of Standard Error of Estimate

X	Y	Y_c	X_c	$(Y-Y_c)^2$	$(X-X_c)^2$
6	9	8.0	4.7	1.00	1.69
2	11	10.6	2.1	0.16	0.01
10	5	5.4	9.9	0.16	0.01
4	8	9.3	6.0	1.69	4.00
8	7	6.7	7.3	0.09	0.49
				$\Sigma(Y-Y_c)^2 = 3.10$	$\Sigma(X-X_c)^2 = 6.20$

Standard Error of Estimates:

$$S_{yx} = \sqrt{\frac{\Sigma(Y-Y_c)^2}{N}} = \sqrt{\frac{3.10}{5}} = +0.7874$$

$$S_{xy} = \sqrt{\frac{\Sigma(X-X_c)^2}{N}} = \sqrt{\frac{6.20}{5}} = +1.11$$

Aliter: S_{yx} and S_{xy} can also be calculated as:

$$\sigma_y = \sqrt{\frac{\Sigma(Y-\bar{Y})^2}{N}} = \sqrt{\frac{\Sigma y^2}{N}} = \sqrt{\frac{20}{5}} = \sqrt{4} = 2$$

$$\sigma_x = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}} = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{40}{5}} = \sqrt{8} = 2.828$$

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}} = \frac{-26}{\sqrt{20\times 40}} = \frac{-26}{\sqrt{800}} = \frac{-26}{28.28} = -0.919$$

$$S_{yx} = \sigma_y\sqrt{1-r^2} = 2\sqrt{1-(-0.919)^2} = 2\times\sqrt{0.155} = 0.7874$$

$$S_{xy} = \sigma_x\sqrt{1-r^2} = 2.828\sqrt{1-(0.919)^2} = 2.828\times\sqrt{0.155} = 1.11$$

Example 43. Given that

$\Sigma X = 15,\ \Sigma Y = 110,\ \Sigma XY = 400,\ \Sigma X^2 = 250,\ \Sigma Y^2 = 3200,\ N = 10$

(*i*) Compute the regression equation of Y on X

(*ii*) Standard Error of Estimate S_{yx}.

Solution: (*i*) byx or $b = \dfrac{N\,\Sigma XY - (\Sigma X)(\Sigma Y)}{N.\Sigma X^2 - (\Sigma X)^2}$

$$= \frac{10(400)-(15)(110)}{10(250)-(15)^2} = 1.033$$

$$\bar{X} = \frac{15}{10} = 1.5,\ \bar{Y} = \frac{110}{10} = 11$$

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$

$$Y - 11 = 1.033(x - 1.5)$$

$$Y - 11 = 1.033X - 1.5495$$

$$Y = 9.4505 + 1.033X$$

(*ii*) Standard error of estimate (S_{yx}) is given by

$$S_{yx} = \sqrt{\frac{\Sigma Y^2 - a\,\Sigma Y - b\,\Sigma XY}{N}}$$

$$= \sqrt{\frac{3200-(9.45)(110)-(1.033)(400)}{10}}$$

$$= \sqrt{\frac{3200-1039.5-413.2}{10}} = \sqrt{\frac{1747.3}{10}}$$

$$= 13.21$$

Example 44. From the following data, find the standard error of the estimate of X on Y and Y on X:

X:	1	2	3	4	5
Y:	6	8	7	6	8

Solution:

X	$(X-\bar{X})$ x	x^2	Y	$(Y-\bar{Y})$ y	y^2	xy
1	–2	4	6	–1	1	2
2	–1	1	8	+1	1	–1
3	0	0	7	0	0	0
4	1	1	6	–1	1	–1
5	2	4	8	+1	1	2
$\Sigma X = 15$		$\Sigma x^2 = 10$	$\Sigma Y = 35$		$\Sigma y^2 = 4$	$\Sigma xy = 2$

$$\bar{X} = \frac{15}{5} = 3 \qquad \bar{Y} = \frac{35}{5} = 7$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{N}} = \sqrt{\frac{10}{5}} = 1.414$$

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{N}} = \sqrt{\frac{4}{5}} = 0.894$$

$$r = \frac{\Sigma xy}{N.\sigma_x.\sigma_y}$$

$$r = \frac{2}{5\times 1.414\times 0.894} = \frac{2}{6.321} = 0.316$$

$$S_{xy} = \sigma_x \sqrt{1-r^2}$$
$$= 1.414\sqrt{1-(0.316)^2} = 1.414 \times 0.949 = 1.342$$
$$S_{yx} = \sigma_y \sqrt{1-r^2}$$
$$= 0.894\sqrt{1-(0.316)^2} = 0.894 \times 0.949 = 0.848$$

Example 45. For 10 observations on price (X) and supply (Y) the following data were obtained:

$$\Sigma X = 130,\ \Sigma Y = 220,\ \Sigma X^2 = 2288$$
$$\Sigma Y^2 = 5506,\ \Sigma XY = 3467,\ N = 10$$

Obtain the standard error of estimate of X on Y and Y on X.

Solution: Given, $N = 10,\ \Sigma X = 130,\ \Sigma Y = 220,\ \Sigma X^2 = 2288$

$$\Sigma Y^2 = 5506,\ \Sigma XY = 3467$$

$$\sigma_x = \sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2} \quad \text{(Formula of S.D.)}$$
$$= \sqrt{\frac{2288}{10} - \left(\frac{130}{10}\right)^2} = \sqrt{228.8 - 169} = \sqrt{59.8} = 7.73$$
$$\sigma_y = \sqrt{\frac{\Sigma Y^2}{N} - \left(\frac{\Sigma Y}{N}\right)^2}$$
$$= \sqrt{\frac{5506}{10} - \left(\frac{220}{10}\right)^2} = \sqrt{550.6 - 484} = \sqrt{66.6} = 8.16$$
$$r = \frac{N.\Sigma XY - \Sigma X.\Sigma Y}{\sqrt{N.\Sigma X^2 - (\Sigma X)^2}\sqrt{N.\Sigma Y^2 - (\Sigma Y)^2}}$$
$$= \frac{10 \times 3467 - (130)(220)}{\sqrt{10 \times 2288 - (130)^2}\sqrt{10 \times 5506 - (220)^2}}$$
$$= \frac{34670 - 28600}{\sqrt{22880 - 16900}\sqrt{55060 - 48400}}$$
$$= \frac{6070}{\sqrt{5980}\sqrt{6660}} = \frac{6070}{6310.84} = 0.961$$

Standard Error of Y on X

$$S_{yx} = \sigma_y \sqrt{1-r^2} = 8.16\sqrt{1-(0.961)^2}$$
$$= 8.16\sqrt{0.07647} = 2.256$$

Standard Error of X on Y

$$S_{xy} = \sigma_x \sqrt{1-r^2}$$
$$= 7.73\sqrt{1-(0.961)^2}$$
$$= 7.73\sqrt{0.07647}$$
$$= 2.137$$

EXERCISE 2.8

1. Find the standard error of estimates:
$\sigma_x = 1.414,\ \sigma_y = 0.894,\ r = 0.316$

[Ans. S_{yx}=.848, S_{xy}=1.342]

2. For a set of 8 pair reading on X and Y, the coefficient of correlation is 0.65 and the standard deviation of Y series is 4.2. Find the standard error of Y on X.

[Ans. S_{yx}=3.1917]

3. From the following data, compute standard error of estimate of the regression of Y on X:
$\Sigma x^2 = 10,\ \Sigma y^2 = 4, \Sigma xy = 2,$ N= 5, where $x = X - \overline{X},\ \ y = Y - \overline{Y}$

[Ans. S_{yx}=0.848]

4. Given the following data:

X:	1	2	3	4	5
Y:	2	4	5	3	6

Obtain the two regression equations and calculate the standard error of estimates.

[Ans. X = 0.74Y + 0.2, Y= 0.7 +1.9, S_{xy} = 1.01, S_{yx} = 1.01]

5. Family income and its percentage spent on food in the case of hundred families gave the following bivariate frequency distribution.

Food Expenditure (in %)	Family Income (Rs.)				
	200—300	300—400	400-500	500—600	600—700
10—15	—	—	—	3	7
15—20	—	4	9	4	3
20—25	7	6	12	5	—
25—30	3	10	19	8	—

Obtain the equations of the two lines of regression. Also compute the standard error of the estimates.

[Hint : See Example 55]

[Ans. $Y = -0.02X + 31.5$; $X = -9.6Y + 666$; $S_{yx} = 4.494$, $S_{xy} = 98.47$]

Explained and Unexplained Variation

The total variation in the dependent variable Y can be split into two:

(a) Explained Variation: The variation in Y which is explained by the variation in X is called explained variation in Y.

(b) Unexplained Variation: The variation in Y which is unexplained by the variation in variable X and is due to some other factors (variables) is called unexplained variation in Y.

Symbolically,

Total variation in Y = Explained variation in Y + Unexplained variation in Y

$$\Sigma(Y-\bar{Y})^2 = \Sigma(Y_c-\bar{Y})^2 + \Sigma(Y-Y_c)^2$$

Where, Y_c = computed (or estimated) value of Y on the basis of regression equation

$\bar{Y}$ = Mean value of Y series

Y = Original value of Y series.

A similar relationship we may have for X variable (Dependent) in terms of Y:

$$\Sigma(X-\bar{X})^2 = \Sigma(X_c-\bar{X})^2 + \Sigma(X-X_c)^2$$

Coefficient of Determination: Based on the above expression, the coefficient of determination (r^2) is defined as the ratio of the explained variation to total variation, *i.e.*,

$$r^2 = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{\Sigma(Y_c-\bar{Y})^2}{\Sigma(Y-\bar{Y})^2}$$

It is clear that the object of coefficient of determination is to determine the percentage variation in Y which is explained by variation in X. For example, let us suppose that the correlation coefficient between X and Y is +0.8, then coefficient of determination $(r^2) = (0.8)^2 = 0.64$. It means that 64% variations in Y are due to variation in X and 36% variations are due to other factors. Thus, explained variations are 64% and unexplained variations are 36%.

Coefficient of Non-Determination: The proportion of unexplained variation to total variation is termed as coefficient of non-determination. It is denoted by k^2, where $k^2 = 1 - r^2$. It is also written as:

$$k^2 = \frac{\text{Unxplained variation}}{\text{Total variation}} = 1 - r^2$$

The square root of k^2 is termed as coefficient of alientation, *i.e.*, $k = \sqrt{k^2} = \sqrt{1-r^2}$

Standard Error of Estimate: Standard error of estimates of Y on X and X on Y can also be calculated as:

$$S_{yx} = \sqrt{\frac{\Sigma(Y-Y_c)^2}{N}} = \sqrt{\frac{\text{Unexplained variation in } Y}{N}}$$

$$S_{xy} = \sqrt{\frac{\Sigma(X-X_c)^2}{N}} = \sqrt{\frac{\text{Unexplained variation in } X}{N}}$$

Example 46. Given: Explained variation = 19.22

Unexplained variation = 19.70

Determine the coefficient of correlation.

Solution: Total variation = Explained variation + Unexplained variation

= 19.22 + 19.70 = 38.92

$$\text{Coefficient of Determination } (r^2) = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{19.22}{38.92} = 0.4938$$

$\Rightarrow$ Coefficient of Correlation $(r) = \sqrt{0.4938} = 0.70$

Example 47. In fitting of a regression of Y on X to a bivariate distribution consisting of 9 observations, the explained and unexplained variations were computed as 24 and 36 respectively. Find: (*i*) coefficient of determination, and (*ii*) standard error of estimate of Y on X.

Solution: Total variation in Y = Explained variation in Y + Unexplained variation in Y

= 24 + 36 = 60

$$\text{Coefficient of Determination } (r^2) = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{24}{60} = \frac{4}{10} = 0.40$$

$$\text{Standard Error of Estimate of Y on X } (S_{yx}) = \sqrt{\frac{\text{Unexplained variation}}{N}}$$

$$= \sqrt{\frac{36}{9}} = \sqrt{4} = 2$$

Example 48. Given the following data:

X:	1	2	3	4	5
Y:	10	20	30	50	40

Calculate:

(*i*) Regression equation of Y on X.

(*ii*) Total variation in Y.

(*iii*) Unexplained variation in Y.

(*iv*) Explained variation in Y.

(*v*) Standard error of estimate.

(*vi*) Coefficient of determination.

Solution:

X	$x = X - \bar{X}$	x^2	Y	$y = Y - \bar{Y}$	y^2	xy
1	−2	4	10	−20	400	40
2	−1	1	20	−10	100	10
3	0	0	30	0	0	0
4	1	1	50	20	400	20
5	2	4	40	10	100	20
$\Sigma X = 15$ $N = 5$	$\Sigma x = 0$	$\Sigma x^2 = 10$	$\Sigma Y = 150$	$\Sigma y = 0$	$\Sigma y^2 = 1000$	$\Sigma xy = 90$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{15}{5} = 3 \qquad \bar{Y} = \frac{\Sigma Y}{N} = \frac{150}{5} = 30$$

$$byx = \frac{\Sigma xy}{\Sigma x^2} = \frac{90}{10} = 9$$

(*i*) **Regression Equation of Y on X**

$$Y - \bar{Y} = byx\,(X - \bar{X})$$
$$Y - 30 = 9\,(X - 3)$$
$$Y - 30 = 9X - 27$$
$$Y = 9X + 3$$

(*ii*) Total variation in $Y = \Sigma(Y - \bar{Y})^2 = \Sigma y^2 = 1000$

(*iii*) Unexplained variation in Y:

X	Y	$Y_c = 9X + 3$	$Y - Y_c$	$(Y - Y_c)^2$
1	10	12	−2	4
2	20	21	−1	1
3	30	30	0	0
4	50	39	11	121
5	40	48	−8	64
				$\Sigma(Y - Y_c)^2 = 190$

Unexplained variation $= \Sigma(Y - Y_c)^2 = 190$

(*iv*) Explained variation in Y = Total variation – Unexplained variation

$= 1000 - 190 = 810$

(*v*) Standard error of estimate $(S_{yx}) = \sqrt{\frac{\Sigma(Y - Y_c)^2}{N}} = \sqrt{\frac{190}{5}} = 6.164$

(*vi*) Coefficient of determination $(r^2) = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{810}{1000} = 0.81$

Example 49. From the following data:

Age of husband (years):	18	19	20	21	22	23	24	25	26	27
Age of wife (years):	17	17	18	18	18	19	19	20	21	21

Obtain the following:

(*i*) The regression of age of husband on the age of wife.

(*ii*) Total variation in the age of husband.

(*iii*) The magnitude of variation in age of the husband, explained by the regression equation.

(*iv*) Standard error of the estimate of age of husband.

Solution: Let Y denote age of husband and X denote age of wife.

Y	A = 23 dy	dy^2	X	A = 19 dx	dx^2	$dxdy$
18	−5	25	17	−2	4	10
19	−4	16	17	−2	4	8
20	−3	9	18	−1	1	3
21	−2	4	18	−1	1	2
22	−1	1	18	−1	1	1
23 = A	0	0	19 = A	0	0	0
24	+1	1	19	0	0	0
25	+2	4	20	+1	1	2
26	+3	9	21	+2	4	6
27	+4	16	21	+2	4	8
$\Sigma Y = 225$ N = 10 $\therefore \bar{Y} = 22.5$	$\Sigma dy = -5$	$\Sigma dy^2 = 85$	$\Sigma X = 188$ N = 10 $\therefore \bar{X} = 18.8$	$\Sigma dx = -2$	$\Sigma dx^2 = 20$	$\Sigma dxdy = 40$

(*i*) $byx = \frac{N.\Sigma dxdy - \Sigma dx.\Sigma dy}{N.\Sigma dx^2 - (\Sigma dx)^2}$

$= \frac{10 \times 40 - (-2)(-5)}{10 \times 20 - (-2)^2} = \frac{400 - 10}{200 - 4} = \frac{390}{196} = 1.989 \cong 1.99$

$\sigma_y^2 = \frac{\Sigma dy^2}{N} - \left(\frac{\Sigma dy}{N}\right)^2$

$= \frac{85}{10} - \left(\frac{-5}{10}\right)^2$

$= 8.5 - 0.25 = 8.25$

Regression Equation of Y on X

$$Y - \bar{Y} = byx(X - \bar{X})$$
$$Y - 22.5 = 1.99(X - 18.8)$$
$$Y - 22.5 = 1.99X - 37.412$$
$$\therefore \quad Y = 1.99X - 14.912$$

(*ii*) Total variation in $Y = \Sigma(Y - \bar{Y})^2 = N.\sigma_y^2$

$= 10 \times 8.25 = 82.5$

(*iii*) **Explained variation in Y :**

X	Y	$Y_c = (1.99X - 14.912)$	$Y_c - \bar{Y}$	$(Y_c - \bar{Y})^2$
17	18	18.9	−3.6	12.96
17	19	18.9	−3.6	12.96
18	20	20.9	−1.6	2.56
18	21	20.9	−1.6	2.56
18	22	20.9	−1.6	2.56
19	23	22.9	0.4	0.16
19	24	22.9	0.4	0.16
20	25	24.9	2.4	5.76
21	26	26.9	4.4	19.36
21	27	26.9	4.4	19.36
				$\Sigma(Y_c - \bar{Y})^2 = 78.40$

Explained variation in Y = 78.40

∴ Magnitude of variation in age of husband (Y) explained by the regression equation = 78.40

Unexplained variation in Y = Total variation – Explained variation
= 82.5 – 78.40 = 4.1

(*iv*) **Standard Error of Estimate (S_{yx})**

$$S_{yx} = \sqrt{\frac{\Sigma(Y - Y_c)^2}{N}} = \sqrt{\frac{\text{Unexplained Variation}}{N}} = \sqrt{\frac{4.1}{10}} = 0.64$$

EXERCISE 2.9

1. The coefficient of correlation (r) between two variables X and Y is + 0.95. What percent variation in Y (dependent variable) remains unexplained by the variation in X (the independent variable).
[**Hint:** $r^2 = 0.9025$, Explained variation = 90.25%]
[Ans. Unexplained variation = 9.75%]
2. If the explained variation is 15.24 and the unexplained variation is 27.09, find the coefficient of determination.
[**Hint:** r^2 = Explained Variation ÷ Total Variation] **[Ans. $r^2 = 0.36$]**
3. Given the bivariate data:

X:	1	5	3	2	1	1	7	3
Y:	6	1	0	0	1	2	1	5

Obtain: (*i*) Regression equation of Y on X
(*ii*) Total variation in Y, and
(*iii*) Explained variation in Y
(*iv*) Standard error of estimate of Y on X.
[Ans. (*i*) Y = 2.874 – 0.304X (*ii*) 38 (*iii*) 3.042 (iv) 2.0299]

4. The coefficient of correlation (r) between consumption expenditure (C) and disposable income (Y) in a study was found to be +0.8. What percentage of variation in C are explained by variations in Y? **[Ans. 64%]**
5. From the following data, find out (i) correlation coefficient (ii) linear regression line of Y on X. Also find the percentage of variation explained by the regression line.

X:	1	2	3	4	5
Y:	2	5	3	8	7

[**Hint:** See Example 59]
[Ans. $r = 0.806$, $Y = 1.3X + 1.1$,
$r^2 = 0.6496 \Rightarrow$ 64.96% variations are explained by the regression line.]
6. Given the following data:
$S_{yx} = 35$, $\Sigma Y^2 = 800$, $\bar{Y} = 5$, $N = 20$
Find (*i*) total variation in Y which are unexplained and explained by X (*ii*) coefficient of determination and (*iii*) coefficient of non-determination.
[Ans. (*i*) 300, 245, 55, (*ii*) 0.1833, (*iii*) 0.8167]

MISCELLANEOUS SOLVED EXAMPLES

Example 50. From the following data, obtain the two regression equations:

X:	1	2	3	4	5	6	7	8	9
Y:	9	8	10	12	11	13	14	16	15

Verify that the coefficient of correlation is the geometric mean of the two regression coefficients.

Solution: (*i*) **Calculation of Regression Equations**

X	$\bar{X}=5$ x	x^2	Y	$\bar{Y}=12$ y	y^2	xy
1	−4	16	9	−3	9	+12
2	−3	9	8	−4	16	+12
3	−2	4	10	−2	4	+4
4	−1	1	12	0	0	0
5	0	0	11	−1	1	0
6	+1	1	13	+1	1	+1
7	+2	4	14	+2	4	+4
8	+3	9	16	+4	16	+12
9	+4	16	15	+3	9	+12
$\Sigma X = 45$	$\Sigma x = 0$	$\Sigma x^2 = 60$	$\Sigma Y = 108$	$\Sigma y = 0$	$\Sigma y^2 = 60$	$\Sigma xy = 57$

$$\overline{X} = \frac{\Sigma X}{N} = \frac{45}{9} = 5 \qquad \overline{Y} = \frac{\Sigma Y}{N} = \frac{108}{9} = 12$$

Since the actual means of X and Y are whole numbers, therefore we should take the deviations from $\overline{X}$ and $\overline{Y}$.

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{57}{60} = +0.95$$

$$byx = \frac{\Sigma xy}{\Sigma x^2} = \frac{57}{60} = +0.95$$

Regression Equation of Y on X

$$Y - \overline{Y} = byx\,(X - \overline{X})$$
$$Y - 12 = +0.95(X - 5)$$
$$Y - 12 = 0.95X - 4.75$$
$$Y = 0.95X + 7.25$$

Regression Equation of X on Y

$$X - \overline{X} = bxy\,(Y - \overline{Y})$$
$$X - 5 = +0.95(Y - 12)$$
$$X - 5 = 0.95Y - 11.4$$
$$X = 0.95Y - 6.4$$

Calculation of Coefficient of Correlation

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$$
$$= \frac{57}{\sqrt{60}\sqrt{60}} = +0.95$$

Verification:

$$r = \sqrt{bxy \cdot byx}$$
$$= \sqrt{0.95 \times 0.95}$$
$$= 0.95$$

Hence the result is verified.

Example 51. The following table gives the ages and blood pressure of 10 women:

Age:	56	42	36	47	49	42	60	72	63	55
Blood Pressure:	147	125	118	128	145	140	155	160	149	150

Estimate the blood pressure of a woman whose age is 45 years.

Solution: Let age be denoted by X and blood pressure be denoted by Y.

X	A = 49 dx	dx^2	Y	A = 140 dy	dy^2	dxdy
56	+7	49	147	+7	49	+49
42	–7	49	125	–15	225	+105
36	–13	169	118	–22	484	+286
47	–2	4	128	–12	144	+24
49 = A	0	0	145	+5	25	0
42	–7	49	140 = A	0	0	0
60	+11	121	155	+15	225	+165
72	+23	529	160	+20	400	+460
63	+14	196	149	+9	81	+126
55	+6	36	150	+10	100	60
$\Sigma X = 522$	$\Sigma dx = 32$	$\Sigma dx^2 = 1202$	$\Sigma Y = 1417$	$\Sigma dy = 17$	$\Sigma dy^2 = 1733$	$\Sigma dxdy = 1275$

For estimating the blood pressure of a woman whose age is 45, we fit a regression equation of Y on X:

$$\overline{X} = \frac{\Sigma X}{N} = \frac{522}{10} = 52.2 \qquad \overline{Y} = \frac{\Sigma N}{N} = \frac{1417}{10} = 141.7$$

$$byx = \frac{N.\Sigma dxdy - \Sigma dx.\Sigma dy}{N.\Sigma dx^2 - (\Sigma dx)^2}$$

$$= \frac{10 \times 1275 - (32)(17)}{10 \times 1202 - (32)^2} = \frac{12750 - 544}{12020 - 1024} = \frac{12206}{10996} = 1.11$$

Regression Equation of Y on X

$$Y - \overline{Y} = byx\,(X - \overline{X})$$
$$Y - 141.7 = 1.11(X - 52.2)$$
$$Y - 141.7 = 1.11X - 57.942$$
$$Y = 1.11X + 83.758$$

For $X = 45$, $Y = 1.11\,(45) + 83.758$

$$= 49.95 + 83.758 = 133.708$$
$$= 133.708 \approx 134$$

Thus, the blood pressure of a woman whose age is 45 = 134.

Example 52. On each of 30 items, two measurements on X and Y are made. The following summations are given:

$$\Sigma X = 15,\ \Sigma Y = -6,\ \Sigma XY = 56,\ \Sigma X^2 = 61 \text{ and } \Sigma Y^2 = 90.$$

Calculate the product moment correlation coefficient and the slope of the regression line of Y on X. How would your results be affected if X is replaced by $U = \frac{X-1}{2}$?

Solution: (*i*) **Coefficient of Correlation**

$$r=\frac{30\times 56+15\times 6}{\sqrt{30\times 61-(15)^2}\sqrt{30\times 90-(-6)^2}}=\frac{1770}{\sqrt{1605}\times\sqrt{2664}}=0.856$$

Since r is independent of change of scale and of change of origin, it would remain same even after making the above mentioned transformation.

(*ii*) **Regression Coefficient of Y on X**

$$byx=\frac{1770}{1605}=1.10$$

We know that byx is independent of change of origin but not of change of scale. If byx denote regression coefficient of Y on X and bvu denote regression coefficient of v on u where, $u=\frac{X-A}{h}$ and $v=\frac{Y-B}{k}$, we know that $byx=\frac{k}{h}bvu$

or $bvu=\frac{h}{k}byx$.

In the example, it is given that $h=2$ and $k=1$, *i.e.*, $bvu=2byx$.

Hence, the new regression coefficient will be two times the old regression coefficient of, *i.e.*, regression coefficient of Y on $U(=\frac{X-1}{2})$ will be equal to

2×1.10, *i.e.*, 2.20.

Example 53. For certain data, $3X+2Y-26=0$ and $6X+Y-31=0$ are the two regression equations. Find the values of means and coefficient of correlation.

Solution: **Calculation of $\overline{X}$ and $\overline{Y}$**

$$3X+2Y-26=0 \quad \text{...}(i)$$

$$6X+Y-31=0 \quad \text{...}(ii)$$

Multiplying (*ii*) by 2 and subtracting (*i*) from (*ii*)

$$12X+2Y-62=0$$

$$3X+2Y-26=0$$

$$-\quad -\quad +$$

$$9X-36=0$$

$$9X=36$$

$\therefore\ X=4$ or $\overline{X}=4$

Putting the value of $X=4$ in (*i*)

$$3(4)+2Y-26=0$$

$$12+2Y-26=0$$

$$2Y=14$$

$$Y=7 \text{ or } \overline{Y}=7$$

$\therefore\ \overline{X}=4,\ \overline{Y}=7$

Calculation of Coefficient of Correlation

Let us take equation (*i*) as Y on X and (*ii*) as X on Y

Regression Equation of Y on X

$$3X+2Y-26=0$$

$$2Y=26-3X$$

$$Y=\frac{26}{2}-\frac{3}{2}X$$

$$\therefore\quad byx=-\frac{3}{2} \quad \text{...}(i)$$

Regression Equation of X on Y

$$6X+Y-31=0$$

$$6X=31-Y$$

$$X=\frac{31}{6}-\frac{1}{6}Y$$

$$\therefore\quad bxy=-\frac{1}{6} \quad \text{...}(ii)$$

Now, $r=\sqrt{byx\,.\,bxy}$

$$=-\sqrt{\left(\frac{-3}{2}\right)\left(\frac{-1}{6}\right)}=-\sqrt{\left(\frac{3}{12}\right)}=-\sqrt{\left(\frac{1}{4}\right)}=\frac{-1}{2}=-0.50$$

Example 54. The following results were worked out from the scores in Statistics and Mathematics in a certain examination:

	Score in Statistics (X)	Score in Mathematics (Y)
Mean:	39.5	47.5
Standard Deviation:	10.8	17.8
Coefficient of correlation = +0.42		

Find both the regression equations. Use these regressions to estimate the value of Y for X=50 and also estimate the value of X for Y=30.

Solution: Given, $\overline{X}=39.5$, $\overline{Y}=47.5$, $\sigma_x=10.8$, $\sigma_y=17.8$, $r=+0.42$

(*i*) **Regression Equation of X on Y**

$$X-\overline{X}=r.\frac{\sigma_x}{\sigma_y}(Y-\overline{Y})$$

$$X-39.5=0.42\times\frac{10.8}{17.8}(Y-47.5)$$

$$X-39.5=0.25(Y-47.5)\ \Rightarrow\ X=0.25Y+27.625$$

Regression Equation of Y on X

$$Y-\overline{Y}=r\cdot\frac{\sigma_y}{\sigma_x}(X-\overline{X})$$

$$Y-47.5=+0.42\times\frac{17.8}{10.8}(X-39.5)$$

$$Y-47.5=0.69\,(X-39.5)\quad\Rightarrow\quad Y=0.69X+20.245$$

(*ii*) For X = 50, Y = 0.69 (50) + 20.245 = 54.745

For Y = 30, X = 0.25 (30) + 27.625 = 35.125

Example 55. Family income and its percentage spent on food in case of hundred families gave the following bivariate distribution:

Food Expenditure (in %)	Family Income (Rs)				
	200—300	300—400	400—500	500—600	600—700
10—15	—		—	3	7
15—20	—	4	9	4	3
20—25	7	6	12	5	—
25—30	3	10	19	8	—

Obtain the equations of two lines of regression. Also compute standard error of estimates.

Solution: **(Landscape Table Given at Page 157)**

Let $dx=\dfrac{X-450}{100}$ and $dy=\dfrac{Y-17.5}{5}$

Regression Coefficient of Y and X

$$byx=\frac{N.\Sigma fdxdy-\Sigma fdx\ \Sigma fdy}{N.\Sigma fdx^2-(\Sigma fdx)^2}\times\frac{i_y}{i_x}$$

$$=\frac{100(-48)-0\times100}{100\times120-(0)^2}\times\frac{5}{100}=\frac{-4800}{12000}\times\frac{1}{20}=\frac{-2}{100}=-0.02$$

Regression Coefficient of X on Y

$$bxy=\frac{N.\Sigma fdxdy-\Sigma fdx\ \Sigma fdy}{N.\Sigma fdy^2-(\Sigma fdy)^2}\times\frac{i_x}{i_y}$$

$$=\frac{100(-48)-0\times100}{100\times200-(100)^2}\times\frac{100}{5}=\frac{-4800-0}{20000-10000}\times20$$

$$=\frac{-4800}{10000}\times20=\frac{-48}{5}=-9.6$$

$$\overline{X}=A+\frac{\Sigma fdx}{N}\times i_x=450+\frac{0}{100}\times100=450$$

$$\overline{Y}=A+\frac{\Sigma fdy}{N}\times i_y=17.5+\frac{100}{100}\times5=22.5$$

Solution 55.

Y ↓ M.V.			X → M.V. / dx / dy	200—300	300—400	400—500	500—600	600—700	f	fdy	fdy^2	fdxdy
				250	350	450	550	650				
				−200	− 100	0	100	200				
			dx	− 2	− 1	0	1	2				
10—15	12.5	−5	−1	—	—	—	−1, 3, −3	−2, 7, −14	10	−10	10	− 17
15—20	17.5	0	0	—	0, 4, 0	0, 9, 0	0, 4, 0	0, 3, 0	20	0	0	0
20—25	22.5	+5	+1	−2, 7, −14	−1, 6, −6	0, 12, 0	1, 5, 5	—	30	30	30	− 15
25—30	27.5	10	+2	−4, 3, −12	−2, 10, −20	0, 19, 0	2, 8, 16	—	40	80	160	− 16
			f	10	20	40	20	10	N = 100	100	200	− 48
			fdx	−20	− 20	0	20	20	0			
			fdx^2	40	20	0	20	40	120			
			fdxdy	− 26	− 26	0	18	− 14	− 48			

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$
$$Y - 22.5 = -0.02(X - 450)$$
$$Y = -0.02X + 9 + 22.5$$
$$Y = -0.02X + 31.5$$

Regression Equation of X on Y

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$
$$X - 450 = -9.6\,(Y - 22.5)$$
$$X - 450 = -9.6Y + 216$$
$$X = -9.6Y + 666$$

Standard Error of Estimates

$$\sigma_x = \sqrt{\frac{\Sigma fdx^2}{N} - \left(\frac{\Sigma fdx}{N}\right)^2} \times i_x$$
$$= \sqrt{\frac{120}{100} - \left(\frac{0}{100}\right)^2} \times 100$$
$$= 109.545$$

$$\sigma_y = \sqrt{\frac{\Sigma fdy^2}{N} - \left(\frac{\Sigma fdy}{N}\right)^2} \times i_y$$
$$= \sqrt{\frac{200}{100} - \left(\frac{100}{100}\right)^2} \times 5$$
$$= \sqrt{2-1} \times 5 = 5$$

$$r^2 = byx \, . \, bxy = (-0.02) \times (-9.6) = 0.192$$
$$S_{yx} = \sigma_y \sqrt{1-r^2} = 5 \times \sqrt{1-0.192} = 5 \times \sqrt{0.808} = 4.494$$
$$S_{xy} = \sigma_x \sqrt{1-r^2} = 109.545 \times \sqrt{1-0.192} = 109.545 \times \sqrt{0.808} = 98.47$$

Example 56. Find the means of X and Y variables and the coefficient of correlation between them from the following two regression equations:

$$2Y - X - 50 = 0$$
$$3Y - 2X - 10 = 0$$

Also calculate the standard error of estimate of Y on X, given that the standard deviation of X is 3.

Solution: **(a) Calculation of $\bar{X}$ and $\bar{Y}$**

$$2Y - X - 50 = 0 \quad \text{...(}i\text{)}$$
$$3Y - 2X - 10 = 0 \quad \text{...(}ii\text{)}$$

Multiplying (*i*) by 2 and subtracting (*ii*) from it,

$$4Y - 2X - 100 = 0$$
$$3Y - 2X - 10 = 0$$
$$- \quad + \quad +$$
$$Y - 90 = 0$$

or $Y = 90$ or $\bar{Y} = 90$

Putting the value of Y in (*i*)

$$2(90) - X - 50 = 0$$
$$180 - X - 50 = 0$$
$$\therefore X = 130 \text{ or } \bar{X} = 130$$
$$\therefore \bar{X} = 130, \bar{Y} = 90$$

(b) Calculation of Correlation Coefficient

Let us assume equation (*i*) as Y on X and equation (*ii*) as X on Y.

Regression of Y on X

$$2Y - X - 50 = 0$$
$$2Y = 50 + X$$
or $$Y = 25 + \frac{1}{2}X$$
$$\therefore \quad byx = \frac{1}{2}$$

Regression of X on Y

$$3Y - 2X - 10 = 0$$
$$3Y - 10 = 2X$$
or $$2X = 3Y - 10$$
$$X = -5 + \frac{3}{2}Y$$
$$\therefore \quad bxy = \frac{3}{2}$$

$$r^2 = byx \, . \, bxy = \frac{1}{2} \times \frac{3}{2}$$
$$\therefore \quad r = \sqrt{\frac{3}{4}} = \sqrt{0.75} = = +0.866$$

(c) We know that $byx = r . \dfrac{\sigma_y}{\sigma_x}$

$$\Rightarrow \quad \frac{1}{2} = 0.866 . \frac{\sigma_y}{3} \qquad [\because \sigma_x = 3 \text{ given}]$$
$$\Rightarrow \quad \sigma_y = 1.732$$

Standard error of estimate of Y on X is

$$S_{yx} = \sigma_y \sqrt{1-r^2}$$
or $$S_{yx} = 1.732\sqrt{1-\frac{3}{4}} = 1.732 \times \frac{1}{2} = 0.866$$

Example 57. A departmental store gives in-service training to its salesmen which is followed by a test. It is considering whether it should terminate the service of any salesmen who does not do well in the test. The following data give the test scores and sales made by nine salesmen during a certain period :

Test Scores	14	19	24	21	26	22	15	20	19
Sales ('00 Rs.)	31	36	48	37	50	45	33	41	39

Calculate the coefficient of correlation between the test scores and the sales. Does it indicate that the termination of services of low test scores is justified ? If the firm wants a minimum sales volume of Rs. 3,000, what is the minimum test score that will ensure continuation of service ? Also estimate the most probable sales volume of a salesman making a score of 28.

Solution: Let X denote the test scores of the salesmen and Y denote their corresponding sales (in '00 Rs.)

Calculations for Regression Lines

X	Y	$x = X - \bar{X}$ $= X - 20$	$y = Y - \bar{Y}$ $= Y - 40$	x^2	y^2	xy
14	31	−6	−9	36	81	54
19	36	−1	−4	1	16	4
24	48	4	8	16	64	32
21	37	1	−3	1	9	−3
26	50	6	10	36	100	60
22	45	2	5	4	25	10
15	33	−5	−7	25	49	35
20	41	0	1	0	1	0
19	39	−1	−1	1	1	1
$\Sigma X = 180$	$\Sigma Y = 360$	$\Sigma x = 0$	$\Sigma y = 0$	$\Sigma x^2 = 120$	$\Sigma y^2 = 346$	$\Sigma xy = 193$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{180}{9} = 20 \qquad \bar{Y} = \frac{\Sigma Y}{N} = \frac{360}{9} = 40$$

byx = Coefficient of regression of Y on X

$$= \frac{\Sigma xy}{\Sigma x^2} = \frac{193}{120} = 1.6083$$

bxy = Coefficient of regression of X on Y

$$= \frac{\Sigma xy}{\Sigma y^2} = \frac{193}{346} = 0.5578$$

Karl Pearson's correlation coefficient r between x and y is given by:

$$r^2 = byx \cdot bxy = 1.6083 \times 0.5578 = 0.8971$$

$$\Rightarrow \quad r = \pm\sqrt{0.8971} = \pm\, 0.9471$$

Since, the regression coefficients are positive, r is also positive.

$$\therefore \quad r = +\,0.9471$$

Aliter:

$$r_{xy} = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}} = \frac{193}{\sqrt{120 \times 346}} = \frac{193}{\sqrt{41520}}$$

$$= \frac{193}{203.7646} = 0.9471$$

Thus, we see that there is a very high degree of positive correlation between the test scores (X) and the sales ('00 Rs.) (Y). This justifies the proposal for the termination of service of those with low test scores.

Regression Equations

To obtain the test score (X) for given sales (Y), we use the equation of the line of regression of X on Y.

The equation of line of regression of X on Y is:

$$X - \bar{X} = bxy\,(Y - \bar{Y})$$

$$\Rightarrow \quad X - 20 = 0.5578\,(Y - 40) = 0.5578Y - 22.312$$

$$\Rightarrow \quad X = 0.5578Y - 22.312 + 20$$

$$\Rightarrow \quad X = 0.5578Y - 2.312 \qquad \text{...}(i)$$

Hence, to ensure the continuation of service, the minimum test score (X) corresponding to a minimum sales volume (Y) of Rs. 3,000 = 30 ('00 Rs.) is obtained on putting $Y = 30$ in (i) and is given by :

$$X = 0.5578 \times 30 - 2.312 = 16.734 - 2.312$$

$$= 14.422 \approx 14$$

To estimate the sales volume (Y) of a salesman with given test score (X), we use the line of regression of Y on X, which is given by :

$$Y - \bar{Y} = byx\,(X - \bar{X})$$

$$\Rightarrow \quad Y - 40 = 1.6083\,(X - 20) = 1.6083X - 32.1660$$

$$\Rightarrow \quad Y = 1.6083X - 32.1660 + 40$$

$$\Rightarrow \quad Y = 1.6083X + 7.8340$$

Hence, the estimated sales volume of a salesman with test score of 28 is (in '00 Rs.)

$$Y = 1.6083 \times 28 + 7.8340$$

$$= 45.0324 + 7.8340$$

$$= 52.8664 \text{ ('00 Rs.)} = \text{Rs. } 5286.64$$

Example 58. In the estimation of regression equation of two variables X and Y, the following results were obtained: $\bar{X} = 90, \bar{Y} = 70, N = 10, \Sigma x^2 = 6360, \Sigma y^2 = 2860, \Sigma xy = 3900$, where, x and y are deviations from their respective means. Obtain the two lines of regression.

Solution: Given, $N = 10, \bar{X} = 90, \bar{Y} = 70, \Sigma x^2 = 6360, \Sigma y^2 = 2860, \Sigma xy = 3900$

$$\therefore \quad byx = \frac{\Sigma xy}{\Sigma x^2} = \frac{3900}{6360} = +\,0.61 \quad \text{Where, } x = X - \bar{X};\ y = Y - \bar{Y}$$

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{3900}{2860} = 1.36$$

Regression Equation of Y on X

$$Y - \bar{Y} = byx\,(X - \bar{X})$$

$$Y - 70 = 0.61\,(X - 90)$$

$$Y - 70 = 0.61X - 54.9$$
$$Y = 0.61X + 15.1$$

Regression Equation of X on Y

$$X - \overline{X} = bxy\,(Y - \overline{Y})$$
$$X - 90 = 1.36\,(Y - 70)$$
$$X - 90 = 1.36Y - 95.2$$
$$X = 1.36Y - 5.2$$

Example 59. From the following data, find out (*i*) correlation coefficient, (*ii*) linear regression line of Y on X. Also find out the percentage of variation explained by the regression line.

X :	1	2	3	4	5
Y :	2	5	3	8	7

Solution :

Calculations of Correlation Coefficient and Regression Equations

X	$\overline{X}=3$ x	x^2	Y	$\overline{Y}=5$ y	y^2	xy
1	–2	4	2	–3	9	6
2	–1	1	5	0	0	0
3	0	0	3	–2	4	0
4	+1	1	8	+3	9	3
5	+2	4	7	+2	4	4
$\Sigma X = 15$ $N = 5$	$\Sigma x = 0$	$\Sigma x^2 = 10$	$\Sigma Y = 25$	$\Sigma y = 0$	$\Sigma y^2 = 26$	$\Sigma xy = 13$

$$\overline{X} = \frac{15}{5} = 3, \overline{Y} = \frac{25}{5} = 5$$

Since the actual means of X and Y are whole numbers, we should take deviations from actual means of X and Y to simplify the calculations.

(*i*) Calculation of Correlation Coefficient

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \times \Sigma y^2}} = \frac{13}{\sqrt{10 \times 26}} = \frac{13}{16.12} = 0.806$$

(*ii*) Calculation of Regression Equations

$$byx = \frac{\Sigma xy}{\Sigma x^2} = \frac{13}{10} = 1.3$$

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{13}{26} = 0.5$$

Regression Equation of Y on X

$$Y - \overline{Y} = byx\,(X - \overline{X})$$
$$Y - 5 = 1.3\,(X - 3)$$
$$Y - 5 = 1.3X - 3.9 \quad \Rightarrow \quad Y = 1.3X + 1.1$$

Regression Equation of X on Y

$$X - \overline{X} = bxy\,(Y - \overline{Y})$$
$$X - 3 = 0.5\,(Y - 5)$$
$$X - 3 = 0.5Y - 2.5 \quad \Rightarrow \quad X = 0.5Y + 0.5$$

Calculation of r^2

$$r^2 = \text{Coefficient of Determination}$$
$$= (0.806)^2 = 0.6496 = 64.96\%$$

This implies that 64.96% variations in Y are explained by the regression line of Y on X

Example 60. Given the regression equation of Y on X and X on Y are respectively. $Y = 2X$ and $6X - Y = 4$ and the second moment of X about origin $\left(i.e., \Sigma X^2 / N\right)$ is 3. (*i*) Find the correlation coefficient and (*ii*) standard deviation of y.

Solution: (*i*) Regression equation of Y on X is $Y = 2X$ $\Rightarrow b_{yx} = 2$

Regression equation of X on Y is $6X - Y = 4$ $\Rightarrow X = \frac{1}{6}Y + \frac{4}{6} \Rightarrow b_{xy} = \frac{1}{6}$

$$\therefore \quad r = \sqrt{2 \times \frac{1}{6}} = \sqrt{0.3333} = 0.578$$

(*ii*) Second moment of X about origin = 3 (given)

$$\Rightarrow \quad \frac{\Sigma X^2}{N} = 3$$

$$\sigma_x^2 = \frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2 = 3 - (\overline{X})^2$$

Solving the two regression equations, we get $\overline{X} = 1, \overline{Y} = 2$

$$\therefore \quad \sigma_x^2 = 3 - 1 = 2 \quad \therefore \quad \sigma_x = \sqrt{2} = 1.414$$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x}$$

i.e., $\quad 2 = 0.578 \cdot \frac{\sigma_y}{\sqrt{2}} \quad$ *i.e.,* $\quad \sigma_y = \frac{2\sqrt{2}}{0.578} = \frac{2.828}{0.578} = 4.89$

Example 61. Find the 'standard error of the estimates'

$$\sigma_x = 1.414, \sigma_y = 8.94, r = 0.316$$

Solution: Given, $r = 0.316, \sigma_x = 1.414, \sigma_y = 8.94$

(*i*) $S_{xy} = \sigma_x \cdot \sqrt{1 - r^2} = 1.414\sqrt{1 - (0.316)^2} = 1.414 \times 0.949 = 1.342$

(*ii*) $S_{yx} = \sigma_y \cdot \sqrt{1 - r^2} = 0.894\sqrt{1 - (0.316)^2} = 0.894 \times 0.949 = 0.848$

Example 62. For a set of 10 pairs of values of X and Y, the regression line of X on Y is $X-2Y+12=0$; mean and standard deviation of Y being 8 and 2 respectively. Later it is known that a pair ($X=3$, $Y=8$) was wrongly recorded and the correct pair detected is ($X=8$, $Y=3$). Find the correct regression line of X on Y.

Solution: In the usual notations we are given, $N=10, \bar{Y}=8, \sigma_y=2$ ...(i)

The equation of the line of regression of X on Y is : $X-2Y+12=0$ (given). Since the lines of regression pass through the point $(\bar{X},\bar{Y})$, we get

$$\bar{X}-2\bar{Y}+12=0 \quad \Rightarrow \quad \bar{X}=2\bar{Y}-12=2\times8-12=4 \quad \text{[Using (i)]}$$

$$X-2Y+12=0 \quad \Rightarrow \quad X=2Y-12 \quad \Rightarrow \quad bxy=2$$

$$bxy=\frac{\frac{\Sigma XY}{N}-\bar{X}\bar{Y}}{\sigma_y^2}=2 \quad \Rightarrow \quad \frac{\Sigma XY}{N}-\bar{X}\bar{Y}=2\times2^2=8$$

$$\Rightarrow \frac{\Sigma XY}{N}-\bar{X}\bar{Y}=8 \quad \Rightarrow \quad \Sigma XY=10[8+4\times8]=10\times40=400$$

$$\sigma_y^2=\frac{\Sigma Y^2}{N}-(\bar{Y})^2 \quad \Rightarrow \quad \Sigma Y^2=N[\sigma_y^2+\bar{Y}^2]=10\ 4+8^2]=680$$

We have $\bar{X}=4, \bar{Y}=8, \Sigma Y^2=680, \Sigma XY=400$

Wrong pair = (X = 3, Y = 8); Correct pair = (X = 8, Y = 3)

Calculation of Correct Values

$$\bar{X}=\frac{\Sigma X}{N} \quad \Rightarrow \quad 4=\frac{\Sigma X}{10} \quad \Rightarrow \quad \Sigma X=40$$

$$\text{Corrected } \Sigma X=40-3+8=45 \quad \Rightarrow \quad \text{Corrected } \bar{X}=\frac{45}{10}=4.5$$

$$\bar{Y}=\frac{\Sigma Y}{N} \quad \Rightarrow \quad 8=\frac{\Sigma Y}{10} \quad \Rightarrow \quad \Sigma Y=80$$

$$\text{Corrected } \Sigma Y=80-8+3=75 \quad \Rightarrow \quad \text{Corrected } \bar{Y}=\frac{75}{10}=7.5$$

$$\text{Corrected } \Sigma Y^2=680-8^2+3^2=625$$

$$\text{Corrected } \sigma_y^2=\frac{\Sigma Y^2}{N}-(\bar{Y})^2=\frac{625}{10}-(7.5)^2=6.25$$

$$\text{Corrected } \Sigma XY=400-24+24=400$$

$$\text{Corrected } bxy=\frac{\frac{\Sigma XY}{N}-\bar{X}\bar{Y}}{\sigma_y^2}=\frac{\frac{400}{10}-(4.5)(7.5)}{6.25}=1$$

Corrected line of regression of X on Y becomes

$$X-\bar{X}=bxy(Y-\bar{Y})$$

$$\Rightarrow \quad X-4.5=1(Y-7.5)$$

$$\Rightarrow \quad X=Y-7.5+4.5 \quad \Rightarrow \quad X=Y-3$$

Example 63. The data in the following table relates the weekly maintenance cost (in Rs.) to the age (in months) of ten machines of similar type in a manufacturing company. Find the least squares regression line of maintenance cost on age and use to predict the maintenance cost for a machine of this type which is 40 months old.

Machine:	1	2	3	4	5	6	7	8	9	10
Age (X):	5	10	15	20	30	30	30	50	50	60
Cost (Y):	190	240	250	300	310	335	300	300	350	395

Solution:

Computation of Regression Equation

X	Y	$X-\bar{X}=x$	$Y-\bar{Y}=y$	x^2	xy
5	190	−25	−107	625	2,675
10	240	−20	−57	400	1,140
15	250	−15	−47	225	705
20	300	−10	3	100	−30
30	310	0	13	0	0
30	335	0	38	0	0
30	300	0	3	0	0
50	300	20	3	400	60
50	350	20	53	400	1,060
60	395	30	98	900	2,940
$\Sigma X=300$	$\Sigma Y=2,970$	$\Sigma x=0$	$\Sigma y=0$	$\Sigma x^2=3,050$	$\Sigma xy=8,550$

$$\bar{X}=\frac{300}{10}=30; \quad \bar{Y}=\frac{2970}{10}=297$$

Since both means are integers deviations have been taken from actual means.

Regression Coefficient of Y on X

$$byx=\frac{\Sigma xy}{\Sigma x^2}=\frac{8,550}{3,050}=2.8033$$

Regression Equation of Y on X is

$$Y-\bar{Y}=byx(X-\bar{X})$$

$$\Rightarrow \quad y-297=2.8033(X-30)=2.8033X-84.099$$

$$\Rightarrow \quad Y=2.8033X+212.901$$

For $X=40$, $Y=2.8033\times40+212.901$

$$=325.033=\text{Rs. }325.$$

Example 64. For certain data, the following regression equations were obtained :

$$4X-5Y+33=0$$

$$20X-9Y-107=0$$

Estimate Y when $X=20$ and X when $Y=20$.

Solution : Let $4X - 5Y + 33 = 0$ be the regression equation of Y on X, while $20X - 9Y - 107 = 0$ be the regression equation of X on Y,

From (*i*) $\quad Y = \frac{4}{5}X + \frac{33}{5} \Rightarrow \quad byx = \frac{4}{5}$ and

From (*ii*) $\quad X = \frac{9}{20}Y + \frac{107}{20} \Rightarrow \quad bxy = \frac{9}{20}$

Now, $\quad r = \sqrt{byx \times bxy} = \sqrt{\frac{4}{5} \times \frac{9}{20}} = \sqrt{\frac{9}{25}} = \frac{3}{5} < 1,$

So our assumption is correct.

$\therefore$ When $X = 20$, $Y_{20} = \frac{4}{5} \times 20 + \frac{33}{5} = \frac{113}{5} = 22.6$ and

When $Y = 20$, $X_{20} = \frac{9}{20} \times 20 + \frac{107}{20} = \frac{287}{20} = 14.35.$

Example 65. Given the following data, find what will be (a) the height of a policeman whose weight is 200 pounds, (b) the weight of a policeman who is 6 ft. tall.

Average height = 68 inches, average weight = 150 pounds, coefficient of correlation between height and weight = 0.6, S.D. of heights = 2.5 inches, S.D. of weights = 20 pounds.

Solution: Let height of policeman be denoted by X and weight of policeman by Y.

We are given:

$\bar{X} = 68''$, $\quad \bar{Y} = 150$ lbs, $\quad \sigma_x = 2.5''$, $\quad \sigma_y = 20$ lbs, $\quad r_{xy} = 0.6$

(*i*) For estimating the height of a policeman whose weight is 200 lbs, we use regression of X on Y as follows:

$$X - \bar{X} = r.\frac{\sigma_x}{\sigma_y}(Y - \bar{Y})$$

$$X - 68 = 0.6 \times \frac{2.5}{20}(Y - 150)$$

$$X - 68 = 0.075\,(Y - 150)$$

$$X - 68 = 0.075Y - 11.25$$

$$X = 0.075Y + 56.75$$

When $Y = 200$, $X = 0.075\,(200) + 56.75 = 71.75$

Thus, the height of a policeman whose weight is 200 lbs shall be 71.75".

(*ii*) For estimating the weight of a policeman whose height is 72" (*i.e.*, 6 ft), we use regression of Y on X as follows:

$$Y - \bar{Y} = r.\frac{\sigma_y}{\sigma_x}(X - \bar{X})$$

$$Y - 150 = 0.6 \times \frac{20}{2.5}(X - 68)$$

$$Y - 150 = 4.8\,(X - 68)$$

$$Y - 150 = 4.8X - 326.4$$

$$Y = 4.8X - 176.4$$

When $X = 72$, $Y = 4.8\,(72) - 176.4 = 169.2$

Thus, the weight of a policeman who is 6 ft. tall should be 169.2 lbs.

Example 66. Prove that regression coefficients are independent of the change of origin but not of scale.

Solution: Change the X and Y variables into new variables in the following manner:

$$U = \frac{x - a}{h},\ V = \frac{y - b}{k}$$

$$bxy = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(Y - \bar{Y})^2} \qquad \text{...(i)}$$

$$X = hu + a \qquad\qquad Y = kv + b$$

$$\Sigma X = h\Sigma u + \Sigma a \qquad\qquad \Sigma Y = k\Sigma v + \Sigma b$$

$$\Sigma X = h\Sigma u + na \qquad\qquad \Sigma Y = k.\Sigma v + nb$$

$$\frac{\Sigma X}{n} = h\cdot\frac{\Sigma u}{n} + a \qquad\qquad \frac{\Sigma Y}{n} = k\cdot\frac{\Sigma v}{n} + b$$

$$\therefore\ \bar{X} = h\bar{u} + a \qquad\qquad \therefore\ \bar{Y} = k\bar{v} + b$$

$$\therefore\ X - \bar{X} = h(u - \bar{u}) \qquad\qquad \therefore\ Y - \bar{Y} = k(v - \bar{v})$$

Substituting in (*i*),

$$bxy = \frac{\Sigma h(u - \bar{u})\cdot k\cdot(v - \bar{v})}{\Sigma k^2(v - \bar{v})^2} = \frac{hk}{k^2}\cdot\frac{\Sigma(u - \bar{u})(v - \bar{v})}{\Sigma(v - \bar{v})^2}$$

$$bxy = \frac{h}{k}\cdot\frac{\Sigma(u - \bar{u})(v - \bar{v})}{\Sigma(v - \bar{v})^2} = \frac{h}{k}\cdot buv$$

Hence the result is proved.

Similarly, we can prove for $byx = \frac{k}{h}bvu$

Example 67. Prove that the mean of two regression coefficients is always greater than the coefficient of correlation.

Solution: We have to prove

$$\frac{byx + bxy}{2} > r$$

$$\frac{r\cdot\frac{\sigma_y}{\sigma_x} + r\cdot\frac{\sigma_x}{\sigma_y}}{2} > r \quad \Rightarrow \quad \frac{r\left(\frac{\sigma_y}{\sigma_x} + \frac{\sigma_x}{\sigma_y}\right)}{2} > r$$

or $$\frac{\frac{\sigma_y}{\sigma_x}+\frac{\sigma_x}{\sigma_y}}{2} > 1 \quad \Rightarrow \quad \frac{\sigma^2_y+\sigma^2_x}{2\sigma_x \cdot \sigma_y} > 1$$

Multiplying both sides by $2\sigma_x . \sigma_y$

$$\Rightarrow \quad \sigma^2_x+\sigma^2_y > 2\sigma_x \cdot \sigma_y \qquad [\because \sigma_x > 0, \sigma_y > 0]$$

$$\Rightarrow \quad \sigma^2_x+\sigma^2_y - 2\sigma_x . \sigma_y \geq 0 \quad \Rightarrow \quad (\sigma_x - \sigma_y)^2 \geq 0$$

As the square of real numbers can never be less than zero. Hence the arithmetic mean of the two regression coefficients is greater than the correlation coefficient.

Example 68. Show that θ, the acute angle between two lines of regression is given by:

$$\tan\theta = \frac{(1-r^2)}{|r|}\frac{\sigma_x \cdot \sigma_y}{\sigma_x^2+\sigma_y^2}$$

Interpret the case when $r = 0, \pm 1$

Solution: Equations of the two lines of regression are:

$$Y-\bar{Y} = r . \frac{\sigma_y}{\sigma_x}(X-\bar{X}) \text{ and } X-\bar{X} = r . \frac{\sigma_x}{\sigma_y}(Y-\bar{Y})$$

We have, m_1 = slope of the line of regression of Y on X = $byx = r . \frac{\sigma_y}{\sigma_x}$

m_2 = slope of the line of regression of X on Y = $\frac{1}{bxy} = \frac{1}{r} . \frac{\sigma_y}{\sigma_x}$

Let θ be the angle between two lines of regression, then

$$\tan\theta = \pm\frac{m_2-m_1}{1+m_1m_2} = \frac{\frac{1}{r}.\frac{\sigma_y}{\sigma_x} - r.\frac{\sigma_y}{\sigma_x}}{1+r.\frac{\sigma_y}{\sigma_x}.\frac{1}{r}.\frac{\sigma_y}{\sigma_x}} = \pm\left(\frac{1-r^2}{r}\right)\left(\frac{\sigma_x \cdot \sigma_y}{\sigma_x^2+\sigma_y^2}\right)$$

Since $r^2 \leq 1$ and σ_x, σ_y are positive.

∴ +ve sign gives the acute angle between the lines.

Hence, $$\tan\theta = \frac{(1-r^2)}{|r|} \cdot \frac{\sigma_x \cdot \sigma_y}{\sigma_x^2+\sigma_y^2}$$

Case I: When $r = 0, \tan\theta = \infty \Rightarrow \theta = \frac{\pi}{2} = 90°$

Thus, the lines of regression are perpendicular to each other.

Case II: When $r = \pm 1, \tan\theta = 0 \Rightarrow \theta = 0$ or π

Thus, the two lines of regression coincide and there will be one regression line.

IMPORTANT FORMULAE

1. Regression Lines

(i) Regression Equation of Y on X

$$Y-\bar{Y} = byx(X-\bar{X})$$

Where, $byx = \frac{N.\Sigma XY - \Sigma X . \Sigma Y}{N.\Sigma X^2 - (\Sigma X)^2}$ **(When we use actual values of X and Y)**

$= \frac{\Sigma xy}{\Sigma x^2}$ **Where,** $x = X - \bar{X}$; $Y = Y - \bar{Y}$

(When deviations are taken from actual means of X and Y)

$= \frac{N \cdot \Sigma dxdy - \Sigma dx \cdot \Sigma dy}{N \cdot \Sigma dx^2 \div (\Sigma dx)^2}$ **Where,** $dx = X - A$, $dy = Y - A$

(When deviations are taken from assumed means of X and Y)

$= r \cdot \frac{\sigma_y}{\sigma_x}$ **(When we use r, σ_y and σ_x)**

(ii) Regression Equation of X on Y

$$X-\bar{X} = bxy(Y-\bar{Y})$$

$bxy = \frac{N.\Sigma XY - \Sigma X . \Sigma Y}{N.\Sigma Y^2 - (\Sigma Y)^2}$ **(When we use actual values of X and Y)**

$= \frac{\Sigma xy}{\Sigma y^2}$ **(When deviations are taken from actual mean)**

$= \frac{N \cdot \Sigma dxdy - \Sigma dx \cdot \Sigma dy}{N \cdot \Sigma dy^2 - (\Sigma dy)^2}$ **(When deviations are taken from assumed mean)**

$= r \cdot \frac{\sigma_x}{\sigma_y}$ **(When we use r, σ_x and σ_y)**

(iii) In Grouped Frequency Distribution

$$byx = \frac{N.\Sigma fdxdy - \Sigma fdx \, \Sigma fdy}{N.\Sigma fdx^2 - (\Sigma fdx)^2} \times \frac{i_y}{i_x}$$

and $$bxy = \frac{N.\Sigma fdxdy - \Sigma fdx \, \Sigma fdy}{N.\Sigma fdy^2 - (\Sigma fdy)^2} \times \frac{i_x}{i_y}$$

Where, $N = \Sigma f$, **stands for the total frequency.**

2. Regression Coefficients

There are two regression coefficients:

(*i*) Regression coefficient of Y on X = $byx = r \cdot \frac{\sigma_y}{\sigma_x}$

(*ii*) Regression coefficient of X on Y = $bxy = r \cdot \frac{\sigma_x}{\sigma_y}$

(*iii*) $r = \sqrt{bxy \times byx}$

3. Standard Error

Standard error of estimate $S_{xy} = \sqrt{\frac{\Sigma(X - X_c)^2}{N}}$ or $S_{xy} = \sigma_x \cdot \sqrt{1 - r^2}$

$$S_{yx} = \sqrt{\frac{\Sigma(Y - Y_c)^2}{N}} \quad or \quad S_{yx} = \sigma_y \cdot \sqrt{1 - r^2}$$

Important Note

(*i*) Regression equation of Y on X is used to estimate the best (average) value of Y for given value of X.

(*ii*) Regression equation of X on Y is used to estimate the best (average) value of X for given value of Y.

QUESTIONS

1. Explain the concept of regression and comment on its utility. Also distinguish between correlation and regression.
2. What are regression coefficients? Explain the properties of regression coefficients.
3. Discuss the difference between correlation and regression.
4. Define the Standard Error of Estimate. How is it computed?
5. What is regression line? Why are there, in general, two regression lines? Under what conditions can there be only one regression line? When the two lines of regression intersect each other at 90°?
6. What would be lines of regression if $r = +1, r = -1, r = 0$. Give interpretation in each case.
7. Explain the meaning of (*i*) Standard Error of Estimate and (*ii*) Coefficient of Determination.
8. How would you identify regression equation of X on Y and Y on X?
9. What is the relationship between correlation and regression coefficients?

Index Numbers–I

3

INTRODUCTION

Economic and business data change from time to time. For instance, prices of all commodities do not remain constant. It is possible that sometimes the price of a commodity rises and sometimes falls. Similarly, output of a commodity sometimes rises and sometimes falls. The measurement of such changes is possible only by means of some statistical methods. Index numbers are such statistical devices which help in the measurement of such changes. In other words, by index numbers we mean the statistical measures with the help of which relative changes in general price level taking place at different points of time can be measured. Application of Index numbers are not limited only to general price levels but rather they help in the relative measurement of every such phenomenon like cost-of-living, output, national income, business activities whose direct measurement is not possible. Index numbers are used to measure the relative changes in some phenomena which we cannot observe directly.

DEFINITION OF INDEX NUMBERS

Some important definitions of index numbers are given below:

1. Index Numbers are a specialized type of averages. **—M. Blair**
2. Index Numbers are devices for measuring differences in the magnitude of a group of related variables. **—Croxten and Cowden**
3. An Index Number is a statistical measure designed to show changes in a variable or group of related variables with respect to time, geographic location or other characteristics. **—Spiegel**

The definitions discussed above specify the following features of index numbers:

(*i*) Relative changes in the aggregates are measured by index numbers (*ii*) Index numbers always present the changes taking place in some variable on an average only (*iii*) By index numbers, positions in base year and current year are compared.

USES OF INDEX NUMBERS

In present times, the importance of index numbers is increasing. Nowadays, they are being used in economics and business fields. To quot **Simpson and Kafka, "Index Numbers are economic barometers"**. The main uses of index numbers are the followings:

(1) To Simplify Complexities: An index number makes possible the measurement of such complex changes whose direct measurement is not possible. In other words, index numbers are used to measure the changes in some quantity which we cannot observe directly.

(2) Helpful in Fixation of Salary and Dearness Allowances: By index numbers, the goverment and other employees can properly make wage and salary fixation. They determine the instalment of dearness allowance for employees on the basis of index numbers only.

(3) Helpful in Predictions: Index numbers give the knowledge as to what changes have occurred in the past. On the basis of these changes alone, predictions about the future are made. Thus, index numbers are economic barometers.

(4) Helpful in Comparison: Index numbers make possible the comparative study of phenomena. By index numbers, the relative changes occurring in the variables are determined. This simplifies the comparison of data on the basis of time and space.

(5) To Measure Purchasing Power of Money: By index numbers, the changes taking place in the purchasing power of money can also measured.

(6) Useful in Business: Index numbers measures the changes taking place in business world and prove very useful in making a comparative study of those changes, *e.g.*, sales, changes in output and value, etc. Thus, index numbers, for a businessman, function like a barometer.

To sum up, index numbers are the sign and guide posts along the business highway that indicate to the businessmen how he should drive or manage his affairs.

LIMITATIONS OF INDEX NUMBERS

Though index numbers are extremely useful statistical devices, yet they have some limitations which are detailed as follows:

(*i*) Index numbers measure relative changes in different phenomena. They do not always hold cent per cent true. Index numbers are true only on the average.

(*ii*) A given type of index numbers is not suitable for all the purposes. Multipurpose index numbers cannot be constructed.

(*iii*) No consideration is given to the changes taking place in the quality of a commodity while constructing the index numbers.

(*iv*) Bias in the selection of base year and selection of representatives sometimes leads to misleading results.

(*v*) Index numbers lack in perfect accuracy because they are mostly constructed on the basis of sample commodities.

TYPES OF INDEX NUMBERS

Index numbers are classified on the basis of the phenomena whose changes they measure. In the economic and business, index numbers can be classified into the following types:

(1) Price Index Numbers: Price index numbers are most popular and commonly used index numbers. These index number measure the changes in prices of some commodities or group of commodities consumed in the given period with reference to the base period. These are of two types:

(*i*) **Wholesale Price Index Number:** It measures the changes in the general price level of a commodity.

(*ii*) **Retail Price Index Number:** It measures the general changes in the retail prices of commodities which are bought and sold in the retail market.

(2) Quantity Index Numbers: Quantity index numbers helps us in measuring and comparing the changes in the physical volume of goods produced or sold or purchased in a given period. Indices of agriculture and industrial production are included in this category.

(3) Value Index Numbers: Value index numbers measure the changes in the value of some commodities or group of commodities consumed or purchased in the given period with reference to base period.

(4) Simple and Aggregative Index Numbers: On the basis of the number of commodities that enter into the construction of an index, index numbers are classified into two categories:

(*i*) **Simple Index Numbers:** When index numbers are constructed for individual commodities, these are termed as simple index numbers.

(*ii*) **Aggregative (or composite) Index Numbers:** When index numbers are constructed for a group of commodities, these are known as aggregative (or composite) index numbers.

(5) Cost of Living Index Number: Cost of living index numbers are also called consumer price index numbers. They help in comparison of average change in consumption and expenditure of the commodity from one time period to another. It shows the average change in consumer expenditure of a particular class of consumers.

(6) Special Purpose Index Numbers: Some index numbers are constructed for some specific purposes or aspects. They measure the average change as compared to the base period of any specific purpose.

PROBLEMS IN THE CONSTRUCTION OF INDEX NUMBERS

A number of problems come up while constructing the index numbers. The important among them are as follows:

(1) Purpose of Index Number: Index numbers are of many types as they are constructed for different purposes. It is very essential to fix the purpose before constructing an index number, because selection of commodities, their prices, fixation of their weights, etc., depend on the very purpose of index numbers. There can be many purposes of an index number-measurement of changes in retail prices or measurement of changes in wholesale prices etc.

(2) Selection of Items: Another important problem in the construction of index numbers is the selection of items. The following things should be considered while making a selection of items: (*i*) only those items should be selected which represent the taste, habit, custom and needs of the related group of people, (*ii*) the selected items should be standardised and of classified feature, (*iii*) in selection of items, their quality too must be considered, (*iv*) the number of items must be enough and they should be of current quality, and (*v*) the selected items must be classified into different groups and sub-groups.

(3) Selection of Prices: After making the selection of items, the next arises the problem of selection of prices. Price can be of both types—retail and wholesale. Thus, whether wholesale or retail prices are to be used, the decision depends upon the purpose of index number. Often, wholesale prices are taken in the construction of index numbers. In addition, prices should be picked up from that place where the related items are traded most.

(4) Selection of Base year: Another important problem in the construction of index numbers is related to the selection of base year. A base year has to be selected for making an index number. The year for which changes are to be determined, is known as **base year**. Index number of base year is always taken as 100. In selecting a base year, the following things are to be kept in mind:

(*i*) Base year should be a normal year and no unusual event like Earthquake, Flood, War, etc., should have taken place in that year,

(*ii*) Base year should not be very far in past,

(*iii*) So far as possible, base year should be close to the current year,

(*iv*) Base year should not be too old or too distint.

(5) Selection of Weights: Another important problem in making of index numbers is to assign weights to different commodities or items. In fact, all commodities included in the construction of index numbers do not have equal importance. Therefore, to have accurate results, commodities are assigned weights according to their importance. There are two ways of assigning weights: (*i*) Quantity, (*ii*) Value. Weights decided in the construction of index numbers should be logical, accurate and rational.

(6) Selection of an Average: The selection of an average is also a significant problem in the preparation of index numbers. Averages can be of several types. Theoretically, any average can be used but in practice, usually arithmetic mean and geometric mean are used. Geometric mean is considered to be the best for the construction of index numbers as this is most suitable for measuring relative changes but due to the difficulties of computation in place of geometric mean, arithmetic mean is most often used in the construction of index numbers.

(7) Section of an Appropriate Formula: Various formulae can be used in the construction of index numbers but it is very essential to select the most suitable out of them. This selection depends upon the purpose of index number and availability of data. **Fisher's formula** which is called as **Fisher's Ideal Index**, is considered to be the best.

METHODS OF CONSTRUCTING INDEX NUMBERS

Different methods are used in constructing the index numbers which, from the view point of convenience of study, have been presented by the following chart:

- Methods of Constructing Index Numbers
  - Unweighted or Simple Index Numbers
    - Simple Aggregative Method
    - Simple Average of Price Relatives Method
  - Weighted Index Numbers
    - Weighted Aggregative Method
      - (a) Laspeyre's Method
      - (b) Paasche's Method
      - (c) Fisher's Method
      - (d) Dorbish and Bowley's Method
      - (e) Marshall and Edgeworth's Method
      - (f) Kelly's Method
    - Weighted Average of Price Relatives Method

SIMPLE OR UNWEIGHTED INDEX NUMBERS

Simple or unweighted index numbers are those ones in the construction of which all commodities are given equal importance. There are two methods of their construction:

(1) Simple Aggregative Method.

(2) Simple Average of Price Relatives Method.

(1) Simple Aggregative Method

It is the simplest method. In this method, sum of current year's prices is divided by the sum of base year's prices and the quotient is multiplied by 100. The following formula is used:

$$P_{01} = \frac{\Sigma p_1}{\Sigma p_0} \times 100$$

Where, Σp_1 = Aggregate of Prices in Current Year

Σp_0 = Aggregate of Prices in Base Year

P_{01} = Price Index

This method can be illustrated with the following examples:

Example 1. Construct price index number for 1990 based on 1981 using Simple Aggregative Method:

Commodity	Price in 1981 (in Rs.)	Price in 1990 (in Rs.)
A	50	80
B	40	60
C	10	20
D	5	10
E	2	8

Solution:

Construction of Price Index Number

Commodity	Price in 1981 (p_0)	Price in 1990 (p_1)
A	50	80
B	40	60
C	10	20
D	5	10
E	2	8
	$\Sigma p_0 = 107$	$\Sigma p_1 = 178$

$$P_{01} = \frac{\Sigma p_1}{\Sigma p_0} \times 100 = \frac{178}{107} \times 100 = 166.36$$

▶ **Merits and Demerits of Simple Aggregate Method**

Simple aggregative method of index number construction is very easy but it can be applied only when the prices of all commodities have been expressed in the same unit. If units are different, the results will be misleading.

(2) Simple Average of Price Relatives Method

In this method, first of all, the price relatives of the commodities or items are found out. To compute the price relatives, price in current year (p_1) is divided by the price in base year (p_0) and then, the quotient is multiplied with 100. In terms of formula,

$$\text{Price Relative} = \frac{\text{Current year's Price}}{\text{Base year's Price}} \times 100 \quad \text{or} \quad P = \frac{p_1}{p_0} \times 100$$

After this, using Arithmetic average or Geometric average, or Median, we find the average of price relatives.

(i) **When Arithmetic mean is used,** then the following formula is used:

$$P_{01} = \frac{\Sigma\left(\frac{p_1}{p_0} \times 100\right)}{N}$$

Where, N = number of items or commodities.

(ii) **When Geometric Mean is used,** then the following formula is used:

$$P_{01} = \text{Antilog}\left(\frac{\Sigma \log P}{N}\right)$$

Where, $P = \frac{p_1}{p_0} \times 100$

(iii) **When Median is used,** the following formula is used :

$$P_{01} = \text{Size of}\left(\frac{N+1}{2}\right)$$

In practice, Arithmetic mean is often used.

The following example illustrates the procedure of the method:

Example 2. The following are the prices of commodities in 1980 and 1985. Construct a price index based on price relatives taking 1980 as base year using *(i)* arithmetic mean, *(ii)* geometric mean and *(iii)* median.

Commodity	A	B	C	D	E
Price in 1980	50	40	80	110	20
Price in 1985	70	60	90	120	20

Solution: *(i)* **Price Index using Arithmetic Mean of Price Relatives**

Commodity	Price in 1980 (p_0)	Price in 1985 (p_1)	Price Relatives (P) $P = \frac{p_1}{p_0} \times 100$
A	50	70	140.0
B	40	60	150.0
C	80	90	112.5
D	110	120	109.1
E	20	20	100.00
$N = 5$			$\Sigma \frac{p_1}{p_0} \times 100 = 611.6$

Using Arithmetic Mean, the formula used is:

$$P_{01} = \frac{\Sigma\left(\frac{p_1}{p_0} \times 100\right)}{N} = \frac{611.6}{5} = 122.32$$

(ii) **Price Index using Geometric Mean of Price Relatives**

Commodity	Price in 1980 p_0	Price in 1985 p_1	$P = \frac{p_1}{p_0} \times 100$	log P
A	50	70	140	2.1461
B	40	60	150	2.1761
C	80	90	112.5	2.0512
D	110	120	109.1	2.0378
E	20	20	100	2.0000
				$\Sigma \log P = 10.4112$

Using Geometric Mean, the formula used is:

$$P_{01} = \text{Antilog}\left(\frac{\Sigma \log P}{N}\right) = \text{Antilog}\left(\frac{10.4112}{5}\right)$$

$$= \text{Antilog}\ (2.0822) = 120.9$$

(iii) **Price Index using Median of Price Relatives**

Commodity	Price Relatives of 1985	Price Relatives arranged in ascending order
A	140	100.0
B	150	109.1
C	112.5	112.5
D	109.1	140.0
E	100	150.0

$$P_{01} = \text{Size of}\left(\frac{N+1}{2}\right)\text{th item} = \text{Size of}\left(\frac{5+1}{2}\right)\text{th item}$$

$$= 3\text{rd item} = 112.5$$

Construction of Simple Index Numbers on the Basis of Average Price

If the simple index numbers are constructed on the basis of average price as base, firstly we compute the average price of each commodity and this is taken as base. The price relatives are then calculated by using the following formula:

$$\text{Price Relative} = \frac{\text{Current's Year Price}}{\text{Average Price}} \times 100$$

After this, using arithmetic mean, we find the average of price relatives. This average gives the index numbers.

The following example illustrate the procedure of this method:

Example 3. Find index number for the three years, taking average price as base by using price relative method:

Price per quintal (Rs.)

Year	A	B	C
1995	3	5	8
1996	5	4	6
1997	7	6	7

Solution: Average price of $A = \frac{3+5+7}{3} = \frac{15}{3} = 5$

Average price of $B = \frac{5+4+6}{3} = \frac{15}{3} = 5$

Average price of $C = \frac{8+6+7}{3} = \frac{21}{3} = 7$

Commodity	Average price	1995		1996		1997	
		p_0	PR	p_1	PR	p_2	PR
A	5	3	$\frac{3}{5}\times 100 = 60$	5	$\frac{5}{5}\times 100 = 100$	7	$\frac{7}{5}\times 100 = 140$
B	5	5	$\frac{5}{5}\times 100 = 100$	4	$\frac{4}{5}\times 100 = 80$	6	$\frac{6}{5}\times 100 = 120$
C	7	8	$\frac{8}{7}\times 100 = 114.29$	6	$\frac{6}{7}\times 100 = 85.71$	7	$\frac{7}{7}\times 100 = 100$
Total of price relatives			274.29		265.71		360
Average of price relatives or price index			91.43		88.57		120

IMPORTANT TYPICAL EXAMPLE

Example 4. From the following data, construct Price Index Number for three years with average price as base by simple average of price relatives method using arithmetic mean:

Rate per rupee

Year	Wheat	Cotton	Oil
I	10 Kg	4 Kg	2 Kg
II	8 Kg	2.5 Kg	2 Kg
III	5 Kg	2.0 Kg	1 Kg

Solution: We are given price per rupee. First convert these prices on a common scale say price per 100 kg. Thus, the price of wheat would be $\frac{100}{10} = 10, \frac{100}{8} = 12.50, \frac{100}{5} = 20$.

Similarly, prices of cotton and oil would be $\frac{100}{4} = 25\ \frac{100}{25} = 40, \frac{100}{2} = 50$ and $\frac{100}{2} = 50, \frac{100}{2} = 50, \frac{100}{1} = 100.$

$$\text{Average Price of Wheat} = \frac{10 + 12.50 + 20}{3} = 14.17$$

$$\text{Average Price of Cotton} = \frac{25 + 40 + 50}{3} = 38.33$$

$$\text{Average Price of Oil} = \frac{50 + 50 + 100}{3} = 66.67$$

Construction of Price Index Number

Commodity	Units	Average Price (p_0)	1st Year		2nd Year		3rd Year	
			Price (p_1)	Price Relatives $\frac{p_1}{p_0}\times 100$	Price (p_2)	Price Relatives $\frac{p_2}{p_0}\times 100$	Price (p_3)	Price Relatives $\frac{p_3}{p_0}\times 100$
Wheat	Per Qtl.	14.17	10	70.57	12.50	88.21	20	141.14
Cotton	Per Qtl.	38.33	25	65.22	40	104.36	50	130.45
Oil	Per Qtl.	66.67	50	75	50	75	100	150.00
Total of Price Relatives				210.79		267.57		421.59
Average of Price Relatives				70.26		89.19		140.53

EXERCISE 3.1

1. Construct the following indices by taking 1997 as the base:
   (*i*) Simple aggregative price index
   (*ii*) Index of average of price relatives

Items	A	B	C	D	E
Prices Rs. (1997):	6	2	4	10	8
Prices Rs. (1998):	10	2	6	12	12
Prices Rs. (1999):	15	3	8	14	16

[Ans. (*i*) 1998 - 140 and 1999 - 186.66 (*ii*) 1998 - 137.32 and 1999-188]

2. Compute a price index from the following by (*i*) simple aggregative method, and (*ii*) average of price relative method by using both arithmetic mean and geometric mean:

Commodity	A	B	C	D	E	F
Price in 1983 (Rs.)	200	300	100	250	400	500
Price in 1988 (Rs.)	250	300	150	350	450	550

[Ans. (*i*) 117.143; (*ii*) 122.92, 121.7]

3. Prepare price index numbers for three years taking the average price as base using simple average of price relative method:

Price per Quintal (Rs.)

Year	Wheat	Cotton	Oil
1995	100	25	30
1996	99	20	25
1997	99	15	30

[Ans. 110.52, 95.97, 93.52]

4. Construct the price index number for three years taking the average price as base by using simple average of price relative method:

Year	Rate per rupee		
	Wheat	Rice	Sugar
I	2 kgm	1 kgm	0.400 kgm
II	1.6 kgm	0.800 kgm	0.400 kgm
III	1 kgm	0.750 kgm	0.250 kgm

[Ans. 79.2, 92.1, 128.7]

5. Find out index number for 1972, 1973 and 1974 taking 1971 as base by using mean, median and geometric mean:

Group	1971	1972	1973	1974
A	8	12	16	20
B	32	40	48	60
C	16	20	32	40

[Hint: See Example 27]

[Ans. P_{01} (AM) = 133.3, 183.3 229.2
P_{01} (M) = 125, 200, 250 , P_{01} (GM) = 132.83, 181.7, 227.1]

◘ WEIGHTED INDEX NUMBERS

In constructing simple index numbers, all commodities are given equal importance but in practice, all commodities don't have equal importance. For example, for a consumer, wheat is more important then vegetable or pulse. Similarly, clothes are more important than a video. To express the relative importance of different commodities, weights on some definite basis are used. *When index numbers are constructed taking into consideration the importance of different commodities, then they are called weighted index numbers*. There are two methods of constructing weighted index numbers.

(I) Weighted Aggregative Method.

(II) Weighted Average of Price Relatives Method.

o (I) Weighted Aggregative Method

In this method, commodities are assigned weights on the basis of the quantities purchased. Many statisticians are divided on the issue that (*i*) Weight should be given on the basis of current year (q_1) (*ii*) On the base year quantity (q_0) (*iii*) On the basis of both base and current years quantities (q_0, q_1). Different statisticians have used different methods of assigning weights. Some of which quantity are the following:

(1) Laspeyre's Method

(2) Paasche's Method

(3) Fisher's Method

(4) Dorbish and Bowley Method

(5) Marshall-Edgeworth's Method

(6) Kelly's Method.

Various methods of constructing weighted index numbers can be illustrated as follows:

▸ **(1) Laspeyre's Method**

Prof. Laspeyre has assigned weights to the commodities on the basis of base year quantities (q_0). Laspeyre's formula is as follows:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

Steps for Calculation

(i) First of all, multiplying current year prices (p_1) with the corresponding base year quantities (q_0), their summation $\Sigma p_1 q_0$ is computed.

(ii) Then, multiplying base year prices (p_0) with base year quantities (q_0), their summation $\Sigma p_0 q_0$ is found out.

(iii) Finally, $\Sigma p_1 q_0$ is divided by $\Sigma p_0 q_0$ and the result is multiplied with 100.

▶ (2) Paasche's Method

Paasche has assigned weights to the commodities on the basis of current year quantities (q_1). Paasche's formula is as follows:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$$

Steps for Calculation

(i) First of all, multiplying current year prices (p_1) with current year quantities (q_1), their summation $\Sigma\, p_1 q_1$ is determined.

(ii) Then, base year prices (p_0) are multiplied with current year quantities (q_1), their summation $\Sigma\, p_0 q_1$ is computed.

(iii) Finally, $\Sigma\, p_1 q_1$ is divided by $\Sigma\, p_0\, q_1$ and the result is multiplied by 100.

▶ (3) Fisher's Method

This is the best method in use for the construction of weighted index numbers. **Irving Fisher** has assigned weights to different commodities on the basis of both the base year as well as current year quantities. Fisher's formula is as follows:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

In fact, Fisher's Index is the geometric mean of Laspeyre's and Paasche's Indices, *i.e.*,

$$P_{01} = \sqrt{L \times P}$$

Steps for Calculation

(i) First of all, current year prices (p_1) are multiplied with current year quantities (q_1) to arrive at $\Sigma\, p_1 q_1$.

(ii) Base year prices (p_0) are multiplied by current year quantities (q_1) to get $\Sigma\, p_0 q_1$.

(iii) Then, current year prices (p_1) are multiplied by base year quantities (q_0) and further summed up to get $\Sigma\, p_1 q_0$.

(iv) Base year prices (p_0) are multiplied by base year quantities (q_0) and summed up to get $\Sigma\, p_0 q_0$.

An Ideal Formula

Index number computed by Fisher's method is known as Fisher's Ideal Index. This is because:

(i) It is based on geometric mean which is the best average considered for computing index numbers.

(ii) This formula provides weighted index number.

(iii) Equal importance is given to base year and current year prices and quantities.

(iv) Fisher's index satisfies many tests of ideal index like - Time and Factor Reversal Tests.

▶ (4) Dorbish and Bowley's Method

Like Fisher, Dorbish and Bowley has assigned weights to different commodities on the basis of current year and base year quantities both. In fact, Dorbish and Bowley's index is the arithmetic average of Laspeyre and Paasche's index. Dorbish and Bowley's formula is as follows:

$$P_{01} = \frac{L + P}{2}$$

Or

$$P_{01} = \frac{\dfrac{\Sigma p_1 q_0}{\Sigma p_0 q_0} + \dfrac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}{2} \times 100$$

The computational process of Dorbish and Bowley's index is exactly alike as Fisher's index.

▶ (5) Marshall-Edgeworth's Method

In this method, the sum of base year and current year quantities are used as weights. Marshall-Edgeworth's formula is as follows:

$$P_{01} = \frac{\Sigma p_1(q_0 + q_1)}{\Sigma p_0(q_0 + q_1)} \times 100$$

Or

$$P_{01} = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

The computational process of Marshall-Edgeworth's Index is same as that of Fisher's Index.

▶ (6) Kelly's Method

In this method, on the basis of fixed quantiy, different commodities are assgined weights. It is not necessary that quantities be related to base year or current year. Kelly's formula is as follows:

$$P_{01} = \frac{\Sigma p_1 q}{\Sigma p_0 q} \times 100$$

Where, q refers to quantity of some period, not necessarily the base year or current year. Kelly's method is also called Fixed weight method.

Different methods of constructing weighted index numbers can be illustrated with the following examples:

Example 5. Construct index number of price from the following data by using:

(*i*) Laspeyre's method

(*ii*) Paasche's method

(*iii*) Fisher's method

(*iv*) Dorbish-Bowley's method

(*v*) Marshall-Edgeworth method

Commodity	1995		1996	
	Price	Quantity	Price	Quantity
A	2	8	4	6
B	5	10	6	5
C	4	14	5	10
D	2	19	2	15

Solution:

Computation of Various Index Nos.

Commodity	1995		1996		L		P	
	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
A	2	8	4	6	16	32	12	24
B	5	10	6	5	50	60	25	30
C	4	14	5	10	56	70	40	50
D	2	19	2	15	38	38	30	30
					$\Sigma p_0q_0 = 160$	$\Sigma p_1q_0 = 200$	$\Sigma p_0q_1 = 107$	$\Sigma p_1q_1 = 134$

(*i*) Laspeyre's Method:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

$$= \frac{200}{160} \times 100 = 125$$

(*ii*) Paasche's Method:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{134}{107} \times 100 = 125.23$$

(*iii*) Fisher's Method:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{200}{160} \times \frac{134}{107}} \times 100 = \sqrt{1.565} \times 100$$

$$= 1.2509 \times 100 = 125.09$$

(*iv*) Bowley's Method:

$$P_{01} = \frac{\dfrac{\Sigma p_1 q_0}{\Sigma p_0 q_0} + \dfrac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}{2} \times 100$$

$$= \frac{\dfrac{200}{160} + \dfrac{134}{107}}{2} \times 100 = \frac{1.25 + 1.252}{2} \times 100$$

$$= \frac{2.502}{2} \times 100 = 125.1$$

(*v*) Marshall-Edgeworth Method:

$$P_{01} = \frac{\Sigma p_1 (q_0 + q_1)}{\Sigma p_0 (q_0 + q_1)} \times 100$$

$$= \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

$$= \frac{200 + 134}{160 + 107} \times 100 = \frac{334}{267} \times 100 = 125.09$$

Example 6. Calculate weighted index number for 1970 using weighted aggregative method from the following data:

Commodity	Quantity	Prices in 1960	Prices in 1970
A	5	10	8
B	4	8	10
C	3	6	9
D	2	5	7

Solution: The quantity consumed neither relates to current year nor to base year. Here quantities are fixed. The appropriate index can be found out by applying Kelly's method.

Commodity	q	p_0	p_1	p_0q	p_1q
A	5	10	8	50	40
B	4	8	10	32	40
C	3	6	9	18	27
D	2	5	7	10	14
				$\Sigma p_0q = 110$	$\Sigma p_1q = 121$

Kelly's Method/Fixed Weight Method

$$P_{01} = \frac{\Sigma p_1 q}{\Sigma p_0 q} \times 100 = \frac{121}{110} \times 100$$

$$= 110$$

IMPORTANT TYPICAL EXAMPLES

Example 7. The table below gives details of price and consumption of 5 commodities for 2005 and 2007. Using an appropriate formula arrive at an index number for 2007 prices with 2005 as base:

Commodity	Price per unit 2005 (Rs.)	Price per unit 2007 (Rs.)	Consumption value 2005 (Rs.)
Rice	40	48	800
Wheat	25	27	400
Oil	95	105	760
Fish	110	120	1100
Milk	80	100	480

Solution: Since, we are given the base year (2005) consumption values (p_0q_0) and current year quantities (q_1) are not given, the appropriate formula for index number is Laspeyre's price index.

Commodity	p_0	q_0	p_1	p_1q_0	p_0q_0
Rice	40	$\frac{800}{40}=20$	48	960	800
Wheat	25	$\frac{400}{25}=16$	27	432	400
Oil	95	$\frac{760}{95}=8$	105	840	760
Fish	110	$\frac{1100}{110}=10$	120	1200	1100
Milk	80	$\frac{480}{80}=6$	100	600	480
Total				4032	3540

$$q_0=\frac{p_0q_0}{p_0}=\frac{\text{Consumption value 2005 (Rs.)}}{\text{Price per cent 2005 (Rs.)}}$$

$$P_{01}=\frac{\Sigma p_1q_0}{\Sigma p_0q_0}\times 100=\frac{4032}{3540}\times 100=113.9$$

Example 8. Given the following data:

Items	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	1	10	2	5
B	1	5	X	2

Find X if the ratio between Laspeyre's and Paasche's index number is L : P :: 28 : 27

Solution:

Items	p_0	q_0	p_1	q_1	$p_0\,q_0$	$p_1\,q_0$	$p_1\,q_1$	$p_0\,q_1$
A	1	10	2	5	10	20	10	5
B	1	5	X	2	5	5X	2X	2
					$\Sigma p_0q_0=15$	$\Sigma p_1q_0=20+5X$	$\Sigma p_1q_1=10+2X$	$\Sigma p_0\,q_1=7$

Laspeyre's Index:

$$L=\frac{\Sigma p_1q_0}{\Sigma p_0q_0}\times 100=\frac{20+5X}{15}\times 100 \quad ...(i)$$

Paasche's Index:

$$P=\frac{\Sigma p_1q_1}{\Sigma p_0q_1}\times 100=\frac{10+2X}{7}\times 100 \quad ...(ii)$$

Dividing (*i*) by (*ii*), we get

$$\frac{L}{P}=\frac{20+5X}{15}\div\frac{10+2X}{7}=\frac{20+5X}{10+2X}\times\frac{7}{15}$$

Given: $\frac{L}{P}=\frac{28}{27}$

$\therefore \quad \frac{28}{27}=\frac{20+5X}{10+2X}\times\frac{7}{15}$

$\Rightarrow \quad \frac{28\times 15}{27\times 7}=\frac{20+5X}{10+2X}$

$\Rightarrow \quad \frac{20}{9}=\frac{20+5X}{10+2X}$

$\Rightarrow \quad 20(10+2X)=9(20+5X)$

$\Rightarrow \quad 200+40X=180+45X$

$\Rightarrow \quad 5X=20$

$\therefore \quad X=4$

Thus, missing figure = 4.

Example 9. From the following data, construct a price index number by using Fisher's Ideal Formula:

Commodity	Base Year		Current Year	
	Price per unit	Expenditure (Rs.)	Price per unit	Expenditure (Rs.)
A	2	40	5	75
B	4	16	8	40
C	1	10	2	24
D	5	25	10	60

Solution: Since we are given the expenditure and the price, we can obtain the quantity by dividing total expenditure by the price for each commodity. We can then apply Fisher's Formula:

Commodities	Base Year		Current Year		$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
	p_0	q_0	p_1	q_1				
A	2	20	5	15	40	100	30	75
B	4	4	8	5	16	32	20	40
C	1	10	2	12	10	20	12	24
D	5	5	10	6	25	50	30	60
					$\Sigma p_0 q_0 = 91$	$\Sigma p_1 q_0 = 202$	$\Sigma p_0 q_1 = 92$	$\Sigma p_1 q_1 = 199$

By Fisher's Formula:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{202}{91} \times \frac{199}{92}} \times 100 = 2.1912 \times 100 = 219.12$$

Example 10. Construct Index Number of prices from the following data by (*i*) Laspeyre's method (*ii*) Paasche's method and (*iii*) Fisher's method.

Commodity	1994		1995	
	Price (Rs.)	Value (Rs.)	Price (Rs.)	Value (Rs.)
A	8	100	10	90
B	10	60	11	66
C	5	100	5	100
D	3	30	2	24
E	2	8	4	20

Solution: Since we are given the total value and the price, we can obtain the quantity figure by dividing the total value by price for each commodity.

Construction of Various Index Numbers

Commodity	1994		1995		$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
	p_0	q_0	p_1	q_1				
A	8	12.5	10	9	100	125	72	90
B	10	6	11	6	60	66	60	66
C	5	20	5	20	100	100	100	100
D	3	10	2	12	30	20	36	24
E	2	4	4	5	8	16	10	20
					$\Sigma p_0 q_0 = 298$	$\Sigma p_1 q_0 = 327$	$\Sigma p_0 q_1 = 278$	$\Sigma p_1 q_1 = 300$

(*i*) **Laspeyre's Method:** $P_{01} = \dfrac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$

$$= \frac{327}{298} \times 100 = 109.73$$

(*ii*) **Paasche's Method:** $P_{01} = \dfrac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$

$$= \frac{300}{278} \times 100 = 107.91$$

(*iii*) **Fisher's Method:** $P_{01} = \sqrt{\dfrac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \dfrac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$

$$= \sqrt{\frac{327}{298} \times \frac{300}{278}} \times 100 = \sqrt{\frac{98100}{82844}} \times 100$$

$$= \sqrt{1.18415} \times 100 = 1.0881 \times 100 = 108.81$$

Example 11. Based on the following data, compute Laspeyre's and Paasche's Price Indices for 1981 and 1982 with 1980 as base:

Items	1980		1981		1982	
	Price (in Rs.)	Qty. (in kg)	Price (in Rs.)	Qty. (in kg)	Price (in Rs.)	Qty. (in kg)
A	2.0	6	4.0	6	4.0	8
B	0.40	40	0.70	40	1.0	36
C	0.50	24	0.20	40	0.25	32

Solution:

Computation of Laspeyre's Price Indices

Items	1980			1981		1982	
	q_0	p_0	$p_0 q_0$	p_1	$p_1 q_0$	p_2	$p_2 q_0$
A	6	2.0	12.0	4.0	24.0	4.0	24.0
B	40	0.40	16.0	0.70	28.0	1.0	40.0
C	24	0.50	12.0	0.20	4.8	0.25	6.0
			$\Sigma p_0 q_0 = 40$		$\Sigma p_1 q_0 = 56.8$		$\Sigma p_2 q_0 = 70$
Index			100		142		175

Laspeyre's Price Index:

$$P_{01}^{L}(1981) = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{56.8}{40} \times 100 = 142$$

$$P_{01}^{L}(1982) = \frac{\Sigma p_2 q_0}{\Sigma p_0 q_0} \times 100 = \frac{70}{40} \times 100 = 175$$

Computation of Paasche's Price Indices

Items	1980			1981				1982			
	q_0	p_0	p_0q_0	q_1	p_0q_1	p_1	p_1q_1	q_2	p_0q_2	p_2	p_2q_2
A	6	2.0	12.0	6	12.0	4.0	24.0	8	16	4	32.0
B	40	0.40	16.0	40	16.0	0.70	28.0	36	14.4	1.00	36
C	24	0.50	12.0	40	20.0	0.20	8.0	32	16.0	0.25	8.0
					$\Sigma p_0q_1 = 48$		$\Sigma p_1q_1 = 60$		$\Sigma p_0q_2 = 46.4$		$\Sigma p_2q_2 = 76$

Paasche's Price Index:

$$P_{01}^{P}(1981) = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{60}{48} \times 100 = 125$$

$$P_{01}^{P}(1982) = \frac{\Sigma p_2 q_2}{\Sigma p_0 q_2} \times 100 = \frac{76}{46.4} \times 100 = 163.79$$

EXERCISE 3.2

1. From the following data, calculate Price Index for 1988 by using:
   (*i*) Laspeyre's Method
   (*ii*) Paasche's Method
   (*iii*) Dorbish and Bowley's Method
   (*iv*) Fisher's Method
   (*v*) Marshall-Edgeworth's Method

Year	A		B		C		D	
	Price	Quantity	Price	Quantity	Price	Quantity	Price	Quantity
1980	24	8	9	3	16	5	10	3
1988	30	10	10	4	20	8	9	4

[Ans. (*i*) 120.67 (*ii*) 120.72 (*iii*) 120.69, (*iv*) 120.7; (*v*) 120.6]

2. From the following data, find the Fisher's Ideal Index No.:

Items	1999		1997	
	Price (Rs.)	Value	Price (Rs.)	Value
A	10	600	6	300
B	2	240	2	200
C	6	360	4	240

[Ans. 143.05]

3. (*i*) Using a suitable formula, construct price index number from the following data:

Commodity	1990		1995	
	Price	Expenditure	Price	Expenditure
A	4	200	10	400
B	3	30	9	18
C	2	10	5	10

[Ans. 253.9]

(*ii*) Find by weighted aggregate method, the index number from the following data:

Commodity	A	B	C	D	E
Price (in Rs.) in 2001:	32	25	90	120	35
Price (in Rs.) in 2005:	50	25	100	140	40
Weight:	8	6	7	3	5

[Hint: $P_{01} = \frac{\Sigma p_1 w}{\Sigma p_0 w}$] **[Ans.** $P_{01} = 119.03$]

4. (*i*) If the ratio between Laspeyre's and Paasche's Index is 52:48, find out the missing figure in the following table:

Commodity	Base Year		Current Year	
	Price	Quantity	Price	Quantity
A	2	10	4	10
B	2	5	?	4

[Ans. 18]

(*ii*) Given that $\Sigma\, p_1\, q_1 = 250$, $\Sigma\, p_0\, q_0 = 150$
Paasche's Index No. = 150, Dorbish-Bowley's Index No. = 145.
Find out (*a*) Fisher's Ideal Index No. and (*b*) Marshal-Edgeworth's Index No.
[Hint: See Example 40] **[Ans.** $\Sigma p_0q_1 = 167$, $\Sigma p_1q_0 = 210$; (*a*) 144.9 (*b*) 145.11]

5. Calculate Price Index Number using weighted aggregative method from the following data:

Commodity	Quantity	Base Year Prices	Current Year Prices
A	5	30	40
B	8	20	30
C	10	10	20

[Hint: Use Kelley's Method] **[Ans.** 156.1]

6. Based on the following data, compute Paasche's Price Index with 1949 as the base:

Items	Prices			Quantities Purchased		
	1949	1950	1960	1949	1950	1960
Foodgrains	5	6	4	500	400	80
Steel	10	8	12	70	60	50
Fruit	4	7	8	100	80	90

[Ans. 117.80, 87.30]

o (II) Weighted Average of Price Relatives Method

In this method, first of all, the price relatives for the current year are calculated on the basis of the base year prices of the commodities. If the weights are given explicitly in the question, then we make use of those weights but if the quantities in the base year (q_0) are given, then by multiplying the base year quantity (q_0) with the base year price (p_0), we compute the value weights (p_0q_0). Whatever be the case, weights (W) are multiplied with the respective price relatives (P) and ΣPW are obtained and then dividing it by the sum of the weights (ΣW). The resulting value gives us the required weighted index number. In terms of formula,

$$\text{Weighted Index Number}\,(P_{01}) = \frac{\Sigma PW}{\Sigma W}, \text{ using A.M.}$$

Where, P = Price relative = $\frac{p_1}{p_0} \times 100$; W = Value in the base year (i.e., p_0q_0) or fixed weights.

If the weighted geometric mean is used, then the formula is

$$P_{01} = \text{Antilog}\left[\frac{\Sigma(\log P).W}{\Sigma W}\right]$$

Note: When the base year's value (p_0q_0) are taken as weights ($W = p_0q_0$), then the weighted average of price relative method gives us Laspeyre's Price Index Number, if arithmetic mean is used.

Example 12. Compute price index from the following data by applying weighted average of price relative method:

Commodity	Base year price (p_0)	Base year quantity (q_0)	Current year price (p_1)
A	6.0	40	8.0
B	3.0	80	3.2
C	2.0	20	3.0

Solution:

Commodity	p_0	q_0	p_1	W (p_0q_0)	$P = \frac{p_1}{p_0} \times 100$	PW
A	6.0	40	8.0	240	$\frac{8}{6} \times 100 = 133.3$	31,992
B	3.0	80	3.2	240	$\frac{3.2}{3.0} \times 100 = 106.7$	25,608
C	2.0	20	3.0	40	$\frac{3}{2} \times 100 = 150.0$	6,000
				ΣW = 520		ΣPW = 63,600

Using Weighted Average of Price Relative Method,

$$P_{01} = \frac{\Sigma PW}{\Sigma W} = \frac{63{,}600}{520} = 122.30$$

o When Prices and Weights of Different Commodities are Given

Example 13. The data in respect of middle class families of a city are given below. Calculate weighted index number for 2006 using (*i*) weighted A.M. of price relatives, (*ii*) weighted G.M. of price relatives:

Items	Weights	2005 Price (Rs.)	2006 Price (Rs.)
Food	30	100	90
Rent	15	20	20
Clothing	20	70	60
Fuel	10	20	15
Misc.	25	40	55

Solution:

Calculation of Weighted Index Numbers

Items	Weights (W)	Price 2005 (p_0)	Price 2006 (p_1)	Price Relatives $P = \frac{p_1}{p_0} \times 100$	WP	log P	W log P
Food	30	100	90	90.0	2700	1.9542	58.626
Rent	15	20	20	100.0	1500	2.0000	30.000
Clothing	20	70	60	85.7	1714	1.9330	38.660
Fuel	10	20	15	75.0	750	1.8751	18.751
Misc.	25	40	55	137.5	3437.5	2.1383	53.457
	ΣW = 100				Σ WP = 10101.5		Σ W log$P$ =199.494

(*i*) $$P_{01}\,(\text{A.M.}) = \frac{\Sigma WP}{\Sigma W} = \frac{10101.5}{100} = 101.015$$

(*ii*) $$P_{01}\,(\text{G.M.}) = \text{Antilog}\left\{\frac{\Sigma W.\log P}{\Sigma W}\right\} = \text{Antilog}\left[\frac{199.494}{100}\right]$$

$$= \text{Antilog}\,[1.9949] = 98.83.$$

Example 14. A department store sells stereo systems, television sets and radios. The percentage distribution of the total sales volume (in rupees) is estimated as 30 per cent stereos, 50 per cent televisions and 20 per cent radios. The price of one stereo, one television and one radio in 1999 was Rs. 20,000, Rs. 15,000 amd Rs. 500 respectively while their respective prices in 2004 were Rs. 25,000, Rs. 20,000 and Rs. 800.

A weighted price index for 2004 with base 1999 is to be computed.

(*i*) Which index number formula is appropriate, why?

(*ii*) Compute the index.

Solution: (*i*) Here, weighted average of price relative method is appropriate because the relative importance or contribution of various products, *i.e.*, stereo systems, televisions and radios, to the total sales volume is to considered.

(ii) **Index Number using Weighted Average of Price Relatives**

Products	Weights (W)	Price in 1999 (p_0)	Price in 2004 (p_1)	$P=\frac{p_1}{p_0}\times 100$	PW
Stereo	30	20,000	25,000	$\frac{25,000}{20,000}\times 100=125$	3,750.00
Television	50	15,000	20,000	$\frac{20,000}{15,000}\times 100=\frac{400}{3}$	6,666.67
Radio	20	500	800	$\frac{800}{500}\times 100=160$	3,200.00
	$\Sigma W = 100$				$\Sigma PW =$ 13,616.67

Hence, the weighted price index for 2004 with base 1999, is given by

$$P_{01}=\frac{\Sigma WP}{\Sigma W}=\frac{13616.67}{100}=136.17$$

Example 15. Calculate index number of prices for 1995 on the basis of 1990 from the data given below:

Commodity	Weight	Price per unit 1990 (Rs.)	Price per unit 1995 (Rs.)
A	40	16	20
B	25	40	50
C	20	12	15
D	15	2	3

If the weights of commodities A, B, C, D are increased in the ratio 1 : 2 : 3 : 4, what will be increase in index number ?

Solution:

Calculations for Index Numbers

Commodity	Weight (W)	Price per unit in Rs. 1990 (p_0)	Price per unit in Rs. 1995 (p_1)	Price Relatives $P=\frac{p_1}{p_0}\times 100$	WP	Increased Weight (W_1)*	W_1P
A	40	16	20	125	5000	$40+\frac{1\times 40}{10}=44$	5500
B	25	40	50	125	3125	$25+\frac{2\times 25}{10}=30$	3750
C	20	12	15	125	2500	$20+\frac{3\times 20}{10}=26$	3250
D	15	2	3	150	2250	$15+\frac{4\times 15}{10}=21$	3150
	$\Sigma W = 100$				$\Sigma WP = 12875$	$\Sigma W_1 = 121$	$\Sigma W_1P = 15650$

$$\text{Original Index Number } (I)=\frac{\Sigma PW}{\Sigma W}=\frac{12875}{100}=128.75$$

Since, the weights of the commodities are increased in the ratio 1 : 2 : 3 : 4, (Total = 10), the increase in weights are :

$$(A):\frac{1}{10}\times 40=4,\quad (B):\frac{2}{10}\times 25=5,\quad (C):\frac{3}{10}\times 20=6,\quad (D):\frac{4}{10}\times 15=6$$

$$\text{New Index Number } (I_1)=\frac{\Sigma W_1 P}{\Sigma W_1}=\frac{15650}{121}=129.34$$

$\therefore$ Increase in the index number $= I_1 - I = 0.59$

When Group Indices and Weights are Given

Example 16. In the construction of a certain cost of living index numbers, the following group index numbers were found. Calculate the cost of living index number by using (i) weighted A.M. and (ii) weighted G.M.

Group	Index Number	Weight
A	352	50
B	200	10
C	230	10
D	160	15
E	190	15

Solution:

Group	Weight (W)	Index Number (P)	WP	$\log P$	$W\cdot\log P$
A	50	352	17600	2.5465	127.3250
B	10	200	2000	2.3010	23.01
C	10	230	2300	2.3617	23.6170
D	15	160	2400	2.2041	33.0615
E	15	190	2850	2.2788	34.1820
	$\Sigma W = 100$		$\Sigma PW = 27150$		$\Sigma W \log P = 241.1955$

(i) $$P_{01}(\text{A.M.})=\frac{\Sigma PW}{\Sigma W}=\frac{27150}{100}=271.50$$

(ii) $$P_{01}(\text{G.M.})=\text{Antilog}\left\{\frac{\Sigma W\log P}{\Sigma W}\right\}=\text{Antilog}\left[\frac{241.1955}{100}\right]$$

$$=\text{Antilog }[2.4119]$$

$$=258.19$$

EXERCISE 3.3

1. Calculate weighted average of price relative index from the following data:

Items	Unit	Base Year Quantity	Base Year Price (Rs.)	Current Year Price (Rs.)
Wheat	Per Qtl.	4 Qtl.	200	250
Sugar	Per Kg.	50 Kgs.	5	7
Milk	Per Litre	40 Litre	5	6
Cloth	Per Metre	20 Metre	10	15
House	Per	1	50	80

[Ans. 131.33]

2. Calculate weighted index number for 2000 by Weighted Aggregative Method and Weighted Average of Relative Method from the following data:

Items	Weights	Prices in 1999 (Rs.)	Prices in 2000 (Rs.)
A	40	16.00	20.00
B	20	40.00	60.00
C	15	0.50	0.50
D	20	5.12	6.25
E	5	2.00	1.50

[Ans. (*i*) $P_{01} = 137.188$, (*ii*) $P_{01} = 123.15$]

3. The price quotation of different commodities for 2001 and 2002 are given below. Calculate the index number for 2002 with 2001 as base year by using (*i*) Simple average of price relatives and (*ii*) Weighted average of price relative.

Commodity	Unit	Weight	Prices (Rs.)	
			2001	2002
A	Kg.	5	2.00	4.50
B	Qt.	7	2.50	3.20
C	Dozen	6	3.00	4.50
D	Kg.	2	1.00	1.80

[Ans. (*i*) 170.5 (*ii*) 164.05]

◘ QUANTITY INDEX NUMBERS

Quantity index numbers are designed to measure the change in physical quantities of goods over a given period. These index numbers represent increase or decrease in physical quantities of goods produced or sold. The method of construction of quantity index is same as that of price index. Just as quantity is taken as weight in case of a price index, similarly, price is taken as weight in case of a quantity index. By interchanging price with quantity and quantity with price in a price index formula, quantity index can be constructed. Quantity index is symbolised as Q_{01}.

Quantity index number can also be simple or weighted:

(A) Simple Quantity Index Numbers

(*i*) **Simple Aggregative Method:** $Q_{01} = \frac{\Sigma q_1}{\Sigma q_0} \times 100$

(*ii*) **Simple Average of Relative Method:** $Q_{01} = \frac{\Sigma\left(\frac{q_1}{q_0} \times 100\right)}{N}$ **(Using AM)**

Simple Average of Relative Method: $Q_{01} = AL\frac{\left[\Sigma \log \frac{q_1}{q_0} \times 100\right]}{N}$ **(Using GM)**

(B) Weighted Quantity Index Numbers

(a) Weighted Aggregate Method

(*i*) **Laspeyre's Quantity Index No.:** $Q_{01} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times 100$

(*ii*) **Paasche's Quantity Index No.:** $Q_{01} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} \times 100$

(*iii*) **Fisher's Quantity Index No.:** $Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} \times 100$

(*iv*) **Bowley's Quantity Index:** $Q_{01} = \frac{\frac{\Sigma q_1 p_0 + \Sigma q_1 p_1}{\Sigma q_0 p_0 + \Sigma q_0 p_1}}{2} \times 100$

(*v*) **Marshall's Quantity Index:** $Q_{01} = \frac{\Sigma q_1 p_0 + \Sigma q_1 p_1}{\Sigma q_0 p_0 + \Sigma q_0 p_1} \times 100$

(b) Weighted Average of Relative Method: $Q_{01} = \frac{\Sigma QW}{\Sigma W}$

Where, $Q = \frac{q_1}{q_0} \times 100$, $W = q_0\ p_0$

Example 17. Compute quantity index for the following by (*i*) Simple Aggregative method (*ii*) Average of Quantity Relative method by using both arithmetic mean and geometric mean:

Items:	A	B	C	D	E	F
Quantity in 1971:	20	30	10	25	40	50
Quantity in 1981:	25	30	15	35	45	55

Solution:

Items	q_0	q_1	$Q = \frac{q_1}{q_0} \times 100$	log Q
A	20	25	125	2.0969
B	30	30	100	2.0000
C	10	15	150	2.1761
D	25	35	140	2.1461
E	40	45	112.5	2.0512
F	50	55	110	2.0414
$N = 6$	$\Sigma q_0 = 175$	$\Sigma q_1 = 205$	$\Sigma \frac{q_0}{q_1} \times 100 = 737.5$	$\Sigma \log Q = 12.5117$

(i) **Simple Aggregative Index**

$$Q_{01} = \frac{\Sigma q_1}{\Sigma q_0} \times 100 = \frac{205}{175} \times 100 = 117.143$$

(ii) **Arithmetic Mean of Quantity Relatives**

$$Q_{01} = \frac{\Sigma\left(\frac{q_1}{q_0} \times 100\right)}{N} = \frac{737.5}{6} = 122.92$$

(iii) **Geometric Mean of Quantity Relatives**

$$Q_{01} = \text{Antilog}\left[\frac{\Sigma \log Q}{N}\right] = \text{Antilog}\left[\frac{12.5117}{6}\right]$$

$$= \text{Antilog}\ [2.0852] = 121.7$$

Example 18. From the following data, construct Quantity Index Number by using *(i)* Laspeyre's formula *(ii)* Paasche's formula and *(iii)* Fisher's Formula:

Commodity	1990		1995	
	Price	Quantity	Price	Quantity
I	8	10	10	11
II	10	9	12	9
III	16	16	20	17

Solution:

Construction of Quantity Index Numbers

Commodity	1990		1995					
	p_0	q_0	p_1	q_1	$q_1 p_0$	$q_0 p_0$	$q_1 p_1$	$q_0 p_1$
I	8	10	10	11	88	80	110	100
II	10	9	12	9	90	90	108	108
III	16	16	20	17	272	256	340	320
					$\Sigma q_1 p_0 = 450$	$\Sigma q_0 p_0 = 426$	$\Sigma q_1 p_1 = 558$	$\Sigma q_0 p_1 = 528$

(i) **Laspeyre's Quantity Index**

$$Q_{01} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times 100 = \frac{450}{426} \times 100 = 105.63$$

(ii) **Paasche's Quantity Index**

$$Q_{01} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} \times 100 = \frac{558}{528} \times 100 = 105.68$$

(iii) **Fisher's Quantity Index**

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} \times 100$$

$$= \sqrt{\frac{450}{426} \times \frac{558}{528}} \times 100 = \sqrt{1.116} \times 100$$

$$= 1.057 \times 100 = 105.7$$

Example 19. Calculate the index number of crime for 2003 with 2002 as base:

	2002	2003	Weights
Robberies	13	8	6
Care thefts	15	22	5
Cycle thefts	249	185	4
Pocket picking	328	259	1
Thefts by servants	497	448	2

Solution:

Index Number of Crime For 2003

	2002	2003	Weights (W)	Crime Relative (R)	RW
Robberies	13	8	6	$\frac{8}{13} \times 100 = 61.54$	369.24
Car thefts	15	22	5	$\frac{22}{15} \times 100 = 146.67$	733.35
Cycle thefts	249	185	4	$\frac{185}{249} \times 100 = 74.30$	297.20
Pocket picking	328	259	1	$\frac{259}{328} \times 100 = 78.96$	78.96
Thefts by servants	497	448	2	$\frac{448}{497} \times 100 = 90.14$	180.28
			$\Sigma W = 18$		$\Sigma RW = 1659.03$

$$\text{Index Number} = \frac{\Sigma RW}{\Sigma W} = \frac{1659.03}{18} = 92.17$$

VALUE INDEX NUMBERS

The value of a commodity is the product of its price and quantity. Therefore, value index is computed by dividing the sum of total value in current year (ΣV_1) by the sum of total value in base year (ΣV_0), and the quotient being multiplied by 100. Its formula is as follows:

$$V_{01} = \frac{\Sigma V_1}{\Sigma V_0} \times 100 = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} \times 100$$

Where V = Value index
ΣV_1 = Total value in current year = $\Sigma p_1 q_1$
ΣV_0 = Total value in base year = $\Sigma p_0 q_0$

The computation of value index is illustrated in the following examples:

Example 20. From the following data, calculate Value Index for the year 1993 and 1994:

Year:	1992	1993	1994
Price (Rs.):	25	30	40
Quantity (Tonnes):	40	50	60

Solution:

1992			1993			1994		
p_0	q_0	$V_0 = p_0 q_0$	p_1	q_1	$V_1 = p_1 q_1$	p_2	q_2	$V_2 = p_2 q_2$
25	40	1000	30	50	1500	40	60	2400

Value Index is given by:

$$V_{01} = \frac{\Sigma V_1}{\Sigma V_0} \times 100$$

For 1993: $$V_{01} = \frac{1500}{1000} \times 100 = 150$$

For 1994: $$V_{01} = \frac{2400}{1000} \times 100 = 240$$

In practice, value indices are not very much used.

EXERCISE 3.4

1. From the following data, calculate (*i*) Price Index by Laspayre's Method, (*ii*) Quantity Index by Fisher, and (*iii*) Value Index:

Commodity	Base Year		Current Year	
	Price (Rs./kg.)	Quantity (kg.)	Price (Rs./kg.)	Quantity (kg.)
A	4	2	6	3
B	5	4	7	6
C	3	6	4	7
D	2	3	3	5

[Hint: $V_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$]

[Ans. (*i*) L = 140.38 (*ii*) F = 140.74 and (*iii*) V_{01} = 198.08]

2. Calculate quantity index numbers from the data given below using (*i*) Laspeyre's method (*ii*) Paasche's method (*iii*) Bowley's method (*iv*) Fisher's method (*v*) Marshall's method (*vi*) Simple average of relatives method.

Commodity	1980		1985	
	Price	Quantity	Price	Quantity
A	4	10	5	12
B	6	8	7	10
C	10	5	12	4
D	3	12	4	4
E	5	7	5	8

[**Ans.** (*i*) 95.70 (*ii*) 93.98 (*iii*) 94.84 (*iv*) 94.83 (*v*) 94.76 (*vi*) 94.52]

3. Compute Quantity Index Number for the following data by (*i*) Simple Aggregative Method, (*ii*) Average of Quantity relative method by using A.M.:

Commodity:	A	B	C	D	E	F
Production (1989)	20	30	10	25	40	50
Production (1999)	25	30	15	35	45	55

[**Ans.** (*i*) 117.43 (*ii*) 122.92]

TESTS OF ADEQUACY OF INDEX NUMBER FORMULAE

Various formulae can be used for the construction of Index numbers. But it is necessary to select an appropriate/suitable formula out of them. Prof. Fisher has given the following tests to select an appropriate formula:

(1) Time Reversal Test - TRT

(2) Factor Reversal Test - FRT

(3) Circular Test.

(1) Time Reversal Test

Illustrating this test, **Prof. Fisher** remarks **"The test is that the formula for calculating an index number should be such that it will give the same ratio between one point of comparison and the other, no matter which of the two is taken as base."** In other words, according to this test, if considering any year as base year, some other year's price index is computed and for another price index, time subscripts are reversed, then the both price indices must be reciprocal to each other, *i.e.*, their product will yield unit.

Time Reversal Test is satisfied when

$$P_{01} = \frac{1}{P_{10}} \text{ or } P_{01} \times P_{10} = 1$$

Where, P_{01} is the price index for the year 1 with 0 as base and P_{10} is the price index for the year 0 with 1 as base.

Time Reversal Test is not satisfied by simple A.M. of price relative, Laspeyre and Paasche's formula. But **this test is satisfied by Fisher's Ideal Index, Marshall-Edgeworth Index and Simple G.M. of Price Relative.**

(i) **Laspeyre's Formula:** According to Laspeyre's Formula (omitting factor 100)

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \quad ...(i)$$

Interchanging time subscripts, *i.e.*, 0 to 1 and 1 to 0

$$P_{10} = \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \quad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times P_{10} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \neq 1$$

Since $P_{01} \times P_{10} \neq 1$, the Laspeyre's formula does not satisfy time reversal test.

(ii) **Paasche's Formula:** According to Paasche's Formula (omitting factor 100)

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \quad ...(i)$$

Interchanging time subscripts, *i.e.*, 0 to 1 and 1 to 0

$$P_{10} = \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0} \quad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times P_{10} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0} \neq 1$$

Since $P_{01} \times P_{10} \neq 1$, the Paasche's formula does not satisfy time reversal test.

(iii) **Fisher's Formula:** According to Fisher's Formula (omitting factor 100)

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \quad ...(i)$$

Interchanging 0 to 1 and 1 to 0

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} \quad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times P_{10} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}}$$

$$= \sqrt{1} = 1$$

Since $P_{01} \times P_{10} = 1$, the Fisher's Formula satisfies time reversal test.

(iv) **Marshall-Edgeworth Formula:** According to Marshall-Edgeworth Formula:

$$P_{01} = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1}$$

Interchanging 0 to 1 and 1 to 0

$$P_{10} = \frac{\Sigma p_0 q_1 + \Sigma p_0 q_0}{\Sigma p_1 q_1 + \Sigma p_1 q_0}$$

$$\therefore \quad P_{01} \times P_{10} = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times \frac{\Sigma p_0 q_1 + \Sigma p_0 q_0}{\Sigma p_1 q_1 + \Sigma p_1 q_0} = 1$$

Since $P_{01} \times P_{10} = 1$, Marshall-Edgeworth satisfies time reversal test.

(2) Factor Reversal Test

This is another test given by Prof. Fisher. According to **Prof. Fisher "Just as our formula should permit the interchange of two times without giving inconsisting results, so it ought to permit interchange of prices and quantities without giving inconsistent results, *i.e.*, the two results multiplied together should give the true value ratio."**

Factor Reversal Test is satisfied when

Price Index × Quantity Index = Value Index

Or

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Like time reversal test, Factor Reversal Test too is not satisfied by Laspeyre's and Paasche's formulae. **Marshall-Edgeworth's formula too does not satisfy factor reversal test. Factor reversal test is satisfied only by Fisher's Ideal formula.** This is shown below:

(i) **Laspeyre's Formula:** According to Laspeyre's Formula (omitting factor 100)

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \quad ...(i)$$

Interchanging *p* to *q* and *q* to *p*

$$Q_{01} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \quad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since $P_{01} \times Q_{01} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, the Laspeyre's formula does not satisfy factor reversal test.

***(ii)* Paasche's Formula:** According to Paasche's Formula (omitting factor 100)

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \qquad ...(i)$$

Interchanging p to q and q to p,

$$Q_{01} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} \qquad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since, $P_{01} \times Q_{01} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, the Paasche's formula does not satisfy factor reversal test.

***(iii)* Fisher's Formula:** According to Fisher's Formula (omitting factor 100)

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \qquad ...(i)$$

Interchanging p to q and q to p

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} \qquad ...(ii)$$

Multiplying (*i*) and (*ii*), we get

$$P_{01} \times Q_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$$

$$= \sqrt{\frac{\Sigma p_1 q_1 \times \Sigma\, p_1 q_1}{\Sigma p_0 q_0 \times \Sigma\, p_0 q_0}} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since, $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, Fisher's formula satisfies factor reversal test.

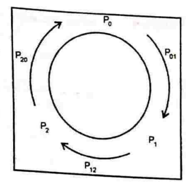
***(iv)* Marshall-Edgeworth Formula:** According to Marshall-Edgeworth Formula:

$$P_{01} = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1}; \; Q_{01} = \frac{\Sigma q_1 p_0 + \Sigma q_1 p_1}{\Sigma q_0 p_0 + \Sigma q_0 p_1}$$

$$\therefore \quad P_{01} \times Q_{01} = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times \frac{\Sigma q_1 p_0 + \Sigma q_1 p_1}{\Sigma q_0 p_0 + \Sigma q_0 p_1} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since, $P_{01} \times Q_{01} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, the test is not satisfied by **Marshall-Edgeworth formula.**

(3) Circular Test



This test is the extension of time reversal test. According to this test, if there are three time periods 0, 1 and 2 and price index of period 1 relative to 0 (P_{01}), price index of period 2 relative to period 1 (P_{12}) and price index of period 0 relative to period 2 (P_{20}) are computed, then the product of these three indices must be unity. Symbolically

$$\mathbf{P_{01} \times P_{12} \times P_{20} = 1}$$

This test is satisfied by the simple aggregative index only.

Circular test is not satisfied by Laspeyre's, Paasche's and Fisher's index number. **Index number based on simple G.M., simple aggregative formula and weighted aggregative formula (with fixed weights) satisfy circular test.**

Example 21. Show that Fisher's formula does not satisfy circular test.

Solution: The circular test is satisfied when

$$P_{01} \times P_{12} \times P_{20} = 1$$

According to Fisher's Ideal Index,

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \qquad ...(i)$$

Interchanging time 0 to 1 and 1 to 2

$$P_{10} = \sqrt{\frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_2 q_2}{\Sigma p_1 q_2}} \qquad ...(ii)$$

Again, interchanging 1 to 2 and 2 to 0

$$P_{20} = \sqrt{\frac{\Sigma p_0 q_2}{\Sigma p_2 q_2} \times \frac{\Sigma p_0 q_0}{\Sigma p_2 q_0}} \qquad ...(iii)$$

Multiplying (*i*), (*ii*) and (*iii*), we have

$$P_{01} \times P_{12} \times P_{20} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times \sqrt{\frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_2 q_2}{\Sigma p_1 q_2}} \times \sqrt{\frac{\Sigma p_0 q_2}{\Sigma p_2 q_2} \times \frac{\Sigma p_0 q_0}{\Sigma p_2 q_0}} \neq 1$$

Since $P_{01} \times P_{12} \times P_{20} \neq 1$, the Fisher's formula does not satisfy circular test.

Example 22. Calculate Fisher's Ideal Index from the following data and show that it satisfies both the time reversal and factor reversal tests:

Commodity	1983		1984	
	Price	**Expenditure**	**Price**	**Expenditure**
A	8	80	10	120
B	10	120	12	96
C	5	40	5	50
D	4	56	3	60
E	20	100	25	150

Solution: Since we are given the expenditure and the price, we can obtain quantity figure by dividing total expenditure by price for each commodity. We can then apply Fisher's Ideal Formula:

Calculation of Fisher's Ideal Index

Commodity	1983		1984					
	p_0	q_0	p_1	q_1	$p_1 q_0$	$p_0 q_0$	$p_1 q_1$	$p_0 q_1$
A	8	10	10	12	100	80	120	96
B	10	12	12	8	144	120	96	80
C	5	8	5	10	40	40	50	50
D	4	14	3	20	42	56	60	80
E	20	5	25	6	125	100	150	120
					$\Sigma p_1q_0 = 451$	$\Sigma p_0q_0 = 396$	$\Sigma p_1q_1 = 476$	$\Sigma p_0q_1 = 426$

By Fisher's Formula:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{451}{396} \times \frac{476}{426}} \times 100 = \sqrt{\frac{214676}{168696}} \times 100$$

$$= \sqrt{1.2726} \times 100 = 112.8$$

Time Reversal Test: $P_{01} \times P_{10} = 1$

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \text{ and } P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}}$$

$$= \sqrt{\frac{451}{396} \times \frac{476}{426} \times \frac{426}{476} \times \frac{396}{451}} = \sqrt{1} = 1$$

Hence, time reversal test is satisfied.

Factor Reversal Test:

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Here,

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \text{ and } Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$$

$$= \sqrt{\frac{451}{396} \times \frac{476}{426} \times \frac{426}{396} \times \frac{476}{451}} = \sqrt{\frac{476 \times 476}{396 \times 396}} = \frac{476}{396} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Hence, the factor reversal test is satisfied.

IMPORTANT TYPICAL EXAMPLES

Example 23. Calculate Price Index for the year 1996 from the following data. Use geometric mean:

Commodity	Average Price 1990 (Rs.)	Average Price 1996 (Rs.)
A	16.1	14.2
B	9.2	8.7
C	15.1	12.5
D	5.6	4.8
E	11.7	13.4
F	100	117

Now reverse the base (taking 1996 as base) and show that the two results are strictly consistent.

Solution: (*i*) Calculationof index number for the year 1996 with 1990 as base year using Geometric Mean of Price Relatives.

Commodity	Average Price 1990 (Rs.) p_0	Average Price 1996 (Rs.) p_1	Price Relatives $\left(\frac{p_1}{p_0} \times 100\right) = P$	log P
A	16.1	14.2	$\frac{14.2}{16.1} \times 100 = 88.20$	1.9455
B	9.2	8.7	$\frac{8.7}{9.2} \times 100 = 94.57$	1.9757
C	15.1	12.5	$\frac{12.5}{15.1} \times 100 = 82.78$	1.9179
D	5.6	4.8	$\frac{4.8}{5.6} \times 100 = 85.71$	1.9331
E	11.7	13.4	$\frac{13.4}{11.7} \times 100 = 114.53$	2.0589
F	100	117	$\frac{117}{100} \times 100 = 117.00$	2.0682
N = 6				$\Sigma \log P = 11.8993$

$$P_{01} \text{ for } 1996 = \text{Antilog}\left(\frac{\Sigma \log P}{N}\right) = \text{Antilog}\left(\frac{11.8993}{6}\right) = \text{Antilog}\,(1.9832) = 96.20$$

(*ii*) Calculation of index number for the year 1990 with 1996 as base year:

Commodity	Average Price 1996 (Rs.) p_0	Average Price 1990 (Rs.) p_1	Price Relatives $\left(\frac{p_1}{p_0} \times 100\right) = P$	log P
A	14.2	16.1	$\frac{16.1}{14.2} \times 100 = 113.38$	2.0545
B	8.7	9.2	$\frac{9.2}{8.7} \times 100 = 105.75$	2.0242
C	12.5	15.1	$\frac{15.1}{12.5} \times 100 = 120.80$	2.0820
D	4.8	5.6	$\frac{5.6}{4.8} \times 100 = 116.67$	2.0669
E	13.4	11.7	$\frac{11.7}{13.4} \times 100 = 87.31$	1.9410
F	117	100	$\frac{100}{117} \times 100 = 85.47$	1.9318
N = 6				$\Sigma \log P = 12.1004$

$$P_{10} \text{ for } 1990 = \text{Antilog}\left(\frac{\Sigma \log P}{N}\right)$$

$$= \text{Antilog}\left(\frac{12.1004}{6}\right)$$

$$= \text{Antilog}\,(2.0167) = 103.928.$$

Two results are strictly consistent as they satisfy the time reversal test *i.e.*

$$P_{01} \times P_{10} = \frac{96.2}{100} \times \frac{103.928}{100} = 1 \text{ [approx.]}$$

Note: If the simple index numbers are computed for the same data relating to two periods using G.M. but with the bases reversed, then the product of the two index number should be equal to unity. This implies that the index number calculates on the basis of G.M. satisfies TRT.

Example 24. Show with the help of the following data the index number calculated on the basis of A.M. does not satisfy the circular test, whereas that by geometric mean method of averaging satisfies it.

Commodity	Price		
	2000	2001	2003
A	20	30	40
B	30	36	45
C	20	30	50
D	12	15	30

Solution: Circular test is satisfied when

$$P_{01} \times P_{12} \times P_{20} = 1$$

Calculation of P_{01}, P_{12} and P_{20}

Commodity	Price			Price Relatives					
	2000	2001	2002	p_{01}	p_{12}	p_{20}	log p_{01}	log p_{12}	log p_{20}
A	20	30	40	150.0	133.3	50.0	2.1761	2.1249	1.6990
B	30	36	45	120.0	125.0	66.7	2.0792	2.0969	1.8241
C	20	30	50	150.0	166.7	40.0	2.1761	2.2219	1.6021
D	12	15	30	125.0	200.0	40.0	2.0969	2.3010	1.6021
N = 4				$\Sigma p_{01} = 545.0$	$\Sigma p_{12} = 625$	$\Sigma p_{20} = 196.7$	8.5283	8.7447	6.7273

Index numbers calculated on the basis of A.M.

$$P_{01} = \frac{545}{4} = 136.25$$

$$P_{01} = \frac{625}{4} = 156.25$$

$$P_{01} = \frac{196.7}{4} = 49.18$$

$$P_{01} \times P_{12} \times P_{20} = \frac{136.25}{100} \times \frac{156.25}{100} \times \frac{49.18}{100} \qquad \text{(Omitting 100)}$$

$$= 1.3625 \times 1.5625 \times 0.4918 \neq 1$$

Thus, the index numbers based on simple arithmetic mean do not satisfy circular test.

Index numbers calculated on the basis of G.M.

$$P_{01} = \text{Antilog}\left[\frac{8.5283}{4}\right] = \text{Antilog}\,[2.1321] = 135.5$$

$$P_{12} = \text{Antilog}\left[\frac{8.7447}{4}\right] = \text{Antilog}\,[2.1862] = 153.6$$

$$P_{20} = \text{Antilog}\left[\frac{6.7273}{4}\right] = \text{Antilog}\,[1.6818] = 48.06$$

$$P_{01} \times P_{12} \times P_{20} = \frac{135.5}{100} \times \frac{153.6}{100} \times \frac{48.06}{100} \qquad \text{(Omitting 100)}$$

$$= 1.355 \times 1.536 \times 0.4806$$

$$= 1$$

Thus, Index numbers based on G.M. satisfies the circular test.

EXERCISE 3.5

1. Calculate Laspeyre's, Paasche's, Marshall-Edgeworth's and Fisher's price index numbers for the following data:

Items	1980		1985	
	Price (Rs)	Expenditure	Price (Rs)	Expenditure
A	5	50	6	72
B	7	84	10	80
C	10	80	12	96
D	4	20	5	30
E	8	56	8	64

Which of them satisfies time reversal and factor reversal tests.

[Ans. 123.10, 120.42, 121.77, 121.96; TRT: Fisher and Marshall; FRT: Fisher]

2. From the following data, calculate quantity index numbers using:
(*i*) Laspeyre's (*ii*) Paasche's and (*iii*) Fisher's methods.

Commodity	1980		1985	
	Price	Value	Price	Value
A	50	350	60	540
B	20	80	30	150
C	24	240	20	300
D	100	600	150	600

Which of the above method satisfy factor reversal test?

[Ans. 103.14, 96.95, 99.99 approx.; Fisher's formula satisfy FRT]

3. With the help of the following data, show that the index number calculated on the basis of A.M. is not reversible while the Index number calculated on the basis of G.M. is reversible. Make comparison between AM and GM.

Commodity	Price in 1998	Price in 1999
A	40	60
B	50	80
C	20	40
D	20	10

[Ans. $P_{01} = 140$, $P_{10} = 94.79$ (using A.M.)
$P_{01} = 124.45$, $P_{10} = 80.334$ (using G.M.)**]**

4. Following are the values:
$\Sigma p_0 q_0 = 425$, $\Sigma p_1 q_0 = 505$, $\Sigma p_1 q_1 = 530$, $\Sigma p_0 q_1 = 470$
Show that Fisher's method, Paasche's method and Marshall method either satisfy time reversal test and factor reversal test or do not satisfy both or one of them.

[Ans. TRT: Fisher and Marshall, Paasche's does not satisfy any one, FRT: Fisher
Marshall does not satisfy FRT**]**

MISCELLANEOUS SOLVED EXAMPLES

Example 25. Compute a price index for the following by (*i*) simple aggregative method (*ii*) average of price relatives method by using both arithmetic mean and geometric mean:

Commodity	A	B	C	D	E	F
Price in 1971 (Rs.):	20	30	10	25	40	50
Price in 1981 (Rs.):	25	30	15	35	45	55

Solution:

Calculation of Price Index

Commodity	p_0	p_1	$P = \frac{p_1}{p_0} \times 100$	log P
A	20	25	125	2.0969
B	30	30	100	2.0000
C	10	15	150	2.1761
D	25	35	140	2.1461
E	40	45	112.5	2.0512
F	50	55	110	2.0414
N = 5	$\Sigma p_0 = 175$	$\Sigma p_1 = 205$	$\Sigma P = 737.5$	$\Sigma \log P = 12.5117$

(*i*) Index number by using Simple Aggregative Method:

$$P_{01} = \frac{\Sigma p_1}{\Sigma p_0} \times 100 = \frac{205}{175} \times 100 = 117.14$$

(*ii*) Index number by using Average of Price Relatives:

(*a*) Using A.M. $P_{01} = \dfrac{\Sigma\left(\dfrac{p_1}{p_0} \times 100\right)}{N} = \dfrac{\Sigma P}{N} = \dfrac{737.5}{6} = 122.92$

(*b*) Using G.M. $P_{01} = \text{Antilog}\left[\dfrac{\Sigma \log P}{N}\right] = \text{Antilog}\left[\dfrac{12.5117}{6}\right]$

$= \text{Antilog}\ [2.0853] = 121.7$

Example 26. Prepare Index Number of prices for three years with average price as base from the data given below by simple average of price relative method using A.M.

Year	Rate per rupee Commodities		
	A	B	C
I	4 kg	2 kg	1 kg
II	2.5 kg	1.6 kg.	1 kg
III	2 kg	1.25 kg.	0.8 kg

Solution: Since we are given quantity prices (*i.e.*, rate per rupee), first we convert these into money prices (*i.e.*, rate per quintal) as shown below:

	First Year	Second Year	Third Year
A	$\frac{100}{4}$ = Rs. 25 per Qtl	$\frac{100}{2.5}$ = Rs. 40 per Qtl	$\frac{100}{2}$ = Rs. 50 per Qtl
B	$\frac{100}{2}$ =Rs. 50 per Qtl	$\frac{100}{1.6}$ =Rs. 62.5 per Qtl	$\frac{100}{1.25}$ =Rs. 80 per Qtl
C	$\frac{100}{1}$ =Rs. 100 per Qtl	$\frac{100}{1}$ =Rs. 100 per Qtl	$\frac{100}{0.8}$ =Rs. 125 per Qtl

Then we determine the average price as follows:

$$\text{Average Price of A} = \frac{25+40+50}{3} = 38\cdot 3$$

$$\text{Average Price of B} = \frac{50+62.5+80}{3} = 64\cdot 2$$

$$\text{Average price of C} = \frac{100+100+125}{3} = 108\cdot 3$$

With average price as base, we compute the price relative (P) and then find the averages:

Commodity	Unit	Average price (100) Base	First year		Second year		Third year	
			Price	P	Price	P	Price	P
A	100 kg	38.3	25	65.3	40	104.4	50	130.5
B	100 kg	64.2	50	77.9	62.5	97.4	80	124.6
C	100 kg	108.3	100	92.3	100	92.3	125	115.4
Total of Relatives				ΣP = 235.5		ΣP = 294.1		ΣP = 370.5
Average of Relatives $\left(\frac{\Sigma P}{N}\right)$				78.5		98.03		123.5

Thus, Index No. for Ist year = 78.5

2nd year = 98.03

3rd year = 123.5

Example 27. Find out index number by using Mean, Median and G.M.:

Group	1985	1986	1987
A	6	12	24
B	9	15	30
C	15	21	36
D	21	27	42
E	24	36	54

Solution:

Construction of Index Number Using Different Averages

Group	1985		1986			1987		
	Price (p_0)	Price Relatives (P)	Price (p_1)	Price Relatives (P)	log P	Price (p_2)	Price Relatives (P)	log P
A	6	100	12	200.00	2.3010	24	400.00	2.6021
B	9	100	15	166.67	2.2217	30	333.33	2.5228
C	15	100	21	140.00	2.1461	36	240.00	2.3802
D	21	100	27	128.57	2.1089	42	200.00	2.3010
E	24	100	36	150.00	2.1761	54	225.00	2.3522
Total	—	500	—	785.24	10.9538	—	1398.33	12.1583
A.M. of Price Relatives (ΣP/N)	—	100	—	157.05	—	—	279.67	—
Median of Price Relatives $(M = \frac{N+1}{2})$	—	100	—	150.00	—	—	240.00	—
G.M. of Price Relatives [Antilog $(\Sigma \frac{\log P}{N})$]	—	100	—	—	155.20	—	—	270.20

Example 28. The price paid and quantities purchased by a household in base and current years are given below. Calculate the additional dearness allowance to be given to the household so as to fully compensate it for the price rise, using both the Laspeyre's and Paasche's index numbers.

Commodity	Base Period		Current Period	
	Price	Quantity	Price	Quantity
A	30	10	40	8
B	12	20	15	18

Solution: Laspeyre's Index Number

$$= \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

$$= \frac{40\times10+20\times15}{30\times10+12\times20}\times100 = \frac{400+300}{300+240}\times100$$

$$= \frac{700}{540}\times100 = 129.63$$

Additional dearness allowance to be paid = 29.63%

Paasche's index number

$$= \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{40 \times 8 + 15 \times 18}{30 \times 8 + 12 \times 18} \times 100 = \frac{320 + 270}{240 + 216} \times 100$$

$$= \frac{590}{456} \times 100 = 129.386$$

Additional dearness allowance to be paid = 29.386%

Example 29. Calculate by suitable method, the index number of price from the following data:

Commodity	1977		1987	
	Price	Quantity	Price	Quantity
A	8	10	10	11
B	10	9	12	9
C	16	16	20	17

Solution: Since we are given the base year and current year price and quantity, Fisher's Ideal Index shall be the most suitable.

Commodity	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
A	8	10	10	11	80	100	88	110
B	10	9	12	9	90	108	90	108
C	16	16	20	17	256	320	272	340
					$\Sigma p_0 q_0 = 426$	$\Sigma p_1 q_0 = 528$	$\Sigma p_0 q_1 = 450$	$\Sigma p_1 q_1 = 558$

Fisher's Ideal Formula is given by

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{528}{426} \times \frac{558}{450}} \times 100 = \sqrt{\frac{294624}{191700}} \times 100$$

$$= \sqrt{1.53690} \times 100 = 1.2397 \times 100 = 123.97$$

Example 30. Calculate Laspeyre's, Paasche's and Fisher's Ideal Index for the following data:

Commodity	1970		1990	
	Price	Expenditure	Price	Expenditure
A	8	100	10	90
B	10	60	11	66
C	5	100	5	100
D	3	30	2	24
E	2	8	4	20

With the help of these data, show which of the above index number satisfies time and factor reversal test.

Solution: Since we are given the expenditure and the price, we can obtain the quantity figure by dividing total expenditure by price for each commodity. We can then apply the formulae.

Commodity	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
A	8	12.5	10	9	100	125	72	90
B	10	6	11	6	60	66	60	66
C	5	20	5	20	100	100	100	100
D	3	10	2	12	30	20	36	24
E	2	4	4	5	8	16	10	20
					$\Sigma p_0 q_0 = 298$	$\Sigma p_1 q_0 = 327$	$\Sigma p_0 q_1 = 278$	$\Sigma p_1 q_1 = 300$

Laspeyre's Method:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{327}{298} \times 100 = 109.73$$

Paasche's Method:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{300}{278} \times 100 = 107.91$$

Fisher's Method:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{327}{298} \times \frac{300}{278}} \times 100 = \sqrt{\frac{98100}{82844}} \times 100$$

$$= \sqrt{1.18415} \times 100$$

$$= 1.0881 \times 100 = 108.81$$

Time Reversal Test

Time reversal test is satisfied when $P_{01} \times P_{10} = 1$

Laspeyre's Index No.

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{327}{298}$$

$$P_{10} = \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} = \frac{278}{300}$$

$$\therefore \quad P_{01} \times P_{10} = \frac{327}{298} \times \frac{278}{300} \neq 1$$

Thus, Laspeyre's Index does not satisfy TRT.

Paasche's Index No.

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{300}{278}$$

$$P_{10} = \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0} = \frac{298}{327}$$

$$\therefore \quad P_{01} \times P_{10} = \frac{300}{278} \times \frac{298}{327} \neq 1.$$

Thus, Paasche's Index does not satisfy TRT.

Fisher's Ideal Index:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{327}{298} \times \frac{300}{278}}$$

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = \sqrt{\frac{278}{300} \times \frac{298}{327}}$$

$$\therefore \quad P_{01} \times P_{10} = \sqrt{\frac{327}{298} \times \frac{300}{278} \times \frac{278}{300} \times \frac{298}{327}} = 1$$

Thus, Fisher's Ideal Index satisfies TRT.

Factor Reversal Test

Factor reversal test is satisfied when $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$.

Laspeyre's Index

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{327}{298}$$

$$Q_{01} = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} = \frac{278}{298}$$

$$\therefore \quad P_{01} \times Q_{01} = \frac{327}{298} \times \frac{278}{298} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Thus, Laspeyre's Index does not satisfy FRT.

Paasche's Index

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{300}{278}$$

$$Q_{01} = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} = \frac{300}{327}$$

$$P_{01} \times Q_{01} = \frac{300}{278} \times \frac{300}{327} \neq \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}.$$

Thus, Paasche's Index does not satisfy FRT.

Fisher's Ideal Index

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{327}{298} \times \frac{300}{278}}$$

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = \sqrt{\frac{278}{298} \times \frac{300}{327}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{327}{298} \times \frac{300}{278} \times \frac{278}{298} \times \frac{300}{327}}$$

$$= \frac{300}{298} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since, $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, Fisher's Ideal Index satisfies FRT.

Example 31. It is stated that Marshall-Edgeworth index is a good approximation to the ideal index number. Verify using the following data:

Commodity	1970		1990	
	Price	Quantity	Price	Quantity
A	5	100	6	50
B	4	80	5	100
C	2.5	60	5	72
D	12.0	30	9	33

Solution:

Commodity	1970		1990					
	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_0 q_1$	$p_1 q_0$	$p_1 q_1$
A	5	100	6	50	500	250	600	300
B	4	80	5	100	320	400	400	500
C	2.5	60	5	72	150	180	300	360
D	12.0	30	9	33	360	396	270	297
					$\Sigma p_0 q_0$ =1330	$\Sigma p_0 q_1$ =1226	$\Sigma p_1 q_0$ =1570	$\Sigma p_1 q_1$ =1457

Fisher's Index:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{1570}{1330} \times \frac{1457}{1226}} \times 100 = 118.44$$

Marshall-Edgeworth Index:

$$P_{01} = \frac{\Sigma p_1 (q_0 + q_1)}{\Sigma p_0 (q_0 + q_1)} \times 100 = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

$$= \frac{1570 + 1457}{1330 + 1226} \times 100 = 118.42$$

The above calculations clearly show that the answer obtained by the Fisher's method and Marshall-Edgeworth's method is the same.

Example 32. Using suitable formula construct the price index number from the following data:

Commodity	1990		1995	
	Price	Expenditure	Price	Expenditure
A	1.0	60.00	1.25	62.50
B	1.50	37.50	2.50	50.00
C	2.00	20.00	3.00	30.00
D	12.00	36.00	18.00	72.00
E	0.10	4.00	0.15	9.00

Check whether it satisfies time reversal and factor reversal test.

Solution: Since we are given the base year and current year price and expenditure, so Fisher's Ideal Formula shall be most suitable.

Commodity	p_0	q_0	p_1	q_1	$p_0 q_0$	$p_1 q_0$	$p_0 q_1$	$p_1 q_1$
A	1.0	60	1.25	50	60	75	50.0	62.50
B	1.50	25	2.50	20	37.50	62.5	30.0	50.00
C	2.00	10	3.00	10	20.00	30.0	20.0	30.00
D	12.0	3	18.00	4	36.00	54.0	48.0	72.00
E	0.10	40	0.15	60	4.00	6.0	6.0	9.00
					$\Sigma p_0 q_0 = 157.5$	$\Sigma p_1 q_0 = 227.5$	$\Sigma p_0 q_1 = 154$	$\Sigma p_1 q_1 = 223.50$

Fisher's Ideal Index:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{227.5}{157.5} \times \frac{223.5}{154}} \times 100 = \sqrt{\frac{50846.25}{24255}} \times 100$$

$$= \sqrt{2.0963} \times 100 = 144.8$$

Time Reversal Test

TRT is said to be satisfied if $P_{01} \times P_{10} = 1$

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{227.5}{157.5} \times \frac{223.5}{154}}$$

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = \sqrt{\frac{154}{223.5} \times \frac{157.5}{227.5}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{227.5}{157.5} \times \frac{223.5}{154} \times \frac{154}{223.5} \times \frac{157.5}{227.5}} = 1$$

Thus, Fisher's formula satisfies time reversal test.

Factor Reversal Test

FRT is said to be satisfied if $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{227.5}{157.5} \times \frac{223.5}{154}}$$

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = \sqrt{\frac{154}{157.5} \times \frac{223.5}{227.5}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{227.5}{157.5} \times \frac{223.5}{154} \times \frac{154}{157.5} \times \frac{223.5}{227.5}}$$

$$\therefore \quad P_{01} \times Q_{01} = \frac{223.5}{157.5} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

As $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$, Fisher's Ideal Formula satisfies factor reversal test.

Example 33. From the data given below, show that Laspeyre's, Paasche's and Fisher's Index numbers do not satisfy the circular test:

Commodity	1995		1996		1997	
	Price	Quantity	Price	Quantity	Price	Quantity
A	1	7	4	13	5	10
B	2	6	9	7	8	4
C	4	8	11	4	10	2

Solution:

Commodity	p_0	q_0	p_1	q_1	p_2	q_2	p_0q_0	p_0q_1	p_0q_2	p_1q_0	p_1q_1	p_1q_2	p_2q_0	p_2q_1	p_2q_2
A	1	7	4	13	5	10	7	13	10	28	52	40	35	65	50
B	2	6	9	7	8	4	12	14	8	54	63	36	48	56	32
C	4	8	11	4	10	2	32	16	8	88	44	22	80	40	20
							51	43	26	170	159	98	163	161	102

Laspeyre's Index Number

$P_{01} \times P_{12} \times P_{20} = 1$ (omitting factor 100 from each index)

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{170}{51} \quad \text{...(i)}$$

$$P_{12} = \frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} = \frac{161}{159} \quad \text{...(ii)}$$

$$P_{20} = \frac{\Sigma p_0 q_2}{\Sigma p_2 q_2} = \frac{26}{102} \quad \text{...(iii)}$$

$$\therefore \quad P_{01} \times P_{12} \times P_{20} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_2}{\Sigma p_2 q_2} = \frac{170}{51} \times \frac{161}{159} \times \frac{26}{102} \neq 1$$

Thus, Laspeyre's index does not satisfy circular test.

Paasche's Index Number

$P_{01} \times P_{12} \times P_{20} = 1$ (omitting factor 100 from each index)

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{159}{43} \quad \text{...(i)}$$

$$P_{12} = \frac{\Sigma p_2 q_2}{\Sigma p_1 q_2} = \frac{102}{98} \quad \text{...(ii)}$$

$$P_{20} = \frac{\Sigma p_0 q_0}{\Sigma p_2 q_0} = \frac{51}{163} \quad \text{...(iii)}$$

$$\therefore \quad P_{01} \times P_{12} \times P_{20} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times \frac{\Sigma p_2 q_2}{\Sigma p_1 q_2} \times \frac{\Sigma p_0 q_0}{\Sigma p_2 q_0} = \frac{159}{43} \times \frac{102}{98} \times \frac{51}{163} \neq 1$$

Thus, Paasche's index does not satisfy circular test.

Fisher's Ideal Index Number

$P_{01} \times P_{12} \times P_{20} = 1$ (omitting factor 100 from each index)

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{170}{51} \times \frac{159}{43}} \quad \text{...(i)}$$

$$P_{12} = \sqrt{\frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_2 q_2}{\Sigma p_1 q_2}} = \sqrt{\frac{161}{159} \times \frac{102}{98}} \quad \text{...(ii)}$$

$$P_{20} = \sqrt{\frac{\Sigma p_0 q_2}{\Sigma p_2 q_2} \times \frac{\Sigma p_0 q_0}{\Sigma p_2 q_0}} = \sqrt{\frac{26}{102} \times \frac{51}{163}} \quad \text{...(iii)}$$

$$P_{01} \times P_{12} \times P_{20} = \sqrt{\frac{170}{51} \times \frac{159}{43} \times \frac{161}{159} \times \frac{102}{98} \times \frac{26}{102} \times \frac{51}{163}} \neq 1$$

$\therefore$ Thus, Fisher's ideal index does not satisfy the circular test. Hence, we conclude that none of these satisfy the circular test.

Example 34. Calculate Fisher's Ideal Index number from given data. Does it satisfy the time reversal and factor reversal test.

Commodity	1995		1996	
	Price	Quantity	Price	Quantity
A	6	50	10	56
B	2	100	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

Solution:

Commodity	p_0	q_0	p_1	q_1	p_0q_0	p_1q_0	p_0q_1	p_1q_1
A	6	50	10	56	300	500	336	560
B	2	100	2	120	200	200	240	240
C	4	60	6	60	240	360	240	360
D	10	30	12	24	300	360	240	288
E	8	40	12	36	320	480	288	432
					Σp_0q_0 =1360	Σp_1q_0 =1900	Σp_0q_1 =1344	Σp_1q_1 =1880

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{1900}{1360} \times \frac{1880}{1344}} \times 100$$

$$= \sqrt{1.397 \times 1.398} \times 100 = \sqrt{1.953} \times 100 = 139.75$$

Time Reversal Test

TRT is said to be satisfied if $P_{01} \times P_{10} = 1$

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{1900}{1360} \times \frac{1880}{1344}}$$

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = \sqrt{\frac{1344}{1880} \times \frac{1360}{1900}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{1900}{1360} \times \frac{1880}{1344} \times \frac{1344}{1880} \times \frac{1360}{1900}}$$

$$= 1$$

Thus, Fisher's Ideal Formula satisfies time reversal test.

Factor Reversal Test

FRT is said to be satisfied if $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{1900}{1360} \times \frac{1880}{1344}}$$

$$Q_{01} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = \sqrt{\frac{1344}{1360} \times \frac{1880}{1900}}$$

$$P_{01} \times Q_{01} = \sqrt{\frac{1880}{1344} \times \frac{1900}{1360} \times \frac{1344}{1360} \times \frac{1880}{1900}}$$

$$\therefore \quad P_{01} \times Q_{01} = \frac{1880}{1360} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Since $P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$. Fisher's ideal formula satisfies factor reversal test.

Example 35. Construct a cost of living index number from the following price relatives for the year 1985 and 1986 with 1982 as base giving weightage to the following groups in the proportion of 30, 8, 6, 4 and 2 respectively:

Group	1982	1985	1986
Food	100	114	116
Rent	100	115	125
Clothing	100	108	110
Fuel	100	105	104
Misc.	100	102	104

Solution:

Construction of Cost of Living Index Number

Group	Price relatives for 1982 (P_0)	Price relatives for 1985 (P_1)	Price relatives for 1986 (P_2)	W	P_1W	P_2W
Food	100	114	116	30	3420	3480
Rent	100	115	125	8	920	1000
Clothing	100	108	110	6	648	660
Fuel	100	105	104	4	420	416
Misc.	100	102	104	2	204	208
				ΣW=50	ΣP_1W =5612	ΣP_2W =5764

$$\text{Index Number for 1985} = \frac{\Sigma P_1 W}{\Sigma W} = \frac{5612}{50} = 112.24$$

$$\text{Index Number for 1986} = \frac{\Sigma P_2 W}{\Sigma W} = \frac{5764}{50} = 115.28$$

Example 36. Calculate the index number of prices for 1972 on the basis of 1971 from the data given below:

Commodity	Weights	Price/Unit in 1971 (Rs.)	Price/Unit in 1972 (Rs.)
Rice	40	16.00	20.00
Wheat	25	40.00	60.00
Linseed	5	0.50	0.50
Gur	20	5.12	6.25
Tobacco	10	2.00	1.50

Solution:

Commodity	W	p_0	p_1	$P = \frac{p_1}{p_0} \times 100$	PW
Rice	40	16	20	125	5000
Wheat	25	40	60	150	3750
Linseed	5	0.50	0.50	100	500
Gur	20	5.12	6.25	122	2440
Tobacco	10	2.00	1.50	75	750
	$\Sigma W = 100$				$\Sigma PW = 12440$

$$P_{01} = \frac{\Sigma PW}{\Sigma W} = \frac{12440}{100} = 124.40$$

Example 37. Calculate Laspeyre's, Paasche's, Fisher's Ideal and Marshall-Edgeworth index for the data:

Commodity	1993		1994	
	Price	Expenditure	Price	Expenditure
A	8	80	10	120
B	10	120	12	96
C	5	40	5	50
D	4	56	3	60
E	20	100	25	150

With the help of these data, show which of the above index satisfies time and factor reversal tests.

Solution: Since we are given the price and expenditure, we can obtain quantity figure by dividing total expenditure by price for each commodity.

Commodity	p_0	q_0	p_1	q_1	$p_1 q_0$	$p_0 q_0$	$p_1 q_1$	$p_0 q_1$
A	8	10	10	12	100	80	120	96
B	10	12	12	8	144	120	96	80
C	5	8	5	10	40	40	50	50
D	4	14	3	20	42	56	60	80
E	20	5	25	6	125	100	150	120
					$\Sigma p_1 q_0 = 451$	$\Sigma p_0 q_0 = 396$	$\Sigma p_1 q_1 = 476$	$\Sigma p_0 q_1 = 426$

(i) Laspeyre's Index:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{451}{396} \times 100 = 113.88$$

(ii) Paasche's Index:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100 = \frac{476}{426} \times 100 = 111.74$$

(iii) Fisher's Ideal Index:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{451}{396} \times \frac{476}{426}} \times 100$$

$$= \sqrt{1.1388 \times 1.1174} \times 100 = 112.80$$

(iv) Marshall-Edgeworth Index:

$$P_{01} = \frac{\Sigma p_1 (q_0 + q_1)}{\Sigma p_0 (q_0 + q_1)} \times 100$$

$$= \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

$$= \frac{451 + 476}{396 + 426} \times 100 = \frac{927}{822} \times 100 = 112.77$$

Time Reversal Test:

$$P_{01} \times P_{10} = 1$$

(i) Laspeyre's Method:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{451}{396}$$

$$P_{10} = \frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} = \frac{426}{476}$$

$$\Rightarrow \quad P_{01} \times P_{10} = \frac{451}{396} \times \frac{426}{476} \neq 1$$

Time Reversal Test is not satisfied by Laspeyre's Method.

(ii) Paasche's Method:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{476}{426}$$

$$P_{10} = \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0} = \frac{396}{451}$$

$$\Rightarrow \quad P_{01} \times P_{10} = \frac{476}{426} \times \frac{396}{451} \neq 1$$

Time Reversal Test is not satisfied by Paasche's Method.

(iii) Fisher's Method:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = \sqrt{\frac{451}{396} \times \frac{476}{426}} \quad \text{...(i)}$$

$$P_{10} = \sqrt{\frac{\Sigma p_0 q_1}{\Sigma p_1 q_1} \times \frac{\Sigma p_0 q_0}{\Sigma p_1 q_0}} = \sqrt{\frac{426}{476} \times \frac{396}{451}} \quad \text{...(ii)}$$

$$\Rightarrow \quad P_{01} \times P_{10} = \sqrt{\frac{451}{396} \times \frac{476}{426} \times \frac{426}{476} \times \frac{396}{451}} = \sqrt{1} = 1$$

$\therefore$ Time Reversal Test is satisfied by Fisher's Method.

(iv) **Marshall-Edgeworth Method:**

$$P_{01}=\frac{\Sigma p_1(q_0+q_1)}{\Sigma p_0(q_0+q_1)}=\frac{\Sigma p_1q_0+\Sigma p_1q_1}{\Sigma p_0q_0+\Sigma p_0q_1}$$

$$=\frac{451+476}{396+426}=\frac{927}{822}$$

$$P_{10}=\frac{\Sigma p_0q_1+\Sigma p_0q_0}{\Sigma p_1q_1+\Sigma p_1q_0}=\frac{426+396}{476+451}=\frac{822}{927}$$

$$\Rightarrow\quad P_{01}\times P_{10}=\frac{927}{822}\times\frac{822}{927}=1$$

Time Reversal Test is satisfied by Marshall-Edgeworth Method.

Factor Reversal Test:

$$P_{01}\times Q_{01}=\frac{\Sigma p_1q_1}{\Sigma p_0q_0}$$

(i) **Laspeyre's Method:**

$$P_{01}=\frac{\Sigma p_1q_0}{\Sigma p_0q_0}=\frac{451}{396}$$

$$Q_{01}=\frac{\Sigma q_1p_0}{\Sigma q_0p_0}=\frac{426}{396}$$

$$\Rightarrow\quad P_{01}\times Q_{01}=\frac{451}{396}\times\frac{426}{396}\neq\frac{\Sigma p_1q_1}{\Sigma p_0q_0}$$

Factor Reversal Test is not satisfied by Laspeyre's formula.

(ii) **Paasche's Method:**

$$P_{01}=\frac{\Sigma p_1q_1}{\Sigma p_0q_1}=\frac{476}{426}$$

$$Q_{01}=\frac{\Sigma q_1p_1}{\Sigma q_0p_1}=\frac{476}{451}$$

$$\Rightarrow\quad P_{01}\times Q_{01}=\frac{476}{426}\times\frac{476}{451}\neq\frac{\Sigma p_1q_1}{\Sigma p_0q_0}$$

Factor Reversal Test is not satisfied by Paasche's formula.

(iii) **Fisher's Method:**

$$P_{01}=\sqrt{\frac{\Sigma p_1q_0}{\Sigma p_0q_0}\times\frac{\Sigma p_1q_1}{\Sigma p_0q_1}}=\sqrt{\frac{451}{396}\times\frac{476}{426}}$$

$$Q_{01}=\sqrt{\frac{\Sigma q_1p_0}{\Sigma q_0p_0}\times\frac{\Sigma q_1p_1}{\Sigma q_0p_1}}=\sqrt{\frac{426}{396}\times\frac{476}{451}}$$

$$\Rightarrow\quad P_{01}\times Q_{01}=\sqrt{\frac{451}{396}\times\frac{476}{426}\times\frac{426}{396}\times\frac{476}{451}}$$

$$=\sqrt{\frac{476\times 476}{396\times 396}}=\frac{476}{396}=\frac{\Sigma p_1q_1}{\Sigma p_0q_0}$$

Factor Reversal Test is satisfied by Fisher's formula.

(iv) **Marshall-Edgeworth Method:**

$$P_{01}=\frac{\Sigma p_1q_0+\Sigma p_1q_1}{\Sigma p_0q_0+\Sigma p_0q_1}$$

$$=\frac{451+476}{396+426}=\frac{927}{822}$$

$$Q_{01}=\frac{\Sigma q_1p_0+\Sigma q_1p_1}{\Sigma q_0p_0+\Sigma q_0p_1}=\frac{426+476}{396+451}=\frac{902}{847}$$

$$\Rightarrow\quad P_{01}\times Q_{01}=\frac{927}{822}\times\frac{902}{847}\neq\frac{\Sigma p_1q_1}{\Sigma p_0q_0}$$

Factor Reversal Test is not satisfied by Marshall-Edgeworth formula.

Example 38. From the following data, calculate:

(i) Price Index (P_{01}) using Marshall-Edgeworth Formula.

(ii) Quantity Index (Q_{01}) using Bowley's Formula.

Items	Current year		Base year	
	Price (Rs.)	Value (Rs.)	Value (Rs.)	Quantity (Rs.)
A	20	200	360	12
B	4	36	64	16
C	14	238	575	23

Solution:

Items	Base Year		Current Year					
	p_0	q_0	p_1	q_1	p_0q_0	p_1q_0	p_0q_1	p_1q_1
A	30	12	20	10	360	240	300	200
B	4	16	4	9	64	64	36	36
C	25	23	14	17	575	322	425	238
Total					Σp_0q_0 = 999	Σp_1q_0 = 626	Σp_0q_1 = 761	Σp_1q_1 =474

(i) Marshall-Edgeworth's Price Index

$$P_{01} = \frac{\Sigma p_1(q_0+q_1)}{\Sigma p_0(q_0+q_1)} \times 100 = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

$$= \frac{626+474}{999+761} \times 100 = \frac{1100}{1760} \times 100 = 62.50$$

(ii) Bowley's Quantity Index (Q_{01})

$$Q_{01} = \frac{L+P}{2} = \frac{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} + \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}{2} \times 100$$

$$= \frac{\frac{761}{999} + \frac{474}{626}}{2} \times 100 = 75.96 \text{ or } 76.00 \text{ (approx.)}$$

Example 39. From the following data, calculate weighted index number for 1997 with 1995 as the base year by using weighted geometric mean:

Group	Weight	Price in 1995	Price in 1997
A	5	2.00	4.50
B	7	2.50	3.20
C	6	3.00	4.50
D	2	1.00	1.80

Solution:

Group	Price in 1995 (p_0)	Price in 1997 (p_1)	P $\left(\frac{p_1}{p_0} \times 100\right)$	log P	W	W.log P
A	2.00	4.50	$\frac{4.50}{2.00} \times 100 = 225$	2.3522	5	11.7610
B	2.50	3.20	$\frac{3.20}{2.50} \times 100 = 128$	2.1072	7	14.7504
C	3.00	4.50	$\frac{4.50}{3.00} \times 100 = 150$	2.1761	6	13.0566
D	1.00	1.80	$\frac{1.80}{1.00} \times 100 = 180$	2.2553	2	4.5106
					$\Sigma W = 20$	$\Sigma W.\log P = 44.0786$

Cost of Living Index Number based on G.M.

$$= \text{Antilog}\left[\frac{\Sigma W \log P}{\Sigma W}\right] = \text{Antilog}\left[\frac{44.0786}{20}\right]$$

$$= \text{Antilog } (2.2039) = 159.9$$

Example 40. Given that $\Sigma p_1 q_1 = 250$, $\Sigma p_0 q_0 = 150$, Paasche's Index Number $=150$, Dorbish-Bowley's Index Number $= 145$.

Find out: Fisher's Ideal Index Number, and Marshall-Edgeworth Index Number.

Solution: Given, $\Sigma p_1 q_1 = 250$, $\Sigma p_0 q_0 = 150$

Paasche's Index Number $= 150$, Dorbish-Bowley Index Number $= 145$

$$P_{01}^{P} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$$

Substituting the given values,

$$150 = \frac{250}{\Sigma p_0 q_1} \times 100$$

or $$\Sigma p_0 q_1 = \frac{25000}{150} = \frac{500}{3} = 167 \text{ (approx.)}$$

And $$P_{01}^{DB} = \frac{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} + \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}{2} \times 100$$

Substituting the given values,

or $$145 = \left[\frac{\Sigma p_1 q_0}{150} + \frac{250}{167}\right] 50$$

or $$\frac{145}{50} = \frac{\Sigma p_1 q_0}{150} + 1.50$$

or $$2.90 - 1.50 = \frac{\Sigma p_1 q_0}{150}$$

$\therefore$ $$1.40 = \frac{\Sigma p_1 q_0}{150}$$

or $$\Sigma p_1 q_0 = 1.40 \times 150$$

$$= 210 \text{ (approx.)}$$

So now,

Fisher's Index Number

$$= \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

$$= \sqrt{\frac{210}{150} \times \frac{250}{167}} \times 100$$

$$= \sqrt{1.40 \times 1.50} \times 100$$

$$= 144.91$$

Marshall-Edgeworth Index Number

$$= \frac{\Sigma p_1 (q_0 + q_1)}{\Sigma p_0 (q_0 + q_1)} \times 100 = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

$$= \frac{210 + 250}{150 + 167} \times 100 = \frac{460}{317} \times 100$$

$$= 145.110 \approx 145$$

IMPORTANT FORMULAE

1. Simple Index Numbers:

(*i*) Simple Aggregative Method:

$$P_{01} = \frac{\Sigma p_1}{\Sigma p_0} \times 100$$

(*ii*) Simple Average of Price Relative Method

(a) $P_{01} = \dfrac{\frac{\Sigma p_1}{p_0} \times 100}{N}$ (using A.M.)

(b) $P_{01} = \text{Antilog}\left[\dfrac{\Sigma \log(p_1/p_0)}{N} \times 100\right]$ using G.M.

2. Weighted Index Numbers:

(a) Weighted Aggregate Method

(*i*) Laspeyre's Method:

$$P_{01} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

(*ii*) Paasche's Method:

$$P_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$$

(*iii*) Fisher's Method:

$$P_{01} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$$

(*iv*) Dorbish-Bowley's Method:

$$P_{01} = \frac{L + P}{2} \text{ or } \frac{1}{2}\left(\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} + \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}\right)$$

(*v*) Marshall-Edgeworth Method:

$$P_{01} = \frac{\Sigma p_1 (q_0 + q_1)}{\Sigma p_0 (q_0 + q_1)} \times 100$$

(*vi*) Kelly's Method:

$$P_{01} = \frac{\Sigma p_1 q}{\Sigma p_0 q} \times 100$$

(b) Weighted Average of Price Relative Method

(*i*) If A.M. is used $P_{01} = \dfrac{\Sigma PW}{\Sigma W}$

(*ii*) If G.M. is used $P_{01} = \text{Antilog}\left[\dfrac{\Sigma(W \log P)}{\Sigma W}\right]$

3. Tests of Adequacy:

Time reversal test is satisfied when:

$$P_{01} \times P_{10} = 1$$

Factor reversal test is satisfied when:

$$P_{01} \times Q_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$$

Circular test is satisfied when:

$$P_{01} \times P_{12} \times P_{20} = 1$$

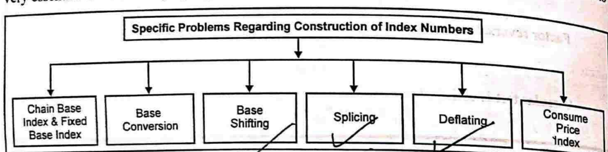
QUESTIONS

1. What are index numbers? Explain the various problems faced in the construction of an index number. What is the utility of index number?
2. Explain the uses and limitations of index numbers.
3. Explain the Laspeyre's, Paasche's and Fisher's formula for computing an index number. Check which of them satisfies the time reversal and factor reversal tests.
4. Discuss the various tests of adequacy of Index Number formulae.
5. Explain the various methods of constructing Index Numbers.
6. Explain (*i*) Time Reversal Test (*ii*) Factor Reversal Test (*iii*) Circular Test. Indicate whether Laspeyre's, Paasche's and Fisher's Ideal Index Numbers satisfy one or other tests.
7. Explain the concept of quantity and value indices.
8. (*i*) Differentiate between Weighted and Unweighted Index Numbers.
   (*ii*) Explain various formulae for calculating index numbers.
   (*iii*) Discuss tests for index numbers.
9. What is Fisher's Ideal Index? Why is it called ideal? Show that it satisfies both the time reversal test as well as factor reversal test.
10. "Index numbers are economic barometers". Explain the statement. What precautions will you take while constructing an index number?

4 Index Numbers–II

◘ INTRODUCTION

In the previous chapter, we have studied the concept of index number, the problems in its construction, methods of constructing them and the tests of an ideal index number. During the process of index numbers' construction, some specific problems like construction of chain index, base shifting, splicing, deflating, construction of a consumer price index, etc., come across us. It is very essential to take them properly into consideration.



Here below, we will study such specific problems:

◘ CHAIN BASE INDEX NUMBERS

Chain base index is that index number in which the year immediately preceeding the one is taken as base year. For example, suppose, we want to construct index numbers for 1990, 91, 92, 93 and we take 1990 as base for 1991, 1991 as base year for 1992 and 1992 as base for 1993, then such type of index is called chain base index.

○ Steps in Construction of Chain Base Index

(*i*) First of all, link relatives are computed using the following formula:

$$\text{Link Relatives} = \frac{\textbf{Current Year's Price}}{\textbf{Previous Year's Price}} \times 100$$

(*ii*) The link relatives are then converted into chain base index using the following formula:

$$\text{Chain Base Index} = \frac{\textbf{Link Relatives of Current Year} \times \textbf{Chain Index of Previous Year}}{\textbf{100}}$$

The construction of chain indices can be illustrated with the following examples:

Example 1. Construct Chain Base Index from the following data:

Year:	1985	1986	1987	1988	1989	1990
Prices:	94	98	102	95	98	100

Solution:

Computation of CBI

Years	Prices	Link Relatives	Chain Base Index
1985	94	100	100
1986	98	$\frac{98}{94} \times 100 = 104.3$	$\frac{104.3 \times 100}{100} = 104.3$
1987	102	$\frac{102}{98} \times 100 = 104.1$	$\frac{104.1 \times 104.3}{100} = 108.6$
1988	95	$\frac{95}{102} \times 100 = 93.1$	$\frac{93.1 \times 108.6}{100} = 101.1$
1989	98	$\frac{98}{95} \times 100 = 103.2$	$\frac{103.2 \times 101.1}{100} = 104.3$
1990	100	$\frac{100}{98} \times 100 = 102$	$\frac{102 \times 104.3}{100} = 106.4$

Example 2. Construct Chain Base Index from the link relatives given below:

Year:	1969	1970	1971	1972	1973
Link Relatives:	100	105	95	175	102

Solution:

Computation of CBI

Year	Link Relatives	Chain Base Index
1969	100	100
1970	105	$\frac{105 \times 100}{100} = 105$
1971	95	$\frac{95 \times 105}{100} = 99.75$
1972	175	$\frac{175 \times 99.75}{100} = 174.56$
1973	102	$\frac{102 \times 174.56}{100} = 178.05$

Example 3. Construct Chain Base Index from the following data:

Group	1970	1971	1972	1973
I	2	3	4	5
II	8	10	12	15
III	4	5	8	10

Solution:

Computation of Chain Base Index Chained to 1970

Group of Commodities	1970 P	1970 L.R.	1971 P	1971 L.R.	1972 P	1972 L.R.	1973 P	1973 L.R.
I	2	100	3	150	4	133.33	5	125
II	8	100	10	125	12	120.00	15	125
III	4	100	5	125	8	160.00	10	125
Total		300		400		413.33		375
Average of L.R.		100		133.33		137.78		125
Chain Base Index		100		$\frac{133.33\times100}{100}$ =133.33		$\frac{137.78\times133.33}{100}$ =183.70		$\frac{125\times183.70}{100}$ =229.62

■ FIXED BASE INDEX

In fixed base index, the base year remains fixed. Fixed Base Index is an index in which the base year is the fixed year. For example, suppose, we want to construct indices for 1990, 91, 92, 93 and we take 1990 alone as base year for all the years, then such type of index is called as fixed base index.

○ Steps in Construction of Fixed Base Index

Fixed base index is computed in the following manner:

(*i*) First of all, the price relatives are found out using the following formula:

$$\text{Price Relatives} = \frac{\text{Current Year's Price}}{\text{Base Year's Price}} \times 100$$

(*ii*) If it is a single commodity case, then these price relatives would be FBIs. But, if it is a multi-commodity case, then by summing up the price relative for each year and dividing the sum by the number of commodity, average of price relatives are found out. These averages are fixed base indices. The method of constructing fixed base index is illustrated by the following examples:

Example 4. Construct index numbers for the following data by taking (a) price of 1975 as base and (b) average of all the prices as base.

Year:	1975	1976	1977	1978	1979	1980	1981	1982
Price:	110	120	160	150	180	200	220	210

Solution:

Calculation of FBI

Year	Price	Index (1975 =100)	Index Base = 168.75*
1975	110	100	$\frac{110}{168.75}\times100=65.2$
1976	120	$\frac{120}{110}\times100=109.1$	$\frac{120}{168.75}\times100=71.1$
1977	160	$\frac{160}{110}\times100=145.5$	$\frac{160}{168.75}\times100=94.8$
1978	150	$\frac{150}{110}\times100=136.4$	$\frac{150}{168.75}\times100=88.9$
1979	180	$\frac{180}{110}\times100=163.6$	$\frac{180}{168.75}\times100=106.7$
1980	200	$\frac{200}{110}\times100=181.8$	$\frac{200}{168.75}\times100=118.5$
1981	220	$\frac{220}{110}\times100=200.0$	$\frac{220}{168.75}\times100=130.4$
1982	210	$\frac{210}{110}\times100=190.9$	$\frac{210}{168.75}\times100=124.4$

$$*\text{Average of all the given prices} = \frac{110+120+160+150+180+200+220+210}{8} = 168.75$$

Example 5. Find out the fixed base index number for 1986, 1987 and 1988 based on 1985.

Commodity	Price in (Rs.) 1985	Price in (Rs.) 1986	Price in (Rs.) 1987	Price in (Rs.) 1988
A	15	30	20	24
B	10	12	16	13
C	12	18	8	10
D	8	8	12	16
E	17	15	16	20

Solution:

Computation of Fixed Base Index Numbers

Commodity	p_0	1985 Price Relatives	p_1	1986 Price Relatives $\frac{p_1}{p_0}\times100$	p_2	1987 Price Relatives $\frac{p_2}{p_0}\times100$	p_3	1988 Price Relatives $\frac{p_3}{p_0}\times100$
A	15	100	30	200.0	20	133.3	24	160.0
B	10	100	12	120.0	16	160.0	13	130.0
C	12	100	18	150.0	8	66.7	10	83.3
D	8	100	8	100.0	12	150.0	16	200.0
E	17	100	15	88.2	16	94.1	20	117.7
Total of Relatives		500		658.2		604.1		691.0

$$\text{Price Index for 1986} = \frac{\Sigma \frac{p_1}{p_0}\times100}{N} = \frac{658.2}{5} = 131.64$$

$$\text{Price Index for 1987} = \frac{\Sigma \frac{p_2}{p_0}\times100}{N} = \frac{604.1}{5} = 120.82$$

$$\text{Price Index for 1988} = \frac{\Sigma \frac{p_3}{p_0}\times100}{N} = \frac{691.0}{5} = 138.2$$

IMPORTANT TYPICAL EXAMPLES

Example 6. From the data given below, calculate the Chain Base Index Numbers.

Year:	1992	1993	1994	1995	1996	1997	1998
Price:	31	22	28	24	30	27	25

Verify that the CBI will be the same as FBI with 1992 as base.

Solution:

Calculation of CBI and FBI

Year	Price	Link relatives	CBI	FBI (Base = 1992)
1992	31	100	100	100
1993	22	$\frac{22}{31}\times 100 = 70.96$	$\frac{100\times 70.96}{100} = 70.96$	$\frac{22}{31}\times 100 = 70.96$
1994	28	$\frac{28}{22}\times 100 = 127.27$	$\frac{127.27\times 70.96}{100} = 90.30$	$\frac{28}{31}\times 100 = 90.3$
1995	24	$\frac{24}{28}\times 100 = 85.71$	$\frac{85.71\times 90.3}{100} = 77.40$	$\frac{24}{31}\times 100 = 77.4$
1996	30	$\frac{30}{24}\times 100 = 125.00$	$\frac{125\times 77.40}{100} = 96.75$	$\frac{30}{31}\times 100 = 96.77$
1997	27	$\frac{27}{30}\times 100 = 90.00$	$\frac{90\times 96.75}{100} = 87.07$	$\frac{27}{31}\times 100 = 87.09$
1998	25	$\frac{25}{27}\times 100 = 92.59$	$\frac{92.59\times 87.07}{100} = 80.61$	$\frac{25}{31}\times 100 = 80.64$

Note: FBI and CBI calculated on the basis of original prices are equal (*i.e.*, same) for a single commodity case. For a multi commodity case, CBI are almost equal to FBI. The slight difference that appear between them are due to approximations made in the calculations.

Example 7. Calculate the fixed base index number and chain base index number from the following data. Are the two results same? If not, why?

Commodity	Price (in rupees)				
	1986	1987	1988	1989	1990
X	2	3	5	7	8
Y	8	10	12	14	18
Z	4	5	7	9	12

Solution: Since base year is not specified, the first year in order of time, *i.e.*, 1986 is taken as base.

(*i*) **Fixed Base Index Numbers (Base year = 1986)**

Commodity	Prices				
	1986 (PR)	1987 (PR)	1988 (PR)	1989 (PR)	1990 (PR)
X	100	$\frac{3}{2}\times 100 = 150$	$\frac{5}{2}\times 100 = 250$	$\frac{7}{2}\times 100 = 350$	$\frac{8}{2}\times 100 = 400$
Y	100	$\frac{10}{8}\times 100 = 125$	$\frac{12}{8}\times 100 = 150$	$\frac{14}{8}\times 100 = 175$	$\frac{18}{8}\times 100 = 225$
Z	100	$\frac{5}{4}\times 100 = 125$	$\frac{7}{4}\times 100 = 175$	$\frac{9}{4}\times 100 = 225$	$\frac{12}{4}\times 100 = 300$
Total of PR	300	400	575	750	925
Average of P.R., *i.e.*, Fixed Base I.No.	100	133.33	191.67	250	308.33

(*ii*) **Chain Base Index Numbers Chained to 1986**

Commodity	1986		1987		1988		1989		1990	
	p_0	LR	p_1	LR	p_2	LR	p_3	LR	p_4	LR
X	2	100	3	150	5	166.70	7	140	8	114.29
Y	8	100	10	125	12	120.00	14	116.67	18	128.57
Z	4	100	5	125	7	140.00	9	128.57	12	133.33
Total of LR		300		400		426.7		385.24		376.19
Average		100		133.33		142.23		128.41		125.40
Chain Indices		100		$\frac{133.33\times 100}{100} = 133.33$		$\frac{133.33\times 142.23}{100} = 189.63$		$\frac{189.63\times 128.41}{100} = 243.50$		$\frac{243.50\times 125.40}{100} = 305.34$

From the above, it is clear that the index numbers obtained by both methods are the same for the first two years and they are different for the remaining years. This is due to the average (combining) of the values for different commodities.

Relative Merits and Demerits of Chain Base Index and Fixed Base Index

(1) Under FBI, base year remains fixed, and all further years are compared on the basis of same fixed year whereas under CBI, base year shifts every year and each year is compared with immediately preceding year as the base.

(2) FBI are constructed on the basis of price relatives whereas CBI are constructed on the basis of link relatives.

(3) FBI indicates the long-term tendency of data whereas CBI depicts the short-term tendency of the data.

(4) Items included in FBI can't be shifted or changed whereas in CBI, every year items or commodities can be shifted or changed.

(5) FBI is easier to compute whereas CBI is difficult to compute.

EXERCISE 4.1

1. Construct Chain Base Index from the following data:

Year:	1979	1980	1981	1982	1983	1984
Prices:	240	300	288	360	480	420

Verify that the CBI will be the same as FBI with 1979 as base.

[**Ans.** 100, 125, 120, 150, 200, 175]

2. Prepare index numbers from the link relatives given below:

Year:	1979	1980	1981	1982	1983	1984	1985
Link relatives:	105	75	71	105	95	90	90

[**Ans.** 100, 75, 53.25, 55.91, 53.11, 47.79, 43.01]

3. The average wholesale prices of three groups of commodities for the year 1995 to 1999 are given below. Compute Chain Base Index number with 1995 as base:

Groups	1995	1996	1997	1998	1999
I	6	9	15	21	24
II	24	30	36	42	54
III	12	15	21	27	36

[**Ans.** 100, 133.3, 189.6, 243.4, 305.2]

4. From the following data, calculate Index Numbers:
   (*i*) taking prices of 1992 as base year,
   (*ii*) taking prices of 1994 as base year,
   (*iii*) taking average prices of five years as base.

Year:	1992	1993	1994	1995	1996
Prices (in Rs):	10	12	15	16	20

[**Ans.** (*i*) 100, 120, 150, 160, 200
(*ii*) 66.67, 80, 100, 106.67, 133.33
(*iii*) 68.49, 82.19, 102.74, 109.59, 136.99]

5. Calculate the fixed base index number and chain base index number from the following data. Are the two results same? If not, why?

Commodity	Price (in Rs.)				
	1992	1993	1994	1995	1996
I	2	3	5	7	8
II	8	10	12	14	18
III	4	5	7	9	12

[**Ans.** FBI: 100, 133.3, 191.67, 250, 308.33
CBI: 100, 133.33, 189.63, 243.49, 305.34]

6. Construct Fixed base and Chain base index numbers from the following data. Are the two results same?

Year	Production	Year	Production
1973	75	1981	120
1974	81	1982	123
1975	90	1983	108
1976	72	1984	96
1977	84	1985	111
1978	87	1986	114
1979	93	1987	117
1980	105	1988	120

[**Ans.** 100, 108, 120, 96, 112, 116, 124, 140, 160, 164, 144, 128, 148, 152, 156, 160
For a single series of index numbers, FBI and CBI are equal.]

BASE CONVERSION

Sometimes, a necessity arises to convert chain base index into fixed base index and vice-versa. The following procedure is taken up for conversion:

(1) Conversion of Chain Base Index into Fixed Base Index.

Its procedure is as follows:

(*i*) For the first year, the FBI will be taken the same as CBI. But if in a question it has been asked to take the first year as a base, the FBI for the first year will then be taken as 100.

(*ii*) For successive years, fixed base indices are computed from chain base index by the following formula:

$$\textbf{Current year's FBI} = \frac{\textbf{Current year's CBI} \times \textbf{Previous year's FBI}}{\textbf{100}}$$

Example 8. From the chain base index number given below, prepare fixed base index numbers:

Year:	1983	1984	1985	1986	1987
CBI:	138	153	156	147	195

Solution:

Conversion of CBI into FBI

Year	CBI	Conversion	FBI
1983	138	—	138
1984	153	$\frac{153 \times 138}{100}$	211.14
1985	156	$\frac{156 \times 211.14}{100}$	329.38
1986	147	$\frac{147 \times 329.38}{100}$	484.19
1987	195	$\frac{195 \times 484.19}{100}$	944.17

(2) Conversion of Fixed Base Index to Chain Base Index.

Its procedure is as follows:

(*i*) For the first year, CBI will be taken the same as FBI.

(*ii*) For successive years, chain base Indices are derived from fixed base Index by the following formula:

$$\text{Current year's CBI} = \frac{\text{Current year's FBI}}{\text{Previous year's FBI}} \times 100$$

Example 9. From the following fixed base index numbers, prepare chain base index numbers:

Year:	1981	1982	1983	1984	1985	1986
FBI:	110	150	180	250	300	440

Solution:

Conversion of FBI into CBI

Year	FBI	Conversion	CBI
1981	110	—	110
1982	150	$\frac{150}{110} \times 100$	136.364
1983	180	$\frac{180}{150} \times 100$	120
1984	250	$\frac{250}{180} \times 100$	138.889
1985	300	$\frac{300}{250} \times 100$	120
1986	440	$\frac{440}{300} \times 100$	146.667

IMPORTANT TYPICAL EXAMPLE

Example 10. From the Chain base index numbers given below prepare fixed base index numbers and verify the answers.

Year:	1971	1972	1973	1974	1975
CBI:	110	160	140	200	150

Solution:

From CBI to FBI

Year	CBI	FBI
1971	110	110
1972	160	$\frac{160 \times 110}{100} = 176.00$
1973	140	$\frac{140 \times 176}{100} = 246.40$
1974	200	$\frac{200 \times 246.40}{100} = 492.80$
1975	150	$\frac{150 \times 492.80}{100} = 739.20$

Verification:

Year	FBI	CBI
1971	110	110
1972	176	$\frac{176}{110} \times 100 = 160$
1973	246.40	$\frac{246.40}{176} \times 100 = 140$
1974	492.8	$\frac{492.8}{246} \times 100 = 200$
1975	739.2	$\frac{739.2}{492.8} \times 100 = 150$

EXERCISE 4.2

1. From the Chain base index numbers given below, prepare fixed base index numbers and verify the answers:

Year:	1974	1975	1976	1977	1978
CBI:	80	110	120	105	95

[Ans. 80, 88, 105.6, 110.9, 105.3]

2. Change the following fixed base index numbers into chain base index numbers:

Year:	1973	1974	1975	1976	1977	1978
FBI:	100	110	175	250	300	400

[Ans. 100, 110.0, 159.1, 142.8, 120.0, 133.3]

3. From the chain base index given below, prepare fixed base index numbers and verify your answers.

Year:	1960	1961	1962	1963	1964
CBI:	90	110	115	120	130

[Ans. 90, 99, 113.8, 136.6, 177.6]

4. From the fixed based index numbers given below, prepare chain base index numbers:

Year:	1970	1971	1972	1973	1974	1975
FBI:	376	392	408	380	392	400

[Ans. 376, 104.28, 104.1, 93.1, 103.2, 102].

BASE SHIFTING, SPLICING, AND DEFLATING OF INDEX NUMBERS

Following are some special problems related to the construction of index numbers:

Base Shifting

Sometimes, it becomes necessary that index numbers are changed by taking some other year as base year rather than the given base year. This operation is called as base-shifting. Base shifting is needed for the following two reasons: (*i*) When the present base year has become rather old (*ii*) When some series are to be compared whose base years are different. In such condition, such series cannot be made fit for comparison unless base years of all the series are the same.

While base-shifting, index number of new base year is taken as 100 and all other old years index numbers are converted on this new base. The following formula is used for this purpose:

$$\textbf{Index No. with New Base} = \frac{\textbf{Index No. with old Base}}{\textbf{Index No. of New Base Year}} \times 100$$

From the following example, the process of base-year shifting is illustrated:

Example 11. The following are index numbers of prices based on 1980, shift the base to 1984.

Year	Index Numbers (1980 = 100)	Year	Index Numbers (1980 = 100)
1980	100	1985	390
1981	120	1986	400
1982	200	1987	420
1983	240	1988	435
1984	300	1989	444

Solution:

Computation of New Index Numbers

Year	Old Index No. (1980 = 100)	New Index No. (1984 = 100)
1980	100	$\frac{100}{300}\times 100 = 33.33$
1981	120	$\frac{120}{300}\times 100 = 40.00$
1982	200	$\frac{200}{300}\times 100 = 66.67$
1983	240	$\frac{240}{300}\times 100 = 80.00$
1984	**300**	**100.00**
1985	390	$\frac{390}{300}\times 100 = 130.00$
1986	400	$\frac{400}{300}\times 100 = 133.33$
1987	420	$\frac{420}{300}\times 100 = 140.00$
1988	435	$\frac{435}{300}\times 100 = 145.00$
1989	444	$\frac{444}{300}\times 100 = 148.00$

Splicing

Sometimes we come across a situation when an index constructed on a given base year is stopped being used and new index is constructed taking the year, in which an index was stopped using, as base year. Thus, we have two types of indices and we have to change them into a single series. The procedure taken up for this is called as splicing. For example, one series is available from 1981 to 1985; whose base year is 1981. This index is stopped using in 1985 and taking 1985 as base year, we have with ourselves a new series of indices from 1985 to 1990. If we are in need of a series of index numbers with continuous time period, then we can make it available with us by the use of splicing. **Thus, splicing is a process by which new series of indices is tied with old index series or old series of indices is tied with new index series.** There may be two types of splicing:

(*i*) **Splicing of new index series to old index series:** The following formula is used for splicing new index series to old index series:

$$\textbf{Spliced new index series to old index series} = \frac{\textbf{New index} \times \textbf{Old index of overlapping year}}{100}$$

(*ii*) **Splicing of old index series to new index series:** The following formula is used for splicing old index series to new index series:

Spliced old index series with new index series

$$= \textbf{Old index} \times \frac{100}{\textbf{Old index of overlapping year}}$$

The following examples illustrate the process of splicing:

Example 12. Given below are two price index series. Splice them on the base 1981 = 100

Year	1981	'82	'83	'84	'85	'86	'87	'88	'89	'90
Index A (1981 = 100)	100	110	120	125	150	—	—	—	—	—
Index B (1985 = 100)	—	—	—	—	100	115	120	130	140	160

Solution:

Splicing of Index B to Index A

Year	Index A (1981 = 100)	Index B (1985 = 100)	Spliced index (1981 = 100)
1981	100	—	100
1982	110	—	110
1983	120	—	120
1984	125	—	125
1985	**150**	**100**	$100\times\frac{150}{100} = 150$
1986	—	115	$115\times\frac{150}{100} = 172.5$

1987	—	120	$120 \times \frac{150}{100} = 180$
1988	—	130	$130 \times \frac{150}{100} = 195$
1989	—	140	$140 \times \frac{150}{100} = 210$
1990	—	160	$160 \times \frac{150}{100} = 240$

Example 13. Given below are two price index series. Splice them on the base 1985 = 100.

Year:	1981	'82	'83	'84	'85	'86	'87	'88	'89	'90
Old Price Index (1981 = 100)	100	110	120	125	150	—	—	—	—	—
New Price Index (1985 = 100)	—	—	—	—	100	115	120	130	140	160

Solution:

Splicing of Old Index Series to New Index Series

Year	Old price index (1981 = 100)	New price index (1985 = 100)	Spliced index (1985 = 100)
1981	100	—	$100 \times \frac{100}{150} = 66.67$
1982	110	—	$110 \times \frac{100}{150} = 73.33$
1983	120	—	$120 \times \frac{100}{150} = 80.00$
1984	125	—	$125 \times \frac{100}{150} = 88.33$
1985	**150**	**100**	$150 \times \frac{100}{150} = 100$
1986	—	115	115
1987	—	120	120
1988	—	130	130
1989	—	140	140
1990	—	160	160

Example 14. A price index series was started in 1994 as base. By 1998 it rose by 25%. The link relative for 1999 was 95. In this year a new series was started. This new series rose by 15 points by next year. During 2004 the price level was only 5% higher than 2002 and in 2002 they were 8% higher than 2000. Splice the two series.

Solution:

Year	Old price index (1994 = 100)	New price index (1999 = 100)	Old price index spliced to new (1999 = 100)
1994	100		$\frac{100}{118.75} \times 100 = 84.2$
1998	125		$\frac{125}{118.75} \times 100 = 105.2$
1999	$\left(\frac{95}{100} \times 125\right) = 118.75$	100	100.00
2000		$100 + 15 = 115$	115.00
2002		$\left(\frac{115 \times 108}{100}\right) = 124.2$	124.20
2004		$\left(\frac{124.2 \times 105}{100}\right) = 130.41$	130.41

Example 15. Given the following values:

A		B	
Year		Year	
1998	$\Sigma p_0 q_0 =$ Rs. 20	2001	$\Sigma p_3 q_3 =$ Rs. 35
1999	$\Sigma p_1 q_0 =$ Rs. 24	2002	$\Sigma p_4 q_3 =$ Rs. 43
2000	$\Sigma p_2 q_0 =$ Rs. 30	2003	$\Sigma p_5 q_3 =$ Rs. 52.5
2001	$\Sigma p_3 q_0 =$ Rs. 40	2004	$\Sigma p_6 q_3 =$ Rs. 55

(*i*) Calculate the price indices in A series with q_0 as weights and in B series with q_3 as weights.

(*ii*) Splice the two series so as to make A a continuous series.

Solution:

Computation of Price Indices and Splicing

Year	Index A (1998 = 100)	Index B (2001 = 100)	Splicing of Series B to A (1998 = 100)
1998	100		100
1999	$\frac{24}{20} \times 100 = 120$		120
2000	$\frac{30}{20} \times 100 = 150$		150
2001	$\frac{40}{20} \times 100 = 200$	100	200
2002		$\frac{43}{35} \times 100 = 122.86$	$\frac{200}{100} \times 122.86 = 245.72$
2003		$\frac{52.5}{35} \times 100 = 150.00$	$\frac{200}{100} \times 150 = 300$
2004		$\frac{55}{35} \times 100 = 157.14$	$\frac{200}{100} \times 157.14 = 314.28$

Deflating of Index Numbers

Often, changes keep taking place in the prices of commodities and the cost of living. Therefore, index numbers of salary or wages need to be revised due to such changes. In other words, wherever, we need to derive real wages from money wages or real income from money income, the index numbers have to be adjusted or revised. In fact, this process is called as Deflating of Index numbers. In other words, **Deflating refers to the correction for price changes in money wages or money income series.** For deflating, the following formulae are used:

$$\textbf{Real Wages (or Deflated Wages)} = \frac{\textbf{Money Wages}}{\textbf{Price Index}} \times 100$$

$$\textbf{Real Income (or Deflated Income)} = \frac{\textbf{Money Income}}{\textbf{Cost of Living Index}} \times 100$$

The real income is also known as deflated income., *i.e.*, income at constant prices.

If real wage index or real income index is to be derived, then first, the above said formulae are used and then indices are computed by taking the first year as base. The following formulae are used to find it:

$$\textbf{Real Wages Index No.} = \frac{\textbf{Real wages of the current year}}{\textbf{Real wages of the base year}} \times 100$$

$$\textbf{Real Income Index No.} = \frac{\textbf{Real income of the current year}}{\textbf{Real income of the base year}} \times 100$$

The following examples illustrate the process of defalting the index numbers:

Example 16. The following data relate to the wages of the people and the general price index number. Calculate:

(*i*) Real Wages and (*ii*) Index Number of Real Wages with 1980 as base.

Year	Wages (in Rs.)	Price Index
1980	800	100
1981	819	105
1982	825	110
1983	876	120
1984	920	125
1985	938	140
1986	924	140

Solution:

Year	Wages (in Rs.)	Price index	Real wages	Real wage index number (1980 =Base year)
1980	800	100	$\frac{800}{100}\times100=800$	=100
1981	819	105	$\frac{819}{105}\times100=780$	$\frac{780}{800}\times100=97.5$

1982	825	110	$\frac{825}{110}\times100=750$	$\frac{750}{800}\times100=93.5$
1983	876	120	$\frac{876}{120}\times100=730$	$\frac{730}{800}\times100=91.25$
1984	920	125	$\frac{920}{125}\times100=736$	$\frac{736}{800}\times100=92$
1985	938	140	$\frac{938}{140}\times100=670$	$\frac{670}{800}\times100=83.75$
1986	924	140	$\frac{924}{140}\times100=660$	$\frac{660}{800}\times100=82.5$

Example 17. The following table gives the per capita income and the cost of living index of a particular community. Calculate the index numbers of real income taking into account the rise in the cost of living:

Year:	1989	1990	1991	1992	1993	1994	1995
Cost of living index (1989=100):	100	104	115	160	210	260	300
Per capita income (in Rs.):	360	400	480	520	550	590	610

Solution:

Construction of Real Income Index Numbers

Year	Per capita income (Rs.)	Cost of living index	Real income (Rs.)	Real income Index Nos. (1989=100)
1989	360	100	$\frac{360}{100}\times100=360$	100
1990	400	104	$\frac{400}{104}\times100=384.61$	$\frac{384.61}{360}\times100=106.84$
1991	480	115	$\frac{480}{115}\times100=417.39$	$\frac{417.39}{360}\times100=115.94$
1992	520	160	$\frac{520}{160}\times100=325$	$\frac{325}{360}\times100=90.28$
1993	550	210	$\frac{550}{210}\times100=261.90$	$\frac{261.90}{360}\times100=72.75$
1994	590	260	$\frac{590}{260}\times100=226.92$	$\frac{226.92}{360}\times100=63.03$
1995	610	300	$\frac{610}{300}\times100=203.33$	$\frac{203.33}{360}\times100=56.48$

Purchasing Power of Money

The concept of deflating can also be used to determine the purchasing power or real value of a rupee. When prices in general are rising, the real value of a rupee is declining. If, for example, the price index in 1992 with base 1990 is 120, the real value of a rupee in 1992 as compared with its value in 1990 $= \frac{1}{120} \times 100 = 0.83$. This implies that a rupee in 1990 is worth only 83 paise in 1992. Thus,

$$\textbf{Purchasing Power of Money} = \frac{1}{\textbf{Cost of Living Index}}$$

Example 18. Table below shows the average wages in rupees per week of a group of industrial workers in the year 1989-1996. The consumer price indices for these year with 1989 as base are also given:

Year:	1989	1990	1991	1992	1993	1994	1995	1996
Average wage of workers (Rs.):	119	133	144	157	175	184	189	194
Consumer price index:	100	107.6	106.6	107.6	116.2	118.9	119.8	120.2

(*i*) Determine the real wage of workers during the year 1989-1996 as compared with their wages in 1989.

(*ii*) Determine the purchasing power of rupee for the year 1996 as compared to the year 1989. What is the significance of this result?

Solution: (*i*) Real Wage $= \frac{\text{Money Wage}}{\text{Consumer Price Index}} \times 100$

Year	Wages (Rs.)	Consumer price index	Real wages
1989	119	100	$\frac{119}{100} \times 100 =$ 119.00
1990	133	107.6	$\frac{133}{107.6} \times 100 =$ 123.61
1991	144	106.6	$\frac{144}{106.6} \times 100 =$ 135.08
1992	157	107.6	$\frac{157}{107.6} \times 100 =$ 145.91
1993	175	116.2	$\frac{175}{116.2} \times 100 =$ 150.60
1994	184	118.9	$\frac{184}{118.9} \times 100 =$ 154.75
1995	189	119.8	$\frac{189}{119.8} \times 100 =$ 157.76
1996	194	120.2	$\frac{194}{120.2} \times 100 =$ 161.40

(*ii*) The purchasing power of rupee $= \frac{100}{120.2} = 0.83$ or 83 paise (approx).

For the year 1996, it means a rupee worth of 1989, is worth only 83 paise in 1996. The purchasing power of rupee has decreased by 17% over the period 1989 to 1996.

Example 19. Given the following data:

Year	Weekly take-home pay (wages)	Consumer Price Index
1999	109.50	112.8
2000	112.20	118.2
2001	116.40	127.4
2002	125.08	138.2
2003	135.40	143.5
2004	138.10	149.8

(*i*) What was the real average weekly wage for each year?

(*ii*) In which year did the employees have the greatest buying power?

(*iii*) What percentage increase in the weekly wages for the year 2004 is required (if any) to provide the same buying power that the employees employed in the year in which they had the highest real wages.

Solution: (*i*) Real average weekly wage can be obtained by the following formula:

$$\text{Real Wage} = \frac{\text{Money Wage}}{\text{Price Index}} \times 100$$

Calculation of Real Wages

Year	Wages (Rs.)	Consumer price index	Real wages
1999	109.50	112.8	$\frac{109.5}{112.8} \times 100 =$ 97.07
2000	112.20	118.2	$\frac{112.2}{118.2} \times 100 =$ 94.92
2001	116.40	127.4	$\frac{116.4}{127.4} \times 100 =$ 91.37
2002	125.08	138.2	$\frac{125.08}{138.2} \times 100 =$ 90.51
2003	135.40	143.5	$\frac{135.4}{143.5} \times 100 =$ 94.36
2004	138.10	149.8	$\frac{138.10}{149.8} \times 100 =$ 92.19

(*ii*) The employees have the greatest buying power in 1999 since the real wage was maximum for the year 1999.

(iii) The percentage increase in prices in 2004 as compared to 1999

$$= \frac{149.8}{112.8} \times 100 = 132.8$$

The weekly money wages in 2004 so as to have the same purchasing power as in 1999

$$= \frac{109.50 \times 132.8}{100} = \text{Rs. } 145.42$$

Thus, the required increase in weekly wages

$$= 145.42 - 138.10 = \text{Rs. } 7.32$$

The percentage increase in weekly wages of 2004

$$= \frac{7.32}{138.10} \times 100 = 5.3\%.$$

Example 20. Calculate the national income at current prices from the following data:

Year:	1999	2000	2001	2002	2003	2004
National income at 1997 prices (Rs. Billion):	80	90	100	105	120	115
Price index (1997 = 100):	100	120	125	140	180	200

Solution:

Calculation of National Income at Current Prices

Year	National income at 1997 prices	Price index (1997=100)	National income at current prices $= \frac{\text{National income at 1997 prices} \times \text{Price index}}{100}$
1999	80	100	$80 \times \frac{100}{100} = 80$
2000	90	120	$90 \times \frac{120}{100} = 108$
2001	100	125	$100 \times \frac{125}{100} = 125$
2002	105	140	$105 \times \frac{140}{100} = 147$
2003	120	180	$120 \times \frac{180}{100} = 216$
2004	115	200	$115 \times \frac{200}{100} = 230$

Example 21. For the following data of a firm, construct the index of average wage and salaries at constant prices (Base year = 1980).

Year:	1980	1981	1982	1983	1984
Average wages and salaries paid (Rs.):	5000	5670	5865	6240	6820
Consumer price index:	100	108	102	104	110

Solution:

Calculation Table

Year	Average wages and salaries paid (Rs.)	Consumer price index	Wages and salaries at constant prices	Index of wages and salaries (1980 = 100)
1980	5000	100	$\frac{5000}{100} \times 100 = 5000$	$\frac{5000}{5000} \times 100 = 100$
1981	5670	108	$\frac{5670}{108} \times 100 = 5250$	$\frac{5250}{5000} \times 100 = 105$
1982	5865	102	$\frac{5865}{102} \times 100 = 5750$	$\frac{5750}{5000} \times 100 = 115$
1983	6240	104	$\frac{6240}{104} \times 100 = 6000$	$\frac{6000}{5000} \times 100 = 120$
1984	6820	110	$\frac{6820}{110} \times 100 = 6200$	$\frac{6200}{5000} \times 100 = 124$

Example 22. The average wage of a rail-road worker per day was Rs. 119 in 2000 and Rs. 245 in 2005. The consumer price index for these year was 95.5 and 123.5 respectively. Show that the money wage increased by about 59% in 2005 as compared as 2000.

Solution: We first calculate a consumer price index with 2000 as base year by dividing given consumer price index by 95.5 and expressing the result as percentage:

Year	C.P.I.	Wages	New C.P.I. (2000 = 100)	Real wages $\left(\frac{\text{Wages}}{\text{C.P.I.}} \times 100\right)$	Money wages index $\left(\frac{\text{Real wages}}{\text{Base year wages}} \times 100\right)$
2000	95.5	119	$\frac{95.5}{95.5} \times 100 = 100$	$\frac{119}{100} \times 100 = 119$	$\frac{119}{119} \times 100 = 100$
2005	123.5	245	$\frac{123.5}{95.5} \times 100 = 129.32$	$\frac{245}{129.32} \times 100 = 189.45$	$\frac{189.45}{119} \times 100 = 159.20$

Although the wages are more than doubled, the money wages increased by only 59.20%.

Example 23. If with rise of 10% in prices, the wages are increased by 20%, calculate the percentage of real wage increase.

Solution: Current real wage percentage is $\frac{120}{110} \times 100 = \frac{1200}{11} = 109.09\%$

Hence, the percentage increase in real wage is $109.09 - 100 = 9.09\%$.

EXERCISE 4.3

Base Shifting

1. The following are the index numbers of wholesale prices of a certain commodity based on 1992:

Year:	1992	1993	1994	1995	1996
Index No. (1992=100):	100	108	120	150	210

Shift the base to 1994 and obtain new index numbers. **[Ans.** 83.33, 90, 100, 125, 175]

2. An index is at 100 in 1981. It rises 4% in 1982, falls 6% in 1983, falls 4% in 1984 and rises 3% in 1985. Calculate the index numbers for five years with 1983 as base.
[**Hint:** See Example 35] **[Ans.** 102.29, 106.38, 100, 96.00, 98.88]

Splicing

3. Given below are the two price index series at different base years. Splice (*i*) 1995-base index series with 1991-base index series, and (*ii*) 1991-base index series with 1995-base index series.

Year:	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Price Index (1991 as base):	100	105	115	123	125	—	—	—	—	—
Price Index (1995 as base):	—	—	—	—	100	104	110	112	120	150

[Ans. (*i*) 100, 105, 115, 123, 125, 130, 137.5, 140, 150, 187.5
(*ii*) 80, 84, 92, 98.4, 100, 104, 110, 112, 120, 150]

4. The following three series of index numbers are given:

Year	Index A (1954 = 100)	Index B (1969 = 100)	Index C (1975 = 100)
1954	100	—	—
1960	120	—	—
1969	200	100	—
1975	—	200	100
1985	—	—	120

Prepare a spliced series of index numbers with base 1975 = 100 as base.
[**Hint:** See Example 38] **[Ans.** 25, 30, 50, 100, 120]

Deflating

5. The following table gives the per capita income and the cost of living index for a particular class of people. Deflate the per capita income by taking into account the changes in the cost of living.

Year	Per capita income (Rs.)	Cost of living index
1979	300	120
1980	320	125
1981	340	150
1982	350	160
1983	375	175

[Ans. 250, 256, 226.27, 218.75, 214.29]

6. Calculate the (*i*) Real Wages and (*ii*) Index of Real Wages from the following information using 1985 as base year:

Year:	1985	1986	1987	1988	1989	1990
Average monthly wages:	200	225	240	280	350	400
Consumer price index:	100	120	125	135	175	225

[Ans. (*i*) 200, 187.5, 192, 207.41, 200, 177.78; (*ii*) 100, 93.75, 96, 103.72, 100, 88.89]

7. The following are the average daily wages in rupees of a group of industrial workers and price indices:

Year:	1999	2000	2001	2002	2003	2004
Daily wages:	80	108	125	147	216	230
Consumer price index:	100	120	125	140	180	200

(*i*) Determine the real wages of workers during the years 1999-2004 as compared with their wages in 1999.

(*ii*) Calculate the purchasing power of rupee for the year 2004 as compared to the year 1999. What is the significance of this result?

[Ans. (*i*) 80, 90, 100, 105, 120, 115 (*ii*) $0.5 = \frac{100}{200}$, purchasing power of rupee has decreased by 50% over the period from 1999 to 2004]

8. Given the following data:

Year:	1995	1996	1997	1998	1999	2000	2001
Monthly pay (Rs.):	10,500	11,000	11,500	12,500	13,500	14,000	14,500
Price index:	115	120	130	138	144	150	160

(*i*) Calculate the real monthly pay for each year.
(*ii*) In which year did the employee have the highest purchasing power?
(*iii*) What percentage increase in the monthly pay for the year 2001 is required (if any) to compensate him with the purchasing power in the year of his highest real pay?

[Ans. (*i*) 9130.43, 9166.66, 8846.15, 9057.97, 9375, 9333.33, 9062.5
(*ii*) 1999, (*iii*) 3.44%]

◘ CONSUMER PRICE INDEX OR COST OF LIVING INDEX NUMBERS

Cost of living index numbers show the direction and magnitude of change taking place in the cost of living of specific group of persons at given time and place. Its purpose is to know how much increase or decrease has taken place in the outlay or subsistence expenditure made by a consumer on his living, therefore these are also known as **consumer price index**. Effect of changes in the prices is not uniform on all the classes of a society because different classes of people consume different commodities and changes in the prices of commodities are different. Therefore, separate cost of living indices are constructed for different classes of people and for different places. **Consumer Price Index Numbers are those index numbers which measure the effects on living conditions of different classes of consumers for any change in the level of prices over a period of time.** Such types of indices are constructed in order to find out how the economic progress of a country has affected the standard of living of a particular class of people.

Uses of Consumer Price Index

The different uses of consumer price index are given below:

(*i*) **To Examine the Effects of Changes in Retail Prices:** It is used to examine the effect of change in the retail prices on the cost of living of a particular class of people.

(*ii*) **Helpful in Policy Formation:** The government may decide its price control, minimum wages, rationing policies in the light of the changes in the cost of living index.

(*iii*) **Fixation of Dearness Allowance:** The amount of dearness allowance and revision of wages of different categories of employees are decided on the basis of consumer price index.

Construction of Consumer Price Index

The procedure of constructing a consumer price index is as follows:

(1) Decision about the Class of People: First of all, it should be ascertained that for which class consumer price index will be constructed, *i.e.*, whether the index is related to industrial workers, teachers, office employees, etc. At the same time, the scope of the index should also be well defined.

(2) Conducting Family Budget Enquiry: After deciding about the specific class, some families from that class should be selected by random sampling and their budgets should be studied to make findings about their items of income-expenditure, quantities of commodities and size of families etc. According to convenience, the items of consumption are divided into five main categories: (*i*) Food, (*ii*) Clothing, (*iii*) Fuel and Lighting, (*iv*) House Rent, (*v*) Miscellaneous.

(3) Obtaining Price Quotations: After selecting the commodities, their retail prices are obtained. Retail prices of the selected commodities are collected from the reliable sources and from those places from where the people of that class buy goods.

(4) To Decide Weight: To express the relative importance of the items of consumption, selective weights are assigned to them. Weights can be given in two ways: (*i*) In the proportion of consumption quantity in the base year (q_0) (*ii*) In the proportion of expenditure made on each commodity in the base year ($p_0 q_0$).

(5) Methods of Constructing Consumer Price Index: After this, consumer price indices are constructed by the following methods:

(*i*) Aggregate Expenditure Method.

(*ii*) Family Budget Method.

(1) Aggregate Expenditure Method

In this method, wages are assigned to items on the base of base year quantities.

$$\textbf{Consumer Price Index } (P_{01}) = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

Aggregative expenditure method is equal to **Laspeyre's Method.**

Steps for Calculation

(*i*) Quantity in base year (q_0) and prices in current year (p_1) are multiplied and their sum ($\Sigma p_1 q_0$) is taken. This is the aggregate expenditure in current year.

(*ii*) Quantity in base year (q_0) and price in base year (p_0) are multiplied and their sum ($\Sigma p_0 q_0$) is taken. This is the aggregate expenditure in base year.

(*iii*) $\Sigma p_1 q_0$ is divided by $\Sigma p_0 q_0$ and the quotient is multiplied by 100.

(2) Family Budget Method

Under this method, the weights are assigned to items on the basis of percentage expenditure on the items.

$$\textbf{Consumer Price Index } (P_{01}) = \frac{\Sigma PW}{\Sigma W}$$

Where, P = Price Relatives $= \frac{p_1}{p_0} \times 100$, W = Total Expenses $= p_0\, q_0$.

If the geometric mean is used, then $P_{01} = AL\left[\frac{\Sigma W \log P}{\Sigma W}\right]$

Family budget method is equal to **Weighted Average of Price Relative Method.**

Steps for Calculation

(*i*) Price relative of current year for each commodity is computed by the following formula:

$$P = \frac{p_1}{p_0} \times 100$$

(*ii*) Price relative of each commodity is multiplied by the expenditure on it or (Value Weight or W) to find out weighted price relatives.

(*iii*) Weighted price relatives are summed up (ΣPW)

(*iv*) ΣW, *i.e.*, summation of weights which is $\Sigma p_0 q_0$ is determined.

(*v*) ΣPW is divided by ΣW and the quotient is multiplied by 100.

Note 1: It should be noted that prices and quantities must be same in units while multiplying. If units of price and quantity are not different, then the unit of quantity must be changed into the unit of price before carrying out multiplication. For example, if the price is per quintal and the quantity purchased in kg. then kgs. must be converted into quintals before carrying out multiplication.

Note 2: The consumer price index numbers (or cost of living index) obtained by both the methods are the same.

Note 3: Aggregate expenditure method should always be preferred to in as much as it proves to be easier than the family budget method so far as calculations are concerned.

Example 24. From the following data, find the Cost of Living Index number of 1990 on the basis of 1980 by (*i*) Aggregate Expenditure Method; and (*ii*) Family Budget Method:

Items	Quantity consumed in 1980	Unit	Prices	
			1980	1990
Wheat	2 Qtls	Qtls	75	125
Rice	20 Kg	Kg	12	16
Sugar	10 Kg	Kg	12	16
Ghee	5 Kg	Kg	10	15
Clothing	25 Meters	Meter	4.5	5
Fuel	40 Kg	Kg	10	12
Rent	One House	House	25	40

Solution: (*i*) Construction of Cost of Living Index Number by Aggregative Expenditure Method

Items	Quantity consumed in 1980 (q_0)	Unit	Prices in 1980 (p_0)	Prices in 1990 (p_1)	p_0q_0	p_1q_0
Wheat	2 Qtls	Qtls	75	125	150	250
Rice	20 Kg	Kg	12	16	240	320
Sugar	10 Kg	Kg	12	16	120	160
Ghee	5 Kg	Kg	10	15	50	75
Clothing	25 Meters	Meter	4.5	5	112.5	125
Fuel	40 Kg	Kg	10	12	400	480
Rent	One House	House	25	40	25	40
					$\Sigma p_0q_0 = 1097.5$	$\Sigma p_1q_0 = 1450$

$$\text{Cost of Living Index Number } (P_{01}) = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100 = \frac{1450}{1097.5} \times 100 = 132.12$$

(*ii*) Computation of Cost of Living Index Number by Family Budget Method

Items	Quantity consumed in 1980 (q_0)	Unit	Prices in 1980 (p_0)	Prices in 1990(p_1)	Price relatives $P = \frac{p_1}{p_0} \times 100$	$W = p_0 q_0$	PW
Wheat	2 Qtls	Qtls	75	125	166.66	150	24999
Rice	20 Kg	Kg	12	16	133.33	240	31999.2
Sugar	10 Kg	Kg	12	16	133.33	120	15999.6
Ghee	5 Kg	Kg	10	15	150.00	50	7500
Clothing	25 Meters	Meter	4.5	5	111.11	112.5	12499.875
Fuel	40 Kg	Kg	10	12	120.00	400	48000
Rent	One House	House	25	40	160.00	25	4000
						$\Sigma W = 1097.5$	$\Sigma PW = 144997.675$

$$\text{Cost of Living Index } (P_{01}) = \frac{\Sigma PW}{\Sigma W} = \frac{144997.675}{1097.5} = 132.12$$

Example 25. Construct Cost of Living Index Number of 1980 on the basis of 1970 by (*i*) Aggregate Expenditure Method, and (*ii*) Family Budget Method:

Items	Quantity consumed in 1970	Unit	Prices in	
			1970	1980
Wheat	2 Qtls	Qtls	50	100
Rice	1 Qtl	Qtls	80	110
Arhar	20 Kg	Kg	1.20	2.80
Sugar	0.5 Qtls.	Kg	2.0	3.00
Salt	10 Kg	Qtls	20	30
Oil	10 Kg	Kg	4	8
Clothing	20 Meter	Meter	3	5
Fuel	4 Qtls	Qtls	12	15
Rent	One House	House	50	75

Solution: Since the unit of price and quantity of two items such as sugar and salt are different, we should convert the unit of quantity into the unit of price before we apply any method.

(*i*) Construction of Cost of Living Index by Aggregate Expenditure Method

Items	Quantity consumed in 1970 (q_0)	Unit	Price in 1970 (p_0)	Price in 1980 (p_1)	p_0q_0	p_1q_0
Wheat	2 Qtls	Qtls	50	100	100	200
Rice	1 Qtl	Qtls	80	110	80	110
Arhar	20 Kg	Kg	1.20	2.80	24	56
Sugar	50 Kg	Kg	2.0	3.00	100	150
Salt	0.1Qtls.	Qtls	20	30	2	3
Oil	10 Kg	Kg	4	8	40	80
Clothing	20 Meter	Meter	3	5	60	100
Fuel	4 Qtls	Qtls	12	15	48	60
Rent	One House	House	50	75	50	75
					$\Sigma p_0q_0 = 504$	$\Sigma p_1q_0 = 834$

$$\text{Cost of Living Index Number } (P_{01}) = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

$$= \frac{834}{504} \times 100 = 165.47 \approx 165.5$$

(ii) Construction of Cost of Living Index by Family Budget mthod

Items	Quantity consumed in 1970	Unit	Prices in 1970 (p_0)	Prices in 1980 (p_1)	Price Relatives $P = p_1/p_0 \times 100$	$W = p_0 q_0$	PW
Wheat	2 Qtls	Qtls	50	100	200	100	20000
Rice	1 Qtl	Qtls	80	110	137.5	80	11000
Arhar	20 Kg	Kg	1.20	2.80	233.33	24	5600
Sugar	50 Kg	Kg	2.0	3.00	150.00	100	15000
Salt	0.10 Qtls.	Qtls	20	30	150.00	2	300
Oil	10 Kg.	Kg	4	8	200	40	8000
Clothing	20 meter	meter	3	5	166.67	60	10000
Fuel	4 Qtls.	Qtls	12	15	125	48	6000
Rent	One House	House	50	75	150	50	7500
						$\Sigma W = 504$	$\Sigma PW = 83400$

$$\text{Cost of Living Index Number } (P_{01}) = \frac{\Sigma PW}{\Sigma W} = \frac{83400}{504} = 165.5$$

IMPORTANT TYPICAL EXAMPLES BASED ON FAMILY BUDGET METHOD

Example 26. An enquiry into the budget of the middle class families in a certain city in India gave the following information:

Items	Food	Fuel	Clothing	Rent	Misc.
Expenses (in %)	35	10	20	15	20
Prices in 2000 (Rs):	150	25	75	30	40
Prices in 2002 (Rs):	145	23	65	30	45

What is the cost of living index number of 2002 as compared with 2000?

Solution: Here the percentage expenses on different items are to be taken as weights (W).

Construction of Cost of Living Index

Items	Weights (W)	Prices in 2000 (p_0)	Prices in 2002 (p_1)	$P = \frac{p_1}{p_0} \times 100$	PW
Food	35	150	145	96.67	3383.45
Fuel	10	25	23	92.00	920.00
Clothing	20	75	65	86.67	1733.40
Rent	15	30	30	100.00	1500.00
Misc.	20	40	45	112.50	2250.00
	$\Sigma W = 100$				$\Sigma PW = 9786.85$

$$\text{Cost of Living Index Numbers } (P_{01}) = \frac{\Sigma PW}{\Sigma W} = \frac{9786.85}{100} = 97.86$$

The percentage decrease in prices in 2002 = 97.86 – 100 = – 2.14.

Thus, there is a decrease of 2.14% in the prices of 2002 as compared to 2000.

Example 27. Construct the Cost of Living Index for 1976 on the basis of 1975 from the following data and give your comments:

Item	Prices in 1975 (Rs.)	Prices in 1976 (Rs.)	Weights
Food	39	47	4
Fuel	8	12	1
Clothing	14	18	3
Rent	12	15	2
Misc.	25	30	1

Solution: Since weights are directly given along with the prices in 1975 and 1976, we first determine price relatives (P) and then multiply it with weights (W).

Calculation of Cost of Living Index

Items	p_0	p_1	$P = \frac{p_1}{p_0} \times 100$	W	PW
Food	39	47	120.51	4	482.04
Fuel	8	12	150.00	1	150.00
Clothing	14	18	128.57	3	385.71
Rent	12	15	125.00	2	250.00
Misc.	25	30	120.00	1	120.00
				$\Sigma W = 11$	$\Sigma PW = 1387.75$

$$\text{Cost of Living Index Numbers } (P_{01}) = \frac{\Sigma PW}{\Sigma W} = \frac{1387.75}{11} = 126.16$$

It shows that the cost of living has gone up by 26.16% in 1976 compared to 1975.

Example 28. Calculate the Cost of Living Index from the following data:

Group	Food	Fuel and Lighting	Clothing	Rent	Misc.
Index:	352	200	230	160	190
Weight:	48	10	8	12	15

Solution: Since we are given index no. for different group items along with their weights, we use family budget method (or weighted average of price relatives) to calculate cost of living index.

Calculation of Cost of Living Index

Items	Index No. (I)	Weight (W)	IW
Food	352	48	16896
Fuel and Lighting	200	10	2000
Clothing	230	8	1840
Rent	160	12	1920
Miscellaneous	190	15	2850
		ΣW=93	ΣIW=25506

$$\text{Cost of Living Index Numbers } (P_{01}) = \frac{\Sigma IW}{\Sigma W} = \frac{25506}{93} = 274.26$$

Example 29. The following table gives the group index numbers and the weights of different heads of expenditure in the calculation of a cost of living index except the index for the group fuel and lighting:

Group:	Food	Clothing	Fuel & Lighting	Rent	Misc.
Index:	132	113	—	128	147
Weights:	65	9	8	10	8

If the cost of living index is 134, find the index number of fuel and lighting.

Solution: Let the index number of Fuel and Lighting be denoted by X

Group	Price Index (P)	Weight (W)	PW
Food	132	65	8580
Clothing	113	9	1017
Fuel and Lighting	X	8	8X
Rent	128	10	1280
Misc.	147	8	1176
		ΣW = 100	ΣPW = 12053 + 8X

$$\text{Cost of Living Index} = \frac{\Sigma PW}{\Sigma W} = 134 \text{ (given)}$$

$$\Rightarrow \quad 134 = \frac{12053 + 8X}{100}$$

$$\Rightarrow \quad 13400 = 12053 + 8X$$

$$\Rightarrow \quad 13400 - 12053 = 8X$$

$$\Rightarrow \quad 1347 = 8X$$

$$\Rightarrow \quad X = \frac{1347}{8} = 168.375$$

Example 30. A textile worker in the city of Delhi earns Rs. 750 p.m. The cost of living index for January 1980 is given as 160. Using the following data, find out the amount he spent on (a) Food and (b) Rent.

Group	Food	Clothing	Rent	Fuel and Lighting	Misc.
Expenditure:	?	125	?	100	75
Group Index:	190	181	140	118	101

Solution: Let expenditure on food be Rs. X and rent be Rs. Y.

Group	Expenditure (W)	Group Index (P)	PW
Food	X	190	190X
Clothing	125	181	22625
Rent	Y	140	140Y
Fuel and Lighting	100	118	11800
Misc.	75	101	7575
	ΣW = 750		Σ PW = 42000 + 190X + 140Y

As textile worker in the city earning Rs. 750 p.m.

$$\therefore \quad X + 125 + Y + 100 + 75 = 750$$

$$\Rightarrow \quad X + Y = 450 \qquad \ldots(i)$$

$$\text{Cost of Living Index} = \frac{\Sigma PW}{\Sigma W} = 160 \text{ (given)}$$

$$160 = \frac{42000 + 190X + 140Y}{750}$$

$$1{,}20{,}000 = 42{,}000 + 190X + 140Y$$

$$\Rightarrow \quad 19X + 14Y = 7800 \qquad \ldots(ii)$$

Solving (*i*) and (*ii*)

Multiply (*i*) by 14 and subtract from (*ii*), we get

$$\begin{array}{rcl} 19X + 14Y & = & 7800 \\ 14X + 14Y & = & 6300 \\ - \quad - & & - \\ \hline 5X & = & 1500 \end{array}$$

$$\Rightarrow \quad X = 300$$

Put X = 300 in (*i*)

$$300 + Y = 450 \quad \Rightarrow \quad Y = 150$$

$$\therefore X = 300,\ Y = 150$$

Thus, Expenditure on Food = Rs. 300

Expenditure on Rent = Rs. 150

EXERCISE 4.4

1. From the following data, calculate consumer price index numbers for 1980 on the basis of 1970 by (i) aggregate expenditure method and (ii) family budget method:

Items	Quantity consumed in 1970	Unit	Prices	
			1970	1980
Wheat	2 Qtls	Qtls	50	75
Rice	25 Kg	Qtls	100	120
Arhar	10 Kg	Qtls	80	120
Ghee (Dalda)	10 Kg	Kg	6.50	7.80
Oil	0.25 Qtls	Kg	2	3
Clothing	50 Meter	Meter	2	2.25
Fuel	4 Qtls.	Qtls	8	10
Rent	1 House	House	20	25

[Ans. 130.6]

2. Construct the consumer price index for 2000 on the basis of 1999 from the following data using:

(*i*) Family Budget method

(*ii*) Aggregative Expenditure method

Commodity:	Rice	Wheat	Pulses	Ghee	Oil
Weights :	40	20	15	20	5
Price (per unit) 1999 (Rs.):	16.00	40.00	0.50	5.12	2.00
Price (per unit) 2000 (Rs.):	20.00	60.00	0.50	6.25	1.50

[**Hint:** See Example 41] [**Ans.** (*i*) P_{01} (FBM) = 123.15, (*ii*) P_{01} (AEM) = 137.188]

3. Construct cost of living index number from the following data for 1986 using (*i*) Family Budget Method; (*ii*) Aggregative Expenditure Method.

Items	Quantity consumed in 1985	Unit	Prices	
			1985	1986
Wheat	4 Qtls	Kg	1.50	1.60
Rice	50 Kg	Qtls.	800	1000
Cloth	40 Meter	Meter	20	25
Oil	20 Litres	Litre	1.80	2
House Rent	One House	House	100	125

[**Ans.** 119.06]

4. The cost of living index uses the following weights— Food 40, Rent 15, Clothing 20, Fuel 10, Miscellaneous 15. During the period 2000-05, the cost of living index raised from 100 to 205.72. Over the same period the percentage rises in prices were : Rent–60, Clothing–180, Fuel–75, Miscellaneous–165. What is the percentage change in the price of food ?

[**Hint :** $\frac{40x+7725}{100} = 105.72$]

[**Ans.** 71.18]

5. An enquiry into the budget of the middle class families in Bombay gave the following information:

Expenses:	Food 35%	Rent 15%	Clothing 20%	Fuel 10%	Misc. 20%
Price in 2004 (Rs.):	150	50	100	20	60
Price in 2005 (Rs.):	174	60	125	25	90

What changes in the cost of living figures in 2005 have taken place as compared to 2004?

[**Ans.** 126.10, Change = 26.10%]

6. In calculation a certain cost of living index number, the following weights were used: Food 15, clothing 3, rent 4, fuel and light 2, misellaneous 1. Calculate the index for a data when the average percentage increases in price of items in the various groups over the base period were 32, 54, 47, 78 and 58 respectively.

Suppose a business executive was earning Rs. 2000 in the base period, what should be his salary in the current period if his standard of living is to remain the same?

[**Ans.** P_{01} = 141.76, Rs. 2835.20]

7. An enquiry into the budgets of the middle class families in a certain city revealed that on an average the percentage expenses of different groups were: Food 45; Rent 15; Clothing 12; Fuel and Light 8 and Miscellaneous 20. The group index numbers for the current year as compared with a fixed base period were respectively 410, 150, 343, 248 and 285. Calculate the cost of living index for the current year. Suppose Mr. X was getting Rs. 240 in the base period and Rs. 430 in the current year. State how much he ought to have received as extra allowance to maintain his former standard of living?

[**Ans.** P_{01} = 325, Extra allowance = Rs. 350]

8. In 2003 for working class people wheat was selling at an average price of Rs. 32 per 10 kg, cloth at Rs. 4 per metre, house rent Rs. 60 per house and other items at Rs. 20 per unit. By 2004 cost of wheat rose by Rs. 8 per 10 kg, house rent by Rs. 30 per house and other items doubled in price. The working class cost of living index for the year 2004 (with 2003 as base) was 160. By how much the cloth rose in during the period ?

[**Hint:** See Example 39]

[**Ans.** Increase in price of cloth is by Rs. 2.60 per metre]

9. The data below show the percentage increase in prices of a few selected food items and weights attached to each of them, calculate index number for the food group:

Food Item	Weightage	Percentage
Rice	33	180
Wheat	11	202
Dal	8	115
Ghee	5	212
Oil	5	175
Spices	3	517
Milk	7	260
Fish	9	426
Vegetables	9	332
Refreshments	10	239

Using the above food index and the information given below, calculate the consumer price index number:

Group	Index	Weight
Food	—	60
Clothing	310	5
Fuel and Light	220	8
Rent and Rates	150	9
Miscellaneous	300	18

[**Ans.** $I_{Food} = 340$; CPI = 304.60]

10. Following information relating to works in an industrial town is given:

Items of Consumption	Consumer Price Index in 2005 (2000 = 100)	Proportion of Expenditure on the item
Food, drinks and tobacco	132	60%
Clothing	154	12%
Fuel and Lighting	147	16%
Housing	178	8%
Miscellaneous	158	4%

Average wage per month in 2000 is Rs. 2000. What should be the dearness allowance expressed as percentage of wages? What should be the average wage per worker per month in 2005 in that town so that the standard of living of the workers does not fall below the 2000 level?

[**Ans.** CPI = 141.76, Dearness allowance as % of wages = 141.76 – 100 = 41.76%

Worker should get = $\frac{2000 \times 141.76}{100}$ = Rs. 2835.20]

MISCELLANEOUS SOLVED EXAMPLES

Example 31. An index is at 100 in 1991. It rises 5% in 1992, falls 6% in 1993, falls 5% in 1994, rises 4% in 1995 and 7% in 1996. Calculate the index numbers for all these years with 1991 as base.

Solution:

Year	Link Relatives	Index Numbers
1991	100	100 = 100
1992	105	$\frac{105 \times 100}{100} = 105$
1993	94	$\frac{94 \times 105}{100} = 98.7$
1994	95	$\frac{95 \times 98.7}{100} = 93.8$
1995	104	$\frac{104 \times 93.8}{100} = 97.6$
1996	107	$\frac{107 \times 97.6}{100} = 104.4$

Example 32. The average wholesale prices of three groups of commodities for the year 1988 to 1992 are given below. Compute chain base index number with 1988 as base:

Group	1988	1989	1990	1991	1992
I	6	9	15	21	24
II	24	30	36	42	54
III	12	15	21	27	36

Solution:

Computation of Chain Base Index

Group	1988		1989		1990		1991		1992	
	P	LR	P	LR	P	LR	P	LR	P	LR
I	6	100	9	150	15	166.67	21	140	24	114.29
II	24	100	30	125	36	120	42	116.67	54	128.57
III	12	100	15	125	21	140	27	128.57	36	133.33
Total		300		400		426.67		385.24		376.19
Average of LR		100		133.33		142.22		128.41		125.4
Chain Index	100		$\frac{100 \times 133.33}{100} = 133.33$		$\frac{133.33 \times 142.22}{100} = 189.62$		$\frac{189.62 \times 128.41}{100} = 243.49$		$\frac{243.49 \times 125.4}{100} = 305.34$	

Example 33. In the following series of index number shift the base from 1980 to 1983.

Year:	1980	1981	1982	1983	1984	1985	1986	1987
Index No.:	100	105	110	125	135	180	195	205

Solution:

Computation of Index Number (Base 1983 = 100)

Year	Index No. (Base 1980 = 100)	Index No. (Base 1983 = 100)
1980	100	$\frac{100}{125}\times100=80$
1981	105	$\frac{105}{125}\times100=84$
1982	110	$\frac{110}{125}\times100=88$
1983	125	100
1984	135	$\frac{135}{125}\times100=108$
1985	180	$\frac{180}{125}\times100=144$
1986	195	$\frac{195}{125}\times100=156$
1987	205	$\frac{205}{125}\times100=164$

Example 34. The price index of cosmetics was 110 in 2001 with base as 1995 and 120 in 2002 with 2001 as base. It further increased by 30% in 2003 in relation to the price index of 2002 and decreased by 10% in 2004 as compared to its level in 2003. Find the index for 2004 with 1995 as base.

Solution:

$$\frac{P_{2001}}{P_{1995}}\times 100 = 110 \quad\Rightarrow\quad \frac{P_{2001}}{P_{1995}} = 1.1$$

Similarly, $\frac{P_{2002}}{P_{2001}} = 1.2,\ \frac{P_{2003}}{P_{2002}} = 1.3,\ \frac{P_{2004}}{P_{2003}} = 0.9$

Price index for 2004 with 1995 as base $= \frac{P_{2004}}{P_{1995}}\times 100$

$$= \frac{P_{2001}}{P_{1995}}\times\frac{P_{2002}}{P_{2001}}\times\frac{P_{2003}}{P_{2002}}\times\frac{P_{2004}}{P_{2003}}\times 100$$

$$= 1.1\times 1.2\times 1.3\times 0.9\times 100$$

$$= 154.44$$

Example 35. An index is at 100 in 1981. It rises 4% in 1982, falls 6% in 1983, falls 4% in 1984 and rises 3% in 1985. Calculate the index numbers for five years with 1983 as base.

Solution:

Computation of Index Numbers (Base 1983 = 100)

Year	Index No. (Base 1981=100)	Index No. (Base 1983=100)
1981	= 100.00	$\frac{100}{97.76}\times100=$ 102.29
1982	100 + 4 = 104.00	$\frac{104}{97.76}\times100=$ 106.38
1983	$\frac{94}{100}\times104=$ 97.76	$\frac{97.76}{97.76}\times100=$ 100.00
1984	$\frac{96}{100}\times97.76=$ 93.85	$\frac{93.85}{97.76}\times100=$ 96.00
1985	$\frac{103}{100}\times93.85=$ 96.67	$\frac{96.67}{97.76}\times100=$ 98.88

Example 36. From the Chain Base Index numbers given below, prepare Fixed Base Index numbers (*i*) when 1991 is not the base year and (*ii*) when 1991 is taken as the base year.

Year:	1991	1992	1993	1994	1995
Index No.:	80	110	120	90	140

Solution:

Conversion from CBI to FBI

Year	CBI	FBI (1991 ≠ 100)	FBI (1991=100)
1991	80	80	100
1992	110	$\frac{110\times80}{100}=88$	$\frac{110\times100}{100}=110$
1993	120	$\frac{120\times88}{100}=105.6$	$\frac{120\times110}{100}=132$
1994	90	$\frac{90\times105.6}{100}=95.04$	$\frac{90\times132}{100}=118.8$
1995	140	$\frac{140\times95.04}{100}=133.05$	$\frac{140\times118.8}{100}=166.32$

Example 37. For the following data of a firm, construct the index of wages at constant prices (Base year = 1986).

Year:	1986	1987	1988	1989	1990	1991
Wages (in Rs.):	300	340	450	460	475	570
Consumer price index.:	100	120	220	230	250	300

Solution:

Year	Wages	Consumer price index	Wages at constant prices	Index of wages (1986 = 100)
1986	300	100	$\frac{300}{100}\times 100 = 300$	100
1987	340	120	$\frac{340}{120}\times 100 = 283.33$	$\frac{283.33}{300}\times 100 = 94.44$
1988	450	220	$\frac{450}{220}\times 100 = 204.55$	$\frac{204.55}{300}\times 100 = 68.18$
1989	460	230	$\frac{460}{230}\times 100 = 200$	$\frac{200}{300}\times 100 = 66.67$
1990	475	250	$\frac{475}{250}\times 100 = 190$	$\frac{190}{300}\times 100 = 63.33$
1991	570	300	$\frac{570}{300}\times 100 = 190$	$\frac{190}{300}\times 100 = 63.33$

Example 38. The following three series of index numbers are given:

Year	Index A (1954=100)	Index B (1969=100)	Index C (1975=100)
1954	100	—	—
1960	120	—	—
1969	200	100	—
1975	—	200	100
1985	—	—	120

Prepare a spliced series of index numbers with base 1975=100.

Solution:

Year	Index A (1954=100)	Index B (1969=100)	Index C (1975=100)
1954	100	$\frac{100}{200}\times 100 = 50$	$\frac{100}{200}\times 50 = 25$
1960	120	$\frac{100}{200}\times 120 = 60$	$\frac{100}{200}\times 60 = 30$
1969	200	100	$\frac{100}{200}\times 100 = 50$
1975	—	200	100
1985	—	—	120

Example 39. In 2003 for working class people wheat was selling at an average price Rs. 120 per 20 kg, cloth Rs. 20 per metre, house rent Rs.300 per house and other items Rs. 100 per unit. By 2004 cost of wheat rose by Rs. 180 per 20 kg, house rent by Rs. 450 and other items doubled in price. The working class cost of living index for the year 2004 with 2003 as base was 160. By how much the cloth rose in price during the period?

Solution: Let the rise in price of cloth be X.

Index Number for 2003

Commodity	Price in 2003 (p_0)	Price in 2004 (p_1)	$\left(\frac{p_1}{p_0}\times 100\right)$
Wheat	120	180	$\frac{180}{120}\times 100 = 150$
Cloth	20	X	$\frac{X}{20}\times 100 = 5X$
House rent	300	450	$\frac{450}{300}\times 100 = 150$
Miscellaneous	100	200	$\frac{200}{100}\times 100 = 200$
N = 4			500 + 5X

$$P_{01} = \frac{\Sigma\left(\frac{p_1}{p_0}\times 100\right)}{N}$$

Hence, $160 = \frac{500+5X}{4}$

$\Rightarrow$ $640 = 500 + 5X$

$\therefore$ $X = \frac{140}{5} = 28$

Hence, the rise in the price of cloth was Rs. 8 per metre. *i.e.*, Rs. 28 – Rs. 20 = Rs. 8

Example 40. In calculating a certain cost of living index, the following weights were used: Food 15, Clothing 3, rent 4, fuel and light 2, miscellaneous 1. Calculate the index for the period when the average percentage increase in prices of items in the various groups over the base period were 32, 54, 47, 78 and 58 respectively. Suppose a business executive was earning Rs. 2,050 in the base period. What should be his salary in the current period if his standard of living is to remain the same?

Solution:

Computation of Cost of Living Index

Group	Average % increase in price	Group Index (I)	Weights (W)	IW
Food	32	132	15	1980
Clothing	54	154	3	462
Rent	47	147	4	588
Fuel and Light	78	178	2	356
Miscellaneous	58	158	1	158
			Σ W = 25	Σ IW = 3544

Cost of Living Index $(P_{01}) = \frac{\Sigma IW}{\Sigma W} = \frac{3544}{25} = 141.76$

For maintaining the same standard, the business executive should get

$$= \frac{2050 \times 141.76}{100}$$

$$= \text{Rs. } 2906.08$$

Example 41. Construct the consumer price index number for 2000 on the basis of 1999 from the following data using:

(*i*) Family Budget Method

(*ii*) Aggregative Expenditure Method

Commodity	Rice	Wheat	Pulses	Ghee	Oil
Weights :	40	20	15	20	5
Price (per unit) 1999 (Rs.):	16.00	40.00	0.50	5.12	2.00
Price (per unit) 2000 (Rs.):	20.00	60.00	0.50	6.25	1.50

Solution: (*i*) **Family Budget Method**

Constructing Consumer Price Index Number

Commodity	Weights (W)	Price (Rs. per unit) 1999 (p_0)	Price (Rs. per unit) 2000 (p_1)	Price Relatives $\frac{p_1}{p_0} \times 100$ (I)	Weighted Relatives (IW)
Rice	40	16.00	20.00	125	5,000
Wheat	20	40.00	60.00	150	3,000
Pulses	15	0.50	0.50	100	1,500
Ghee	20	5.12	6.25	122	2,440
Oil	5	2.00	1.50	75	375
	$\Sigma W = 100$				$\Sigma IW = 12,315$

Cost of Living Index for 2000 $= \frac{\Sigma IW}{\Sigma W} = \frac{12315}{100}$

$$= \frac{12315}{100}$$

$$= 123.15$$

Thus, there is an increase of 23.15% in prices of 2000 with that of 1999.

(*ii*) **Aggregative Expenditure Method**

Consumer Price Index Number

Commodity	Weights (q_0)	Price (Rs.) 1999 (p_0)	Price (Rs.) 2000 (p_1)	(p_0q_0)	Weighted Relatives (p_1q_0)
Rice	40	16.00	20.00	640.00	800.00
Wheat	20	40.00	60.00	800.00	1200.00
Pulses	15	0.50	0.50	7.50	7.50
Ghee	20	5.12	6.25	102.40	125.00
Oil	5	2.00	1.50	10.00	7.50
	$\Sigma W = 100$			$\Sigma p_0q_0 = 1,559.90$	$\Sigma p_1q_0 = 2,140.00$

Consumer Price Index for 2000 $= \frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times 100 = \frac{2140}{1559.90} \times 100 = 137.188$

Thus, there is increase of 37.188% in prices of 2000 with that of 1999.

Example 42. During a certain period the cost of living index goes up from 110 to 200 and the salary of a worker is also raised from Rs. 500 to 900. Does the worker really gain, and if so, by how much in real terms?

Solution: To maintain the same standard of living, the salary of the worker should be

$$= \frac{\text{Salary at the beginning of period} \times \text{CLI at the end of the period}}{\text{CLI at the beginning of the period}}$$

$$= \frac{500 \times 200}{110} = \text{Rs. } 909.09$$

The actual salary received by the worker at the end of the period = Rs. 900.

Loss = 900 – 909.09 = – 9.09

Thus, the worker really losses. If the worker has got an additional increase of $\frac{909.09 - 900}{900} \times 100 = 1.01\%$, then his standard of living may have remained the same.

Example 43. Mr. Ashok was getting Rs. 400 in the base year and Rs. 800 in the current year. If consumer price index is 350, what extra amount is required for maintaining the earlier standard of living?

Solution: For former standard of living,

Ashok should get $= \frac{400 \times 350}{100} = 1400$

∴ Amount required for maintaining the same standard of living

= Rs. 1,400 – Rs. 800

= Rs. 600

Example 44. In a working class consumer price index of a particular town, the weights according to different groups of items were as follows:

Food 55, Fuel 15, Clothing 10, Rent 12 and Miscellaneous 8. In October 1999, the dearness allowance was fixed by a mill of the town at 182 per cent of workers' wages which fully compensated for the rise in the prices of food and rent but did not compensate for anything else. Another mill of the same town paid dearness allowance of 46.5 per cent which compensated for the rise in fuel and miscellaneous groups. It is known that the rise in food is double the rise in fuel and the rise in miscellaneous group is double the rise in rent.

Find the rise in food, fuel, rent and miscellaneous groups.

Solution: Let rise in fuel be x, $\therefore$ the rise in food is $2x$.

Let rise in rent by y, $\therefore$ rise in Misc. group is $2y$.

The first mill compensated fully for the rise in food and rent but not anything else, by paying 182% D.A.

Items	Index (I)	Weights (W)	$W \times I$
Food	$2x$	55	$110x$
Fuel	100	15	1500
Clothing	100	10	1000
Rent	y	12	$12y$
Misc.	100	8	800
		$\Sigma W = 100$	$\Sigma WI = 3300 + 110x + 12y$

$$\text{Index} = \frac{3300+110x+12y}{100} = 282$$

$$\Rightarrow \quad 110x+12y=24900 \quad \ldots(i)$$

Similarly, the second mill compensated fully for fuel and misc. group by paying 46.5% D.A.

Items	Index (I)	Weights (W)	$W \times I$
Food	100	55	5500
Fuel	x	15	$15x$
Clothing	100	10	1000
Rent	100	12	1200
Misc.	$2y$	8	$16y$
		$\Sigma W = 100$	$\Sigma WI = 7700 + 15x + 16y$

$$\text{Index} = \frac{7700+15x+16y}{100} = 146.5$$

$$\Rightarrow \quad 15x+16y=6950 \quad \ldots(ii)$$

Solving (*i*) and (*ii*) for x and y, we get

$x = 199.37$ and $y = 247.44$

Hence, the rise is as follows:

Food $2x = 2 \times 199.37 = 398.74$

Fuel $x = 199.37$

Rent $y = 247.44$

Misc. $2y = 494.88$

IMPORTANT FORMULAE

1. Chain Base Index:

$$\text{Chain Base Index} = \frac{\text{Average of LR of Current year} \times \text{Chain Base Index of Previous year}}{100}$$

2. Conversion of Chain Base Index to Fixed Base Index:

$$\text{Fixed Base Index} = \frac{\text{Current year's C.B.I.} \times \text{Previous year's F.B.I.}}{100}$$

3. Conversion of Fixed Base Index to Chain Base Index:

$$\text{Chain Base Index Number} = \frac{\text{Current year's F.B.I.}}{\text{Previous year's F.B.I.}} \times 100$$

4. Base Shifting:

$$\text{New Index Number} = \frac{\text{Old Index Number of the Current year}}{\text{Old Index Number of the New year}} \times 100$$

5. Splicing of Index Numbers:

Spliced Index Number

$$= \frac{\text{Index Number of Current year} \times \text{Old Index Number of New Base year}}{100}$$

6. Deflating of Index Number:

$$\text{Real Income} = \frac{\text{Money Income}}{\text{Cost of Living Index Number}} \times 100$$

7. Cost of Living Index Numbers/Cousumer's Price Index

(i) **Aggregative Expenditure Method:**

$$\text{Consumer Price Index} = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$$

(ii) **Family Budget Method:**

$$\text{Consumer Price Index} = \frac{\Sigma PW}{\Sigma W}$$

(iii) **When Group Indices and Weights are given:**

$$\text{Consumer Price Index} = \frac{\Sigma IW}{\Sigma W}$$

QUESTIONS

1. Distinguish between fixed base and chain base methods of constructing index number and give their relative merits.
2. Write a short note on 'Chain Base Index'.
3. Explain the following terms: *(i)* Base Shifting *(ii)* Splicing and *(iii)* Deflating.
4. What is the cost of living index number? Discuss its uses. Give formulae you will use in the construction of cost of living index.
5. Discuss the uses and construction of Consumer Price Index.
6. Using suitable examples, explain the operations of base shifting, splicing and deflating in the context of index numbers.

Time Series Analysis–I

5

INTRODUCTION

Time series has immense importance in the economic and business fields. Most of the economic and business data like prices, national income, population, imports, exports, production, consumption, sales, profits, etc., are collected on the basis of time (such as days, months, years). These data are subject to regular and irregular changes constantly with the passage of time. For example, sometimes prices show upward trend and sometimes the declining trend. For the measurement of such changes in the values of the data over time, the study of time series is extremely useful.

MEANING OF TIME SERIES

The set of data collected on the basis of time (such as days, months, years) is called as time series. In other words, when data are observed on the basis of time (days, months or years), these are known as time series. Under time series, there are two types of variables: *(i)* Independent variable: This represent the time, *(ii)* Dependent variables: This represents changes taking place in the value of data (population, sales, production, etc.) with the passage of time. The following examples illustrate the meaning of time series:

TABLE I		TABLE II	
Year	**Production (in tonnes)**	**Year**	**Population (in crores)**
1975	50	1931	27.9
1976	52	1941	31.9
1977	53	1951	36.1
1978	55	1961	43.9
1979	60	1971	54.5

In the above said table I, production of some commodity is shown between time period 1975 to 1979. Table II contains the population figures of a country in different years. Both of them are the examples of time series.

DEFINITION OF TIME SERIES

Some important definitions of time series are as follows:

1. A set of data depending on the time is called time series. —**Kenny and Keeping**
2. A time series consists of data arranged chronological. —**Croxton and Cowden**
3. A time series is a set of observations taken at speficied times, usually at equal intervals. —**Speigel**

UTILITY OF TIME SERIES

The study of time series has great importance in economic and business world which is illustrated by the following points:

(1) To Study the Past Behaviour of the Data: With the help of time series, changes occurred in the past are studied. Only by analyzing the various sorts of changes occurred in the past, economists and businessmen can frame their present policies by taking advantage of the past experience.

(2) To Forecast Future Behaviour: With the help of time series anticipation of changes going to occur in the future becomes possible because studies about past prove to be very useful for forecasting about future.

(3) Estimation of Trade Cycles: Cyclical fluctuations in a time series give idea about the changes taking place in the business like boom, recession, depression and recovery. Businessman can apply this knowledge to rationalise his course of action, by way of which potential losses can be avoided.

(4) Comparison with other Time Series: Time series analysis is also important for the comparison of various time series. By comparing different time series together, their cause and effect can be more elaborately analyzed.

(5) Study of Present Variations: Time series analysis is also helpful in studying the present variations in different economic variables like national income, export-import, price, output, etc.

(6) Universal Utility: Time series analysis benefits all classes like businessmen, farmers, consumers, economists and government and accordingly they plan and direct their activities.

COMPONENTS OF TIME SERIES

Many types of changes collectively exert influence on time series. Such changes are called as components of time series. A time series has the following four important components:

(1) Secular Trend/Trend - T

(2) Seasonal Variations - S

(3) Cyclical Variations - C

(4) Irregular Variations - I

(1) Secular Trend/Trend-T: Secular trend refers to the general tendency of the data to grow or decline over a long period of time. Any time series shows various fluctuations from time to time but in long period, that series has the increasing or declining trend in one direction. This is secular trend. Secular trend includes or incorporates not short period variations but long-period variations. For example, if we study the industrial output statistics since 1951, we will find that except for one or two years the production of industrial goods has been subject to general rise. Similarly, prices of the commodities, money supply, bank deposits, population, etc., are subject to the upward trend. Similarly, since 1951, there has been persistent decline in death rate per thousand in India. Death rate in some year has risen due to abnormal causes but secular trend is towards decline. Thus, secular trend shows persistent growth or decline in a time series. Quite often, time series exhibit secular trend due to population growth, technological reforms, capital formation, improvements in business organisations, etc.

Secular trend is usually of two types:

(a) **Linear Trend:** When long-term rise or fall in a time series takes place by a constant amount, then that is called a linear trend. This is also known as straight line trend. This is represented by the following equation:

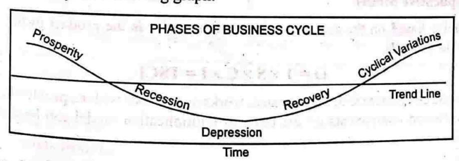
$$Y = a + bX$$

(b) **Parabolic Trend:** The trend is said to be parabolic when long-term rise or fall in a time series is not taking place at a definite rate. It has many forms but most prominent of them is the second Degree Parabolic or Quadratic trend. Its equation is as follows:

$$Y = a + bX + cX^2$$

(2) Seasonal Variations-S: Seasonal variations refer to periodic variations in time series which occur regularly within a period of 12 months. These variations are of regular type and they are easy to be anticipated. Most of the variations taking place in economic and business world are of this type. Prices of foodgrains tend to be low at the time of harvesting and high at sowing time. During winter, woollens clothes rise in price and become cheap during summer. In India, utensils sale shoots up on the occasion of Diwali, sale of essential items tends to be high in the first week of the month and tends to be low in last week of the month. Similarly production, consumption, prices of commodities, interest rates, etc., move up and down all the year due to seasonal variations. Seasonal variations are most perceptible during different months or weeks of the year. Seasonal variations are affected by the climate and customs. During summer, demand for fans, ice, coca-cola tends to be greater compared with other seasons. Demand for college text-books tends to be high during July-August each year. The time period of seasonal variations generally remains definite.

(3) Cyclical Variations-C: Cyclical variations refer to the oscillatory variations in a time series which have a duration anywhere between 2 to 10 years. Cyclical variations arise due to trade cycles. **"Business cycles have been described as successive waves of expansion and contraction that occur at about the same time in many economic activities."—Burns and Michell.** Cyclical variations have four phases (1) Prosperity, (2) Recession, (3) Depression and (4) Recovery. These phases are illustrated by the following graph:



Above graph makes it clear that business activities (Production, Prices, Employment, Sales, etc.) tend to be at their height during prosperity, then recession or the fall starts and gradually, there comes the lowest limit of decline in trade activities. This is the stage of depression. This is always the sequence of cyclical fluctuations but each cycle and its length of phase is different. According to an estimate, a trade cycle takes complete turn from 2 to 10 years and then it repeats itself. Thus, the sequence goes on.

(4) Irregular Variations-I: Irregular variations refer to those short term variations in time series which occur irregularly due to certain accidental causes. These are also known as random fluctuations. Irregular variations take place due to accidental causes like war, earthquake, floods, industrial strikes, etc. They don't happen in definite order or in a systematic way. Their measurement and anticipation becomes very difficult. The decline in industrial output due to the strike in a factory is an example of irregular variations.

ANALYSIS OR DECOMPOSITION OF TIME SERIES

Time series is composed of four components viz. Trend (T), Seasonal variations (S), Cyclical variations (C) and Irregular variations (I). There is always some sort of relationship in these four components. To study the influence of these components on time series, they are measured separately, which is called as Analysis or Decomposition of time series. According to **Prof. Speigel: "The analysis of time series consists in decomposition of a time series into its basic components."**

Models of Analysing Time Series: Analysis of time series is based on two models:

(1) Additive model **(2) Multiplicative model**

(1) Additive Model

This model is based on the assumption that time series is the sum of the four components. According to the formula:

$$O = T + S + C + I$$

This model treats all the constituents as residuals on the basis of which, by deducting trend from the original data, short term fluctuations can be determined. Similarly, cyclical variations and irregular variations can be determined by deducting seasonal variations from short-term variations. On the basis of additive model, the analysis of various components is illustrated as given below:

$$O - T = S + C + I$$
$$O - T - S = C + I$$
$$O - T - S - C = I$$

(2) Multiplicative Model

This model is based on the assumption that a time series is the product of four components. According to the formula:

$$O = T \times S \times C \times I = TSCI$$

Whatever the component is to be separated, works as a divisor with respect to original data (O). Analysis of different components on the basis of multiplicative model can be expressed in the following forms:

$$\frac{O}{T} = SCI$$
$$\frac{O}{T \times S} = CI$$
$$\frac{O}{T \times S \times C} = I$$

Usually, multiplicative model is most often used for time series analysis.

METHODS OF MEASURING TREND

The main methods of measuring trend in a time series are as follows:

(1) **Freehand Curve Method.**
(2) **Semi-Average Method.**
(3) **Moving Average Method.**
(4) **Least Square Method.** Important

Let us study these methods in detail.

(1) Freehand Curve Method

11 12.30 $\boxed{12:30 \text{ to } 2:30}$ 12:30 to 4:30

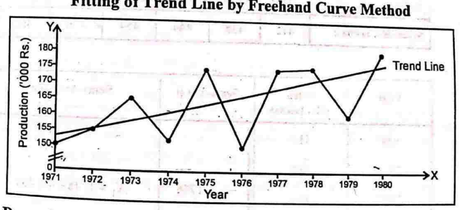
This is the simplest method of trend-fitting. In this method first of all, the original data of the time series is plotted on a graph paper. Thereafter, taking care of the fluctuations of data, a smooth curve is drawn which passes through the mid-points of the fluctuations of time series. Infact, this curve is called as **freehand trend curve.** This method is also called as **trend fitting by inspection.**

The procedure of this method can be illustrated by the following example:

Example 1. Fit a trend line to the following data by the freehand curve method:

Year:	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
Production ('000 Rs.):	150	155	165	152	174	150	174	175	160	180

Solution: **Fitting of Trend Line by Freehand Curve Method**



Merits and Demerits of Freehand Curve Method

Merits:

(i) This method is simple.
(ii) This method is flexible.
(iii) No mathematical formula is used in this method.
(iv) This method is also used for forecasting about future.

Demerits:

(i) This method is based on subjective judgements. So bias may affect the findings.
(ii) There is lack of accuracy in this method.
(iii) Long-run movement (trend) obtained from this method is not definite. This is because a number of curves can be fitted from the same original data.

EXERCISE 5.1

1. Fit a trend line to the following data by freehand curve method:

Year:	1986	1987	1988	1989	1990	1991	1992	1993	1994
Sales ('000 Units):	22	28	24	30	18	26	20	32	16

(2) Semi-Average Method

In this method, first of all time series is divided into two equal parts and thereafter, separate arithmetic mean is calculated for each part. The two values of arithmetic means is plotted on graph corresponding to the time periods. Joining the two points, straight line thus obtained is called as trend line.

The semi-average method can be applied in case of two situations:

(1) When the number of years in a series is even: When the given number of years in a series is even like 4, 6, 8, etc., then the series can be easily divided into two equal parts. In this situation, trend-fitting process can be illustrated with the following example:

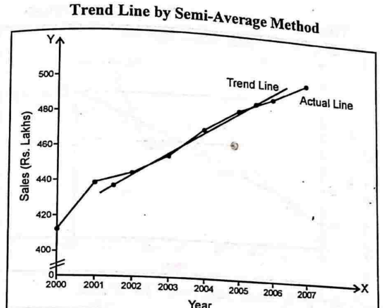
Example 2. Fit a trend line by the method of semi-average to the data given below:

Year:	2000	2001	2002	2003	2004	2005	2006	2007
Sales (Rs. lakhs):	412	438	444	454	470	482	490	500

Solution:

Year	Sales (Rs. lakhs)	Semi-Total	Semi-Average	Middle Year
2000	412			
2001	438	⟶ 1748	$\overline{X}_1 = 1748 \div 4 = 437$	⟶ 2001.5
2002	444			
2003	454			
2004	470			
2005	482	⟶ 1942	$\overline{X}_2 = 1942 \div 4 = 485.5$	⟶ 2005.5
2006	490			
2007	500			

Trend Line by Semi-Average Method



(2) When the number of years in a series is odd: When the number of years in a series is odd like 5, 7, 9, then there will be a problem in dividing the series into two equal parts. In such case, the mid-year figure is to be dropped. For example, if 1981 to 1989 (*i.e.*, 9 years) figures are given, then we will delete 1985, *i.e.*, 5th year and its corresponding figure and we will make 4-4 years' parts, *i.e.*, 1981 to 1984 and 1986 to 1989. The remaining process will be the same as before. Trend fitting in this case can be illustrated by the following example:

Example 3. Fit a trend line by the method of semi-averages to the data given below:

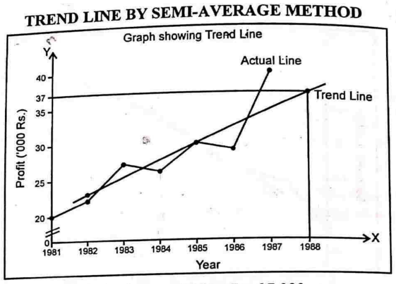
Year:	1981	1982	1983	1984	1985	1986	1987
Profit ('000 Rs.):	20	22	27	26	30	29	40

Also estimate the profit for the year 1988.

Solution: Since there are 7 years, the middle year 1984 will be left out and the arithmetic average of the two parts will be calculated as given below:

Year	Profit ('000 Rs.)	Semi-total	Semi-average	Middle year
1981	20			
1982	22	⟶ 69	$\overline{X}_1 = 69 \div 3 = 23$	⟶ 1982
1983	27			
1984	26		Omitted	
1985	30			
1986	29	⟶ 99	$\overline{X}_2 = 99 \div 3 = 33$	⟶ 1986 / 1985
1987	40			

TREND LINE BY SEMI-AVERAGE METHOD



The estimated profit for the year 1988 = Rs. 37,000.

▶ Merits and Demerits of Semi-Average Method

Merits:

(*i*) This is an easy method.
(*ii*) This method is free from bias.
(*iii*) Trend values thus obtained are definite.
(*iv*) Less time and effort is involved in drawing the trend line.

Demerits:

(*i*) This method is based on straight line trend assumption which does not always hold true.
(*ii*) This method is affected by extreme values.
(*iii*) This method ignores the effect of cyclical fluctuations.

Notwithstanding above said demerits, this method is more suitable in comparison with freehand trend method.

EXERCISE 5.2

1. Fit a trend line by the method of semi-average to the data given below:

Year:	1981	1982	1983	1984	1985	1986	1987
Production ('000 tonnes):	30	26	24	38	40	30	46

2. Fit a trend line to the following data by the semi-average method:

Year:	1983	1984	1985	1986	1987	1988
Profits ('000 Rs.):	80	82	85	70	89	95

3. Fit a trend line for the following data by semi-average method:

Year:	1980	1981	1982	1983	1984	1985	1986	1987
Production ('000 units):	12	14	16	20	20	31	28	15

Also estimate the value for the year 1988.

(3) Moving Average Method

Moving average method is very widely used in practice. Under this method, moving averages are calculated. In moving average computations, one has to decide what moving year average - 3 year, 4 year, 5 year, 7 year should be taken up. The period of moving average depends upon the periodicity of data and there is no specific rule for that. The period is determined by plotting the data on the graph paper and noticing the average time interval of successive peaks or troughs. However, it is essential to consider while selecting the period of moving average that after how many years most of the fluctuations occur in the data. Moving average method is studied in two different situations:

(*i*) **Odd Period Moving Average**
(*ii*) **Even Period Moving Average.**

▶ (*i*) Odd Period Moving Average

When period of the moving average is odd, say 3 years, then following steps are to be taken for the computation of moving average:

(1) First of all, add up the values corresponding to first 3 years in the time series and put the sum before the middle year (*i.e.*, 2nd year).

(2) Thereafter, leaving the first year value, add up second, third and fourth year values and put the sum in front of middle year (*i.e.*, 3rd year). Carry this process further till we reach the last value of the series.

(3) Moving totals thus obtained are to be divided by the period of the moving average and show the trend values of different years.

Similarly, five-yearly, seven-yearly moving averages can be obtained.

The computation procedure of 3-yearly moving averages can be illustrated with the following example:

Example 4. From the following data, calculate trend values using 3-yearly moving average:

Year:	1981	1982	1983	1984	1985	1986	1987
Production:	412	438	446	454	470	483	490

Solution:

Three-yearly Moving Average

Year	Production	Three-yearly moving totals	Three yearly moving average (Trend values)
1981	412	—	—
1982	438	⟶ 412+438+446=1296	1296÷3=432
1983	446	⟶ 438+446+454=1338	1338÷3=446
1984	454	⟶ 446+454+470=1370	1370÷3=457
1985	470	⟶ 454+470+483=1407	1407÷3=469
1986	483	⟶ 470+483+490=1443	1443÷3=481
1987	490	—	—

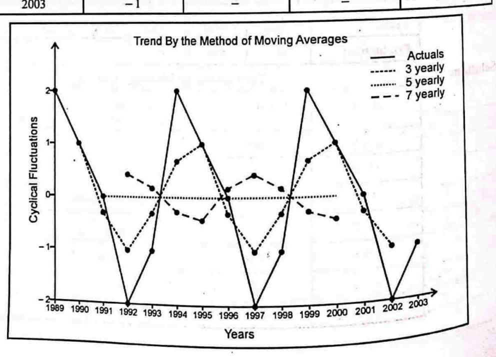
Example 5. From the following data calculate 3-yearly, 5-yearly and 7-yearly moving averages and plot the data on the graph:

Year:	1989	1990	1991	1992	1993	1994	1995	1996	1997
Cyclical fluctuations:	+2	+1	0	−2	−1	+2	+1	0	−2
Year:	1998	1990	2000	2001	2002	2003			
Cyclical fluctuations:	−1	+2	+1	0	−2	−1			

Solution:

Calculation of 3-Yearly, 5-Yearly and 7-Yearly Moving Averages

Year	Cyclical fluctuations	3-yearly moving averages	5-yearly moving averages	7-yearly moving averages
1989	+2	–	–	–
1990	+1	+1.00	–	–
1991	0	−0.33	0	–
1992	−2	−1.00	0	+0.43
1993	−1	−0.33	0	+0.14
1994	+2	+0.67	0	−0.28
1995	+1	+1.00	0	−0.43
1996	0	−0.33	0	+0.14
1997	−2	−1.00	0	+0.43
1998	−1	−0.33	0	+0.14
1999	+2	+0.67	0	−0.28
2000	+1	+1.00	0	−0.43
2001	0	−0.33	0	–
2002	−2	−1.00	–	–
2003	−1	–	–	–



(ii) Even Period Moving Average

When moving average peirod is even, say 4 years, then moving averages have to be centered. It can be computed by two methods:

(a) **First Method**

(b) **Second Method.**

Let us study these methods in detail.

(a) **First Method:** The computation procedure of 4 yearly moving average is as follows:

(i) First of all, add up first 4 values corresponding to the first 4 years and put the sum in between second and third year. Thereafter the next total (*i.e.*, from 2nd to 5th year total) is to be put in between 3rd and 4th year. Carry on this process till the last value of the series.

(ii) Now add up the 1st and 2nd 4-year totals and put them in front of 3rd year. Similarly, add up 2nd and 3rd 4-year total and put them in front of 4th year. Carry on this process till the last value.

(iii) 8 years' totals thus obtained are to be divided by 8. These values are 4-yearly moving averages and show the trend values for different years.

The computation procedure is made clear by the following example:

Example 6. Calculate the trend values using 4-yearly moving average from the following data:

Year:	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Sales (in crore):	7	8	9	11	10	12	8	6	5	10

Solution:

Computation of 4-Yearly Moving Average (Trend) by Centering the Totals

Year (1)	Sales (in crore) (2)	4-yearly moving totals (3)	2 period moving totals of 4-yearly moving totals (4)	4-yearly moving average (centred) (Trend values) (5) = (4) ÷ 8
1970	7	—	—	—
1971	8	—	—	—
		⟶ 7+8+9+11=35		
1972	9		⟶ 35 + 38 = 73	73 ÷ 8 = 9.125
		⟶ 8+9+11+10=38		
1973	11		⟶ 38 + 42 = 80	80 ÷ 8 = 10.00
		⟶ 9+11+10+12=42		
1974	10		⟶ 42 + 41 = 83	83 ÷ 8 = 10.375
		⟶ 11+10+12+8=41		

1975	12		⟶ 41 + 36 = 77	77 ÷ 8 = 9.625
		⟶ 10+12+8+6=36		
1976	8		⟶ 36 + 31 = 67	67 ÷ 8 = 8.375
		⟶ 12+8+6+5=31		
1977	6		⟶ 31 + 29 = 60	60 ÷ 8 = 7.5
		⟶ 8+6+5+10=29		
1978	5	—	—	—
1979	10	—	—	—

***(b)* Second Method:** There is an alternative method of constructing 4-yearly centered moving averages, the procedure of which is given below:

(*i*) First of all, add up the 4-values corresponding to the first 4 years and put the sum in between second and third year. Thereafter, the next total (*i.e.*, from second to fifth year total) is to be put in between 3rd and 4th year. Carry on this process till the last value of the series.

(*ii*) 4-yearly moving totals thus obtained are divided by 4 to obtain 4-yearly uncentered moving averages.

(*iii*) Now add up 1st and 2nd 4-yearly moving averages and divide it by 2. Put this average in front of 3rd year. Similarly, add up 2nd and 3rd 4-yearly moving averages and divide it by 2 and put this average in front of 4th year. Carry on this process till the last value. These values so obtained are 4-yearly centered moving averages and show the trend values for different years.

The computation procedure is made clear from the following examples:

Example 7. Calculate trend values using 4-yearly moving average from the following data:

Year:	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Sales (in crore):	7	8	9	11	10	12	8	6	5	10

Solution:

Computation of 4-yearly Moving Average by Centering the Average

Year	Sales (in crore)	4-yearly moving totals	4-yearly moving average (not centered)	2 period moving total of col. (4)	4-yearly moving average (centred) (Trend values)
(1)	(2)	(3)	(4) = (3) ÷ 4	(5)	(6)
1970	7	—	—	—	—
1971	8	—	—	—	—
		⟶ 35	8.75		
1972	9			⟶ 18.25	9.125
		⟶ 38	9.50		
1973	11			⟶ 20.00	10.00

		⟶ 42	10.50		
1974	10			⟶ 20.75	10.375
		⟶ 41	10.25		
1975	12			⟶ 19.25	9.625
		⟶ 36	9.00		
1976	8			⟶ 16.75	8.375
		⟶ 31	7.75		
1977	6			⟶ 15.00	7.500
		⟶ 29	7.25		
1978	5	—	—	—	—
1979	10	—	—	—	—

Note: Answers obtained by using Method -I and Method -II are the same.

Example 8. The figures of quarterly income of municipal corporation (in Rs. lakhs) for 2 years are given below:

Year:	Ist Quarter	IInd Quarter	IIIrd Quarter	IVth Quarter
1995	74	56	48	69
1996	83	52	49	81

Using a four-quarterly moving average, estimate the trend values.

Solution:

Calculating of Four Quarterly Moving Average

Years	Quarters	Value	4-Quarterly Moving totals	2-Quarterly totals of 4-Quarterly Moving totals	4-Quarterly Moving Average (or Trend values)
1995	I	74			
	II	56			
			⟶ 247		
	III	48		⟶ 503	⟶ 62.875
			⟶ 256		
	IV	69		⟶ 508	⟶ 63.5
			⟶ 252		
1996	I	83		⟶ 505	⟶ 63.125
			⟶ 253		
	II	52		⟶ 518	⟶ 64.75
			⟶ 265		
	III	49			
	IV	81			

Use of Weights in the Calculation of Moving Averages

Weights may also be used in the calculation of moving averages. Its objective is to assign different importance to the values of different years. The following examples illustrates the procedure of computing weighted moving averages.

Example 9. Calculate 3 yearly weighted moving averages with weights, 1, 4 and 1 from the following data:

Year:	1971	1972	1973	1974	1975	1976	1977
Values :	2	6	1	5	3	7	2

Solution:

Computation of 3-yearly Weighted Moving Average

Year (1)	Values (Y) (2)	3-yearly Weighted Moving Totals (3)	ΣW (4)	3-yearly Weighted M.A.* (3) ÷ (4) (5)
1971	2	—		
1972	6	$(2\times1+6\times4+1\times1)=27$	6	$27\div6=4.5$
1973	1	$(6\times1+1\times4+5\times1)=15$	6	$15\div6=2.5$
1974	5	$(1\times1+5\times4+3\times1)=24$	6	$24\div6=4.0$
1975	3	$(5\times1+3\times4+7\times1)=24$	6	$24\div6=4.0$
1976	7	$(3\times1+7\times4+2\times1)=33$	6	$33\div6=5.5$
1977	2	—		

* Weighted Moving Average $= \dfrac{\Sigma WY}{\Sigma W}$

Example 10. For the following data, verify that the 5-yearly weighted moving average with weights, 1, 2, 2, 2, 1 respectively is equivalent to 4 years centered moving average:

Year:	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Sales (Rs. in lakhs)	2	6	1	5	3	7	2	6	4	8

Solution:

Calculation of 5-yearly Weighted Moving Average

Year	Sales	5-yearly Weighted Moving Totals	ΣW	5-yearly Weighted M.A.
1989	2	—	—	
1990	6	—	—	
1991	1	$(1\times2)+2(6+1+5)+(1\times3)=29$	8	$29\div8=3.63$
1992	-5	$(1\times6)+2(1+5+3)+(1\times7)=31$	8	$31\div8=3.88$
1993	3	$(1\times1)+2(5+3+7)+(1\times2)=33$	8	$33\div8=4.13$
1994	7	$(1\times5)+2(3+7+2)+(1\times6)=35$	8	$35\div8=4.38$
1995	2	$(1\times3)+2(7+2+6)+(1\times4)=37$	8	$37\div8=4.63$
1996	6	$(1\times7)+2(2+6+4)+(1\times8)=39$	8	$39\div8=4.88$
1997	4		—	
1998	8		—	

Computation of 4-yearly Centered Moving Average

Year (1)	Sales. (2)	4-yearly moving totals (3)	2-yearly moving totals of Col (3) (4)	4-yearly centered moving average (Trend Values) (5)
1989	2	—	—	
1990	6	→ 14	—	
1991	1	→ 15	→ 29	$29\div8=3.63$
1992	5	→ 16	→ 31	$31\div8=3.88$
1993	3	→ 17	→ 33	$33\div8=4.13$
1994	7	→ 18	→ 35	$35\div8=4.38$
1995	2	→ 19	→ 37	$37\div8=4.63$
1996	6	→ 20	→39	$39\div8=4.88$
1997	4			—
1998	8		—	—

Hence, both the results are the same.

Period of Moving Average

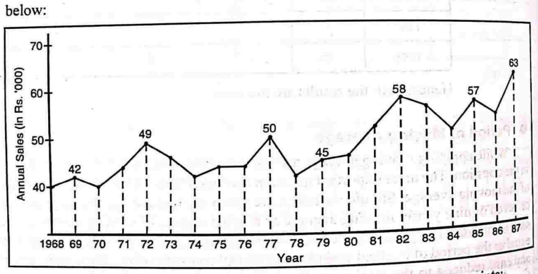
While applying moving average method, sometimes the period of moving average is not given in the question. The most important question that arises here is about the determination of the period of the moving average. Should we take three yearly moving average or four yearly moving average or seven or nine yearly moving average or moving average of some other period, is a question not easy to answer? The general conclusion in this regard is that the period of moving average should be equal to the period of cyclical variations so that all types of cyclical fluctuations are eliminated or in any case reduced to the minimum. But sometimes the period of cycle in a series is not uniform. Sometimes the cycle may complete in five years, at others in seven years and at still others in eight or nine years. Under such circumstances, the average duration of the cycle should be calculated and this should be taken as the period of moving average. The duration of the cycle can be found out by plotting original data on a graph paper and reading the time distances between various peaks or troughs the average of these time distances would give the average duration of the cycle and this should be taken as the period of moving average.

The following example would make the procedure clear:

Example 11. Determining the period of moving average, find trend values by moving averages for the following data:

Year	Annual Sales (in Rs. '000)	Year	Annual Sales (in Rs. '000)
1968	40	1978	42
1969	42	1979	45
1970	40	1980	46
1971	44	1981	52
1972	49	1982	58
1973	46	1983	56
1974	42	1984	51
1975	44	1985	57
1976	44	1986	54
1977	50	1987	63

Solution: We know that the appropriate period of moving average is the period of the cyclic variation. The given data does not reveal a regular cycle of any fixed period. To determine the appropriate period of moving average, we first plot the data as given below:



If we examine the graph carefully, we have the peaks at the following points:

Year:	1969	1972	1977	1979	1982	1985	1987
Peak Value:	42	49	50	45	58	57	63
Period:	3	5	2	3	3	2	

Thus, the data shows 6 cycles with varying periods 3, 5, 2, 3, 3 and 2 respectively. The appropriate period of moving average is given by the arithmetic mean of periods of different cycles shown by the data. Hence the period of the moving average is given by:

$$\text{Average Cyclical Period} = \frac{3+5+2+3+3+2}{6} = \frac{18}{6} = 3$$

Computation of Three Yearly Moving Average

Year	Sales	3-Yearly moving totals	3-Yearly moving averages (trend values)	Year	Sales	3-Yearly moving totals	3-Yearly moving averages (trend values)
1968	40	—	—	1978	42	140	140/3 = 46.66
1969	42	122	122/3 = 40.66	1979	48	136	136/3 = 45.33
1970	40	126	126/3 = 42.00	1980	46	146	146/3 = 48.66
1971	44	133	133/3 = 44.33	1981	52	156	156/3 = 52.00
1972	49	139	139/3 = 46.33	1982	58	166	166/3 = 55.33
1973	46	137	137/3 = 45.66	1983	56	165	165/3 = 55.00
1974	42	132	132/3 = 44.00	1984	51	164	164/3 = 54.66
1975	44	130	130/3 = 43.33	1985	57	162	162/3 = 54.00
1976	44	138	138/3 = 46.00	1986	54	174	174/3 = 58.00
1977	50	136	136/3 = 45.33	1987	63	—	—

Example 12. Determining the period of moving averages, find the trend values by moving average method for the following data:

Year	Value	Year	Value	Year	Value
1	390	6	396	11	459
2	381	7	387	12	438
3	372	8	381	13	435
4	405	9	435	14	492
5	420	10	474	15	510

Solution: Since the peaks of the given data occur at the years 1, 5, 10 and 15, the data exhibits a regular cyclic movement with period 5. Hence the period of the moving average for determining the trend values is also 5 years, viz., the period of the cyclic variations.

Computation of Five Yearly Moving Average

Year	Value	5-Yearly moving totals	5-Yearly moving averages (Trend Values)
1	390	—	—
2	381	—	—
3	372	1968	393.6

4	405	1974	394.8
5	420	1980	396.0
6	396	1989	397.8
7	387	2019	403.8
8	381	2073	414.6
9	435	2136	427.2
10	474	2187	437.4
11	459	2241	448.2
12	438	2298	459.6
13	435	2334	466.8
14	492	—	—
15	510	—	—

Measurement of Short-term Fluctuations

Time series is the mixture of both trend and short-term fluctuations.. Therefore, if trend component is eliminated from the original data, we can find out short-term fluctuations. We can use either additive model or multiplicative model to eliminate trend.

The following examples make clear the measurement of short-term fluctuations:

Example 13. Using three yearly moving averages, determine the trend and short-term fluctuations. Plot the original and trend values on the same graph paper:

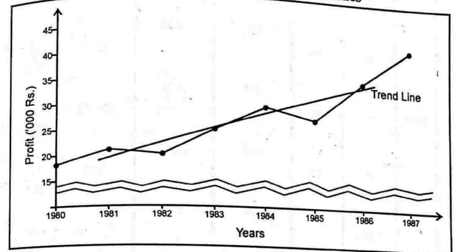
Year:	1980	1981	1982	1983	1984	1985	1986	1987
Profit ('000 Rs.):	18	21	20	25	29	27	35	42

Solution: **Calculation of Trend and Short-term Fluctuations using Additive Model**

Year (1)	Profit ('000 Rs.) (Y) (2)	3-yearly moving totals (3)	3-yearly moving averages (Trend Values–T) (4)	Short-term Fluctuations (Y–T) (5)
1980	18	—	—	—
1981	21	59	19.667	1.333
1982	20	66	22.00	–2
1983	25	74	24.667	0.333
1984	29	81	27.000	2
1985	27	91	30.333	–3.333
1986	35	104	34.667	0.333
1987	42	—	—	—

Deducting trend values from the original series (Y), the residual left is short-term fluctuations. These fluctuations have been shown in above said column (5).

Graph of the Original and Trend Values



Example 14. Eliminate trend by 'moving average method':

Year	Ist Quarter	IInd Quarter	IIIrd Quarter	IVth Quarter
2001	40	35	38	40
2002	42	37	39	38
2003	41	35	38	42

Solution:

Year A	Quarter B	Given Values C	4-quarterly moving totals D	4-quarterly moving average E	4-quarterly moving average centred F	Given figure percentage of moving average or trend eliminated values $G = \frac{C \times 100}{F}$
2001	I	40				
	II	35				
			153	38.25		
	III	38			38.5	$\frac{38}{38.5} \times 100 = 98.7$
			155	38.75		
	IV	40			39	$\frac{40}{39} \times 100 = 102.56$
			157	39.25		
2002	I	42			39.375	$\frac{42 \times 100}{39.375} = 106.66$
			158	39.5		
	II	37			39.25	$\frac{37 \times 100}{39.25} = 94.2675$
			156	39.0		

	III	39			38.875	$\frac{39\times100}{38.875}=100.32$
			155	38.75		
	IV	38			38.5	$\frac{38}{38.5}\times100=98.70$
			153	38.25		
2003	I	41			38.125	$\frac{41}{38.125}\times100=107.54$
			152	38.00		
	II	35			38.5	$\frac{35}{38.5}\times100=90.9$
			156	39.00		
	III	38				
	IV	42				

After eliminating the trend, we are left with short-term fluctuations.

▶ Merits and Demerits of Moving Average Method

Merits:

(*i*) This method is easy to understand and simple to use.

(*ii*) This method is flexible, *i.e.*, if number of years are added in a series, previous calculations are not affected.

(*iii*) This method is most suitable for eliminating cyclical fluctuations.

(*iv*) This method has great practical usefulness.

Demerits:

(*i*) It is difficult to ascertain the proper period of moving averages and if proper period is not ascertained, results will be misleading and inaccurate.

(*ii*) The second defect of this method is that some beginning years' and some terminal years' trend values remain beyond the scope of calculations.

(*iii*) The limitations of arithmetic mean affect this method adversely.

(*iv*) If periodicity in the series is not clearly visible, this method should not be used.

EXERCISE 5.3

1. Find trend values for the following data, by using 5-yearly moving averages. Also plot the actual data and trend values on a graph:

Year:	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Production:	672	679	690	702	712	802	807	809	816	821

[Ans. 691.0, 717.0, 742.6, 766.4, 789.2, 811.0]

2. Calculate trend values by 4-yearly moving average for the following data:

Year:	1981	1982	1983	1984	1985	1986	1987	1988
Production ('000):	13	18	15	21	23	20	24	27

[Ans. 18, 19.5, 20.875, 22.75]

3. Assume a 4-yearly cycle, eliminate the trend values from the following data by moving average method. Use additive model.

Year:	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
Values:	100	105	115	90	95	85	80	65	75	70	75	80

[Ans. 101.88, 98.75, 91.88, 84.38, 78.75, 74.38, 71.88, 73.13
13.12, – 8.75, 3.12, 0.62, 1.25, – 9.38, 3.12, – 3.13]

4. For the following data, verify that the 5-yearly weighted moving average with weights 1, 2, 2, 2, 1 respectively is equivalent to 4 years centered moving averages.

Year:	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Sales (Lakhs Rs.):	5	3	7	6	4	8	9	10	8	9	9

5. Estimate the trend values using the data given below by taking a four quarterly moving average:

Year/Quarter	I	II	III	IV
2002	78	62	56	71
2003	84	64	61	82
2004	92	70	63	85

[Ans. 67.5, 68.5, 69.375, 71.375, 73.750, 75.500, 76.50, 77.125]

6. Determining the period of moving averages, find the trend values by moving averages for the following data:

Year:	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983
Sales ('000 Rs.):	26	29	35	47	51	26	32	37	46	53
Year:	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Sales ('000 Rs.):	25	30	36	46	54	28	31	36	46	54

(Hint: 5-yearly Moving Average) **[Ans.** 37.6, 37.6, 38.2, 38.6, 38.4, 38.8, 38.6, 8.2, 38, 38, 38.2, 38.8, 39, 39, 39, 39]

(4) Least Square Method

This is the best method of trend-fitting in a time series and is most used in practice. This is a mathematical method and a trend line in this method is fitted or obtained in such a way that following two conditions are fulfilled:

(1) $\Sigma(Y - Y_c) = 0$, *i.e.*, the sum of the deviations of the actual values of Y and computed trend values (Y_c) is zero.

(2) $\Sigma(Y - Y_c)^2$ is least, *i.e.*, the sum of the squares of the deviations of the actual and computed trend values from this line is least.

Trend line thus fitted under this method is called as the **Line of Best Fit.**

Least square method can be used to fit **straight line trend** or **parabolic trend** or **exponential trend**.

▶ (A) Fitting of Straight Line Trend

A straight line trend can be expressed by the following equation:

$$Y = a + bX$$

Where, Y=Trend values, X=Unit of time

a is the Y-intercept and b is the slope of the line.

In the above equation, to determine two constants, a and b, the following two normal equations are solved:

$$\Sigma Y = Na + b\,\Sigma X \quad ...(i)$$
$$\Sigma XY = a\,\Sigma X + b\,\Sigma X^2 \quad ...(ii)$$

After determining the equation Y = a + bX, we find the trend values related to different years and plot them on the graph paper which show a straight line trend.

There are two methods of computing straight line trend by using least square method:

(1) Direct Method

(2) Short-cut Method

(1) Direct Method

The procedure to compute straight line trend in this method is as follows:

(i) Any year is taken as the year of origin. Usually first year or before that is taken as zero, deviations of other years are marked on 1, 2, 3 ..., etc. Time deviations are denoted by X:

(ii) Then $\Sigma X, \Sigma Y, \Sigma XY$ and ΣX^2 are computed.

(iii) The values computed are put in the following normal equations:

$$\Sigma Y = Na + b\,\Sigma X \quad ...(i)$$
$$\Sigma XY = a\,\Sigma X + b\,\Sigma X^2 \quad ...(ii)$$

The values of a and b are determined by solving the above said two normal equations.

(iv) Finally, the calculated values of a and b are put in $Y = a + bX$ and trend values are computed. The following example makes clear the procedure of this method:

Example 15. Fit a straight line trend by the method of least square (taking 1978 as year of origin) to the following data:

Year:	1979	1980	1981	1982	1983	1984
Production (lakh tons):	5	7	9	10	12	17

Also obtain the trend values.

Solution:

Fitting of Straight Line Trend

Year	Production (Y)	Deviations from 1978 (X)	XY	X^2
1979	5	1	5	1
1980	7	2	14	4
1981	9	3	27	9
1982	10	4	40	16
1983	12	5	60	25
1984	17	6	102	36
N = 6	$\Sigma Y = 60$	$\Sigma X = 21$	$\Sigma XY = 248$	$\Sigma X^2 = 91$

The straight line trend is defined by the equation:

$$Y = a + bX$$

Two normal equations are

$$\Sigma Y = Na + b\,\Sigma X$$
$$\Sigma XY = a\,\Sigma X + b\,\Sigma X^2$$

Substituting the values, we get

$$60 = 6a + 21b \quad ...(i)$$
$$248 = 21a + 91b \quad ...(ii)$$

Solving the two equations (i) and (ii)

Multiplying (i) by 7 and (ii) by 2 and then subtracting

$$420 = 42a + 147b$$
$$496 = 42a + 182b$$
$$-\quad -\quad -$$
$$-76 = -35b$$
$$b = \frac{-76}{-35} = 2.17$$

By substituting the value of 'b' in equation (i), we get

$$60 = 6a + 21b$$

or $$60 = 6a + 21(2.17)$$

or $$6a = 14.43$$

$\therefore$ $$a = 2.40$$

Hence, the trend equation is

Y = 2.40 + 2.17X; origin = 1978, X unit = 1 Year.

Computation of Trend Values

For 1979, X = 1, Y = 2.40 + 2.17 (1) = 4.57

For 1980, X = 2, Y = 2.40 + 2.17 (2) = 6.74

For 1981, X = 3, Y = 2.40 + 2.17 (3) = 8.91

For 1982, X = 4, Y = 2.40 + 2.17 (4) = 11.08

For 1983, X = 5, Y = 2.40 + 2.17 (5) = 13.25

For 1984, X = 6, Y = 2.40 + 2.17 (6) = 15.42

Example 16. Fit a straight line trend by the method of least square (taking 1980 as year of origin) from the following data:

Year:	1980	1981	1982	1983	1984	1985	1986
Production (in '000 units):	125	128	133	135	140	141	143

Estimate the values for the years 1987 and 1989.

Solution:

Fitting of Straight Line Trend

Year	Production (Y)	Deviations from 1980 (X)	XY	X^2
1980	125	0	0	0
1981	128	1	128	1
1982	133	2	266	4
1983	135	3	405	9
1984	140	4	560	16
1985	141	5	705	25
1986	143	6	858	36
N=7	ΣY=945	ΣX=21	ΣXY=2,922	ΣX^2=91

The straight line trend is given by:

$$Y = a + bX \quad ...(i)$$

Two normal equations are

$$\Sigma Y = Na + b\,\Sigma X$$

$$\Sigma XY = a\,\Sigma X + b\,\Sigma X^2$$

Substituting the values, we get

$$945 = 7a + 21b \quad ...(ii)$$

$$2922 = 21a + 91b \quad ...(iii)$$

Solving the two equations. Multiplying equation (*ii*) by 3, we will get

$$2835 = 21a + 63b \quad ...(iv)$$

Subtracting (*iv*) from (*iii*) we get,

$$2922 = 21a + 91b$$

$$2835 = 21a + 63b$$

$$- \quad - \quad -$$

$$87 = 28b$$

$$b = \frac{87}{28} = 3.107$$

By substituting the value of b in equation (*ii*), we get

From $945 = 7a + 21b$

$$7a = 879.753$$

$\therefore$ $a = 125.679$

Thus, the straight line trend is

$$Y = 125.679 + 3.107X$$

Origin = 1980, X unit = 1 Year.

Estimation for 1987 and 1989

For 1987, X = 7

$$Y = 125.679 + 3.107(7) = 125.679 + 21.749$$

$$Y = 147.428$$

For 1989, X = 9

$$Y = 125.679 + 3.107(9) = 125.679 + 27.963$$

$$Y = 153.642$$

EXERCISE 5.4

1. Fit a straight line trend by the method of least square (taking 1981 as year of origin) to the following data:

Year:	1981	1982	1983	1984	1985
Value:	15	21	25	33	40

Also obtain the trend values.

[Ans. Y = 14.4+6.2X, 14.4, 20.6, 26.8, 33, 39.2]

2. Fit a straight line trend by the method of least square to the following data:

Year:	1969	1970	1971	1972	1973
Sales (in lakhs):	45	56	78	46	75

Also obtain the trend values.
(Take the year 1968 as the working origin)

[Ans. Y = 45+5X; 50, 55, 60, 65, 70]

3. Fit a straight line trend by the method of least square to the following data:

Year:	1980	1981	1982	1983	1984	1985	1986
Production (in '000 Qtls.):	80	90	92	83	94	99	92

(Take the year 1980 as the working origin)

[Ans. Y = 84 + 2X]

(2) Short-Cut Method

The process of computation in this method to find straight line trend is as follows:

(1) First of all, middle-year is taken as the year of origin with value zero and deviations for other years are computed. Sum of the deviations will be zero, *i.e.*, $\Sigma X = 0$. Since deviations above middle-year will be –1, –2, –3, etc., and deviations after middle year will be 1, 2, 3, ..., etc., and deviations above and below middle year will balance out. This is made clear by the following example:

Year:	1952	1953	1954	1955	1956	1957	1958
X:	–3	–2	–1	0	+1	+2	+3

(2) ΣY, ΣXY and ΣX^2 are computed.

(3) For computing the values of a, b, we need not have normal equations but they are found by the following formulae:

$$a = \frac{\Sigma Y}{N}; \quad b = \frac{\Sigma XY}{\Sigma X^2}$$

(4) Finally, the calculated values of a, b are put in the equation $Y = a + bX$ and with its help, trend values are computed.

Short-cut method is studied in two cases:

(*i*) **When number of years is odd.**

(*ii*) **When number of years is even.**

(*i*) When Number of Years is Odd

When number of years is odd like 5, 7, 9..., etc. then the computation of straight line trend can be illustrated with the following examples:

Example 17. Fit a straight line trend by the method of least squares to the following data and also show on graph paper:

Year:	1993	1994	1995	1996	1997	1998	1999
Production (in '000 units):	80	90	92	83	94	99	92

Solution:

Year	Production (Y)	Deviations from 1996 (X)	(XY)	(X^2)
1993	80	–3	–240	9
1994	90	–2	–180	4
1995	92	–1	–92	1
1996	83	0	0	0
1997	94	1	+94	1
1998	99	2	+198	4
1999	92	3	+276	9
N = 7	$\Sigma Y = 630$	$\Sigma X = 0$	$\Sigma XY = 56$	$\Sigma X^2 = 28$

The equation of the straight line trend is:

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$a = \frac{\Sigma Y}{N}$$

$\therefore$

$$b = \frac{\Sigma XY}{\Sigma X^2}$$

Substituting the values, we get

$$a = \frac{\Sigma Y}{N} = \frac{630}{7} = 90$$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{56}{28} = 2$$

Thus, $Y_c = 90 + 2X$

Origin = 1996, X unit = One Year.

Computation of Trend Values

For 1993, $X = -3$, $Y_c = 90 + 2\,(-3) = 84$

For 1994, $X = -2$, $Y_c = 90 + 2\,(-2) = 86$

For 1995, $X = -1$, $Y_c = 90 + 2\,(-1) = 88$

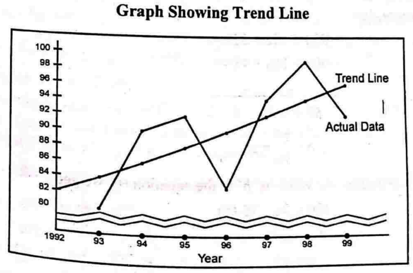
For 1996, $X = 0$, $Y_c = 90 + 2\,(0) = 90$

For 1997, $X = +1$, $Y_c = 90 + 2\,(1) = 92$

For 1998, $X = +2$, $Y_c = 90 + 2\,(2) = 94$

For 1999, $X = +3$, $Y_c = 90 + 2\,(3) = 96$

Graph Showing Trend Line



Example 18. Fit a straight line trend to the following data by the method of least square taking (*i*) 1994 as origin, and (*ii*) 1997 as origin. Estimate the sales for 2002.

Year:	1995	1996	1997	1998	1999
Sales (Rs. lakh):	45	56	78	46	75

Solution: **(*i*) Fitting of Straight Line Trend by Least Square Method (Origin 1994 = 0)**

Year	Sales (Y)	Deviations from 1994 (X)	XY	X^2
1995	45	1	45	1
1996	56	2	112	4
1997	78	3	234	9
1998	46	4	184	16
1999	75	5	375	25
N = 5	$\Sigma Y = 300$	$\Sigma X = 15$	$\Sigma XY = 950$	$\Sigma X^2 = 55$

The equation of the straight line trend is

$$Y = a + bX$$

Two normal equations are

$$\Sigma Y = Na + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

Substituting the values, we get

$$300 = 5a + 15b \quad ...(i)$$

$$950 = 15a + 55b \quad ...(ii)$$

Solving the two equations. Multiplying equation (*i*) by 3 and substracting it from equation (*ii*)

$$950 = 15a + 55b$$

$$900 = 15a + 45b$$

$$- \quad - \quad -$$

$$50 = 10b$$

$$\therefore \quad b = \frac{50}{10} = 5$$

By substituting the value of 'b' in the equation (*i*), we get

$$300 = 5a + 15\,(5)$$

$$5a = 225$$

$$\Rightarrow \quad a = 45$$

Hence the trend equation is :

Y = 45 + 5X, Origin = 1994, X unit = 1 year.

(*ii*) Fitting of Straight Line Trend by Least Square Method (Origin 1997 = 0)

Year	Sales (Y)	Deviations from 1997 (X)	XY	X^2
1995	45	− 2	− 90	4
1996	56	− 1	− 56	1
1997	78	0	0	0
1998	46	+ 1	46	1
1999	75	+ 2	150	4
N = 5	$\Sigma Y = 300$	$\Sigma X = 0$	$\Sigma XY = 50$	$\Sigma X^2 = 10$

The equation of the straight line trend is

$$Y = a + bX$$

$$\text{Since, } \Sigma X = 0,\ a = \frac{\Sigma Y}{N},\ b = \frac{\Sigma XY}{\Sigma X^2}$$

Substituting the values, we get

$$a = \frac{300}{5} = 60; \qquad b = \frac{50}{10} = 5$$

Thus, the straight line trend is

$Y = 60 + 5X$, Origin = 1997, X unit = 1 year.

Estimation of Trend Value for 2000

(a) **Taking 1994 as orgin**

For 2000, X = 6, Y = 45 + 5(6) = Rs. 75 lakh.

(b) **Taking 1997 as origin**

For 2000, X = 3, Y = 60 + 5(3) = Rs. 75 lakh.

Note : It is clear that the value of 'b' remains the same whether we use 1994 as origin or 1997 as origin but the value of 'a' will differ. Change of origin does not affect the value of 'b' which is the slope of the trend line but it affects the value of 'a' (Y-intercept)

(*ii*) When Number of Years is Even

When given number of years is even (6, 10, 12..., etc.), in such a case, the selection of the middle year becomes a problem. In such a case, mean of the two middle years is taken as year of origin and corresponding, deviations for other years are found out. Deviations will be –2.5, – 1.5, – 0.5, 0.5, 1.5 and 2.5. To simplify the computation process, these deviations are divided by $\frac{1}{2}$ or multiplied by 2. The remaining steps are the same as before.

Example 19. Fit a straight line trend by the method of least square to the following data and estimate the profits for the year 1970.

Year:	1961	1962	1963	1964	1965	1966	1967	1968
Profits (Rs. crore):	80	90	92	83	94	99	92	104

Solution:

Fitting of Straight Line Trend

Year	Profits (Y)	Deviations from 1964.5	Deviations Multiplied by 2 (X)	XY	X^2
1961	80	–3.5	–7	–560	49
1962	90	–2.5	–5	–450	25
1963	92	–1.5	–3	–276	9
1964	83	–0.5	–1	–83	1
1965	94	+0.5	1	94	1
1966	99	+1.5	3	297	9
1967	92	+2.5	5	460	25
1968	104	+3.5	7	728	49
N = 8	$\Sigma Y = 734$		$\Sigma X = 0$	$\Sigma XY = 210$	$\Sigma X^2 = 168$

The equation of the straight line trend is:

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore \quad a = \frac{\Sigma Y}{N} \text{ and}$$

$$b = \frac{\Sigma XY}{\Sigma X^2}$$

Substituting the values, we get

$$a = \frac{734}{8} = 91.75$$

$$b = \frac{210}{168} = 1.25 \text{ (} b \text{ gives half-yearly increment)}$$

The straight line trend is:

$Y = 91.75 + 1.25X$; Origin = 1964.5, X unit = $\frac{1}{2}$ year.

Estimation for 1970

For 1970, X = 11,

$$Y = 91.75 + 1.25\ (11)$$

$$= \text{Rs. } 105.5 \text{ lakh}$$

Thus, the estimated profits for the year 1970 are Rs. 105.5 lakh.

Alternative Method: The same result will be obtained if we don't multiply the deviations by 2. But in that case the computation will be more lengthy as could be seen below:

Year	Profits (Y)	Deviations from 1964.5 (X)	XY	X^2
1961	80	– 3.5	– 280	12.25
1962	90	– 2.5	– 225	6.25
1963	92	– 1.5	– 138	2.25
1964	83	– 0.5	– 41.5	0.25
1965	94	+ 0.5	47.0	0.25
1966	99	+ 1.5	148.5	2.25
1967	92	+ 2.5	230	6.25
1968	104	+ 3.5	364	12.25
N = 8	ΣY = 734	ΣX = 0	ΣXY = 105	$\Sigma X^2 = 42$

The equation of the straight line trend is:

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore \quad a = \frac{\Sigma Y}{N} \text{ and } \quad b = \frac{\Sigma XY}{\Sigma X^2}$$

Substituting the values, we get

$$a = \frac{734}{8} = 91.75$$

$$b = \frac{105}{42} = 2.5 \text{ (} b \text{ gives yearly increment)}$$

Thus, Y = 91.75 + 2.5X, Origin = 1964.5, X unit = 1 year.

Estimation for 1970

For 1970, X = 5.5

$$Y = 91.75 + 2.5\ (5.5) = \text{Rs. } 105.5 \text{ lakh.}$$

Note: In the second method, the value of b gives annual increment rather than 6 monthly increment as in the first method discussed above. It is clear from the example that in the first method, the value of b is half of what we obtained from the second method (b was 1.25 in the first case and 2.50 in the second case).

Example 20. Fit a straight line trend by least square method to the data given below. Also find an estimate for the year 2006. What is the annual and monthly increase in sales?

Year:	2000	2001	2002	2003	2004	2005
Sales (Rs. lakhs):	28	32	29	35	40	50

Solution:

Fitting of Straight Line Trend

Year	Sales (Y)	Deviations from 2002.5	Deviations Multiplied by 2 (X)	XY	X^2
2000	28	–2.5	–5	–140	25
2001	32	–1.5	–3	–96	9
2002	29	–0.5	–1	–29	1
2003	35	+0.5	+1	+35	1
2004	40	+1.5	+3	+120	9
2005	50	+2.5	+5	+250	25
N = 6	$\Sigma Y = 214$		$\Sigma X = 0$	$\Sigma XY = 140$	$\Sigma X^2 = 70$

The equation of the straight line trend is:

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore \quad a = \frac{\Sigma Y}{N} \quad \text{and} \quad b = \frac{\Sigma XY}{\Sigma X^2}$$

Substituting the values, we get

$$a = \frac{214}{6} = 35.67$$

$$b = \frac{140}{70} = 2$$

Thus, $Y = 35.67 + 2X$; Origin = 2002.5, X unit $= \frac{1}{2}$ year.

Estimation for 2006

For 2006, $X = +7$

$$Y = 35.67 + 2\,(7)$$
$$= 49.67$$

Thus, the estimated sales for 2006 is Rs. 49.67 lakh.

Annual increase in sales $= 2 \times b = 2 \times 2 = 4$

Monthly increase in sales $= \frac{b}{6} = \frac{2}{6} = 0.33$ lakh

Or $= \frac{\text{Annual Increase}}{12}$

$= \frac{4}{12} = 0.33$ lakh.

IMPORTANT TYPICAL EXAMPLES

o When there is a gap in the given years

Example 21. Below are given the figures of production (in thousand quintals) of a sugar factory:

Year:	1980	1982	1983	1984	1985	1986	1987
Production:	77	88	94	95	91	98	90

(*i*) Fit a straight line trend by the method of least square and also show the trend line on the graph paper.

(*ii*) Eliminate the trend using additive model. What components of the time series are left over?

(*iii*) What is the monthly increase in the production of sugar?

Solution: (*i*) Since there is a gap in the given data. It is not necessary that deviations taken from the middle year would be zero.

Fitting of Straight Line Trend

Year	Production (Y)	X	XY	X^2
1980	77	–4	–308	16
1982	88	–2	–176	4
1983	94	–1	–94	1
1984	95	0	0	0
1985	91	+1	91	1
1986	98	+2	196	4
1987	90	+3	270	9
N = 7	$\Sigma Y = 633$	$\Sigma X = -1$	$\Sigma XY = -21$	$\Sigma X^2 = 35$

The equation of the straight line trend is

$$Y = a + bX$$

Since ΣX is not equal to zero, the values of a and b will be found out by solving the following two normal equations:

$$\Sigma Y = Na + b\Sigma X \quad \ldots(i)$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 \quad \ldots(ii)$$

Substituting the values, we get

$$633 = 7a + (-1)\,b \quad \ldots(iii)$$

$$-21 = -a + 35b \quad \ldots(iv)$$

Multiplying equation (*iv*) by 7 and adding (*iii*) and (*iv*)

$$633 = 7a - b$$

$$\underline{-147 = -7a + 245b}$$

$$486 = 244b$$

$$\Rightarrow \qquad b = \frac{486}{244} = 1.991 = 2 \text{ (approx.)}$$

Putting the value of b in equation (*iii*), we get

$$633 = 7a + (-1)(2) \qquad \Rightarrow \qquad 635 = 7a$$

Thus, $a = \frac{635}{7} = 90.71$

So, the equation of the straight line trend is

$Y = 90.71 + 2X$, Origin = 1984

Computation of Trend Values

For 1980, $X = -4$, $Y = 90.71 + 2(-4) = 82.71$

For 1982, $X = -2$, $Y = 90.71 + 2(-2) = 86.71$

For 1983, $X = -1$, $Y = 90.71 + 2(-1) = 88.71$

For 1984, $X = 0$, $Y = 90.71 + 2(0) = 90.71$

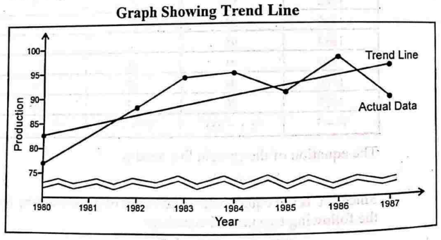
For 1985, $X = +1$, $Y = 90.71 + 2(1) = 92.71$

For 1986, $X = +2$, $Y = 90.71 + 2(2) = 94.71$

For 1987, $X = +3$, $Y = 90.71 + 2(3) = 96.71$

The straight line can be fitted by plotting trend values on the graph paper.

Graph Showing Trend Line



(*ii*) **Elimination of Trend using Additive Model**

Year	Y	Yc	Elimination (Y–Yc)
1980	77	82.71	–5.71
1982	88	86.71	+1.29
1983	94	88.71	+5.29
1984	95	90.71	+4.29
1985	91	92.71	–1.71
1986	98	94.71	+3.29
1987	90	96.71	–6.71

After eliminating the trend, cyclical and irregular components are left over, since seasonal variations will be absent as the data given is annual.

(*iii*) Annual increase in the production of sugar = b = 2,000 quintals.

$\therefore$ Monthly increase in the production of sugar $= \frac{b}{12} = \frac{2{,}000}{12}$

$= 166.67$ quintals.

o To Convert Annual Trend Equation to Monthly/Quarterly Trend Equation (Odd Period)

In certain problems, we require to convert the annual trend equation to monthly trend equation. For converting the annual trend equation to monthly trend equation, divide the value of 'a' by 12; and the value of 'b' by 144. For converting annual trend equation to quarterly trend equation divide the value of a by 4 and the value of b by 16.

Example 22. Below are given the annual production (in thousand tons) of a fertiliser factory:

Year:	1977	1978	1979	1980	1981	1982	1983
Production:	70	75	90	91	95	98	100

(*i*) Fit a straight line trend by the method of least squares and tabulate the trend values.

(*ii*) Convert your annual trend equation into a monthly and quarterly trend equation.

(*iii*) What is the rate of growth of production per month?

Solution: **Fitting of Straight Line Trend**

Year	Production (Y)	Production (X)	XY	X^2
1977	70	–3	–210	9
1978	75	–2	–150	4
1979	90	–1	–90	1
1980	91	0	0	0
1981	95	+1	+95	1
1982	98	+2	+196	4
1983	100	+3	+300	9
N = 7	$\Sigma Y = 619$	$\Sigma X = 0$	$\Sigma XY = 141$	$\Sigma X^2 = 28$

The straight line trend is given by

$$Y = a + bX$$

Since, $\Sigma X = 0$, $\quad a = \frac{\Sigma Y}{N} = \frac{619}{7} = 88.43$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{141}{28} = 5.04$$

$\therefore \quad Y = 88.43 + 5.04X$; Origin = 1980.

Computation of Trend Values

For 1977, X = –3, Y = 88.43 + 5.04 (–3) = 73.31
For 1978, X = –2, Y = 88.43 + 5.04 (–2) = 78.35
For 1979, X = –1, Y = 88.43 + 5.04 (–1) = 83.39
For 1980, X = 0, Y = 88.43 + 5.04 (0) = 88.43
For 1981, X = +1, Y = 88.43 + 5.04 (1) = 93.47
For 1982, X = +2, Y = 88.43 + 5.04 (2) = 98.51
For 1983, X = +3, Y = 88.43 + 5.04 (3) = 103.55

(ii) Monthly Trend Equation

$$Y_c = \frac{a}{12} + \frac{b}{144} X$$

$$\therefore\ Y_c = \frac{88.43}{12} + \frac{5.04}{144} X$$

$$Y_c = 7.369 + 0.035X$$

Quarterly Trend Equation

$$Y_c = \frac{a}{4} + \frac{b}{16} X$$

$$Y_c = \frac{88.43}{4} + \frac{5.04}{16} X$$

$$Y_c = 22.1075 + 0.315X$$

(iii) Rate of growth per month

$$= \frac{b}{12} = \frac{5.04}{12} = 0.42 \text{ thousand tons.}$$

To Convert Annual Trend Equation to Monthly/Quarterly Trend Equation (Even Period)

For converting annual trend equation to monthly trend equation in case of even number of years, we divide the value of 'a' by 12 and the value of 'b' by 6 × 12 (= 72). For converting the annual trend equation into quarterly trend equation, divide the value of a by 4 and the value of b by 8.

Example 23: Fit a straight line trend by the method of least square to the following data :

Year:	2000	2001	2002	2003	2004	2005
Profits (Rs. lakh):	10	20	30	56	40	60

Convert annual trend equation into monthly trend equation. What is the rate of growth of profit per month?

Solution:

Year	Profits (Y)	X	XY	x^2
2000	10	–5	–50	25
2001	20	–3	–60	9
2002	30	–1	–30	1
2003	56	+1	56	1
2004	40	+3	120	9
2005	60	+5	300	25
$N = 6$	$\Sigma Y = 216$	$\Sigma X = 0$	$\Sigma XY = 336$	$\Sigma x^2 = 70$

(i) The equation of the straight line trend is

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore\ a = \frac{\Sigma Y}{N} = \frac{216}{6} = 36$$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{336}{70} = 4.8$$

$$\therefore\ Y = 36 + 4.8X; \quad \text{Origin} = 2002.5\,; \quad X \text{ unit} = \frac{1}{2} \text{ year.}$$

(ii) Monthly Trend Equation will be as:

$$Y_c = \frac{a}{12} + \frac{b}{72} X$$

$$Y_c = \frac{36}{12} + \frac{4.8}{72} X$$

$$\therefore\ Y_c = 3 + 0.066X$$

(iii) Monthly Increase in Profit $= \frac{b}{6} = \frac{4.8}{6} = 0.8$ lakh.

Shifting the Trend Origin

While computing trend values, a certain year (or unit of time) is assumed as point of origin. Sometimes it may be necessary to shift the origin of the trend equation to some other year in the series. For this, there is no need of making all calculations again but there will be the following adjustment to shift the origin of the trend equation:

$$Y = a + b\,(X \pm k)$$

Where 'k' is the number of time units shifted. If the origin is shifted forward in time k will be positive and if origin is shifted backward in time, k will be negative.

Note: Shifting of origin affects the values of 'a' (Y-intercept). The value of 'b' remains the same as the slope of the trend line is the same irrespective of the year of origin.

Example 24. You are given the following trend equation: Y = 45 + 5X (Origin : 1990, X unit = 1 year)
Shift the origin to (i) 1988 (ii) 1993

Solution: Y = 45 + 5X (Origin : 1990)

(i) **Shifting Origin to 1988:** 1988 – 1990 = – 2 shift the origin by 2 years backwards. Replace X by X – 2 in the above trend equation.

$$Y = 45 + 5\,(X - 2)$$
$$= 45 + 5X - 10$$
$$\therefore\ Y = 35 + 5X, \text{ Origin : } 1988$$

(ii) **Shifting Origin to 1993 :** 1993 – 1990 = + 3 shift the origin by 3 years forward. Replace X by X + 3 in the above trend equation.

$$Y = 45+5\,(X+3) = 45 + 5X + 15$$

$$\therefore \quad Y = 60 + 5X, \text{ Origin : } 1993$$

Example 25. Sales of a particular commodity is given as:

Year:	1995	1996	1997	1998	1999
Sales (lakhs Rs.):	45	56	78	46	75

(*i*) Fit a straight line trend to the above data assuming 1997 as the year of origin. Estimate the value for 2000.

(*ii*) How would you shift the year of origin to 1995 in the above problem? Explain.

(*iii*) Convert your annual trend equation into monthly trend equation.

Solution:

Year	Sales (Y)	X	XY	X^2
1995	45	–2	–90	4
1996	56	–1	–56	1
1997	78	0	0	0
1998	46	+1	46	1
1999	75	+2	150	4
N = 5	$\Sigma Y = 300$	$\Sigma X = 0$	$\Sigma XY = 50$	$\Sigma X^2 = 10$

(*i*) The equation of the straight line trend is

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore \quad a = \frac{\Sigma Y}{N} = \frac{300}{5} = 60$$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{50}{10} = 5$$

$$\therefore \quad Y = 60 + 5X, \text{ Origin} = 1997, X \text{ unit} = 1 \text{ year.}$$

Estimation for 2000, X = 3, Y = 60 + 5(3) = Rs. 75 lakh.

(*ii*) **Shifting Origin to 1995:** 1995 – 1997 = –2, shift the origin by 2 years backwards. Replace X by X – 2 in above trend equation.

$$Y = 60 + 5\,(X-2) = 60 + 5X - 10$$

$$\therefore \quad Y = 50 + 5X; \quad \text{Origin} = 1995; \quad X \text{ unit} = 1 \text{ year.}$$

(*iii*) **Monthly Trend Equation will be as follows:**

$$Y = \frac{60}{12} + \frac{5}{144} X = 5 + 0.034X$$

$$\therefore \quad Y = 5 + 0.034X$$

EXERCISE 5.5

Odd Period

1. Fit the straight line trend by method of least square to the following data and show on graph paper:

Year:	1981	1982	1983	1984	1985	1986	1987
Production ('000 tons):	77	88	94	85	91	98	90

[Ans. $Y = 89+2X$; 83, 85, 87, 89, 91, 93, 95]

2. Fit a trend line to the following data by the least square method and estimate the production for 1995 and 2000:

Year:	1990	1992	1994	1996	1998
Production ('000 tons):	18	21	23	27	16

What is the monthly increase in production.

[Ans. $Y = 21 + 0.1X$, $Y_{1995} = 21.1$ ('000 tons); $Y_{2000} = 21.6$ ('000 tons), 0.0083 thousand]

3. Below are figures of sales ('000) tons:

Year:	1991	1992	1993	1994	1995	1996	1997
Sales ('000 tons):	12	10	14	11	13	15	16

Fit a straight line trend and calculate trend values:

[Ans. $Y = 13 + 0.75X$, 10.75, 11.50, 12.25, 13.0, 13.75, 14.50, 15.25]

4. The sales of a particular commodity is given as:.

Year:	1980	1981	1982	1983	1984	1985	1986
Sales ('000 tons):	20	23	22	25	26	29	30

(*i*) Fit a straight line trend to the above data assuming 1983 as the year of origin. Estimate the value for 2000.

(*ii*) How would you shift the year of origin to 1979 in the above problem? Explain.

[Ans. $Y = 25 + 1.642X$; $Y_{2000} = 52.914$; $Y = 18.432 + 1.642X$]

5. Calculate the trend values by the method of least square from the data given below:

Year:	1992	1995	1997	1998	2000	2001	2003
Value:	75	67	68	65	50	54	41

[Ans. $Y = 60 - 3X$, 78, 69, 63, 60, 54, 51, 45]

6. Calculate the trend values by least square method for the following time series:

Year:	1980	1981	1982	1983	1984	1985	1986	1987	1988
Sales (000' Rs.):	53	79	76	66	69	87	79	95	104

Also calculate the trend values by taking 4-years moving average period.

[Hint: See Example 39]

[Ans. (*i*) $Y = 78.67 + 4.65X$, 60.07, 64.72, 69.37, 74.02, 78.67, 83.32, 87.97, 92.62, 97.27

(*ii*) - -70.50, 73.50, 74.875, 78.875, 86.875, - -]

7. The following data give the value of sales of a company for the years 1988 to 1998.

Year (X):	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Sales (in lakh Rs.) (Y):	50.0	36.5	43.0	44.5	38.9	38.1	32.6	38.7	41.7	41.1	33.8

(*i*) Use the method of least square to fit a straight line trend to the data given above. Compute the trend values for 1991 and 1996 (Take $X = 0$ for the year 1993 and the unit of X as 1 year)

(*ii*) Construct a five year moving average and compare the trend values for the year 1991 and 1996.

[**Ans.** (*i*) Linear trend equation : $Y = 39.9 - 0.767X$; $x = X - 1993$
Trend values for 1991 and 1996 are 41.4 (lacs Rs.) and 37.6 (lacs Rs.) respectively.
(*ii*) 5-yearly M.A. values for 1990 to 1996 are (in lakhs Rs.):
42.6, **40.2**, 39.4, 38.6, 38.0, 38.4, **37.6**]

8. Use the method of least square to fit a straight line trend for the following data. (Take 1986 as origin).

Year:	1984	1985	1986	1987	1988
Production (in '000 Qtls.):	28	38	46	40	56

Also compute the trend values. What are the annual and monthly increase in production?
[**Ans.** $Y = 41.6 + 5.8X$; 30, 35.8, 41.6, 47.4, 53.2; Annual Increase = 5.8 thousands Qtls.
Monthly Increase = 0.483 thousand Qtls.]

Even Period

9. Fit a straight line trend by the method of least square. Also estimate the production for the year 1995. What is the monthly increase in production?

Year:	1989	1990	1991	1992	1993	1994
Production (in lakh tonnes):	25	40	47	59	68	80

[**Ans.** $Y = 53.17 + 5.3X$, 90.27, 0.883]

10. Fit a linear trend to the following time series by the method of least square and also obtain the trend values:

Year:	1954	1955	1956	1957	1958	1959
Production (in crore of Rs.):	7	10	12	14	17	24

[**Ans.** $Y = 14 + 1.54X$; 6.3, 9.38, 12.46, 15.54, 18.62, 21.7]

11. Fit a straight line trend by the method of least square to the following data:

Year:	1988	1989	1990	1991	1992	1993
Sales (in lakhs):	12	18	26	35	29	42

Estimate the sales for the year 1995.

[**Ans.** $Y = 27 + 2.74X$; $Y_{1995} = 51.66$]

12. The weights of a calf taken at weekly intervals are given below. Fit a straight line using the method of least square and calculate the average rate of growth per week.

Age (*t*):	1	2	3	4	5	6	7	8	9	10
Weight (*Y*):	52.5	58.7	65.0	70.2	75.4	81.1	87.2	95.5	102.2	108.4

Also estimate the weight of the calf when age is (*i*) 12 (*ii*) 15 weeks.
[**Ans.** (*i*) $Y = 79.62 + 3.08X$, (*ii*) $2 \times 3.08 = 6.16$, (*iii*) 119.66, 138.14]

Gap in the Given Years

13. Below are given figures of production (in thousand tons) of a sugar factory:

Year:	1986	1988	1989	1990	1991	1992	1995
Production:	77	88	94	85	91	98	90

(*i*) Fit a straight line trend by the least square method and tabulate the trend values.
(*ii*) Eliminate the trend. What components of the time series are thus left over?
(*iii*) What is the monthly increase in the production of sugar?

[**Ans.** $Y = 88.803 + 1.38X$; 83.28, 86.04, 87.42, 88.80, 90.18, 91.56, 95.73;
Monthly increase = 0.115 thousand qtl.]

14. The following time series shows the cost/unit (Rs.) of a product for the period 1981–1995:

Years:	1981	1982	1984	1985	1987	1990	1991	1992	1993	1995
Cost/Unit (Rs.):	332	317	357	392	402	405	410	427	405	438

Calculate the trend values by the least square method and estimate the cost per unit for 1997.
[**Ans.** $Y = 388.5 + 7.509X$; $Y_{1997} = 456.081$]

Conversion of Annual Trend Equation to Monthly/Quarterly Trend Equation

15. The following data relate to sales of Reliance Ltd.:

Year:	2000	2001	2002	2003	2004
Sales (Rs. lakhs):	40	80	120	200	160

(*i*) Fit a straight line trend by the method of least square and tabulate the trend values.
(*ii*) Eliminate the trend using additive model. What components of the time series are thus left over?
(*iii*) Estimate the likely sales for the year 2006.
(*iv*) What is half yearly, quarterly and monthly increase in the sales?
(*v*) Convert the trend equation into (*a*) on monthly basis, (*b*) on quarterly basis.
(*vi*) Shift the origin (*a*) to 2004, (*b*) to 2000.

[**Ans.** (*i*) $Y = 120 + 36X$; (*ii*) only cyclical and irregular variations,
(*iii*) $Y_{2006} = 264$, (*iv*) 18, 9, 3 (*v*) (*a*) $Y = 10 + 0.25X$, (*b*) $Y = 30 + 2.25X$
(*vi*) (*a*) $Y = 192 + 36X$, (*b*) $Y = 48 + 36X$]

Measurement of Short Term Fluctuations/Elimination of Trend

16. Following data relate to sales of Bharat Ltd.

Year:	2000	2001	2002	2003	2004	2005
Sales (Rs. lakhs):	10	20	30	56	40	60

(*i*) Fit a straight line trend by the method of least square and tabulate the trend values.
(*ii*) Eliminate the trend using additive model. What components of the time series are thus left over?
(*iii*) Estimate the likely sales for the year 2006.
(*iv*) What is annual increase in the sales?
(*v*) What is monthly increase in the sales?
(*vi*) By what year the company's expected sales would have equalled to its target of 84 lakhs.

[Hint: See Example 38]

[**Ans.** (*i*) Y = 36 + 4.8X (*ii*) – 2, – 1.6, – 1.2, 15.2, – 10.4, O, C & I are left over. (*iii*) 69.6 (*iv*) 4.8 × 2 = 9.6 (*v*) 0.8 lakh (*vi*) Target is expected to be attained in the year 2007.5 (*i.e.*, 2002.5 + 10/2)]

17. Fit a straight line trend to the following data by the method of least square after summing the quarterly data to yearly data:

Year	Q_1	Q_2	Q_3	Q_4
1993	10	13	14	12
1994	12	14	15	13
1995	13	15	18	14
1996	15	18	21	18
1997	15	22	23	20

Also find out short-term fluctuations for the given years using additive model.

[Hint: See Example 36]

[**Ans.** Y= 63+8X; Trend: 47, 55, 63, 71, 79; Short-term fluctuations: 2, –1, –3, 1, 1]

Shifting of Origin

18. Production of a particular commodity is given as:

Year:	1997	1998	1999	2000	2001	2002	2003
Production of Steel (in tonnes):	60	72	75	65	80	85	95

(*i*) Fit a straight line trend to the above data assuming 2000 as the year of origin. Estimate the value for 2004.
(*ii*) How would you shift the year of origin to 1997 in the above problem? Explain.
(*iii*) Convert your annual trend equation to monthly trend equation.

[**Ans.** (*i*) Y = 76 + 4.857X, Y_{2004} = 95.428, (*ii*) Y = 61.429 + 4.857X: Origin 1997; (*iii*) Y = 6.33 + 0.0337X]

19. The trend of the annual sales of ABC Co. Ltd. is described by the following equation:
Y = 30 + 3.6X, Origin = 2001, X unit = 1 year, Y unit = annual sales.
Convert the annual trend equation into monthly trend equation. [**Ans.** Y = 2.5 + 0.025X]

20. Given the following trend equation:
Y_t = 84.26 + 5.8 X (origin = 1978, X unit = 1 year)
Shift the origin to (*i*) 1985 and (*ii*), 1969

[**Ans.** (*i*) Y = 124.86 + 5.8X, Origin : 1985 (*ii*) Y = 32.06 + 5.8 X, Origin: 1969]

21. The annual trend equation is:
Y = 50 + 2X , Origin: 2000, X unit = 1 year.
Shift the origin to 1997. [**Ans.** Y = 44+2X, Origin: 1997]

22. The trend of annual sales of a company is described by the following equation:
Y = 15 + 0.5X
(Origin: 1987, X unit = 1 year, Y unit = Annual Production)
Convert the equation to quarterly trend equation. [**Ans.** Y = 3.75 + 0.03125X]

(B) Fitting of Second Degree Parabolic Trend or Quadratic Trend

There may be many such situations in economic and business fields in which a straight line trend may not represent the long-term tendency of the time series data. In such cases, a second degree parabolic trend or quadratic trend is fitted. The equation of the second degree parabolic trend or the quadratic trend is:

$$Y = a + bX + cX^2$$

Where, a is the Y-intercept, b is the slope of the curve at the origin and c is the rate of change in slope.

Under the method of least square, the values of the constants a, b and c are obtained by solving the following three normal equations:

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$
$$\Sigma X^2Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

Short-cut Method: If the time deviations are taken from the middle year (or arithmetic mean of two middle years), the values of ΣX and ΣX^3 would be zero (i.e. $\Sigma X = 0$ and $\Sigma X^3 = 0$) and the equations are reduced to the following:

$$\Sigma Y = Na + c\Sigma X^2 \qquad \text{...(i)}$$
$$\Sigma XY = b\Sigma X^2 \qquad \text{...(ii)}$$
$$\Sigma X^2Y = a\Sigma X^2 + c\Sigma X^4 \qquad \text{...(iii)}$$

When further solved, these equations gives the following values of the constants a, b and c.

From equation (*ii*) we get $b = \dfrac{\Sigma XY}{\Sigma X^2}$

From equation (*i*), we get $a = \dfrac{\Sigma Y - c\Sigma X^2}{N}$

and from equation (*iii*), we get $c = \dfrac{N\Sigma X^2Y - (\Sigma X^2)(\Sigma Y)}{N \cdot \Sigma X^4 - (\Sigma X^2)^2}$

After finding the values of a, b and c in the above manner, the trend equation can be fitted to obtain the trend values of the given time series by substituting the respective values of X therein.

Note: In practice, short-cut method is widely used for fitting second degree parabolic trend.

Procedure

The computation of second degree parabolic trend/quadratic trend involves the following steps:

(*i*) Find the time deviation of each year from the middle year and denote it by 'X'.

(*ii*) Then $\Sigma X, \Sigma Y, \Sigma XY, \Sigma X^2Y, \Sigma X^2, \Sigma X^3$ and ΣX^4 are computed.

(*iii*) The values computed are put in the above formulae of b, a and c.

(*iv*) Finally, the calculated values of a, b and c are put in $Y = a + bX + cX^2$ and trend values are computed.

Type I: Odd Number of Years

Example 26. Fit a second degree parabolic trend ($Y = a + bX + cX^2$) to the following data:

Year:	1983	1984	1985	1986	1987
Production (in crore Rs.):	5	7	4	9	12

Also compute the trend values. Predict the value for 1988.

Solution: **Fitting of Second Degree Parabolic Trend**

Year	Production (Y)	Deviations from 1985 (X)	XY	X^2Y	X^2	X^3	X^4
1983	5	–2	–10	20	4	–8	16
1984	7	–1	–7	7	1	–1	1
1985	4	0	0	0	0	0	0
1986	9	+1	9	9	1	+1	1
1987	12	+2	24	48	4	+8	16
N = 5	$\Sigma Y = 37$	$\Sigma X = 0$	$\Sigma XY = 16$	$\Sigma X^2Y = 84$	$\Sigma X^2 = 10$	$\Sigma X^3 = 0$	$\Sigma X^4 = 34$

The second degree parabolic trend is given by the equation

$$Y = a + bX + cX^2$$

The three normal equations are:

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$
$$\Sigma X^2Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

Substituting the values, we get

$$37 = 5a + b(0) + c(10) \Rightarrow 37 = 5a + 10c \quad \text{....}(i)$$
$$16 = a(0) + b(10) + c(0) \Rightarrow 16 = 10b \quad \text{....}(ii)$$
$$84 = a(10) + b(0) + c(34) \Rightarrow 84 = 10a + 34c \quad \text{....}(iii)$$

From (*ii*) equation, $16 = 10b \Rightarrow b = \dfrac{16}{10} = 1.6$

Multiplying (*i*) by 2 and subtracting it from (*iii*)

$$84 = 10a + 34c$$
$$74 = 10a + 20c$$
$$\overline{10 = \quad 14c}$$
$$\therefore c = \frac{10}{14} = 0.71$$

Putting the value of c in (*i*), we get

$$37 = 5a + 10(0.71)$$
$$\Rightarrow \quad 5a = 29.9$$
$$\therefore \quad a = 5.98 \approx 6.0$$

Thus, the second degree parabolic trend is

$Y = 6 + 1.6X + 0.71X^2$, Origin = 1985, X unit = 1 year

Computation of Trend Values

For 1983, $X = -2$, $Y = 6 + 1.6(-2) + 0.71(-2)^2 = 5.64$

For 1984, $X = -1$, $Y = 6 + 1.6(-1) + 0.71(-1)^2 = 5.11$

For 1985, $X = 0$, $Y = 6 + 1.6(0) + 0.71(0)^2 = 6.00$

For 1986, $X = +1$, $Y = 6 + 1.6(1) + 0.71(1)^2 = 8.31$

For 1987, $X = +2$, $Y = 6 + 1.6(2) + 0.71(2)^2 = 12.04$

Predicted value for 1988 is given by:

For 1988, $X = +3$, $Y = 6 + 1.6(3) + 0.71(3)^2 = 17.19$

Type II: Even Number of Years

Example 27. The price of a commodity during 1994-99 are given below. Fit a second degree parabola to the following data. Calculate the trend values and estimate the price of the commodity in 2001.

Year:	1994	1995	1996	1997	1998	1999
Price:	100	107	128	140	181	192

Solution:

Fitting of Second Degree Parabola

Year	Price (Y)	Deviations from 1996.5	Deviations multiplied by 2 (X)	XY	X^2Y	X^2	X^3	X^4
1994	100	–2.5	–5	–500	2500	25	–125	625
1995	107	–1.5	–3	–321	963	9	–27	81
A 1996	128	–0.5	–1	–128	128	1	–1	1
1997	140	+0.5	+1	+140	140	1	+1	1
1998	181	+1.5	+3	+543	1629	9	+27	81
1999	192	+2.5	+5	+960	4800	25	+125	625
N = 6	$\Sigma Y = 848$		$\Sigma X = 0$	ΣXY =694	ΣX^2Y =10160	$\Sigma X^2 = 70$	$\Sigma X^3 = 0$	ΣX^4 =1414

The second degree parabolic trend is given by the equation

$$Y = a + bX + cX^2$$

The three normal equations are:

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$
$$\Sigma X^2Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

Substituting the values, we get

$$848 = 6a + b(0) + 70c \quad \Rightarrow 848 = 6a + 70c \qquad ...(i)$$
$$694 = a(0) + b(70) + c(0) \Rightarrow 694 = 70b \qquad ...(ii)$$
$$10160 = 70a + b(0) + 1414c \Rightarrow 10160 = 70a + 1414c \qquad ...(iii)$$

From equation (*ii*), $694 = 70b \Rightarrow b = \frac{694}{70} = 9.914$

Multiplying (*i*) by 70 and (*iii*) by 6 and subtracting (*i*) from (*iii*)

$$60960 = 420a + 8484c$$
$$59360 = 420a + 4900c$$
$$\overline{1600 = \qquad 3584c}$$

$$\therefore \qquad c = \frac{1600}{3584} = 0.4464$$

Putting the value of *c* in (*i*), we get

$$848 = 6a + 70\,(0.4464)$$
$$\Rightarrow \qquad 848 = 6a + 31.248$$
$$\therefore \qquad a = \frac{816.752}{6} = 136.125$$

Thus, the second degree parabola is

$$Y = 136.125 + 9.914X + 0.4464X^2, \qquad \text{Origin : 1996.5,} \qquad \text{X unit} = \frac{1}{2}\text{ year.}$$

Computation of Trend Values

For 1994, $X = -5$, $Y = 136.125 + 9.914(-5) + 0.4464(-5)^2 = 97.715$

For 1995, $X = -3$, $Y = 136.125 + 9.914(-3) + 0.4464(-3)^2 = 110.4006$

For 1996, $X = -1$, $Y = 136.125 + 9.914(-1) + 0.4464(-1)^2 = 126.6574$

For 1997, $X = +1$, $Y = 136.125 + 9.914(1) + 0.4464(1)^2 = 146.4854$

For 1998, $X = +3$, $Y = 136.125 + 9.914(3) + 0.4464(3)^2 = 169.884$

For 1999, $X = +5$, $Y = 136.125 + 9.914(5) + 0.4464(5)^2 = 196.855$

Predicted value for 2001:

For 2001, $X = +9$, $Y = 136.125 + 9.914(9) + 0.4464(9)^2 = 261.5094$

Thus, the likely price of the commodity for the year 2001 is 261.5094.

EXERCISE 5.6

1. Fit a parabolic trend $Y = a + bX + cX^2$ to the following data:

Year:	1981	1982	1983	1984	1985	1986	1987	1988	1989
Output (in '000 units):	2	6	7	8	10	11	11	10	9

Also compute the trend values. Predict the value for 1990.
(Take the year 1985 as working origin)

[Ans. $Y = 10.02 + 0.85X - 0.27X^2$, **Trend values: 2.3, 5.04, 7.24, 8.9, 10.02, 10.60, 10.64, 10.14, 9.1,** $Y_{1990} = 7.52$**]**

2. The prices of a commodity during 1981–1986 are given below. Fit a second degree parabola to the following data. Calculate the trend values and estimate the price of the commodity in 1987.

Year:	1981	1982	1983	1984	1985	1986
Price:	110	114	120	138	152	218

[Ans. $Y = 124.15 + 9.6X + 1.53X^2$, **Trend values: 114.40, 109.12, 116.08, 135.28, 166.72, 210.40,** $Y_{1987} = 266.32$**]**

3. Fit a second degree parabola ($Y = a + bX + cX^2$) for the following data:

X:	1	2	3	4	5
Y:	25	28	33	39	46

[Hint: See Example 40]

[Ans. $Y = 22.78 + 1.46X + 0.64X^2$**]**

4. Fit a parabolic curve of the second degree to the data given below and estimate the value for 1999.

Year:	1993	1994	1995	1996	1997	1998
Sales (in '000 Rs.):	10	12	13	10	8	11

[Ans. $Y = 11.031 - 0.143X - 0.3125X^2$, $Y_{1999} = 1995.5$]

5. Sales of a particular commodity is given as:

Year:	1981	1982	1983	1984	1985	1986	1987
Sales ('000 tons):	13	13	22	21	54	60	83

(i) Fit a second degree parabola ($Y = a + bX + cX^2$) to the above data assuming 1984 as the year of origin and estimate the value for 1988.

(ii) How would shift the year of origin to 1980 in the above problem? Explain.

[Ans. (i) $Y = 30 + 12X + 2X^2$, Origin: 1984; 110 (ii) $Y = 14 - 4X + 2X^2$]

6. Calculate trend values for the following data using second degree equation:

Year:	1984	1985	1986	1987	1988	1989	1990
Output ('000 tons):	100	107	128	140	181	192	200

Also make a forecast for 1993.

[Ans. $Y = 149.10 + 18.68X + 0.16X^2$, Trend values: 94.49, 112 38, 130.58, 149.10, 167.94, 187.10, 206.58, Forecasted value = 266.94]

(C) Fitting of Exponential Trend

If the time series is increasing or decreasing by a constant percentage rather than constant absolute amount, the fitting of exponential trend is considered appropriate. Such tendency is found in many economic and business data.

The equation of the exponential trend is

$$Y = ab^X$$

where a is Y-intercept and b the slope of the curve at the origin of X.

In the logarithmic form, the above equation is written as under:

$$\log Y = \log a + X \log b$$

When plotted on a semi-logarithmic graph, the curve gives a straight line (or called logarithmic straight line). However, on an arithmatic scale chart, the curve gives a non-linear trend.

Under the method of least square, the values of the constants a and b are obtained by solving the following two normal equations:

$$\Sigma \log Y = N \cdot \log a + \log b \Sigma X$$

$$\Sigma(X \log Y) = \log a \Sigma X + \log b \Sigma X^2$$

Short-cut Method: If the time deviations are taken from the middle year (or arithmetic mean of two middle years), the value of ΣX would be zero (*i.e.*, $\Sigma X = 0$) and the normal equations are reduced to the following:

$$\Sigma \log Y = N \cdot \log a \quad \text{or} \quad \log a = \frac{\Sigma \log Y}{N} \quad \Rightarrow a = \text{Antilog}\left[\frac{\Sigma \log Y}{N}\right] \quad ...(i)$$

$$\Sigma(X \log Y) = \log b \Sigma X^2 \quad \text{or} \quad \log b = \frac{\Sigma(X \log Y)}{\Sigma X^2} \quad \Rightarrow b = \text{Antilog}\left[\frac{\Sigma X \log Y}{\Sigma X^2}\right] \quad ..(ii)$$

After obtaining the values of a and b in the above manner and substituting their values in the equation, we can fit the trend equation under this method and with such an equation we can obtain the trend values of the time series and predict the values for the future year.

Note: The annual rate of growth in case of exponential trend is obtained as $r = (\text{Antilog } b - 1) \times 100$.

Procedure

The computation of exponential trend involves the following steps:

(i) Find the time deviation of each year from the middle year and denote it by 'X'.

(ii) Obtain the logarithms of the variable Y.

(iii) Multiply $\log Y$ by the corresponding time deviation 'X' and obtain $X \log Y$.

(iv) Square up the time deviation X and obtain ΣX^2.

(v) The values computed are put in the above formulae of $\log a$ and $\log b$.

(vi) Finally in calculated values of $\log a$ and $\log b$ are put in $\log Y = \log a + X \log b$.

(vii) Take the antilogs of these logs to arrive at the actual trend values.

Type I: Odd Number of Years

Example 28. Fit an exponential trend $Y = ab^X$ to the following data:

Year:	1990	1991	1992	1993	1994	1995	1996
Sales (Rs. crore):	12	10	14	18	20	24	30

Estimate the sales for the year 1997.

Solution:

Fitting of Exponential Trend

Year	Sales (Y)	X	$\log Y$	$X \log Y$	X^2
1990	12	–3	1.0792	–3.2376	9
1991	10	–2	1.0000	–2.0000	4
1992	14	–1	1.1461	–1.1461	1
1993	18	0	1.2553	0	0
1994	20	+1	1.3010	1.3010	1
1995	24	+2	1.3802	2.7604	4
1996	30	+3	1.4771	4.4313	9
N = 7		$\Sigma X = 0$	$\Sigma \log Y = 8.6389$	$\Sigma X \log Y = 2.1090$	$\Sigma X^2 = 28$

The equation of the exponential trend is

$$Y = ab^X$$

Taking logarithms of both sides, we have

$$\log Y = \log a + X \log b \quad ...(i)$$

$$...(ii)$$

Since, $\Sigma X = 0$

$$\therefore \quad \log a = \frac{\Sigma \log Y}{N} = \frac{8.6389}{7} = 1.2341$$

$$\log b = \frac{\Sigma X \log Y}{\Sigma X^2} = \frac{2.1090}{28} = 0.075$$

Thus, the equation of the exponential trend in logarithmic form is:

$$\log Y = 1.2341 + 0.075X \quad \text{or} \quad Y_c = \text{Antilog } [1.2341 + 0.075X]$$

Here, origin = 1993, X unit = 1 year.

Estimation for 1997

For 1997, $X = +4$, $Y_c = \text{Antilog } [1.2341 + 0.075X]$

$$\Rightarrow Y_c = \text{Antilog } [1.5341] = 34.20 \text{ crore}$$

Thus, the estimated sales for 1997 is 34.20 crore.

Example 29. The sales of a company (in lakhs of Rs.) for the seven years are given below:

Year:	1990	1991	1992	1993	1994	1995	1996
Sales:	32	47	65	88	132	190	275

Find out the trend values by using the equation $Y = ab^X$ and annual rate of growth.

Solution: **Fitting of Exponential Trend**

Year	Sales (Y)	X	log Y	X log Y	X^2
1990	32	−3	1.5052	−4.5153	9
1991	47	−2	1.6721	−3.3442	4
1992	65	−1	1.8129	−1.8129	1
1993	88	0	1.9445	0	0
1994	132	+1	2.1206	2.1206	1
1995	190	+2	2.2788	4.5576	4
1996	275	+3	2.4393	7.3179	9
N = 7		$\Sigma X = 0$	$\Sigma \log Y = 13.7733$	$\Sigma X \log Y = 4.3237$	$\Sigma X^2 = 28$

The equation of the exponential trend is

$$Y = ab^X$$

In the logarithms form, the equation is written as

$$\log Y = \log a + X \log b$$

Since, $\Sigma X = 0$

$$\therefore \quad \log a = \frac{\Sigma \log Y}{N} = \frac{13.7733}{7} = 1.9676$$

$$\log b = \frac{\Sigma X \log Y}{\Sigma X^2} = \frac{4.3237}{28} = 0.154$$

$$\Rightarrow \quad b = \text{Antilog } [0.154] = 1.4256$$

Thus, the exponential trend equation in logarithmic form is:

$\log Y = 1.9676 + 0.154X$ Origin: 1993; X unit = 1 year

Computation of Trend Values

For 1990, $X = -3$, $Y_c = \text{Antilog}[1.9676 + 0.154(-3)] = \text{Antilog}[1.5056] = 32.03$

For 1991, $X = -2$, $Y_c = \text{Antilog}[1.9676 + 0.154(-2)] = \text{Antilog}[1.6596] = 45.66$

For 1992, $X = -1$, $Y_c = \text{Antilog}[1.9676 + 0.154(-1)] = \text{Antilog}[1.8136] = 65.10$

For 1993, $X = 0$, $Y_c = \text{Antilog}[1.9676 + 0.154(0)] = \text{Antilog}[1.9676] = 92.81$

For 1994, $X = +1$, $Y_c = \text{Antilog}[1.9676 + 0.154(1)] = \text{Antilog}[2.1216] = 132.31$

For 1995, $X = +2$, $Y_c = \text{Antilog}[1.9676 + 0.154(2)] = \text{Antilog}[2.2756] = 188.62$

For 1996, $X = +3$, $Y_c = \text{Antilog}[1.9676 + 0.154(3)] = \text{Antilog}[2.4296] = 268.91$

Computation of Annual Growth Rate

$$r = (\text{Antilog } b - 1) \times 100 \quad \text{or} \quad r = (b - 1) \times 100$$

$$= [\text{Antilog}(0.154) - 1] \times 100 \qquad = (1.4256 - 1) \times 100$$

$$= (1.4256 - 1) \times 100 = 0.4256 \times 100 = 42.56\% \qquad = 42.56\%$$

⚬ Type II: Even Number of Years

Example 30. Fit an exponential trend ($Y = ab^X$) to the following data:

Year:	1941	1951	1961	1971	1981	1991
Population of India (in crore):	31.9	36.1	43.9	54.8	68.3	84.4

Also predict the population for 2001.

Solution: **Fitting of Exponential Trend**

Year	Population (Y)	X	log Y	X log Y	X^2
1941	31.9	−5	1.5038	−7.5190	25
1951	36.1	−3	1.5575	−4.6725	9
1961	43.9	−1	1.6425	−1.6425	1
1971	54.8	+1	1.7388	1.7388	1
1981	68.3	+3	1.8344	5.5032	9
1991	84.4	+5	1.9263	9.6315	25
N = 6		$\Sigma X = 0$	$\Sigma \log Y = 10.2033$	$\Sigma X \log Y = 3.0395$	$\Sigma X^2 = 70$

The equation of the exponential trend is

$$Y = ab^X$$

In the logarithmic form, the equation is written as

$$\log Y = \log a + X \log b$$

Since, $\Sigma X = 0$

$$\therefore \quad \log a = \frac{\Sigma \log Y}{N} = \frac{10.2033}{6} = 1.70$$

$$\log b = \frac{\Sigma X \log Y}{\Sigma X^2} = \frac{3.0395}{70} = 0.043$$

Thus, the exponential trend equation in the logarithmic form is:

$\log Y = 1.70 + 0.043X$; Origin: 1966; X unit = 5 years.

or $Y = \text{Antilog}\,[1.70 + 0.043X]$

Further $\log a = 1.70 \Rightarrow a = \text{Antilog}\,[1.70] = 50.12$

$\log b = 0.043 \Rightarrow b = \text{Antilog}\,[0.043] = 1.10$

Thus, the equation of the exponential trend is

$$Y = 50.12(1.10)^X$$

Prediction for 2001

For 2001, $X = +7$, $Y = \text{Antilog}\,[1.70 + 0.043(7)]$

$= \text{Antilog}\,[2.001] = 100.2$ crore.

Example 31. Fit a logarithmic straight line to the following data:

Year:	1991	1992	1993	1994	1995	1996
Production ('000 tonnes):	64	70	75	82	88	95

Solution:

Fitting of Logarithmic Straight Line

Year	Production (Y)	X	$\log Y$	$X \log Y$	X^2
1991	64	–5	1.8062	–9.031	25
1992	70	–3	1.8451	–5.5353	9
1993	75	–1	1.8751	–1.8751	1
1994	82	1	1.9138	1.9138	1
1995	88	3	1.9445	5.8335	9
1996	95	5	1.9777	9.8885	25
N = 6		$\Sigma X = 0$	$\Sigma \log Y = 11.3624$	$\Sigma X \log Y = 1.1944$	$\Sigma X^2 = 70$

The equation of the exponential trend is

$$Y = ab^X$$

In the logarithmic form, the equation is written as

$$\log Y = \log a + X \log b \qquad \text{[Logarithmic Straight Line Form]}$$

Since, $\Sigma X = 0$

$$\therefore \quad \log a = \frac{\Sigma \log Y}{N} = \frac{11.3624}{6} = 1.8937$$

$$\log b = \frac{\Sigma X \log Y}{\Sigma X^2} = \frac{1.1944}{70} = 0.017$$

Thus, the equation of the logarithmic straight line is:

$\log Y = 1.8937 + 0.017X$; Origin: 1993.5; X unit $= \frac{1}{2}$ year.

or $Y = \text{Antilog}\,[1.8937 + 0.017X]$

Merits and Demerits of Least Square Method

Merits:

(*i*) This method is far better than moving average method because the trend values for all the years are obtained. Not even a single initial or terminal trend values is left over in this method.

(*ii*) It results in a mathematical equation which may be used for forecasting.

(*iii*) It is widely used method of fitting a curve to the given data. The results obtained are reliable and appropriate.

Demerits:

(*i*) The computation process in this method is complex which is not easily understandable.

(*ii*) This method does not have the attribute of flexibility. If some figures are added to or subtracted from the original data, all computations have to be redone.

(*iii*) It is difficult to select an appropriate type of equation in this method. Results based on inappropriate selection of equation are likely to be misleading.

EXERCISE 5.7

1. Fit an exponential trend $Y = ab^X$ to the following data and calculate the trend values. Also estimate the trend for 1992.

Year:	1985	1986	1987	1988	1989
Sales (Rs. crore):	100	105	112	120	130

[**Ans.** $\log Y = 2.0527 + 0.0286X$ or $Y = 112.90(1.07)^X$,
Trend values: 98.97, 105.71, 112.90, 120.59, 128.80,
Sales for 1992 = 156.93 crore]

2. The population figures of India is given below:

Census Year:	1911	1921	1931	1941	1951	1961	1971
Population (in crore):	25.0	25.1	27.9	31.9	36.1	43.9	54.7

Fit an exponential trend $Y = ab^X$ to the above data by the method of least square and find the trend values. Estimate the population in 1981. **[Ans.** $Y = 33.60\ (1.142)^X$, $Y_{1981} = 57.15$]

3. In the following exponential trend equation (origin: 2001, X-unit= 1 year, Y - annual profits) shift the origin to 2005.

$$Y_c = 15(1.7)^X$$

[Ans. $Y = 12.528\ (1.7)^X$, Origin 2005]

4. The consumption of electricity in the agriculture sector during the period 1991–1999 was recorded as under:

Year:	1991	1992	1993	1994	1995	1996	1997	1998	1999
Consumption ('000 lakh tonnes):	50	65	70	85	82	75	65	90	95

Find : (*i*) The second degree polynomial equation.
(*ii*) The exponential trend equation.

[Ans. (*i*) $Y_t = 77.4 + 3.9X + (-0.32)X^2$; (*ii*) $Y_t = (70.13)\ (1.053)^X$]

MISCELLANEOUS SOLVED EXAMPLES

Example 32. Calculate trend values from the following data by using four yearly moving average:

Year	Value	Year	Value
1972	41	1979	60
1973	61	1980	67
1974	55	1981	73
1975	48	1982	78
1976	53	1983	76
1977	67	1984	84
1978	62		

Solution: Four yearly moving averages can be obtained as follows:

Year (1)	Y (2)	4-yearly moving totals (3)	2 period moving total of cols. (3) (4)	4-yearly moving average (Trend values) (5) = (4) ÷ 8
1972	41	—	—	—
1973	61	—	—	—
1974	55	205	422	52.75
1975	48	217	440	55.00
1976	53	223	453	56.625
1977	67	230	472	59.00
1978	62	242	498	62.25
1979	60	256	518	64.75
1980	67	262	540	67.50
1981	73	278	572	71.50
1982	78	294	605	75.625
1983	76	311	—	—
1984	84	—	—	—

Example 33. Determine the trend for the following data by
(*i*) moving average of length 4, and
(*ii*) the least square method.

Year	Quarters			
	I	II	III	IV
1991	210	191	216	200
1992	218	197	230	211
1993	246	215	235	225

Solution: (*i*)

Moving Average of Length 4

Year	Quarter	Given Figures	Four-figure moving total	Two-figure moving total	Four-figure moving Average
1991	I	210			
	II	191			
			817		
	III	216		1642	$\frac{1642}{8} = 205.25$
			825		
	IV	200		1656	207.00
			831		
1992	I	218		1676	209.50
			845		
	II	197		1701	212.625
			856		
	III	230		1740	217.50
			884		
	IV	211		1786	223.25
			902		
1993	I	246		1809	226.125
			907		
	II	215		1828	228.50
			921		
	III	235			
	IV	225			

(ii)

Least Square Method

Year/ Quarter	t	Y	$X=\frac{t-6.5}{0.5}$	X^2	XY	Trend Value
1991 I	1	210	−11	121	−2310	200.88
II	2	191	−9	81	−1719	203.66
III	3	216	−7	49	−1512	206.44
IV	4	200	−5	25	−1000	209.22
1992 I	5	218	−3	9	−654	212.00
II	6	197	−1	1	−197	214.78
III	7	230	1	1	230	217.56
IV	8	211	3	9	633	220.34
1993 I	9	246	5	25	1230	223.12
II	10	215	7	49	1505	225.90
III	11	235	9	81	2115	228.68
IV	12	225	11	121	2475	231.46
		$\Sigma Y=2594$	$\Sigma X=0$	$\Sigma X^2=572$	$\Sigma XY=796$	

Let $Y=a+bX$ be the trend equation, where a and b be are calculated from the normal equations.

$$\Sigma Y=Na+b\Sigma X \quad ...(i)$$

$$\Sigma XY=a\Sigma X+b\Sigma X^2 \quad ...(ii)$$

Putting respective values from the above tables, we get

$$2594=12a+0$$

or $$a=\frac{2594}{12}=216.17$$

Again $$796=0+b(572)$$

or $$b=\frac{796}{572}=1.39$$

$\therefore$ Trend equation is $Y=216.17+1.39X$ ...(iii)

For finding trend values we put $X=-11,-9,.....$, in order and the corresponding trend values have been shown in the last column of the above table.

Example 34. From the given data, compute 'trend' and short-term fluctuations by moving average method (Periodicity = 4 years)

Years:	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Production: (in Qtls.)	462	510	525	470	515	550	575	560	580	620	645

Solution:

Computation of Trend

Year	Y	4-yearly moving totals	2-period moving total of cols.(3) (centered)	4-yearly moving average (centered)
(1)	(2)	(3)	(4)	(5)
1980	462	—	—	—
1981	510	1967	—	—
1982	525	2020	3987	498.375
1983	470	2060	4080	510.000
1984	515	2110	4170	521.250
1985	550	2200	4310	538.750
1986	575	2265	4465	558.125
1987	560	2335	4600	575.000
1988	580	2405	4740	592.500
1989	620	—	—	—
1990	645	—	—	—

Computation of Short-term fluctuations

The short-term fluctuations for the given years are obtained by deducting the trend values computed by using 4-yearly moving average from the actual values of the series using additive model.

Years	Actual Values (Y)	Trend Values (Yc)	Short-term fluctuations
1980	462	—	—
1981	510	—	—
1982	525	498.375	26.625
1983	470	510.000	−40.00
1984	515	521.250	−6.25
1985	550	538.750	11.25
1986	575	558.125	16.875
1987	560	575.000	−15.00
1988	580	592.500	−12.50
1989	620	—	—
1990	645	—	—

Example 35. Calculate the quarterly trend values by the method of least square for the following quarterly data for five years given below:

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1994	60	80	72	68
1995	68	104	100	88
1996	80	116	108	96
1997	108	152	136	124
1998	160	184	172	164

Solution:

Calculation of Annual Trend by the Method of Least Square

Year	Yearly Total	Yearly Average (Y)	Deviations from Mid-Year i.e., 1996 X	XY	X^2	Trend Value Y_c
1994	280	70	−2	−140	4	64
1995	360	90	−1	−90	1	88
1996	400	100	0	0	0	112
1997	520	130	+1	130	1	136
1998	680	170	+2	340	4	160
N = 5		$\Sigma Y = 560$	$\Sigma X = 0$	$\Sigma XY = 240$	$\Sigma X^2 = 10$	

Equation of straight line is

$$Y = a + bX$$

$$\therefore \quad a = \frac{\Sigma Y}{N} = \frac{560}{5} = 112$$

and $$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{240}{10} = 24$$

Trend Equation $Y_c = 112 + 24X$; Origin = 1996 ...(*i*)

The trend values have been obtained by the equation (*i*).

Yearly increment = 24

$$\therefore \quad \text{Quarterly increment} = \frac{24}{4} = 6$$

Calculation of Quarterly Trend Values

Now, we calculate the quarterly trend values. Consider the year 1994. Trend value is 64. This is the value for the middle of the year 1994, *i.e.*, middle second and third quarter. The quarterly increment is 6. Therefore, the trend value for the second quarter of 1994 would be 64 – 3 or 61 and for the third quarter it would be 64 + 3 or 67. The value for the first quarter of 1994 would be 61 – 6 = 55 and for the last quarter 67 + 6 = 73. Similarly, trend values of the various quarters of other years can be calculated. These values are tabulated below:

Quarterly Trend Values

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1994	55	61	67	73
1995	79	85	91	97
1996	103	109	115	121
1997	127	133	139	145
1998	151	157	163	169

Example 36. Fit a straight line trend to the following data by the method of least square after summing the quarterly data to yearly data:

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1983	10	13	14	12
1984	12	14	15	13
1985	13	15	18	14
1986	15	18	21	18
1987	15	22	23	20

Also find out short-term fluctuations for the given years using additive model.

Solution:

Year	Yearly total (Y)	X	XY	X^2
1983	49	−2	−98	4
1984	54	−1	−54	1
1985	60	0	0	0
1986	72	+1	72	1
1987	80	+2	160	4
N = 5	$\Sigma Y = 315$	$\Sigma X = 0$	$\Sigma XY = 80$	$\Sigma X^2 = 10$

The equation of the straight line trend is

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$\therefore \quad a = \frac{\Sigma Y}{N} = \frac{315}{5} = 63$$

and $$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{80}{10} = 8$$

Thus, $Y_c = 63 + 8X$; Origin = 1985; X unit = 1 year.

Computation of Trend Values

For 1983, $X = -2$, $Y_c = 63 + 8(-2) = 47$

For 1984, $X = -1$, $Y_c = 63 + 8(-1) = 55$

For 1985, $X = 0$, $Y_c = 63 + 8(0) = 63$

For 1986, $X = 1$, $Y_c = 63 + 8(1) = 71$

For 1987, $X = 2$, $Y_c = 63 + 8(2) = 79$

Calculation of Short-term fluctuations

Year	Y	Yc	Short-term fluctuations using additive model $(Y - Y_c)$
1983	49	47	+2
1984	54	55	−1
1985	60	63	−3
1986	72	71	+1
1987	80	79	+1

Example 37. Fit a straight line trend by the method of least square and estimate the value for 2001

Year:	1951	1961	1971	1981	1991
Population (crore):	34	50	67	75	85

Solution:

Fitting of Straight Line Trend

Year	Y	X	XY	X^2
1951	34	−20	−680	400
1961	50	−10	−500	100
1971	67	0	0	0
1981	75	+10	750	100
1991	85	+20	1700	400
N = 5	$\Sigma Y = 311$	$\Sigma X = 0$	$\Sigma XY = 1270$	$\Sigma X^2 = 1000$

The straight line trend is given by

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$a = \frac{\Sigma Y}{N} \quad \text{and} \quad b = \frac{\Sigma XY}{\Sigma X^2}$$

$$a = \frac{311}{5} = 62.2; \quad b = \frac{1270}{1000} = 1.27$$

$\therefore$ $Y = 62.2 + 1.27X$; Origin = 1971

Estimation for 2001

For 2001, $X = +30$, $\therefore$ $Y = 62.2 + 1.27\,(30)$

$Y_{2001} = 100.3$ crore

Example 38. Following data relate to sales of Bharat Ltd.

Year:	2000	2001	2002	2003	2004	2005
Sales (Rs. lakhs):	10	20	30	56	40	60

(i) Fit a straight line trend by the method of least square and tabulate the trend values.

(ii) Eliminate the trend using additive model. What components of the time series are thus left over?

(iii) Estimate the likely sales for the year 2006.

(iv) What is annual increase in the sales?

(v) What is monthly increase in the sales?

(vi) By what year the company's expected sales would have equalled to its target of 84 lakh.

Solution: (i) **Fitting of Striaght Line Trend**

Year	Deviations from 2002.5	X	Sales X^2	Y	XY	Trend Values $Y_c = a + bX$	Short-term fluctuations $Y - Y_c$
2000	−2.5	−5	25	10	−50	12.0	−2.0
2001	−1.5	−3	9	20	−60	21.6	−1.6
2002	−0.5	−1	1	30	−30	31.2	−1.2
2003	+0.5	1	1	56	56	40.8	15.2
2004	+1.5	3	9	40	120	50.4	−10.4
2005	+2.5	5	25	60	300	60.0	0
N = 6		$\Sigma X = 0$	$\Sigma X^2 = 70$	$\Sigma Y = 216$	$\Sigma XY = 336$		$\Sigma(Y - Y_c) = 0$

The straight line trend is given by

$$Y = a + bX$$

Since, $\Sigma X = 0$

$$a = \frac{\Sigma Y}{N} = \frac{216}{6} = 36$$

and

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{336}{70} = 4.8$$

Hence, the annual trend equation is given by:

$$Y = 36 + 4.8X$$

[Origin = 2002.5, X unit = Half year, Y unit = Annual sales in lakh of Rs.]

Trend values for different years are shown in the 7th column of the table.

(ii) Eliminates trend values as shown in the 8th column of the table. After eliminating the trend only cyclical and irregular variations are left since seasonal variations are absent as the annual data is given.

(iii) Likely sales for the year 2006 : For 2006, $X = 7$

$$Y_{2006} = 36 + 4.8 \times 7 = 69.6$$

(iv) Annual increase in sales $= 4.8 \times 2 = 9.6$

(v) Monthly increase in the sales $= \frac{b}{6} = \frac{4.8}{6} = 0.8$ lakh

or $\frac{\text{Anual Increase}}{12} = \frac{9.6}{12} = 0.8$ lakh

(vi) $84 = 36 + 4.8X \Rightarrow X = \frac{(84-36)}{4.8} = 10$ half years from origin (*i.e.*, 2002.5)

Hence, target is expected to be attained in the year 2007.5 (*i.e.*, 2002.5 + 10/2).

Example 39. The production of cement by a firm is given below:

Year:	1980	1981	1982	1983	1984	1985	1986	1987	1988
Production (tonnes):	4	5	5	6	7	8	9	8	10

Calculate the trend values by taking 3-yearly moving average.

Also calculate the trend values of the above data by least square method.

Solution: (i) **Calculation of 3-yearly moving average**

Year	Production (Y)	3-yearly moving totals	3-yearly moving average (trend)
1980	4	—	—
1981	5	14	4.67
1982	5	16	5.33
1983	6	18	6.00
1984	7	21	7.00
1985	8	24	8.00
1986	9	25	8.33
1987	8	27	9.00
1988	10	—	—

(ii) **Fitting of Linear Trend Equation**

Year	Y	X	XY	X^2
1980	4	−4	−16	16
1981	5	−3	−15	9
1982	5	−2	−10	4
1983	6	−1	−6	1
1984	7	0	0	0
1985	8	+1	+8	1
1986	9	+2	+18	4
1987	8	+3	+24	9
1988	10	+4	+40	16
N = 9	$\Sigma Y = 62$	$\Sigma X = 0$	$\Sigma XY = 43$	$\Sigma X^2 = 60$

The linear trend equation is given by

$$Y = a + bX$$

Since $\Sigma X = 0$, $\therefore\ a = \frac{\Sigma Y}{N} = \frac{62}{9} = 6.89$; $b = \frac{\Sigma XY}{\Sigma X^2} = \frac{43}{60} = 0.72$

Thus, $Y = 6.89 + 0.72X$; Origin: 1984; X unit = 1 year.

Computation of Trend Values

For 1980, $X = -4$, $Y = 6.89 + 0.72(-4) = 4.01$

For 1981, $X = -3$, $Y = 6.89 + 0.72(-3) = 4.73$

For 1982, $X = -2$, $Y = 6.89 + 0.72(-2) = 5.45$

For 1983, $X = -1$, $Y = 6.89 + 0.72(-1) = 6.17$

For 1984, $X = 0$, $Y = 6.89 + 0.72(0) = 6.89$

For 1985, $X = +1$, $Y = 6.89 + 0.72(1) = 7.61$

For 1986, $X = +2$, $Y = 6.89 + 0.72(2) = 8.33$

For 1987, $X = +3$, $Y = 6.89 + 0.72(3) = 9.05$

For 1988, $X = +4$, $Y = 6.89 + 0.72(4) = 9.77$

Example 40. Fit the equation of the form $Y = a + bX + cX^2$ to the data given below:

X:	1	2	3	4	5
Y:	25	28	33	39	46

Solution: **Fitting of Parabolic Trend**

X	Y	t $t = X-3$	t^2	t^3	t^4	tY	t^2Y
1	25	−2	4	−8	16	−50	100
2	28	−1	1	−1	1	−28	28
3	33	0	0	0	0	0	0
4	39	1	1	1	1	39	39
5	46	2	4	8	16	92	184
N = 5	$\Sigma Y = 171$	$\Sigma t = 0$	$\Sigma t^2 = 10$	$\Sigma t^3 = 0$	$\Sigma t^4 = 34$	$\Sigma tY = 53$	$\Sigma t^2Y = 351$

Let the second degree trend equation between Y and t be:

$$Y = a + bt + ct^2 \text{ where } t = X - 3$$

The three normal equations are:

$\Sigma Y = Na + b\Sigma t + c\Sigma t^2 \Rightarrow 171 = 5a + 10c$ ...(i)

$\Sigma tY = a\Sigma t + b\Sigma t^2 + c\Sigma t^3 \Rightarrow 53 = 10b$ ...(ii)

$\Sigma t^2Y = a\Sigma t^2 + b\Sigma t^3 + c\Sigma t^4 \Rightarrow 351 = 10a + 34c$ ...(iii)

From (ii), $b = \frac{53}{10} = 5.3$

Multiplying (*i*) by 2 and subtracting from (*iii*), we get

$$351 = 10a + 34c$$
$$342 = 10a + 20c$$
$$9 = 14c \Rightarrow c = \frac{9}{14} = 0.64$$

Putting the value of c in (*i*), we get

$$a = \frac{171 - 10 \times 0.64}{5} = \frac{171 - 6.4}{5} = 32.92$$

Thus, $Y = 32.92 + 5.3t + 0.64\,t^2$ where $t = X - 3$

Hence, the second degree trend equation of Y on X becomes

$$Y = 32.92 + 5.3\,(X - 3) + 0.64\,(X - 3)^2$$
$$Y = 32.92 + 5.3X - 15.9 + 0.64(X^2 - 6X + 9)$$
$$Y = 22.78 + 1.46X + 0.64X^2$$

Example 41. Find average of quarterly trend values for the given years from the data given below:

Year	Quarters I	II	III	IV
1990	4	6	5	4
1991	5	8	6	6
1992	3	5	2	3
1993	6	9	4	5
1994	6	8	5	5

Solution: We compute the following table to find the trend equation:

Year	Yearly total	Quarterly average (Y)	Deviations from 1992 (X)	XY	X^2
1990	19	4.75	−2	−9.5	4
1991	25	6.25	−1	−6.25	1
1992	13	3.25	0	0.00	0
1993	24	6.00	1	6.00	1
1994	24	6.00	2	12.00	4
N = 5		$\Sigma Y = 26.25$	$\Sigma X = 0$	$\Sigma XY = 2.25$	$\Sigma X^2 = 10$

The values of a and b with equation $Y = a + bX$ are

$$\therefore \quad a = \frac{\Sigma Y}{N} = \frac{26.25}{5} = 5.25$$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{2.25}{10} = 0.225$$

Therefore, $Y_c = 5.25 + 0.225X$; Origin = 1992; X unit = 1 year.

Y_c denotes average of quarterly trend values.

Year:	1990	1991	1992	1993	1994
Average of quarterly trend values :	4.800	5.025	5.250	5.475	5.700

Example 42. The total annual fertiliser consumption in tonnes during 1995-2001 in XYZ village of Karnataka state was recorded as given below:

Year:	1995	1996	1997	1998	1999	2000	2001
Consumption:	50	56	60	68	70	75	78

(*i*) Fit a straight-line trend by the method of least squares and compute the trend quantities.

(*ii*) What has been the annual increase in fertiliser consumption?

(*iii*) Eliminate the trend variations from the fertiliser consumption data using multiplicative method.

Solution: (*i*)

Fitting of Straight Line Trend

Year	Consumption (Y)	X	XY	X^2
1995	50	−3	−150	9
1996	56	−2	−112	4
1997	60	−1	−60	1
1998	68	0	0	0
1999	70	+1	+70	1
2000	75	+2	+150	4
2001	78	+3	+234	9
N = 7	$\Sigma Y = 457$	$\Sigma X = 0$	$\Sigma XY = 132$	$\Sigma X^2 = 28$

The straight line trend is given by

$$Y = a + bX$$

Since, $\Sigma X = 0$, $a = \dfrac{\Sigma Y}{N} = \dfrac{457}{7} = 65.29$

$$b = \frac{\Sigma XY}{\Sigma X^2} = \frac{132}{28} = 4.71$$

$\therefore \quad Y = 65.29 + 4.71X$; Origin = 1998; X unit = 1 year.

Computation of Trend Values:

For 1995, X = – 3, Y = 65.29 + 4.71 (– 3) = 51.16
For 1996, X = – 2, Y = 65.29 + 4.71 (– 2) = 55.87
For 1997, X = – 1, Y = 65.29 + 4.71 (– 1) = 60.58
For 1998, X = 0, Y = 65.29 + 4.71 (0) = 65.29
For 1999, X = 1, Y = 65.29 + 4.71 (1) = 70.00
For 2000, X = 2, Y = 65.29 + 471 (2) = 74.71
For 2001, X = 3, Y = 65.29 + 4.71 (3) = 79.42

(*ii*) Annual increase in fertiliser consumption = *b* = 4.71 tonnes.

(*iii*) **Elimination of Trend Using Multiplicative Model**

Year	Y	Trend (T)	Elimination of Trend / Detrended values
1995	50	51.16	$\frac{50}{51.16}\times 100 = 97.73$
1996	56	55.87	$\frac{56}{55.87}\times 100 = 100.23$
1997	60	60.58	$\frac{60}{60.58}\times 100 = 99.04$
1998	68	65.29	$\frac{68}{65.29}\times 100 = 104.15$
1999	70	70.00	$\frac{70}{70}\times 100 = 100$
2000	75	74.71	$\frac{75}{74.71}\times 100 = 100.39$
2001	78	79.42	$\frac{78}{79.42}\times 100 = 98.21$

IMPORTANT FORMULAE

1. Components of Time Series:

(*i*) Secular Trend or Trend – T
(*ii*) Seasonal Variations – S
(*iii*) Cyclical Variations – C
(*iv*) Irregular Variations – I

2. Models of Analysis of Time Series:

(*i*) **Additive Model:**

O = T + S + C + I

(*ii*) **Multiplicative Model:**

O = T × S × C × I

3. Methods of Measuring Trend:

(*i*) Freehand Curve Method
(*ii*) Method of Semi-Averages
(*iii*) Method of Moving Averages
(*iv*) Method of Least Square.

4. Fitting of Linear Trend by Least Square Method:

$$Y = a + bX$$

Two Normal Equations:

$$\Sigma Y = Na + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

If $\Sigma X = 0$, then $a = \frac{\Sigma Y}{N}$ and $b = \frac{\Sigma XY}{\Sigma X^2}$

5. Fitting of Quadratic Trend by Least Square Method:

$$Y = a + bX + cX^2$$

Three Normal Equations:

$$\Sigma Y = Na + b\Sigma X + c\Sigma X^2$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2 + c\Sigma X^3$$

$$\Sigma X^2 Y = a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4$$

If $\Sigma X = 0$, $\Sigma X^3 = 0$, then

$$b = \frac{\Sigma XY}{\Sigma X^2}, \quad a = \frac{\Sigma Y - c\Sigma X^2}{N}$$

$$c = \frac{N \cdot \Sigma X^2 Y - (\Sigma X^2)(\Sigma Y)}{N \cdot \Sigma X^4 - (\Sigma X^2)^2}$$

6. Fitting of Exponential Trend by Least Square Method:

$$Y = ab^X$$

$$\log Y = \log a + X \log b$$

Two Normal Equations:

$$\Sigma \log Y = N \cdot \log a + \log b\Sigma X$$

$$\Sigma X \log Y = \log a\Sigma X + \log b\Sigma X^2$$

If $\Sigma X = 0$, then

$$\log a = \frac{\Sigma \log Y}{N}$$

$$\log b = \frac{\Sigma X \log Y}{\Sigma X^2}$$

QUESTIONS

1. What is time series? Explain any one method of measuring trend in a time series.
2. What is time series? What is the need for analysis of time series?
3. Explain briefly the components of time series.
4. Distinguish between secular trend, seasonal variations and cyclical fluctuations.
5. Explain the meaning and importance of time series.
6. Explain briefly the additive and multiplicative models of time series. Which of these models is more popular in practice?
7. Explain briefly the various methods of determining a trend in a time series. Explain merits and demerits of each method.
8. Compare the moving average and least square methods of measuring trend in a given time series. Which method is better and why?
9. Distinguish between seasonal and cyclical fluctuations with suitable examples.
10. Explain the utility of time series analysis to a businessman and an economist. Also state the different components in a time series.
11. Define a time series. Explain the components of time series.
12. Explain the procedure of fitting linear trend, quadratic trend and exponential trend using least square method.

Or

Discuss least square method of fitting linear, quadratic and exponential trend.

13. What is time series? State its utility in business.

Time Series Analysis–II

6

INTRODUCTION

For a trader, along with trend analysis, the knowledge about seasonal variations is also very useful. With its help he can, on the one hand, make short-term planning for his business activities and on the other hand, he can immune himself from the effects of short-period variations. Therefore, analysis of seasonal variations is very important. In this chapter, we will discuss the methods of measuring seasonal variations.

MEASUREMENT OF SEASONAL VARIATIONS

The main methods of measuring seasonal variations are as follows:

(1) **Method of Simple Averages**

(2) **Method of Moving Average**

(3) **Ratio to Moving Average**

(4) **Ratio to Trend Method**

(5) **Link Relatives Method**

Let us consider them in detail.

(1) Method of Simple Averages

This is the simplest method of measuring seasonal variations. This method is used in those situations where trend is assumed to be absent in the data. This method involves the following steps:

(1) The given data is arranged monthwise or quarterwise for different years.

(2) The totals of each month or quarter for different years are obtained and then dividing the sum by 12 or 4, the average of each month or quarter is computed.

(3) The average of the monthly average or quarterly average is then computed.

(4) Taking the general average as base, seasonal indices for each month or quarter are computed by using the following formulae:

(*i*) **When monthly data is given:**

$$\text{Seasonal Index for Jan.} = \frac{\text{Average of Jan.}}{\text{General Average}} \times 100$$

$$\text{Seasonal Index for Feb.} = \frac{\text{Average of Feb.}}{\text{General Average}} \times 100$$

Similarly, seasonal indices for other months can also be computed:

***(ii)* When quarterly data is given:**

$$\text{Seasonal Index for I Quarter} = \frac{\text{Average of I Quarter}}{\text{General Average}} \times 100$$

$$\text{Seasonal Index for II Quarter} = \frac{\text{Average of II Quarter}}{\text{General Average}} \times 100$$

Similarly, seasonal indices for III and IV quarters can also be computed.

The following examples illustrate the procedure of this method:

Example 1. Assuming that trend is absent, determine if there is any seasonality in the data given below:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1982	37	41	33	35
1983	37	39	36	36
1984	40	43	33	31

What are the seasonal indices for various quarters?

Solution:

Computation of Seasonal Indices

Years	Q_1	Q_2	Q_3	Q_4
1982	37	41	33	35
1983	37	39	36	36
1984	40	43	33	31
Total	114	123	102	102
Average	38	41	34	34

$$\text{Average of Average} = \frac{38+41+34+34}{4} = 36.75$$

(Grand Average)

$$\text{Seasonal Index} = \frac{\text{Quarterly Average}}{\text{General Average}} \times 100$$

$$\text{Seasonal Index for 1st Quarter} = \frac{38}{36.75} \times 100 = 103.40$$

$$\text{Seasonal Index for 2nd Quarter} = \frac{41}{36.75} \times 100 = 111.56$$

$$\text{Seasonal Index for 3rd Quarter} = \frac{34}{36.75} \times 100 = 92.52$$

$$\text{Seasonal Index for 4th Quarter} = \frac{34}{36.75} \times 100 = 92.52$$

Example 2. Compute the seasonal indices for the following time series data by using method of simple averages:

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1981	15	16	18	18	23	23	20	28	29	33	33	38
1982	23	22	28	27	31	28	22	28	32	37	34	44
1983	25	25	35	36	36	30	30	34	38	47	41	53

Solution:

Calculation of Seasonal Variation Index

Month	Years			3-yearly Totals	Monthly Average	Seasonal Index
	1981	1982	1983			
Jan.	15	23	25	63	21	$\frac{21}{30} \times 100 = 70$
Feb.	16	22	25	63	21	$\frac{21}{30} \times 100 = 70$
Mar.	18	28	35	81	27	$\frac{27}{30} \times 100 = 90$
Apr.	18	27	36	81	27	$\frac{27}{30} \times 100 = 90$
May	23	31	36	90	30	$\frac{30}{30} \times 100 = 100$
June	23	28	30	81	27	$\frac{27}{30} \times 100 = 90$
July	20	22	30	72	24	$\frac{24}{30} \times 100 = 80$
Aug.	28	28	34	90	30	$\frac{30}{30} \times 100 = 100$
Sept.	29	32	38	99	33	$\frac{33}{30} \times 100 = 110$
Oct.	33	37	47	117	39	$\frac{39}{30} \times 100 = 130$
Nov.	33	34	41	108	36	$\frac{36}{30} \times 100 = 120$
Dec.	38	44	53	135	45	$\frac{45}{30} \times 100 = 150$
	Totals				360	
	Average of Monthly Averages $= \frac{360}{12} = 30$ (or General Average)					

$$\text{Seasonal Index} = \frac{\text{Monthly Average}}{\text{General Average}} \times 100$$

$$\text{Seasonal Index for Jan.} = \frac{\text{Average of Jan.}}{\text{General Average}} \times 100 = \frac{21}{30} \times 100 = 70$$

	W	221		→ 1077	135	+86	+68.4375
			→ 548				
1979	S	56		→ 1126	141	–85	–74.8125
			→ 578				
	M	172		→ 1170	146	+26	+25.4375
			→ 592				
	A	129		→ 1195	149	–20	–19.0625
			→ 603				
	W	235		→ 1235	154	+81	+68.4375
			→ 632				
1980	S	67		→ 1271	159	–92	–74.8125
			→ 639				
	M	201		→ 1345	168	+33	+25.4375
			→ 706				
	A	136	—	—	—	—	–19.0625
	W	302	—	—	—	—	+68.4375

Calculation of Seasonal Variations
(from short-term fluctuations)

Years	Summer	Monsoon	Autumn	Winter
1976	—	—	–11	+42
1977	–50	+12	–14	+64
1978	–73	+30	–32	+86
1979	–85	+26	–20	+81
1980	–92	+33	—	—
Total	–300	101	–77	+273
Average	–75	+ 25.25	–19.25	+ 68.25
Seasonal Variations	–75 – (– 0.1875) = – 74.8125	25.25 – (– 0.1875) = + 25.4375	–19.25 – (– 0.1875) = – 19.0625	68.25 – (– 0.1875) = + 68.4375

$$\text{General Average} = \frac{-75+25.25-19.25+68.25}{4}$$

$$= \frac{-0.75}{4} = -0.1875$$

EXERCISE 6.2

1. Find the seasonal fluctuations by the method of moving averages from the following data:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1983	74	76	74	80
1984	82	68	50	62
1985	70	74	70	82

[**Ans.** 6.25, 1.375, –8.50, 0.875]

2. Find the seasonal fluctuations by using the following data:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1983	16.00	13.50	14.70	17.00
1984	15.90	12.20	15.60	18.00
1985	16.30	11.90	16.90	19.20
1986	17.10	13.20	15.00	18.70

[**Ans.** 0.754, – 3.33, 0.087, 2.4]

(3) Ratio to Moving Average Method

This is the most popular method of measuring seasonal variations. It is based on the multiplicative model of time series. The following steps are taken up under this method:

(1) Obtain the trend values by the moving averages method. If given data are quarterly, then 4-quarterly moving averages are found out. As against it, if given data are monthly, then 12-monthly moving averages are computed.

(2) After this, using multiplicative model each figure relating to the time-periods of original data is divided by the corresponding trend value and the quotient is multiplied by 100 to get ratio-to-moving average:

$$\textbf{Ratio to Moving Average} = \frac{O}{T} \times 100$$

Where, O = Original Value
T = Moving Average

(3) Next, arithmetic averages are computed after arranging the ratio-to-moving averages related to different periods in a separate table.

(4) All averages relating to ratio-to-moving averages are summed up and treated to get a general average.

(5) Finally, making the general average as base, the seasonal indices for quarters are found by the following formula:

$$\textbf{Seasonal Indices} = \frac{\textbf{Quarterly Average}}{\textbf{General Average}} \times 100$$

This method can be illustrated with the following examples:

Example 4. From the following data, calculate seasonal indices by the Ratio to Moving Average Method.

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1985	68	62	61	63
1986	65	58	61	61
1987	68	63	63	67

Solution:

Calculation of Seasonal Indices by Ratio-to-Moving Average Method

Year (1)	Quarter (2)	Values (3) (O)	4 Quarterly Moving Totals (4)	2 Period Totals Centralized Totals	4 Quarterly Moving Average (or Trend Values) (T)	Ratio to Moving Average ($\frac{O}{T} \times 100$)
1985	I	68	—	—	—	—
	II	62	—	—	—	—
			→ 254			
	III	61		→ 505	63.125	96.6
			→ 251			
	IV	63		→ 498	62.25	101.2
			→ 247			
1986	I	65		→ 494	61.75	105.3
			→ 247			
	II	58		→ 492	61.50	94.3
			→ 245			
	III	61	—	→ 493	61.625	99.0
			→ 248			
	IV	61		→ 501	62.625	97.4
			→ 253			
1987	I	68		→ 508	63.50	107.1
			→ 255			
	II	63		→ 516	64.50	97.7
			→ 261			
	III	63	—	—	—	—
	IV	67	—	—	—	—

Calculation of Seasonal Indices

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1985	—	—	96.6	101.2
1986	105.3	94.3	99.0	97.4
1987	107.1	97.7	—	—
Totals	212.4	192	195.6	198.6
A. Average	106.2	96.0	97.8	99.3
Seasonal Indices	106.4	96.2	98.0	99.5

$$\text{General Average} = \frac{106.2+96+97.8+99.3}{4} = 99.825$$

Calculation of Seasonal Indices:

Seasonal Indices for

$$\text{I Quarter} = \frac{106.2}{99.825} \times 100 = 106.4$$

$$\text{II Quarter} = \frac{96}{99.825} \times 100 = 96.2$$

$$\text{III Quarter} = \frac{97.8}{99.825} \times 100 = 98$$

$$\text{IV Quarter} = \frac{99.3}{99.825} \times 100 = 99.5$$

Example 5. Calculate seasonal indices for each quarter from the following percentages of wholesale prices to their moving averages:

Year	Quarters			
	I	II	III	IV
1987	—	—	85.71	90.25
1988	128.12	91.71	96.10	103.90
1989	112.33	100.35	78.13	97.88
1990	105.26	103.50	—	—

Solution:

Calculation of Seasonal Indices

Year	Q_1	Q_2	Q_3	Q_4
1987	—	—	85.71	90.25
1988	128.12	91.71	96.10	103.90
1989	112.33	100.35	78.13	97.88
1990	105.26	103.50	—	—
Totals	345.71	295.56	259.94	292.03
Average	115.23	98.52	86.64	97.34
Seasonal Indices	115.89	99.08	87.13	97.89

$$\text{General Average} = \frac{115.23+98.52+86.64+97.34}{4} = \frac{397.73}{4} = 99.43$$

Calculation of Seasonal Indices

$$\text{Seasonal Indices for I Quarter} = \frac{115.23}{99.43} \times 100 = 115.89$$

$$\text{II Quarter} = \frac{98.52}{99.43} \times 100 = 99.08$$

$$\text{III Quarter} = \frac{86.64}{99.43} \times 100 = 87.13$$

$$\text{IV Quarter} = \frac{97.34}{99.43} \times 100 = 97.89$$

EXERCISE 6.3

1. Calculate the seasonal index for the data given below by Ratio-to-Moving Average Method.

Years	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1980	68	62	61	63
1981	65	58	56	61
1982	68	63	63	67
1983	70	59	56	62
1984	60	55	51	58

[**Ans.** 107.0, 96.4, 94.5, 102.1]

2. Eliminate trend by moving average method:

Year	Ist Quarter	IInd Quarter	IIIrd Quarter	IVth Quarter
2001	40	35	38	40
2002	42	37	39	38
2003	41	35	38	42

[**Ans.** 98.7, 102.56, 106.66, 94.2675, 94.2675, 98.70, 107.54, 90.9]

(4) Ratio to Trend Method

Under this method, the following steps are taken up for the measurement of seasonal variations:

(1) Obtain the trend values season-wise (quarterly) by the method of least squares.

(2) By dividing the each value of original data (O) relating to all the periods by the corresponding trend value (T), ratio-to-trend is computed. Symbolically,

$$\textbf{Ratio to Trend} = \frac{O}{T} \times 100$$

(3) The arithmetic mean of each quarterly or monthly period ratio-to-trend is computed.

(4) General average is computed by summing up all the averages relating to the quarters or months. Hereafter, seasonal indices are computed by using the following formula:

$$\textbf{Seasonal Index} = \frac{\textbf{Quarterly Average}}{\textbf{General Average}} \times 100$$

This method can be illustrated with the following example:

Example 6. Find out seasonal index by ratio-to-trend method from the data given below:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1987	30	40	36	34
1988	34	52	50	44
1989	40	58	54	48
1990	54	76	68	62
1991	80	92	86	82

Solution: First we have to determine the trend values for yearly data by fitting a straight line trend by the method of least squares.

Year	Yearly Total	Quarterly Average (Y)	Deviations from 1989 (X)	XY	X^2	Trend values (Yc)
1987	140	35	–2	–70	4	32
1988	180	45	–1	–45	1	44
1989	200	50	0	0	0	56
1990	260	65	+1	+65	1	68
1991	340	85	+2	+170	4	80
N = 5		$\Sigma Y = 280$	$\Sigma X = 0$	$\Sigma XY = 120$	$\Sigma X^2 = 10$	

The equation of the straight line trend is

$$Y = a + bX$$

Since, $\Sigma X = 0$, $\quad a = \frac{\Sigma Y}{N}, \; b = \frac{\Sigma XY}{\Sigma X^2}$

$\therefore \quad a = \frac{280}{5} = 56$ and $b = \frac{120}{10} = 12$

$\therefore \quad Y = 56 + 12X$

For 1987, $X = -2$, $Y = 56 + 12(-2) = 56 - 24 = 32$

Other trend values can be found by adding the value of b in the preceding trend values.

Yearly increment $= b = 12$

Thus, quarterly increment $= \frac{12}{4} = 3$

Calculation of Quarterly Trend Values: Consider 1987: Trend value for the middle quarter, *i.e.*, half of the 2nd and half of the 3rd is 32. Quarterly increment is 3. So the trend value of 2nd quarter is 32 – 3/2, *i.e.*, 30.5 and for 3rd quarter is 32 + 3/2, *i.e.*, 33.5. Trend value for the 1st quarter is 30.5 – 3, *i.e.*, 27.5 and of 4th quarter is 33.5 + 3, *i.e.*, 36.5. We thus get quarterly trend values which are given below:

Quarterly Trend Values

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1987	27.5	30.5	33.5	36.5
1988	39.5	42.5	45.5	48.5
1989	51.5	54.5	57.5	60.5
1990	63.5	66.5	69.5	72.5
1991	75.5	78.5	81.5	84.5

The given values (O) are to be expressed as the percentages of the corresponding trend values.

Quarterly Values as % of Trend Values

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1987	109.1	131.1	107.5	93.1
1988	86.1	122.4	109.9	90.7
1989	77.7	106.4	93.9	79.3
1990	85.0	114.3	97.8	85.5
1991	106.0	117.1	105.5	97.0
Total	463.9	591.3	514.6	445.6
Average	92.78	118.26	102.92	89.12
Seasonal Indices	$\frac{92.78}{100.77}\times 100$ = 92.0	$\frac{118.26}{100.77}\times 100$ = 117.4	$\frac{102.92}{100.77}\times 100$ = 102.1	$\frac{89.12}{100.77}\times 100$ = 88.4

$$\text{General Average} = \frac{92.78+118.26+102.92+89.12}{4} = \frac{403.08}{4} = 100.77$$

EXERCISE 6.4

1. Using 'Ratio-to-Trend' method, determine the quarterly seasonal indices for the following data:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1984	60	80	72	68
1985	68	104	100	88
1986	80	116	108	96
1987	108	152	136	124
1988	160	184	172	164

[Ans. 92.03, 117.33, 102.10, 88.44]

(5) Link Relatives Method

This is another method of measuring the seasonal variations. The steps involved in this method are as follows:

(1) Calculate the link relatives of the seasonal figures - monthly or quarterly. For this, the following formula is used:

$$\text{Link Relatives} = \frac{\text{Current season's figure}}{\text{Previous season's figure}} \times 100$$

(2) Then the average of the link relatives for each month or quarter is computed.

(3) The average of the link relatives are then converted into chain relatives. For this, the following formula is used:

Chain Relatives =

$$\frac{\text{Average of LR of the current season's figure} \times \text{Chain Relatives of the previous season's figure}}{100}$$

(4) The chain relative for the Ist term is calculated on the basis of chain relatives of the last term. For this, the following formula is used:

Chain Relatives of the Ist term =

$$\frac{\text{Chain Relatives of the last seasonal's figure} \times \text{Average of LR of the 1st season}}{100}$$

(5) Theoretically, chain relative of the first period should be 100 but sometime, due to the influence of the trend, this can be more than or less than 100. The difference in this case be found out by deducting 100 from the revised chain relative of the first term. This difference is divided by the number of periods and the quotient is multiplied by 1, 2, 3, etc. Values thus obtained are subtracted from the chain relative of 2nd term, chain relative of the 3rd term and the chain relative of the 4th term.

(6) Finally, arithmetic mean of the adjusted or corrected chain relatives is computed. By taking general average as base, seasonal indices are computed by the following formula:

$$\textbf{Seasonal Indices} = \frac{\textbf{Corrected Chain Relatives}}{\textbf{General Average}} \times 100$$

This method is illustrated by the following examples:

Example 7. Calculate seasonal indices from the following data by using link relatives method:

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1984	20	40	60	80
1985	30	30	40	90
1986	40	60	30	120
1987	50	50	70	150

Solution:

Calculation of Seasonal Indices by Link Relative Method

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1984	—	200	150	133.3
1985	37.5	100	133.3	225
1986	44.4	150	50	400
1987	41.7	100	140	214.3
Total of L.R.	123.6	550	473.3	972.6
Average of L.R.	41.2	137.5	118.3	243.2
Chain Relatives	100	$\frac{100\times137.5}{100}$ = 137.5	$\frac{137.5\times118.3}{100}$ = 162.7	$\frac{162.7\times243.2}{100}$ = 395.7
Corrected Chain Relatives	100	137.5 − 15.75 = 121.75	162.7 − 31.5 = 131.2	395.7 − 47.25 = 348.45
Seasonal Indices	$\frac{100\times100}{175.35}$ = 57.03	$\frac{121.75\times100}{175.35}$ = 69.43	$\frac{131.2\times100}{175.35}$ = 74.82	$\frac{348.45\times100}{175.35}$ = 198.72

Chain relative of the first quarter $= \frac{41.2\times395.7}{100} = 163$

(on the basis of the last quarter)

The difference between these chain relatives = 163 – 100 = 63

Difference per quarter $= \frac{63}{4} = 15.75$

Adjusted (or corrected) chain relatives are obtained by subtracting 1×15.75, 2×15.75, 3×15.75 from the chain relatives of the 2nd, 3rd and 4th quarters respectively.

Average of corrected chain relatives $= \frac{100+121.75+131.2+348.45}{4} = 175.35$

(or General Average)

Seasonal variation indices have been calculated as follows:

$$\text{Seasonal variation indices} = \frac{\text{Corrected Chain Relative}}{\text{General Average}} \times 100$$

Example 8. Calculate seasonal indices by link relative method from the following data:

Link Relatives

Quarter/Year	1991	1992	1993	1994	1995
I	—	80	88	80	83
II	120	117	129	125	117
III	133	113	111	115	120
IV	83	89	93	96	79

Solution:

Calculation of Seasonal Indices by Link Relative Method

Year	Q_1	Q_2	Q_3	Q_4
1991	—	120	133	83
1992	80	117	113	89
1993	88	129	111	93
1994	80	125	115	96
1995	83	117	120	79
Total of L.R.	331	608	592	440
Average of L.R.	82.75	121.6	118.4	88
Chain Relatives	100	$\frac{121.6\times100}{100}$ = 121.6	$\frac{121.6\times118.4}{100}$ = 143.97	$\frac{143.97\times88}{100}$ = 126.69
Adjusted Chain Relatives	100	121.6 − 1.21 = 120.39	143.97 − 2 × 1.21 = 141.55	126.69 − 3 × 1.21 = 123.06
Seasonal Indices	$\frac{100}{121.25}\times100$ = 82.47	$\frac{120.39}{121.25}\times100$ = 99.29	$\frac{141.55}{121.25}\times100$ = 116.74	$\frac{123.06}{121.25}\times100$ = 101.5

In the above table, the correction factor has been calculated as follows:

Chain relative of the first quarter = 100

(on the basis of the first quarter)

Chain relative of the first quarter $= \frac{82.75\times126.69}{100} = 104.84$

(on the basis of the last quarter)

The difference between these two chain relatives = 104.84 – 100 = 4.84

Difference per quarter $= \frac{4.84}{4} = 1.21$

Adjusted (or corrected) chain relatives are obtained by subtracting 1×1.208, 2×1.208, 3×1.208 from the chain relatives of the 2nd, 3rd and 4th quarters respectively.

Average of corrected chain relatives $= \frac{100+120.39+141.55+123.06}{4} = 121.25$

(or General Average)

Seasonal indices have been calculated as follows:

$$\text{Seasonal indices} = \frac{\text{Corrected Chain relative}}{\text{General Average}} \times 100$$

EXERCISE 6.5

1. Apply method of link relatives to the following data and calculate seasonal indices:

Quarter	Quarterly Figures				
	1992	1993	1994	1995	1996
I	6.0	5.4	6.8	7.2	6.6
II	6.5	7.9	6.6	5.8	7.4
III	7.8	8.4	9.3	7.5	8.0
IV	8.7	7.3	6.4	8.5	7.1

[**Ans.** 88.09, 94.44, 113.05, 104.43]

2. Apply the method of link relatives to the following data and obtain seasonal indices:

Link Relatives

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1987	—	120	133	83
1988	80	117	113	89
1989	88	129	111	92
1990	80	125	115	96
1991	83	117	120	79

[**Ans.** 82.5, 99.4, 116.8, 101.4]

◘ DESEASONALISATION OF DATA

There are two objectives of studying seasonal variations: (*a*) to measure them and (*b*) to eliminate them from the given series. Elimination of the seasonal effects from the series is termed as deseasonalisation of data. If multiplicative model is assumed, the following formula will be used for deseasonalisation:

$$\text{Deseasonalised value} = \frac{\text{O}}{\text{Seasonal Index}} \times 100 \quad \text{where, O = Time Series Data}$$

Example 9. Deseaonalise the following data with the help of the seasonal index given against:

Quarter:	1st	2nd	3rd	4th
Sales (Rs. '000):	40	35	38	40
Seasonal Index:	108.26	92.07	96.98	102.69

Solution:

Computation of Deseasonalised Values

Quarter	Sales (O) (Rs. '000)	Seasonal Index (S.I.)	Deseasonalised Value $\left(\frac{\text{O}}{\text{S.I.}} \times 100\right)$
1st	40	108.26	$\frac{40}{108.26} \times 100 = 36.95$
2nd	35	92.07	$\frac{35}{92.07} \times 100 = 38.01$
3rd	38	96.98	$\frac{38}{96.98} \times 100 = 39.18$
4th	40	102.69	$\frac{40}{102.69} \times 100 = 38.95$

EXERCISE 6.6

1. Deseasonalise the following data with the help of the seasonal index given against:

Month:	Jan.	Feb.	March	April	May	June
Sales (Rs. '000):	360	400	550	360	350	550
Seasonal Index:	120	80	110	90	70	100

[**Ans.** 300, 500, 500, 400, 500, 500]

2. Find the seasonal index from the following table by ratio to moving average method. Also deseasonalize the data.

Quarter	1983	1984	1985	1986	1987
I	40	42	41	45	44
II	35	37	35	36	38
III	38	39	38	36	38
IV	40	38	42	41	42

[**Ans.** (*i*) 108.26, 92.07, 96.98, 102.69 (*ii*) 36.95, 38.01, 39.18, 38.95; 38.80, 40.19, 40.21, 37.00; 37.87, 38.01, 39.18, 40.90; 41.57, 39.10, 37.12, 39.93; 40.64, 41.27, 39.18, 40.90]

USES OF SEASONAL INDICES

Example 10. The seasonal indices of sales of a firm during 1996 are as under:

Month	Seasonal Index	Month	Seasonal Index
Jan.	106	July	93
Feb.	105	Aug.	89
Mar.	101	Sept.	92
Apr.	104	Oct.	102
May	98	Nov.	106
June	96	Dec.	108

If the firm is expecting total sales of Rs. 42,00,000 during 1996, estimate the sales for the individual in months of 1996.

Solution: Estimated Monthly Sales = Average Monthly Sales × Seasonal Effect.

Where, Average monthly sales $= \dfrac{\text{Annual Sales}}{12} = \dfrac{42{,}00{,}000}{12} = 3{,}50{,}000$

and Seasonal Effect or S.E. $= \dfrac{\text{Seasonal Index}}{100}$

On the basis of the above formulae, the estimates of the monthly sales are computed as follows:

Estimation of Monthly Sales

Month	Seasonal Index	Seasonal Effect (S.E.)	Estimated monthly sales
January	106	$\frac{106}{100} = 1.06$	$3{,}50{,}000 \times 1.06 = 3{,}71{,}000$
February	105	1.05	$3{,}50{,}000 \times 1.05 = 3{,}67{,}500$
March	101	1.01	$3{,}50{,}000 \times 1.01 = 3{,}53{,}500$
April	104	1.04	$3{,}50{,}000 \times 1.04 = 3{,}64{,}000$
May	98	0.98	$3{,}50{,}000 \times 0.98 = 3{,}43{,}000$
June	96	0.96	$3{,}50{,}000 \times 0.96 = 3{,}36{,}000$
July	93	0.93	$3{,}50{,}000 \times 0.93 = 3{,}25{,}500$
August	89	0.89	$3{,}50{,}000 \times 0.89 = 3{,}11{,}500$
September	92	0.92	$3{,}50{,}000 \times 0.92 = 3{,}22{,}000$
October	102	1.02	$3{,}50{,}000 \times 1.02 = 3{,}57{,}000$
November	106	1.06	$3{,}50{,}000 \times 1.06 = 3{,}71{,}000$
December	108	1.08	$3{,}50{,}000 \times 1.08 = 3{,}78{,}000$
Total	1200	12.00	42,00,000

Example 11. The seasonal indices of the sale of garments of a particular type in a certain shop are given below:

	Quarter	Seasonal Index
I	Jan.—March	97
II	April—June	85
III	July—Sept.	83
IV	Oct.—Dec.	135

If the total sales in the first quarter of a year be worth Rs. 15,000, determine how much worth of garments of this type should be kept in stock by the shop owner to meet the demand in each of the three quarters of the year.

Solution:

Calculation of Estimated Stocks

Quarters	Seasonal Index	Estimated value of stock (in Rs.)
Jan.—March	97	15,000
April—June	85	$15{,}000 \times \frac{85}{97} = \text{Rs. } 13{,}144$
July—Sept.	83	$15{,}000 \times \frac{83}{97} = \text{Rs. } 12{,}835$
Oct.—Dec.	135	$15{,}000 \times \frac{135}{97} = \text{Rs. } 20{,}876$

Aliter: The deseaonalised sale for the first quarter is given by

$$S = \frac{\text{Observed Value}}{\text{Seasonal Effect}} = \frac{15{,}000}{0.97} = 15463.9175$$

On the basis of this deseasonalised sale of the Ist quarter, the stocks of the remaining three quarters will be computed as under:

Quarter	Seasonal Index	Seasonal Effect (S.E.)	Estimated value of stock
II	85	0.85	$15463.9175 \times 0.85 = 13{,}144$ app.
III	83	0.83	$15463.9175 \times 0.83 = 12{,}835$ app.
IV	135	1.35	$15463.9175 \times 1.35 = 20{,}876$

Example 12. Given the following ratio of observed to trend values (%), calculate the seasonal indices. If the annual sales for 1991 are expected to be Rs. 2000 lakh, what are the likely sales for the individual quarters.

Year	Q_1	Q_2	Q_3	Q_4
1987	80	95	80	110
1988	101	104	90	110
1989	100	95	90	100
1990	115	110	100	120

Solution: We rewrite the given percentage to the trend in the following form.

Computation of Seasonal Indices

Year	Ist Quarter	2nd Quarter	3rd Quarter	4th Quarter
1987	80	95	80	110
1988	101	104	90	110
1989	100	95	90	100
1990	115	110	100	120
Total	396	404	360	440
Quarterly Average	99	101	90	110
Seasonal Indices	$\frac{99}{100}\times100=99$	$\frac{101}{100}\times100=101$	$\frac{90}{100}\times100=90$	$\frac{110}{100}\times100=110$

$$\text{General Average } = \frac{99+101+90+110}{4} = 100$$

Expected annual sales for 1991 is 2,000 lakh.

$$\text{Expected quarterly sales} = \frac{\text{Annual Sales}}{4} = \frac{2{,}000}{4} = 500$$

The estimates of the quarterly sales are computed as

Quarters	S.I.	Expected Sales
Q_1	99	$500\times\frac{99}{100}$ = Rs. 495 lakh
Q_2	101	$500\times\frac{101}{100}$ = Rs. 505 lakh
Q_3	90	$500\times\frac{90}{100}$ = Rs. 450 lakh
Q_4	110	$500\times\frac{110}{100}$ = Rs. 550 lakh

Example 13. Calculate the seasonal index number from the following data:

Ratios of Observed to Trend Values%

Year/Quarter	I	II	III	IV
1999	108	130	107	93
2000	86	120	110	91
2001	92	118	104	88
2002	78	100	94	78
2003	82	110	98	86
2004	106	118	105	98

If the sales of a good X by a firm in the first quarter of 2005 is worth Rs. 92,000, determine how much worth of the good should be kept in stock by the firm to meet the demand in each of the remaining three quarters of 2005 by using the seasonal index numbers calculated above.

Solution:

Calculation of Seasonal Index by Ratio to Trend Method

Year	Ist Quarter	IInd Quarter	IIIrd Quarter	IVth Quarter
1999	108	130	107	93
2000	86	120	110	91
2001	92	118	104	88
2002	78	100	94	78
2003	82	110	98	86
2004	106	118	105	98
Total	552	696	618	534
Average/S.I.	92	116	103	89

Total of averages $92 + 116 + 103 + 89 = 400$. Since, the total is 400, no adjustment is required. The actual sales in the first quarter is worth Rs. 92,000. So, the sales without considering the seasonal, effect, *i.e.*, deseasonalised sales will be

$$\frac{92{,}000\times100}{92} = 1{,}00{,}000$$

Quarter	Deseasonalised sales	Seasonal index	Estimated stock
II	1,00,000	116	$1{,}00{,}000\times\frac{116}{100}=1{,}16{,}000$
III	1,00,000	103	$1{,}00{,}000\times\frac{103}{100}=1{,}03{,}000$
IV	1,00,000	89	$1{,}00{,}000\times\frac{89}{100}=89{,}000$

Example 14. The trend equation for quarterly sales of a firm is estimated to be as follows: $Y = 20 + 2X$, where Y is sales per quarter in millions of rupees, the unit of X is one quarter and the origin is the middle of the first quarter (Jan.–March) of 2001. The seasonal indices of sales for the four quarters are as follows:

Quarter:	I	II	III	IV
Seasonal indices :	120	105	85	90

Estimate the acutal sales for each quarter of 2006.

Solution: First quarter of 2001 is the origin

X = 0 in the first quarter of 2001

X = 20 in the first quarter of 2006

X = 21 in the second quarter of 2006

X = 22 in the third quarter of 2006

X = 23 in the fourth quarter of 2006

2006 Quarter	X	Yc	Seasonal index	Estimated actual sales
I	20	$20 + 2 \times 20 = 60$	120	$60 \times 1.20 = 72.0$
II	21	$20 + 2 \times 21 = 62$	105	$62 \times 1.05 = 65.1$
III	22	$20 + 2 \times 22 = 64$	85	$64 \times 0.85 = 54.4$
IV	23	$20 + 2 \times 23 = 66$	90	$66 \times 0.90 = 59.4$

Example 15. The seasonal indices of the sales of garments of a particular type in a certain shop are given below:

Quarter:	I	II	III	IV
Seasonal index:	97	85	83	135

If the total sales in the first quarter of a year be worth Rs. 15,000 and sales are expected to rise by 4% in each quarter, determine how much worth of garments of this type be kept in stock by the shopowner to meet the demand for each of the three quarters of the year.

Solution:

Computation of Stock

Quarter	Expected sales	Seasonal index	Estimated stock
I	15,000	97	$\frac{15,000 \times 97}{100} = 14,550$
II	$\frac{15,000 \times 104}{100} = 15,600$	85	$\frac{15,600 \times 85}{100} = 13,260$
III	$\frac{15,600 \times 104}{100} = 16,224$	83	$\frac{16,224 \times 83}{100} = 13,466$
IV	$\frac{16,224 \times 104}{100} = 16,873$	135	$\frac{16,873 \times 135}{100} = 22,779$

Note: Since, sales are expected to rise by 4% in each quarter, the expected sales of previous quarter will become the base for calculating sales of the next quarter.

EXERCISE 6.7

1. A company estimates its sales for a particular year to be Rs. 36,00,000. The seasonal indices for sales are as follows:

Month	Seasonal Index	Month	Seasonal Index
January	80	July	100
February	90	August	105
March	95	September	100
April	130	October	110
May	140	November	70
June	120	December	60

Using this information, calculate estimates of monthly sales of the company. (Assume that there is no trend.)

[Ans. 2.40, 2.30, 2.85, 3.90, 4.20, 3.60, 3.00, 3.15, 3.00, 3.30, 2.10, 1.80 lakhs]

2. The quarterly seasonal indices of the sales of a popular brand of colour television of a company in Delhi are given below:

Quarter :	I	II	III	IV
Seasonal index :	130	90	75	105

If the total sales for the first quarter of 1997 is Rs. 6,50,000, estimate the worth of television to be kept in store to meet the demand in other quarters. Assume that there is no trend.

[Ans. 4,50,000; 3,75,000 and 5,25,000]

3. The seasonal indices of the sale of garments of a particular type in a store are given below:

Quarter :	I	II	III	IV
Seasonal index :	98	89	83	130

If the total sales for the first quarter of a year be worth Rs. 10,000, find how much worth of garments of this type should be kept in stock to meet the demand in each of the remaining quarters.

[Ans. Garments to be kept (in Rs.): IInd Quarter : 9081.63; IIIrd Quarter: 8469.39; IV Quarter: 13265.30]

4. Calculate seasonal index number from the following data of sales of goods X:

Year	I	II	III	IV
2001	108	130	107	93
2002	86	120	110	91
2003	92	118	104	88
2004	78	100	94	78
2005	82	110	98	86
2006	106	118	105	98

If sales of goods X in the first quarter of 2007 are worth Rs. 20,000, determine how much worth of goods should be kept in stock by the firm to meet the demand in each of the remaining three quarters of 2007 by using the seasonal index numbers calculated above.

[Ans. 25217.39, 22391.30, 19347.82]

5. On the basis of quarterly sales (in Rs. lakh) of a certain commodity for the year 1994-95, the following calculations were made:

Seasonal Variations

Quarter :	I	II	III	IV
Seasonal Index :	80	90	120	110

$Y = 20 + 0.5X$ with origin : Ist quarter of 1994

X unit = one quarter ; Y_c = Quarterly sales (Rs. lakhs)

Estimate quarterly sales for each of the four quarters of 1995, using the multiplicative model.

[Ans. Estimated quarterly sales for the four quarters of 1995 (in Rs. lakh) are 17.60, 20.25, 27.60, 25.85 respectively]

6. Given the following ratio of observed values to trend values (%), calculate the seasonal index:

Year	Quarter I	Quarter II	Quarter III	Quarter IV
1987	105	100	90	115
1988	110	105	95	115
1989	110	95	95	105
1990	115	110	100	125

If the annual sales for 1991 are expected to be Rs. 20,000 lakh, what are the likely sales for the individe quarters?

[Ans. 102.38, 97.62, 90.48, 109.52, 511.9, 488.1, 452.4, 547.6]

IMPORTANT POINTS

- **Methods of Measuring Seasonal Variations/Seasonal Indices:**
  (1) Simple Averages Method
  (2) Moving Averages Method
  (3) Ratio to Moving Average Method
  (4) Ratio to Trend Method
  (5) Link Relatives Method

QUESTIONS

1. What is a seasonal index? Explain the different methods of estimating it.
2. Discuss the ratio-to moving average and the ratio-to trend method of measuring seasonal variations. Compare the two methods.
3. Explain any method of estimating the seasonal index for a time series based on quarterly data.
4. Describe, step by step, the moving average method of determining seasonal index.
5. Explain briefly the various methods of isolating seasonal fluctuations in time series.
6. What are seasonal variations? How would you construct a seasonal index using ratio to trend method? What are the uses and limitations of seasonal indices.

Probability

7

INTRODUCTION

In day-to-day life, we all make use of the word 'probability'. But generally people have no definite idea about the meaning of probability. For example, we often hear or talk phrases like, "Probability it may rain today"; "it is likely that the particular teacher may not come for taking his class today"; "there is a chance that the particular student may stand first in the university examination"; "it is possible that the particular company may get the contract which it bid last week"; "most probably I shall be returning within a week"; "it is possible that he may not be able to join his duty". In all the above statments, the terms - possible, probably, likely, chance, etc., convey the same meaning, *i.e.*, the events are not certain to take place. In other words, there is involved an element of uncertainty or chance in all these cases. A numerical measure of uncertainty is provided by the theory of probability. The aim of the probability theory is to provide a measure of uncertainty. The theory of probability owes its origin to the study of games of chance like games of cards, tossing coins, dice, etc. But in modern times, it has great importance is decision making problems.

SOME BASIC CONCEPTS

Before we give definition of the word probability, it is necessary to define the following basic concepts and terms widely used in its study:

(1) An Experiment

When we conduct a trial to obtain some statistical information, it is called an experiment.

Examples:

(*i*) Tossing of a fair coin is an experiment and it has two possible outcomes: Head (H) or Tail (T).

(*ii*) Rolling a fair die is an experiment and it has six possible outcomes: appearance of 1 or 2 or 3 or 4 or 5 or 6 on the upper most face of a die.

(*iii*) Drawing a card from a well shuffled pack of playing cards is an experiment and it has 52 possible outcomes. 52

(2) Events

The possible outcomes of a trial/experiment are called events. Events are generally denoted by capital letters A, B, C, etc.

Examples:

(*i*) If a fair coin is tossed, the outcomes - head or tail are called events.

(*ii*) If a fair die is rolled, the outcomes 1 or 2 or 3 or 4 or 5 or 6 appearing up are called events.

(3) Exhaustive Events

The total number of possible outcomes of a trial/experiment are called exhaustive events. In other words, if all the possible outcomes of an experiment are taken into consideration, then such events are called exhaustive events.

Examples: (*i*) In case of tossing a die, the set of six possible outcomes, *i.e.*, 1, 2, 3, 4, 5 and 6 are exhaustive events.

(*ii*) In case of tossing a coin, the set of two outcomes, *i.e.*, H and T are exhaustive events.

(*iii*) In case of tossing of two dice, the set of possible outcomes are $6 \times 6 = 36$ which are given below:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4.6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

(4) Equally-Likely Events

The events are said to be equally-likely if the chance of happening of each event is equal or same. In other words, events are said to be equally likely when one does not occur more often than the others.

Examples: (*i*) If a fair coin is tossed, the events H and T are equally-likely events.

(*ii*) If a die is rolled, any face is as likely to come up as any other face. Hence, the six outcomes - 1 or 2 or 3 or 4 or 5 or 6 appearing up are equally likely events.

(5) Mutually Exclusive Events

Two events are said to be mutually exclusive when they cannot happen simultaneously in a single trial. In other words, two events are said to be mutually exclusive when the happening of one excludes the happening of the other in a single trial.

Example: (*i*) In tossing a coin, the events Head and Tail are mutually exclusive because both cannot happen simultaneously in a single trial. Either head occurs or tail occurs. Both cannot occur simultaneously. The happening of head excludes the possibility of happening of tail.

(*ii*) In tossing a die, the events 1, 2, 3, 4, 5 and 6 are mutually exclusive because all the six events cannot happen simultaneously in a single trial. If number 1 turns up, all the other five (*i.e.*, 2, 3, 4, 5, or 6) cannot turn up.

(6) Complementary Events

Let there be two events A and B. A is called the complementary event of B and B is called the complementary event of A if A and B are mutually exclusive and exhaustive.

Examples: (*i*) In tossing a coin, occurrence of head (H) and tail (T) are complementary events.

(*ii*) In tossing a die, occurrence of an even number (2, 4, 6) and odd number (1, 3, 5) are complementary events.

(7) Simple and Compound Events

In case of simple events, we consider the probability of happening or not happening of single events.

Example: If a die is rolled once and A be the event that face number 5 is turned up, then A is called a simple event.

In case of compound events, we consider the joint occurrences of two or more events.

Example: If two coins are tossed simultaneously and we shall be finding the probability of getting two heads, then we are dealing with compound events.

(8) Independent Events

Two events are said to be independent if the occurence of one does not affect and is not affected by the occurence of the other.

Example: (*i*) In tossing a die twice, the event of getting 4 in the 2nd throw is independent of getting 5 in the first throw.

(*ii*) In tossing a coin twice, the event of getting a head in the 2nd throw is independent of getting head in the 1st throw.

(9) Dependent Events

Two events are said to be dependent when the occurence of one does affect the probability of the occurence of the other events.

Example: (*i*) If a card is drawn from a pack of 52 playing cards and is not replaced, this will affect the probability of the second card being drawn.

(*ii*) The probability of drawing a king from a pack of 52 cards is $\frac{4}{52}$ or $\frac{1}{13}$. But if the card drawn (king) is not replaced in the pack, the probability of drawing again a king is $\frac{3}{51}$.

DEFINITION OF PROBABILITY

The probability is defined in the following three different ways:

(1) **Classical or Mathematical Definition**

(2) **Empirical or Relative Frequency Definition**

(3) **Subjective Approach.**

(1) Classical or Mathematical Definition

This is the oldest and simplest definition of probability. This definition is based on the assumption that the outcomes or results of an experiment are equally likely and mutually exclusive.

According to Laplace, **"Probability is the ratio of the favourable cases to the total number of equally likely cases".** From this definition, it is clear that in order to calculate the probability of an event, we have to find the number of favourable cases and it is to be divided by the total number of cases. For example, if a bag contains 6 green and 4 red balls, then the probability of getting a green ball will be 6/4 + 6 = 6/10 because the total number of balls are 10 and the number of green balls is 6.

Symbolically,

$$P(A) = p = \frac{\textbf{Number of Favourable Cases}}{\textbf{Total Number of Equally Likely Cases}} = \frac{m}{n}$$

Where, P(A) = Probability of occurrence of an event A

m = Number of favourable cases

n = Total number of equally likely cases

Similarly,

$$P(\bar{A}) = q = 1 - P(A) = 1 - \frac{m}{n}$$

Where, $P(\bar{A}) = q$ = Probability of non-occurrence of an event A.

From the above definition, it is clear that the sum of the probability of happening of an event called success (p) and the probability of non-happening of an event called failure (q) is always one (1), *i.e.*, $p + q = 1$. If p is known, we can find q and if q is known, then we can find p. In practice, the value of p lies between 0 and 1, *i.e.*, $0 \le p \le 1$. To quote **Prof. Morrison, "If an event can happen in m ways and fail to happen in n ways, then probability of happening is $\frac{m}{m+n}$ and that of its failure to happen is $\frac{n}{m+n}$".**

▶ Limitations of Classical Definition

Following are the main limitations of classical definition of probability:

(1) If the various outcomes of the random experiment are not equally-likely, then we cannot find the probability of the event using classical definition.

(2) The classical definition also fails when the total number of cases are infinite.

(3) If the actual value of N is not known, then the classical definition fails.

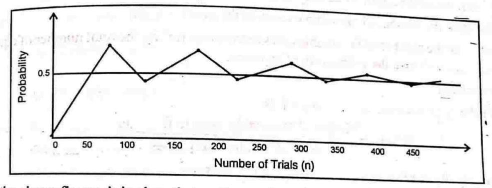
(2) Empirical or Relative Frequency Definition

This definition of probability is not based on logic but past experience and experiments and present conditions. If vital statistics gives the data that out of 100 newly born babies, 55 of them are girls, then the probability of the girl birth will be 55/100 or 55%. According to **Croxten and Cowden, "Probability is the limit of the relative frquency of success in infinite sequences of trials".** To quote **Kenny and Keeping, "If event has occured *r* times in a series of *n* independent trials, all are made under the same identical conditions, the ratio *r/n* is the relative frequency of the occurence of the event. The limit of *r/n* as *n* tends to infinity is the probability of the occurence of the event".**

Symbolically,

$$\boxed{P(A) = \textbf{Limit}_{n \to \infty} \frac{r}{n}}$$

For example, if a coin is tossed 100 times and the heads turn up 55 times, then the relative frequency of head will be $\frac{55}{100} = 0.55$. Similarly, if a coin is tossed 1000 times and if the head turns up 495 times, then the relative frequency will be $\frac{495}{1000} = 0.495$. In 10,000 tosses, the head turns up 5085 then the relative frequency will be 0.5085. Thus as we go on increasing the number of trials, there is a tendency that the relative frequency of head would approach to 0.50. The following figure illustrate the idea:



From the above figure, it is clear that as the number of trials increases, the probability of head tends to approach 0.5 and when the number of trials is infinite, *i.e.*, $n \to \infty$, the probability of getting head is equal to 0.5.

(3) Subjective Approach

According to this approach, probability to an event is assigned by an individual on the basis of evidence available to him. Hence probability is interpreted as a measure of degree of belief or confidence that a particular individual reposes in the occurrence of an event. But the main problem here is that different persons may differ in their degree of confidence even when same evidence is offered.

IMPORTANCE OF PROBABILITY

The theory of probability has its origin in the games of chance related to gambling such as tossing a die, tossing a coin, drawing a card from a deck of 52 cards and drawing a ball of a particular colour from a bag. But in modern times, it is widely used in the field of statistics, economics, commerce and social sciences that involve making predictions in the face of uncertainty. The importance of probability is clear from the following points:

(1) Probability is used in making economic decision in situations of risk and uncertainty by sales managers, production managers, etc.

(2) Probability is used in theory of games which is further used in managerial decisions.

(3) Various sampling tests like Z-test, t-test and F-test are based on the theory of probability.

(4) Probability is the backbone of insurance companies because life tables are based on the theory of probability.

Thus, probability is of immense utility in various fields.

PROBABILITY SCALE

The probability of an event always lies between 0 and 1, *i.e.*, $0 \le p \le 1$. If the event cannot take place, *i.e.*, impossible event, then its probability will be zero, *i.e.*, P(E) = 0 and if the event is sure to occur, then its probability will be one, *i.e.*, P (E) = 1.

Calculation of Probability of an Event

The following steps are to be followed while calculating the probability of an event:

(1) Find the total number of equally likely cases, *i.e.*, n

(2) Obtain the number of favourable cases to the event, *i.e.*, m

(3) Divide the number of favourable cases to the event (m) by the total number of equally likely cases (n). This will give the probability of an event.

Symbolically,

Probability of occurrence of an event E is:

$$P(E) = \frac{\text{Number of favourable cases to E}}{\text{Total number of equally likely cases}} = \frac{m}{n}$$

Similarly, Probability of non-occurrence of event E is:

$$P(\bar{E}) = 1 - P(E)$$

The following examples will illustrate the procedure:

Example 1. Find the probability of getting a head in a tossing of a coin.

Solution: When a coin is tossed, there are two possible outcomes - Head or Tail.

Total number of equally likely cases = $n = 2$

Number of cases favourable to H = $m = 1$

$\therefore \quad P(H) = \frac{m}{n} = \frac{1}{2}$

Example 2. What is the probability of getting an even number in a throw of an unbiased die?

Solution: When a die is tossed, there are 6 equally likely cases, *i.e.*, 1, 2, 3, 4, 5, 6.

Total number of equally likely cases = $n = 6$

Number of cases favourable to even points (2, 4, 6) = $m = 3$

$\therefore \quad$ Probability of getting an even number = $\frac{3}{6} = \frac{1}{2}$

Example 3. What is the probability of getting a king in a draw from a pack of cards?

Solution: Number of exhaustive cases = $n = 52$

There are 4 king cards in an ordinary pack.

$\therefore \quad$ Number of favourable cases = $m = 4$

$\therefore \quad$ Probability of getting a king = $\frac{4}{52} = \frac{1}{13}$

Example 4. From a bag containing 5 red and 4 black balls. A ball is drawn at random. What is the probability that it is a red ball?

Solution: Total No. of balls in the bag = 5 + 4 = 9

No. of red balls in the bag = 5

$\therefore \quad$ Probability of getting a red ball = $\frac{5}{9}$

Example 5. A bag contains 5 black and 10 white balls. What is the probability of drawing (*i*) a black ball, (*ii*) a white ball?

Solution: Total number of balls = 5 + 10 = 15

(*i*) $P(\text{black ball}) = \frac{\text{No. of black balls}}{\text{Total No. of balls}} = \frac{5}{15} = \frac{1}{3}$

(*ii*) $P(\text{white ball}) = \frac{\text{No. of white balls}}{\text{Total No. of balls}} = \frac{10}{15} = \frac{2}{3}$

Example 6. In a lottery, there are 10 prizes and 90 blanks. If a person holds one ticket, what are the chances of

(*i*) getting a prize

(*ii*) not getting a prize

Solution: Total No. of tickets = 10 + 90 = 100

(*i*) Probability of getting a prize:

No. of prizes = 10

$\therefore$ No. of favourable cases = 10

Total No. of cases = 100

Required Probability = $\frac{10}{100} = \frac{1}{10} = 0.1$

(*ii*) The probability of not getting a prize:

No. of Blanks = 90

$\therefore$ Number of favourable cases = 90

Total Number of cases = 100

Required Probability = $\frac{90}{100} = 0.9$

Example 7. What is the probability of getting a number greater than 4 with an ordinary die?

Solution: Number greater than 4 in a die are 5 and 6.

$\therefore \quad$ Number of favourable cases = 2

Total number of cases = 6