UNIT – III

Bivariate Linear Correlation

(1) Performance of students in the Preliminary examination & the Final examination are given below. Calculate Karl Pearson's Coefficient of Correlation.

Marks in Prelims	Marks in Final Exam
70	80
50	55
65	70
63	90
60	70
81	80
45	62
86	85

Solution:

Marks in Prelims (x)	Marks in Final Exam (y)	x^2	y^2	xy
70	80	4900	6400	5600
50	55	2500	3025	2750
65	70	4225	4900	4550
63	90	3969	8100	5670
60	70	3600	4900	4200
81	80	6561	6400	6480
45	62	2025	3844	2790

86	85	7396	7225	7310
$\sum x = 520$	**$\sum y = 592$**	**$\sum x^2 = 35176$**	**$\sum y^2 = 44794$**	**$\sum xy = 39350$**

$$\bar{x} = \frac{\sum x}{n} = \frac{520}{8} = 65, \quad \bar{y} = \frac{\sum x}{n} = \frac{592}{8} = 74$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}\sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}}$$

$$= \frac{\frac{39350}{8} - (65) \cdot (74)}{\sqrt{\frac{35176}{8} - (65)^2}\sqrt{\frac{44794}{8} - (74)^2}}$$

$$= \frac{4918.75 - 4810}{\sqrt{4397 - 4225}\sqrt{559.25 - 5476}}$$

$$= \frac{108.75}{\sqrt{172}\sqrt{123.25}}$$

$$= \frac{108.75}{(13.115)\,(11.102)}$$

$$= \frac{108.75}{145.59}$$

$$= 0.75$$

$$r = 0.75$$

(2) The following data gives profits for shop I and shop II. Compute Karl Pearson's Coefficient of Correlation.

Shop I	Shop II
55	60
57	65
72	80
83	92
50	62
70	83
42	60
75	82

Solution:

Shop I (x)	Shop II (y)	x^2	y^2	xy
55	60	3025	3600	3300
57	65	3249	4225	3705
72	80	5184	6400	5760
83	92	6889	8464	7636
50	62	2500	3844	3100
70	83	4900	6889	5810
42	60	1764	3600	2520
75	82	5625	6724	6150
504	**584**	**33136**	**43746**	**37981**

n = 8

$$\bar{x} = \frac{\sum x}{n} = \frac{504}{8} = 63, \quad \bar{y} = \frac{\sum x}{n} = \frac{584}{8} = 73$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}\sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}}$$

$$= \frac{\frac{37981}{8} - (63)\,(73)}{\sqrt{\frac{33136}{8} - (63)^2}\sqrt{\frac{43746}{8} - (73)^2}}$$

$$= \frac{4747.625 - 4599}{\sqrt{4142 - 3969}\sqrt{5468.25 - 5329}}$$

$$= \frac{148.625}{\sqrt{173}\sqrt{139.25}}$$

$$= \frac{148.625}{(13.15)\,(11.80)}$$

$$= \frac{148.625}{155.21}$$

$$= 0.96$$

(3) Calculate Karl Pearson's Correlation coefficient for the following data:
$n = 20$, $\Sigma x = 240$, $\Sigma y = 480$, $\Sigma x^2 = 4720$, $\Sigma y^2 = 15200$, $\Sigma xy = 7060$.

Solution:

$$\bar{x} = \frac{\sum x}{n} = \frac{240}{20} = 12 \quad \bar{y} = \frac{\sum x}{n} = \frac{480}{20} = 24$$

$$r = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}\sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}}$$

$$= \frac{\frac{7060}{20} - (12)\,(24)}{\sqrt{\frac{4720}{20} - (12)^2}\sqrt{\frac{15200}{20} - (24)^2}}$$

$$= \frac{353 - 288}{\sqrt{236 - 144}\sqrt{760 - 576}}$$

$$= \frac{65}{\sqrt{92}\sqrt{184}}$$

$$= \frac{65}{9.59 \times 13.56}$$

$$= \frac{65}{130.0404}$$

$$= 0.50$$

(4) Calculate Karl Pearson's Correlation coefficient for the following data:

$n = 10$, $\sum (X - \bar{X})(Y - \bar{Y}) = 55$, $\sum (X - \bar{X})^2 = 60$, $\sum (Y - \bar{Y})^2 = 70$

Solution:

$$\text{Cov}(X, Y) = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n} = \frac{55}{10} = 5.5$$

$$\sigma_x = \sqrt{\frac{\sum (X - \bar{X})^2}{n}} = \sqrt{\frac{60}{10}} = \sqrt{6} = 2.45$$

$$\sigma_y = \sqrt{\frac{\sum (Y - \bar{Y})^2}{n}} = \sqrt{\frac{70}{10}} = \sqrt{7} = 2.65$$

$$r = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{5.5}{(2.45)(2.65)} = \frac{5.5}{6.48} = 0.85$$

(5) Compute Karl Pearson's Coefficient of Correlation from the following data:

Price (in Rs.)	Demand (in units)
43	80
56	65
60	55
55	60
53	62
63	52
43	50
49	80

Solution:

Price (in Rs.) (x)	Demand (in units) (y)	x^2	y^2	xy
43	80	1849	6400	3440
56	65	3136	4225	3640
60	55	3600	3025	3300
55	60	3025	3600	3300
53	62	2809	3844	3286
63	52	3969	2704	3276
43	50	1849	2500	2150
49	80	2401	6400	3920
422	**504**	**22638**	**32698**	**26312**

n = 8

$\bar{x} = \frac{\sum x}{n} = \frac{422}{8} = 52.75$, $\bar{y} = \frac{\sum y}{n} = \frac{504}{8} = 63$

$$r = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}\sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}}$$

$$= \frac{\frac{26312}{8} - (52.75)\ (63)}{\sqrt{\frac{22638}{8} - (52.75)^2}\sqrt{\frac{32698}{8} - (63)^2}}$$

$$= \frac{3289 - 3323.25}{\sqrt{2829.75 - 2782.5625}\sqrt{4087.25 - 3969}}$$

$$= \frac{-34.25}{\sqrt{47.1875}\sqrt{118.25}}$$

$$= \frac{-34.25}{(6.87)\ (10.87)}$$

$$= \frac{-34.25}{74.699}$$

$$= -0.46$$

(6) Ten Students are ranked in a personality contest according to performance in a first round 1 (R1) and round 2 (R2). Find Spearman's Rank Correlation coefficient.

R1	R2
8	7
2	4
3	3
5	5
9	8
10	9
1	2
4	1
7	10
6	6

Solution:

R1	R2	d = R1 – R2	d^2
8	7	1	1
2	4	– 2	4
3	3	0	0
5	5	0	0
9	8	1	1
10	9	1	1
1	2	– 1	1
4	1	3	9

7	10	– 3	9
6	6	0	0
			26

$= \sum d^2$

$$n = 10$$

$$R = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \times 26}{10((10)^2 - 1)}$$

$$= 1 - \frac{6 \times 26}{10(100 - 1)}$$

$$= 1 - \frac{26}{165}$$

$$= 1 - 0.1575$$

$$= 0.8425$$

(7) The following data gives temperature in two cities for a week. Compute Spearman's Rank Correlation Coefficient:

City A	City B
38	39
33	38
39	45
44	52
40	39
39	39
40	45

Solution:

City A	City B	R1	R2	d	d^2
38	39	6	5	1	1
33	38	7	7	0	0
39	45	4.5	2.5	2	4
44	52	1	1	0	0
40	39	2.5	5	– 2.5	6.25
39	39	4.5	5	– 0.5	0.25
40	45	2.5	2.5	0	0
				$\sum d^2 =$	**11.5**

$$\text{Rank Frequency} = \frac{m(m^2 - 1)}{12}$$

$$4.5 \rightarrow 2 \quad \frac{2 \times 3}{12} = 0.5$$

$$2.5 \rightarrow 2 \quad \frac{2 \times 3}{12} = 0.5$$

$$2.5 \rightarrow 2 \quad \frac{2 \times 3}{12} = 0.5$$

$$5 \rightarrow 3 \quad \frac{3 \times 8}{12} = 2$$

$$\underline{\underline{\mathbf{3.5}}}$$

$$n = 7 \quad C.F. = 3.5$$

$$R = 1 - \frac{6(\sum d^2 + C.F.)}{n(n^2 - 1)}$$

$$= 1 - \frac{6(11.5 + 3.5)}{7((7)^2 - 1)}$$

$$= 1 - \frac{6 \times 15}{10(49 - 1)}$$

$$= 1 - \frac{15}{56}$$

$$= 1 - 0.268$$

$$= 0.732$$

(8) The following data gives production (in tonnes) for two shifts in a factory for 8 days. Compute Spearman's Rank Correlation Coefficient:

Shift I	Shift II
46	42
63	46
46	49
47	49
40	46
46	43
52	44
50	48

Solution:

Shift I	Shift II	R1	R2	d	d^2
46	42	6	8	– 2	4
63	46	1	4.5	– 3.5	12.25
46	49	6	1.5	4.5	20.25
47	49	4	1.5	2.5	6.25
40	46	8	4.5	3.5	12.25
46	43	6	7	– 1	1
52	44	2	6	– 4	16
50	48	3	3	0	0
				$\sum d^2 =$	**72**

$$\text{Rank Frequency} = \frac{m(m^2 - 1)}{12}$$

$$6 \rightarrow 3 \qquad \frac{3 \times 8}{12} = 2$$

$$4.5 \rightarrow 2 \quad \frac{2 \times 3}{12} = 0.5$$

$$1.5 \rightarrow 2 \quad \frac{2 \times 3}{12} = 0.5$$

$$\underline{\underline{\mathbf{3}}}$$

n = 8 C.F. = 3

$$R = 1 - \frac{6(\sum d^2 + C.F.)}{n(n^2 - 1)}$$

$$= 1 - \frac{6(72+3)}{8(8^2-1)}$$

$$= 1 - \frac{6 \times 75}{8(64-1)}$$

$$= 1 - \frac{450}{504}$$

$$= 1 - 0.893$$

$$= 0.107$$

Bivariate Linear Regression

(1) The following data gives the marks obtained by 10 students in two tests. Obtain the two regression lines and hence find:

(i) Most likely marks in test 1 of a student who has scored 75 marks in test 2.

(ii) Most likely marks in test 2 of a student who has scored 60 marks in test 1.

Test 1	**Test 2**
60	70
85	81
72	78
65	70
61	67
75	80
59	65
65	65
48	49
60	65

Solution:

Test 1 (x)	**Test 2 (y)**	x^2	y^2	xy
60	70	3600	4900	4200
85	81	7225	6561	6885
72	78	5184	6084	5616
65	70	4225	4900	4550
61	67	3721	4489	4087

$n = 10$

$\bar{x} = \frac{\sum x}{n}$

$= \frac{650}{10}$

$\bar{x} = 65$

75	80	5625	6400	6000
59	65	3481	4225	4225
65	65	4225	4225	4225
48	49	2304	2401	2352
60	65	3600	4225	3900
650	**690**	**43190**	**48410**	**45650**

$$\bar{y} = \frac{\sum y}{n} = \frac{690}{10}$$

$$\bar{y} = 69$$

$$b_{xy} = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\frac{\sum y^2}{n} - (\bar{y})^2} = \frac{\frac{45650}{10} - (65)(69)}{\frac{48410}{10} - (69)^2} = \frac{4565 - 4485}{4841 - 4761} = \frac{80}{80} = 1$$

$$b_{yx} = \frac{\frac{\sum xy}{n} - \bar{x} \cdot \bar{y}}{\frac{\sum x^2}{n} - (\bar{x})^2} = \frac{\frac{45650}{10} - (65)(69)}{\frac{43190}{10} - (65)^2} = \frac{4565 - 4485}{4319 - 4225} = \frac{80}{94} = 0.85$$

To estimate x when y is 75.

We use regression line of x on y.

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$x - 65 = 1\,(75 - 69)$$
$$x = 65 + 6$$
$$x = 71$$

To estimate y when x is 60.

We use regression line of y on x.

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$y - 69 = 0.85\,(60 - 65)$$
$$y = 69 + 0.85\,(-5)$$
$$y = 69 - 4.25$$
$$y = 64.75$$

(2) The following data gives expenditure on Research and Development (X) and Profits (Y) of a company. Obtain:

(i) Probable profit when R & D expenditure is Rs. 150,000.

(ii) Probable expenditure on R & D when the company enjoys profit of Rs. 400,000.

R & D expenditure ('000 Rs.)	Profit ('000 Rs.)
22	40
30	42
25	45
31	47
27	48
33	40
28	38
28	28

Solution:

R & D expenditure ('000 Rs.) (x)	Profit ('000 Rs.) (y)	x^2	y^2	xy
22	40	484	1600	880
30	42	900	1764	1260
25	45	625	2025	1125
31	47	961	2209	1457
27	48	729	2304	1296
33	40	1089	1600	1320
28	38	784	1444	1064

n = 8

$\bar{x} = \frac{\sum x}{n}$

$= \frac{224}{8}$

$= 28$

$\bar{y} = \frac{\sum y}{n}$

$= \frac{328}{8}$

28	28	784	784	784
224	**328**	**6356**	**13730**	**9186**

$\bar{y} = 41$

$$b_{xy} = \frac{\frac{\sum xy}{n} - \bar{x}\cdot\bar{y}}{\frac{\sum y^2}{n} - (\bar{y})^2}$$

$$= \frac{\frac{9186}{8} - (28)\ (41)}{\frac{13730}{8} - (41)^2}$$

$$= \frac{1148.25 - 1148}{1716.25 - 1681}$$

$$= \frac{0.25}{35.25} = 0.007$$

$$b_{yx} = \frac{\frac{\sum xy}{n} - \bar{x}\cdot\bar{y}}{\frac{\sum x^2}{n} - (\bar{x})^2}$$

$$= \frac{\frac{9186}{8} - (28)\ (41)}{\frac{6356}{8} - (28)^2}$$

$$= \frac{1148.25 - 114.8}{794.5 - 784}$$

$$= \frac{0.25}{10.5} = 0.24$$

To estimate x when y is 400000.

We use regression line of x on y.

$$x - \bar{x} = b_{xy}\ (y - \bar{y})$$

$$x - 28 = 0.007\ (400 - 41)$$

$$x = 28 + 0.007\ (359)$$

$$x = 28 + 2.513$$

$$x = 30.513$$

To estimate y when x is 150000.

We use regression line of y on x.

$$y - \bar{y} = b_{yx}\ (x - \bar{x})$$

$$y - 41 = 0.24\ (150 - 28)$$

$$y = 41 + 0.24\ (122)$$

$$y = 41 + 29.28$$

$$y = 70.28$$

(3) From the following data on the heights and weights of 1000 students, find:

(i) The weight of a student whose height is 150 cms.

(ii) The height of a student who weighs 70 kgs.:

	Mean	**Standard Deviation**
Height (in cms.)	170	9
Weight (in kg)	55	4
r = 0.8		

Solution:

$$b_{xy} = r \times \frac{\sigma_x}{\sigma_y} = 0.8 \times \frac{9}{4} = 0.8 \times 2.25 = 1.8$$

To find x when y is 70,

We use regression line of x on y.

$$x - \bar{x} = b_{xy} (y - \bar{y})$$

$$x - 170 = 1.8 (70 - 55)$$

$$x = 170 + 1.8 (15)$$

$$x = 170 + 27$$

$$x = 197$$

$$b_{yx} = r \times \frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{4}{9} = 0.8 \times 0.44 = 0.35$$

To find y when x is 150.

We use regression line of y on x.

$$y - \bar{y} = b_{yx} (x - \bar{x})$$

$$y - 55 = 0.35 (150 - 170)$$

$$y = 55 + 0.35 (- 20)$$

$$y = 55 - 7$$

$$y = 48$$

(4) The following data gives the Marks in Preliminary Exam (X) and marks in Annual Exam (Y).

	X	Y
Average	65	60
S. D.	10	15
Coefficient correlation 0.70		

Obtain the two regression lines and estimate (i) the marks of a student in annual exam who has scored 70 in Preliminary Exam (ii) the marks of a student in Preliminary Exam who has scored 65 in annual exam.

Solution:

$$b_{xy} = r \times \frac{\sigma_x}{\sigma_y} = 0.7 \times \frac{10}{15} = 0.7 \times 0.67 = 0.467$$

To find x when y is 65.

We use regression line of x on y.

$$x - \bar{x} = b_{xy} (y - \bar{y})$$
$$x - 65 = 0.467 (65 - 60)$$
$$x = 65 + 0.467 (5)$$
$$x = 65 + 2.335$$
$$x = 67.335 \approx 67 \text{ marks}$$

$$b_{yx} = r \times \frac{\sigma_y}{\sigma_x} = 0.7 \times \frac{15}{10} = 0.7 \times 1.5 = 1.05$$

To find y when x is 70.

We use regression line of y on x.

$$y - \bar{y} = b_{yx} (x - \bar{x})$$
$$y - 60 = 1.05 (70 - 65)$$
$$y = 60 + 1.05 (5)$$
$$y = 60 + 5.25$$
$$y = 65.25 \approx 65 \text{ marks}$$

(5) From the regression equations given below
$3x + 2y - 26 = 0$
$6x + y - 31 = 0$
Obtain:
(i) Mean of x and y
(ii) y when x = 4
(iii) x when y = 7

Solution:

$$3x + 2y = 26 \qquad \text{............ (1)}$$

Multiplying by 2,

$$6x + 4y = 52 \qquad \text{............ (2)}$$

$$6x + y = 31 \qquad \text{............ (3)}$$

Subtracting (2) by (3) we get,

$$\begin{array}{rl} 6x + 4y &= 52 \\ -\quad 6x + y &= 31 \\ \hline 3y &= 21 \\ y &= 7 \end{array} \qquad \begin{array}{rl} \text{Put } y \text{ in (1)} & \\ 3x + 2(7) &= 26 \\ 3x + 14 &= 26 \\ 3x &= 12 \\ x &= 4 \end{array}$$

$$(1)\ \bar{x} = 4, \quad \bar{y} = 7$$

Let us assume,

$$3x + 2y = 26 \text{ as regression line of y on x.}$$

$$2y = -3x + 26$$

$$y = -\frac{3}{2}x + 13$$

$$b_{yx} = \text{coefficient of } x = -\frac{3}{2}.$$

And, $6x + y = 31$ as regression line of x on y.

$$6x = -y + 31$$

$$x = -\frac{1}{6}y + 31$$

$$b_{xy} = \text{coefficient of } y = -\frac{1}{6}.$$

$$r = \pm\sqrt{b_{yx} \cdot b_{xy}}$$

$$= -\sqrt{-\frac{3}{2} \cdot -\frac{1}{6}}$$

$$= -\sqrt{\frac{1}{4}}$$

$$= -0.5$$

To find x when y = 7.

We use regression line of x on y.

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - 4 = -\frac{1}{6}(7 - 7)$$

$$x = 4 - \frac{1}{6}(0)$$

$$x = 4 - 0$$

(III) $\boxed{x = 4}$

To find y when x = 4.

We use regression line of y on x.

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$y - 7 = -\frac{3}{2}(4 - 4)$$

$$y = 7 - \frac{1}{6}(0)$$

$$y = 7 - 0$$

(II) $\boxed{y = 7}$

(6) From the regression equations given below
$2x + y - 8 = 0$
$x + 5y - 13 = 0$
Obtain:
(i) Mean of X and Y
(ii) Coefficient of correlation

Solution:

$$2x + y = 8 \quad \text{............ (1)}$$
$$x + 5y = 13 \quad \text{............ (2)}$$

Multiplying (1) by 5,

$$10x + 5y = 40 \quad \text{............ (3)}$$

Subtracting (3) by (2) we get,

$$\begin{array}{rl} 10x + 5y & = 40 \\ x + 5y & = 13 \\ (-) \quad (-) & \quad (-) \\ \hline 9x & = 27 \\ x & = 3 \end{array}$$

Put x in (1)

$$2(3) + y = 8$$
$$6 \times y = 8$$
$$y = 8 - 6$$
$$y = 2$$

(1) $\boxed{\bar{x} = 3, \quad \bar{y} = 2}$

Let us assume,

$2x + y = 8$ as regression line of x on y.

$$2x = -y + 8$$
$$x = -\frac{1}{2}y + 4$$

b_{xy} = coefficient of y $= -\frac{1}{2}$.

And, $x + 5y = 13$ as regression line of y on x.

$$5y = -x + 13$$
$$y = -\frac{1}{5}x + \frac{13}{5}$$

b_{yx} = coefficient of x $= -\frac{1}{5}$.

$$r = \pm\sqrt{b_{xy} \cdot b_{yx}}$$

$$= -\sqrt{-\frac{1}{5} \cdot -\frac{1}{2}}$$

$$= -\sqrt{0.1}$$

$$\boxed{r \;= -\,0.316}$$

(7) Two random variables have regression lines with equations:
$4x - y - 23 = 0$
$3x - 2y + 4 = 0$
Obtain:
(i) Mean of x and y
(ii) The coefficient of correlation and σ_y, if $\sigma_x^2 = 12$

Solution:

$$4x - y = 23 \quad \text{............ (1)}$$
$$3x - 2y = -4 \quad \text{............ (2)}$$

Multiplying by (2),

$$8x - 2y = 46 \quad \text{............ (3)}$$

Subtracting (3) by (2),

$$\begin{array}{rcl} 8x - 2y &=& 46 \\ 3x - 2y &=& -4 \\ (-) \quad (+) & & (+) \\ \hline 5x &=& 50 \\ x &=& 10 \end{array} \qquad \begin{array}{l} \text{Put x in (1)} \\ 4(10) - y = 23 \\ 40 - y = 23 \\ y = 40 - 23 \\ y = 17 \end{array}$$

$$\boxed{\bar{x} = 10, \quad \bar{y} = 17}$$

Let us assume,

$$4x - y = 23 \text{ as regression line of x on y.}$$
$$4x = y + 23$$
$$x = \frac{1}{4}y + \frac{23}{4}$$
$$b_{xy} = \text{coefficient of } y = \frac{1}{4}.$$

And, $$3x - 2y = -4 \text{ as regression line of y on x.}$$
$$2y = 3x + 4$$
$$y = \frac{3}{2}x + 2$$
$$b_{yx} = \text{coefficient of } x = \frac{3}{2}.$$
$$r = \pm\sqrt{b_{yx} \times b_{xy}}$$

$$= \sqrt{\frac{3}{2} \times \frac{1}{4}}$$

$$= \sqrt{1.5 \times 0.25}$$

$$= \sqrt{0.375}$$

$$r = 0.6124$$

$$b_{xy} = r \times \frac{\sigma_x}{\sigma_y}$$

$$\frac{1}{4} = 0.6124 \times \frac{\sqrt{12}}{\sigma_y}$$

$$\frac{1}{4} = \frac{2.1214}{\sigma_y}$$

$$\sigma_y = 2.1214 \times 4$$

$$\sigma_y = 8.4856$$

Time Series Analysis

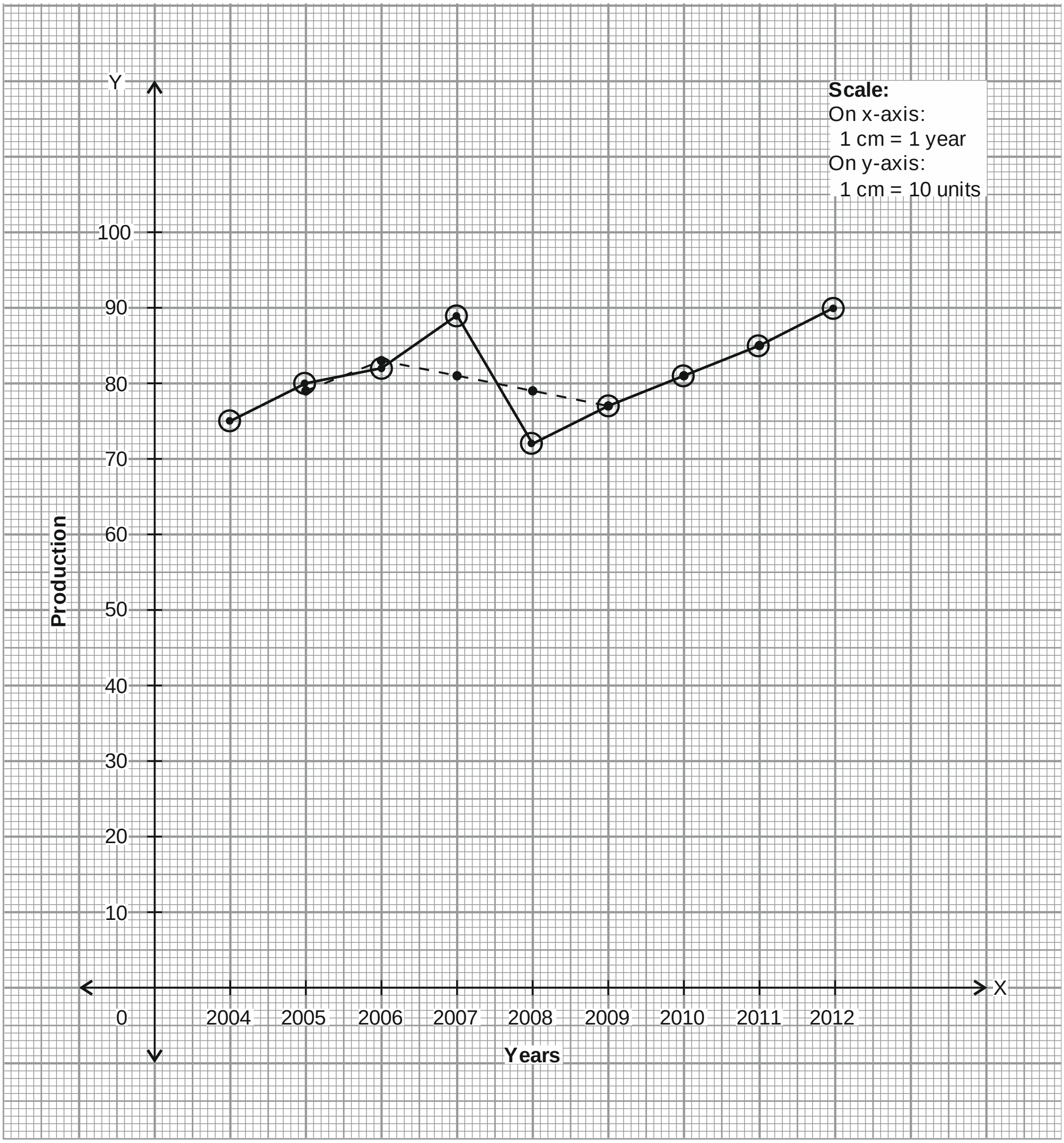
(1) Compute 3 yearly moving averages for the following time series and plot the original values and trend values on a graph.

Years	Production (in 1000 units)
2004	75
2005	80
2006	82
2007	89
2008	73
2009	77
2010	81
2011	85
2012	90

Solution:

Years	Production (in 1000 units)	3 yearly moving totals	3 yearly moving averages
2004	75	–	–
2005	80	237	79
2006	82	251	83.67
2007	89	244	81.33
2008	73	239	79.67
2009	77	231	77
2010	81	243	81
2011	85	256	85.33

2012	90	–	–



Y
Scale:
On x-axis:
1 cm = 1 year
On y-axis:
1 cm = 10 units
100
90
80
70
60
50
40
30
20
10
0
Production
X
2004
2005
2006
2007
2008
2009
2010
2011
2012
Years

(2) Calculate trend values by method of moving averages of length 5 for the following time series:

Years	Turnover (in crores)
2001	52
2002	59
2003	63
2004	68
2005	72
2006	80
2007	82
2008	84
2009	89
2010	90
2011	91
2012	92

Solution:

Years	Turnover (in crores)	5 yearly moving totals	5 yearly moving averages
2001	52	–	–
2002	59	–	–
2003	63	314	62.8
2004	68	342	68.4
2005	72	365	73
2006	80	386	77.2
2007	82	407	81.4
2008	84	425	85
2009	89	436	87.2
2010	90	446	89.2
2011	91	–	–
2012	92	–	–

(3) Calculate the 4 yearly moving averages for the following Time Series.

Years	production (Rs. in 00s)
2003	32
2004	27
2005	25
2006	30
2007	35
2008	33
2009	28
2010	29
2011	30
2012	32
2013	30

Solution:

Years	production (Rs. in 00s)	4 yearly moving totals	Centered totals	4 yearly moving averages
2003	32	–	–	–
2004	27	–	–	–
		114		
2005	25		231	28.875
		117		
2006	30		240	30
		123		
2007	35		249	31.125
		126		
2008	33		251	31.375
		125		

2009	28		245	30.625
		120		
2010	29		239	29.875
		119		
2011	30		240	30
		121		
2012	32	–	–	–
2013	30	–	–	–

(4) The following data gives rainfall in cms of a city for 8 years

Years	Rainfall
2008	50
2009	63
2010	85
2011	90
2012	75
2013	68
2014	53
2015	60

Fit a trend line by method of Least Squares and estimate rainfall for the year 2016.

Solution:

Years	Rainfall	x = 2(yr – 2011.5)	x^2	xy	y = 68 – 0.27x
2008	50	– 7	49	– 350	69.89
2009	63	– 5	25	– 315	69.35
2010	85	– 3	9	– 255	68.81
2011	90	– 1	1	– 90	68.27
2012	75	1	1	75	67.73
2013	68	3	9	204	67.19
2014	53	5	25	265	66.65
2015	60	7	49	420	66.11
Total	**544**	**0**	**168**	**– 46**	

Normal Equations:

$$\sum y = na + b\sum x$$

$$544 = 8a + b(0)$$

$$a = \frac{544}{8} = 68$$

$$\sum xy = a\sum x + b\sum x^2$$

$$-46 = a(0) + b(168)$$

$$\therefore b = \frac{-46}{168} = -0.27$$

Trend Line $y = a + bx = 68 + (-0.27)x$

$= 68 - 0.27x$

For year 2016, $x = 2\,(\text{Year} - 2011.5) = 2\,(2016 - 2011.5)$

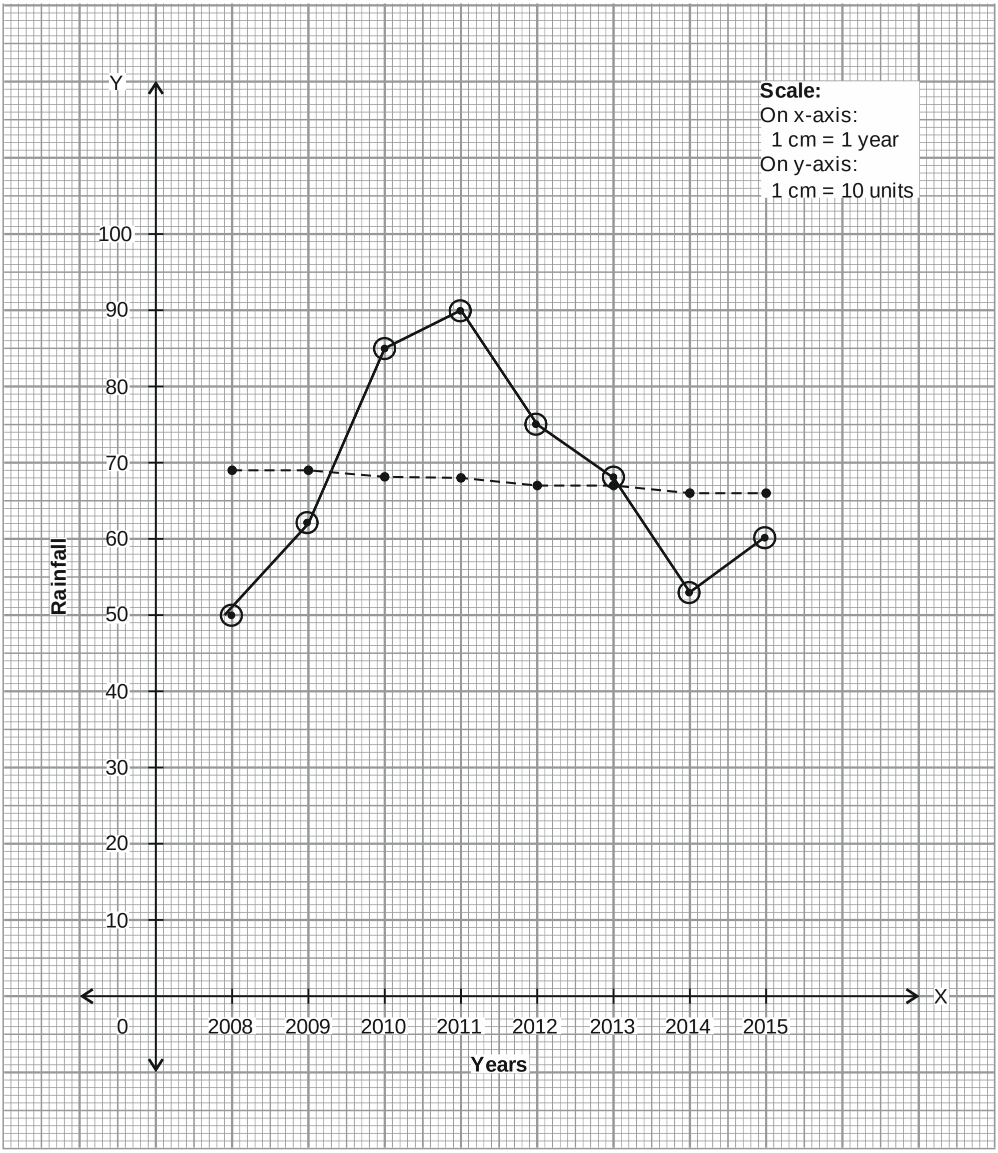
$= 9$

Trend value $y = 68 - 0.27x$

$= 68 - 0.27\,(9)$

$= 68 - 2.43$

$= 65.57$


Scale:
On x-axis:
1 cm = 1 year
On y-axis:
1 cm = 10 units
Y
X
Rainfall
Years
0
10
20
30
40
50
60
70
80
90
100
2008
2009
2010
2011
2012
2013
2014
2015

(5) Given below are traffic offences committed in a city during 2003 to 2011. Fit a straight line and estimate traffic offences for 2010.

Years	**Traffic Offences ('00s)**
2003	120
2004	125
2005	132
2006	147
2007	150
2008	123
2009	135
2010	141
2011	142

Solution:

Years	**Traffic Offences ('00s)**	**x = Year - 2007**	**x^2**	**xy**	**y = 135 + 1.97x**
2003	120	– 4	16	– 480	127.12
2004	125	– 3	9	– 375	129.09
2005	132	– 2	4	– 264	131.06
2006	147	– 1	1	– 147	133.03
2007	150	0	0	0	135
2008	123	1	1	123	136.97
2009	135	2	4	270	138.94
2010	141	3	9	423	140.91
2011	142	4	16	568	142.88
Total	**1215**	**0**	**60**	**118**	

Normal Equations:

$$\sum y = na + b\sum x$$

$$1215 = 9a + (0)\ b$$

$$\therefore a = \frac{1215}{9} = 135$$

$$\sum xy = a\sum x + b\sum x^2$$

$$118 = a(0) + b(60)$$

$$\therefore b = \frac{118}{60} = 1.97$$

Trend Line $y = a + bx = 135 + 1.97x$

For year 2010, $x = \text{Year} - 2007 = 2010 - 2007 = 3$

Trend value $y = 135 + 1.97x$

$= 135 + 1.97(3)$

$= 135 + 5.91$

$= 140.91$

(6) Determine the Seasonal Indices by Method of Simple Averages from the following Time Series:

Years	**Sale (1000 units)**			
	Q_1	**Q_2**	**Q_3**	**Q_4**
2006	120	110	105	150
2007	140	140	125	180
2008	135	170	150	200
2009	140	180	160	220
2010	135	230	180	240

Solution:

Years	**Sale (1000 units)**				
	Q_1	**Q_2**	**Q_3**	**Q_4**	
2006	120	110	105	150	
2007	140	140	125	180	
2008	135	170	150	200	
2009	140	180	160	220	
2010	135	230	180	240	
Total	670	830	720	990	
Average	134	166	144	198	**642**

$$\text{Grand Average} = \frac{642}{4} = 160.5$$

$$\text{Seasonal Indices} = \frac{\text{Quarterly Average}}{\text{Grand Average}} \times 100$$

$$Q_1 = \frac{134}{160.5} \times 100 \quad Q_2 = \frac{166}{160.5} \times 100 \quad Q_3 = \frac{144}{160.5} \times 100 \quad Q_4 = \frac{198}{160.5} \times 100$$

$$= 83.49 \qquad = 103.43 \qquad = 89.72 \qquad = 123.66$$

(7) Compute the seasonal components by the method of seasonal indices:

Years	**Demand (in lakh Rs.)**			
	Q_1	**Q_2**	**Q_3**	**Q_4**
2005	129	135	158	144
2006	130	132	160	165
2007	143	145	142	171
2008	128	123	155	155
2009	110	105	175	130

Solution:

Years	**Demand (in lakh Rs.)**				
	Q_1	**Q_2**	**Q_3**	**Q_4**	
2005	129	135	158	144	
2006	130	132	160	165	
2007	143	145	142	171	
2008	128	123	155	155	
2009	110	105	175	130	
Total	640	640	790	765	
Average	128	128	158	153	**567**

$$\text{Grand Average} = \frac{567}{4} = 141.75$$

$$\text{Seasonal Indices} = \frac{\text{Quarterly Average}}{\text{Grand Average}} \times 100$$

$$Q_1 = \frac{128}{141.75} \times 100 \quad Q_2 = \frac{128}{141.75} \times 100 \quad Q_3 = \frac{158}{141.75} \times 100 \quad Q_4 = \frac{153}{141.75} \times 100$$

$$= 90.29 \qquad = 90.29 \qquad = 111.46 \qquad = 107.94$$

Index Numbers

(1) Construct simple and weighted price index numbers for the following data by average of price relatives and aggregative method.

Comm.	Price		Weights
	Base year	Current year	
A	50	60	20
B	73	80	10
C	20	35	25
D	45	50	15
E	60	79	30

Solution:

Comm.	Price		Weights (w)	$i = \frac{P_1}{P_0} \times 100$	iw	P_0w	P_1w
	Base year (P_0)	Current year (P_1)					
A	50	60	20	120	2400	1000	1200
B	73	80	10	109.59	1095.9	730	800
C	20	35	25	175	4375	500	875
D	45	50	15	111.11	1666.65	675	750
E	60	79	30	131.67	3950.1	1800	2370
Total	**248**	**304**	**100**	**647.37**	**13487.65**	**4705**	**5995**

Simple Average of Price Relatives:

$$I = \frac{\sum i}{k} = \frac{647.37}{5} = 129.474$$

Simple Aggregate Method:

$$I = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{304}{248} \times 100 = 122.58$$

Weighted Average of Price Relatives:

$$I = \frac{\sum iw}{\sum w} = \frac{13487.65}{100} = 134.8765$$

Weighted Aggregate Method:

$$I = \frac{\sum P_1 w}{\sum P_0 w} \times 100 = \frac{5995}{4705} \times 100 = 127.42$$

(2) Construct Laspeyre's, Paasche's, Fisher's, Dorbish Boweley's & Marshall Edgeworth Price Index Numbers from the following Data:

Comm.	**Base Year**		**Current Year**	
	Price	**Qty.**	**Price**	**Qty.**
A	45	20	25	22
B	45	25	35	30
C	62	15	20	20
D	35	10	15	12
E	30	5	30	7

Solution:

Comm.	**Base Year**		**Current Year**		**p_0q_0**	**p_0q_1**	**p_1q_0**	**p_1q_1**
	Price (p_0)	**Qty. (q_0)**	**Price (p_1)**	**Qty. (q_1)**				
A	45	20	25	22	900	990	500	550
B	45	25	35	30	1125	1350	275	1050
C	62	15	20	20	930	1240	300	400
D	35	10	15	12	350	420	150	180
E	30	5	30	7	150	210	150	210
				Total	**3455**	**4210**	**1975**	**2390**

Laspeyre's Price Index Number:

$$I_L = \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$$

$$= \frac{1975}{3455} \times 100 = 57.16$$

Paasche's Price Index Number:

$$I_P = \frac{\sum p_1q_1}{\sum p_0q_1} \times 100$$

$$= \frac{2390}{4210} \times 100 = 56.76$$

Fisher's Price Index Number:

$$I_F = \sqrt{I_L \times I_P}$$

$$= \sqrt{57.16 \times 56.76}$$

$$= \sqrt{3244.4016} = 56.96$$

Dorbish Bowley's Price Index Number:

$$I_{DB} = \frac{I_L + I_P}{2}$$

$$= \frac{57.16 + 56.76}{2}$$

$$= 56.96$$

Marshall Edgeworth's Price Index Number:

$$I_{ME} = \frac{\sum p_1q_0 + \sum p_1q_1}{\sum p_0q_0 + \sum p_0q_1} \times 100$$

$$= \frac{1975 + 2390}{3455 + 4210} \times 100$$

$$= \frac{4365}{7655} \times 100 = 56.95$$

(3) Construct Laspeyre's, Paasche's, Fisher's and Dorbish Bowley's price Index Numbers from the following Data and test them for Factor Reversal.

Comm.	**Base Price**	**Current Price**	**Base Quantity**	**Current Quantity**
A	35	40	15	18
B	45	50	25	27
C	50	52	10	12
D	25	30	15	16
E	30	30	20	20

Solution:

Comm.	**Base Price** p_0	**Current Price** p_1	**Base Quantity** q_0	**Current Quantity** q_1	p_0q_0	p_0q_1	p_1q_0	p_1q_1
A	35	40	15	18	525	630	600	720
B	45	50	25	27	1125	1215	1250	1350
C	50	52	10	12	500	600	520	624
D	25	30	15	16	375	400	450	480
E	30	30	20	20	600	600	600	600
				Total	**3125**	**3445**	**3420**	**3774**

Laspeyre's Price Index Number:

$$I_L = \frac{\sum p_1q_0}{\sum p_0q_0} \times 100$$

$$= \frac{3420}{3125} \times 100 = 109.44$$

Paasche's Price Index Number:

$$I_P = \frac{\sum p_1q_1}{\sum p_0q_1} \times 100$$

$$= \frac{3774}{3445} \times 100 = 109.55$$

Fisher's Price Index Number:

$$I_F = \sqrt{I_L \times I_P}$$

$$= \sqrt{109.44 \times 109.55}$$

$$= \sqrt{11989.152} = 109.49$$

Dorbish Bowley's Price Index Number:

$$I_{DB} = \frac{I_L + I_P}{2}$$

$$= \frac{109.44 + 109.55}{2}$$

$$= \frac{218.99}{2} = 109.495$$

(4) Construct Laspeyre's, Paasche's, Fisher's and Marshall Edgeworth price index numbers and check which satisfies Time Reversal Test.

Comm.	Base Year		Current Year	
	Price	Qty.	Price	Qty.
A	15	15	18	16
B	20	10	22	15
C	18	14	20	18
D	12	13	14	7

Solution:

Comm.	Base Year		Current Year		p_0q_0	p_0q_1	p_1q_0	p_1q_1
	Price p_0	Qty. q_0	Price p_1	Qty. q_1				
A	15	15	18	16	225	240	270	288
B	20	10	22	15	200	300	220	330
C	18	14	20	18	252	324	280	360
D	12	13	14	7	156	84	182	98
				Total	**883**	**948**	**952**	**1076**

Laspeyre's Price Index Number:

$$I_L = \frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_0q_1}{\sum p_1q_1}$$

$$= \frac{952}{853} \times \frac{948}{1076} = 1.007$$

Paasche's Price Index Number:

$$I_P = \frac{\sum p_1q_1}{\sum p_0q_1} \times \frac{\sum p_0q_0}{\sum p_1q_0}$$

$$= \frac{1076}{948} \times \frac{833}{952} = 0.993$$

Fisher's Price Index Number:

$$I_F = \sqrt{\frac{\sum p_1q_0}{\sum p_0q_0} \times \frac{\sum p_1q_1}{\sum p_0q_1}} \times \sqrt{\frac{\sum p_0q_1}{\sum p_1q_1} \times \frac{\sum p_0q_0}{\sum p_1q_0}}$$

$$= \sqrt{\frac{952}{853} \times \frac{1076}{948}} \times \sqrt{\frac{948}{1076} \times \frac{833}{952}}$$

$$= 1$$

Marshall Edgeworth's Price Index Number:

$$I_{ME} = \frac{\sum p_1(q_0 + q_1)}{\sum p_0(q_0 + q_1)} \times \frac{\sum p_0(q_1 + q_0)}{\sum p_1(q_1 + q_0)} = 1$$

$\therefore$ Fisher's and Marshall Edgeworth's Price Index Number satisfies Time Reversal Test with value equal to 1.

(5) Calculate chain base index number for the following data.

Years	**Profits (in lakh Rs.)**				
	2010	**2011**	**2012**	**2013**	**2014**
A	40	45	48	45	50
B	25	28	29	28	32
C	35	32	36	37	41

Solution:

$$\text{Link Relative} = \frac{\text{Current Year Price}}{\text{Previous Year Price}} \times 100$$

Link Relatives:

Years	**A**	**B**	**C**	**Average L. R.**	**CBI**
2010	100	100	100	100	100
2011	112.5	112	91.43	105.31	105.31
2012	106.67	103.57	112.5	107.58	113.29
2013	93.75	96.55	102.78	97.69	110.67
2014	111.11	114.29	110.81	112.07	124.03

Chain Base Index Number:

$$\text{CBI} = \frac{\text{Link Relative of Current Year} \times \text{CBI of Previous Year}}{100}$$

(6) Construct chain base index number for the series given below:

Years	Cost of Living Index Number
2002	100
2003	126
2004	129
2005	136
2006	140
2007	145
2008	151
2009	160
2010	162
2011	170
2012	175

Solution:

Years	Cost of Living Index Number	Link Relative	Chain Base Index Number
2002	100	100	100
2003	126	126	126
2004	129	102.38	128.99
2005	136	105.43	135.99
2006	140	102.94	139.98
2007	145	103.57	144.97
2008	151	104.14	150.97

2009	160	105.96	159.97
2010	162	101.25	161.97
2011	170	104.94	169.97
2012	175	102.94	174.97

$$\text{Link Relative} = \frac{\text{Current Year Price}}{\text{Previous Year Price}} \times 100$$

Chain Base Index Number:

$$\text{CBI} = \frac{\text{Link Relative of Current Year} \times \text{CBI of Previous Year}}{100}$$

(7) Shift the base year of the following series to 2008.

Years	Index Nos.
2006	100
2007	115
2008	120
2009	132
2010	150
2011	160
2012	164
2013	170
2014	182

Solution:

Years	Index Nos.	New Index No.
2006	100	83.33
2007	115	95.83
2008	120	100
2009	132	110
2010	150	125
2011	160	133.33
2012	164	136.67
2013	170	141.67
2014	182	151.67

$$\text{New Index No.} = \frac{\text{Old Index No.}}{\text{Index No. of New Base Year}} \times 100$$

(8) Construct cost of living index number for the following data by family budget method.

Group	**Weights (W)**	**Index No. (I)**
Food	40	185
Fuel & Lighting	25	120
Clothing	25	130
House Rent	5	140
Miscellaneous	5	120

Solution:

Group	**Weights (W)**	**Index No. (I)**	**iw**
Food	40	185	7400
Fuel & Lighting	25	120	3000
Clothing	25	130	3250
House Rent	5	140	700
Miscellaneous	5	120	600
	100		**14950**

Cost of Living Index Number by Family Budget Method:

$$= \frac{\sum iw}{\sum w} = \frac{14950}{100} = 149.5$$

(9) Construct cost of living index numbers for the year 2010 by aggregative expenditure method.

Commodity	Price per unit		Quantity
	2010	2015	in 2010
A	60	65	25
B	80	90	20
C	20	25	10
D	40	45	25
E	50	55	10

Solution:

Commodity	Price per unit		Quantity	p_1q_0	p_0q_0
	2010 p_0	2015 p_1	in 2010 q_0		
A	60	65	25	1625	1500
B	80	90	20	1800	1600
C	20	25	10	250	200
D	40	45	25	1125	1000
E	50	55	10	550	500
				5350	**4800**

Cost of Living Index Number:

$$= \frac{5350}{4800} \times 100 = 111.458$$

(10) Construct cost of living index numbers for 2000 by family budget method.

Group	Price		Expenses
	2000	2005	(in %)
Food	90	120	40
Light & Fuel	30	40	10
Clothing	40	40	15
Rent	40	40	15
Miscellaneous	20	25	20

Solution:

Group	Price		Expenses	$I = \frac{p_1}{p_0} \times 100$	iw
	2000 p_0	2005 p_1	(in %) w		
Food	90	120	40	133.33	5333.2
Light & Fuel	30	40	10	133.33	1333.3
Clothing	40	40	15	100	1500
Rent	40	40	15	100	1500
Miscellaneous	20	25	20	125	2500
			100		**12166.5**

Cost of Living Index Number by Family Budget Method:

$$= \frac{\sum iw}{\sum w} = \frac{12166.5}{100} = 121.665$$

Discrete Probability Distributions

(1) The probability that a student selected at random from a group of students is Gujarati is 1/3. If 5 students are selected from that group find the probability that

(i) At least one is a Gujarati

(ii) 3 or more are Gujarati

(iii) Exactly 4 are Gujarati

Solution:

$$n = 5,\ p = \frac{1}{3},\ q = \frac{2}{3}$$

$$P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x} = {}^{5}C_{x} \left(\frac{1}{3}\right)^{x} \left(\frac{2}{3}\right)^{5-x}$$

(i)

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$$

$$= 1 - {}^{5}C_{0} \left(\frac{1}{3}\right)^{0} \left(\frac{2}{3}\right)^{5} = 1 - \left[1 \times 1 \times \frac{32}{243}\right]$$

$$= 1 - \frac{32}{243} = \frac{211}{243} = 0.8683$$

(ii)

$$P(3 \text{ or more}) = 1 - P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[{}^{5}C_{0}\left(\frac{1}{3}\right)^{0}\left(\frac{2}{3}\right)^{5} + {}^{5}C_{1}\left(\frac{1}{3}\right)^{1}\left(\frac{2}{3}\right)^{4} + {}^{5}C_{2}\left(\frac{1}{3}\right)^{2}\left(\frac{2}{3}\right)^{3}\right]$$

$$= 1 - \left[(1)\,(1)\left(\frac{32}{243}\right) + (5)\left(\frac{1}{3}\right)\left(\frac{16}{81}\right) + \left(\frac{5 \times 4}{2}\right)\left(\frac{1}{9}\right)\left(\frac{8}{27}\right)\right]$$

$$= 1 - \left[\frac{32}{243} + \frac{80}{243} + \frac{80}{243}\right]$$

$$= 1 - \frac{192}{243} = \frac{51}{243}$$

$$= \underline{\underline{0.2099}}$$

(iii)

$$P(X = 4) = {}^{5}C_{4}\left(\frac{1}{3}\right)^{4}\left(\frac{2}{3}\right)^{1} = (5)\left(\frac{1}{81}\right)\left(\frac{2}{3}\right) = \frac{10}{243} = \underline{\underline{0.04115}}$$

(2) It is observed that, on an average one person out five likes coffee. If 4 persons are selected at random find the probability that

(i) At the most one likes coffee

(ii) Exactly 2 likes coffee

(iii) At least 2 likes coffee

Solution:

$$n = 4,\ p = \frac{1}{5},\ q = \frac{4}{5}$$

$$P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x} = {}^{4}C_{x} \left(\frac{1}{5}\right)^{x} \left(\frac{4}{5}\right)^{4-x}$$

(i)

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= {}^{4}C_{0} \left(\frac{1}{5}\right)^{0} \left(\frac{4}{5}\right)^{4} + {}^{4}C_{1} \left(\frac{1}{5}\right)^{1} \left(\frac{4}{5}\right)^{3}$$

$$= (1)\ (1) \left(\frac{256}{625}\right) + (4) \left(\frac{1}{5}\right) \left(\frac{64}{125}\right)$$

$$= \frac{256}{625} + \frac{256}{625} = \frac{512}{625} = \underline{\underline{0.8192}}$$

(ii)

$$P(X = 2) = {}^{4}C_{2} \left(\frac{1}{5}\right)^{2} \left(\frac{4}{5}\right)^{2}$$

$$= (6) \left(\frac{1}{25}\right) \left(\frac{16}{25}\right) = \frac{96}{625} = \underline{\underline{0.1536}}$$

(iii)

$$P(X \geq 2) = 1 - P(X < 2)$$

$$= 1 - [P(X = 0) + P(X = 1)]$$

$$= 1 - [P(X = 0) + P(X = 1)]$$

$$= 1 - \left[\frac{256}{625} + \frac{256}{625}\right] = 1 - \frac{512}{625}$$

$$= 1 - 0.8192$$

$$= \underline{\underline{0.1808}}$$

(3) If mean and variance of binomial distribution is 4 & 2 respectively. Find

(i) Probability of 6 successes

(ii) Probability of less than 2 successes

Solution:

$$\text{Mean} = np = 4 \quad \text{Var} = npq = 2$$

$$npq = 2$$

$$4q = 2$$

$q = \frac{1}{2}$, $p = \frac{1}{2}$, $n = 8$

$$P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x}$$

(i) $$P(X = 6) = {}^{8}C_{6} \left(\frac{1}{2}\right)^{6} \left(\frac{1}{2}\right)^{2} = \left(\frac{8 \times 7}{2}\right) \left(\frac{1}{64}\right) \left(\frac{1}{4}\right) = \frac{7}{64} = \underline{\underline{0.109375}}$$

(ii) $$P(X < 2) = P(X = 0) + P(X = 1)$$

$$= {}^{8}C_{0} \left(\frac{1}{2}\right)^{0} \left(\frac{1}{2}\right)^{8} + {}^{8}C_{1} \left(\frac{1}{2}\right)^{1} \left(\frac{1}{2}\right)^{7}$$

$$= (1)\,(1) \left(\frac{1}{256}\right) + (8) \left(\frac{1}{2}\right)\left(\frac{1}{128}\right)$$

$$= \frac{1}{258} + \frac{8}{256}$$

$$= \frac{1 + 8}{256}$$

$$= \frac{9}{256}$$

$$= \underline{\underline{0.03515625}}$$

(4) The average number of incoming calls in at a telephone switch board per minute is 3. Find the probability that during a given minute,

(i) 2 or more calls are received

(ii) At least one call is received

(iii) Exactly 4 calls are received

(Hint $e^{-3} = 0.0498$)

Solution:

$m = 3$,

$$P(X = x) = \frac{e^{-m} m^x}{x!} = \frac{e^{-3}\, 3^x}{x!} \qquad x = 0, 1, 2, 3$$

(i) $P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)]$

$$= 1 - \left[\frac{e^{-3}\,(3)^0}{0!} + \frac{e^{-3}\, 3^1}{1!}\right]$$

$$= 1 - \left(\frac{0.0498\ (1)}{(1)} + \frac{0.0498\ (3)}{1}\right)$$

$$= 1 - (0.0498 + 0.1494)$$

$$= 1 - 0.1992 = \underline{\underline{0.8008}}$$

(ii) $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$

$$= 1 - \frac{e^{-3}\,(3)^0}{0!}$$

$$= 1 - 0.0498$$

$$= \underline{\underline{0.9502}}$$

(iii) $P(X = 4) = \frac{e^{-3}\,(3)^4}{4!} = \frac{(0.0498)\ (81)}{4 \times 3 \times 2}$

$$= (0.006225)\ (27)$$

$$= \underline{\underline{0.168075}}$$

(5) It is observed that on an average 5 items manufactured on a machine per day are defectives. Find the probability that on a particular day there are

(i) No defectives

(ii) Exactly 3 defective

(iii) At most two defectives

(Hint $e^{-5} = 0.00673$)

Solution:

$m = 5$,

$$P(X = x) = \frac{e^{-m} m^x}{x!} = \frac{e^{-5}\, 5^x}{x!} \qquad x = 0, 1, 2, 3, 4, 5$$

(i) $P(X = 0) = \frac{e^{-5}\,(5)^0}{0!} = \frac{0.00673 \times 1}{1} = \underline{\underline{0.00673}}$

(ii) $P(X = 3) = \frac{e^{-5}\,(5)^3}{3!} = \frac{0.00673 \times 125}{6} = \frac{0.84125}{6} = \underline{\underline{0.140208}}$

(iii) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$$= \frac{e^{-5}\,(5)^0}{0!} + \frac{e^{-5}\,(5)^1}{1!} + \frac{e^{-5}\,(5)^2}{2!}$$

$$= \frac{0.00673 \times 1}{1} + \frac{0.00673 \times 5}{1} + \frac{0.00673 \times 25}{2}$$

$$= 0.00673 + 0.03365 + 0.084125$$

$$= \underline{\underline{0.124505}}$$

(6) From the past experience it is found that demand for certain product follows Poisson distribution with mean 4 units per week. If certain shop keeps 6 units during particular week, find probability that demand will exceed supply during that week. (Hint $e^{-4} = 0.0183$)

Solution:

$$P(X > 6) = ?$$

$$P(X = x) = \frac{e^{-m}\,m^x}{x!} = \frac{e^{-4}\,4^x}{x!} \qquad x = 0, 1, \ldots\ldots$$

$$P(X > 6) = 1 - P(X \leq 6)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)]$$

$$= 1 - \left[\frac{e^{-4}\,(4)^0}{0!} + \frac{e^{-4}\,(4)^1}{1!} + \frac{e^{-4}\,(4)^2}{2!} + \frac{e^{-4}\,(4)^3}{3!} + \frac{e^{-4}\,(4)^4}{4!} + \frac{e^{-4}\,(4)^5}{5!} + \frac{e^{-4}\,(4)^6}{6!}\right]$$

$$= 1 - \left[\frac{0.0183\,(1)}{1} + \frac{0.0183\,(4)}{1} + \frac{0.0183\,(16)}{2} + \frac{0.0183\,(64)}{6} + \frac{0.0183\,(256)}{24} + \frac{0.0183\,(1024)}{120} + \frac{0.0183\,(4096)}{720}\right]$$

$$= 1 - (0.0183 + 0.0732 + 0.1464 + 0.1952 + 0.1952 + 0.15616 + 0.1041)$$

$$= 1 - 0.888567$$

$$= \underline{\underline{0.111433}}$$

(7) The probability that a person is allergic to a certain drug is 0.001. out of 2000 individuals who are administered the drug, find the probability that

(i) Exactly 3 get the allergy

(ii) More than 2 get the allergy

(iii) None get the allergy

(Hint $e^{-2} = 0.1353$)

Solution:

$$m = np = 2000 \times 0.001 = 2$$

$$P(X = x) = \frac{e^{-m}\, m^x}{x!} = \frac{e^{-2}\, 2^x}{x!} \qquad x = 0, 1, \ldots\ldots, 2000$$

(i) $$P(X = 3) = \frac{e^{-2}\, 2^3}{3!} = \frac{0.1353 \times 8}{6} = \underline{\underline{0.1804}}$$

(ii) $$P(X > 2) = 1 - P(X \leq 2)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[\frac{e^{-2}\, 2^0}{0!} + \frac{e^{-2}\, 2^1}{1!} + \frac{e^{-2}\, 2^2}{2!}\right]$$

$$= 1 - \left[\frac{0.1353 \times 1}{1} + \frac{0.1353 \times 2}{1} + \frac{0.1353 \times 4}{2}\right]$$

$$= 1 - [0.1353 + 0.2706 + 0.2706]$$

$$= 1 - 0.6765$$

$$= \underline{\underline{0.3235}}$$

(iii) $$P(X > 2) = \frac{e^{-2}\, 2^0}{0!} = \frac{0.1353 \times 1}{1} = \underline{\underline{0.1353}}$$

(8) The probability that a car passing through a particular junction meets with an accident is 0.00005. Among 10000 cars that pass through that junction on a day, what is the probability that

(i) Exactly 2 cars meet with accident

(ii) No cars meet with accident

(iii) At least one car meets with accident

(Hint $e^{-0.5} = 0.607$)

Solution:

$$m = np = 10000 \times 0.00005 = 0.5$$

$$P(X = x) = \frac{e^{-m}\, m^x}{x!} = \frac{e^{-0.5}\, 0.5^x}{x!} \qquad x = 0, 1, \ldots\ldots, 10000$$

(i) $P(X = 2) = \frac{e^{-0.5}\,(0.5)^2}{2!} = \frac{0.607 \times 0.25}{2} = \underline{\underline{0.075875}}$

(ii) $P(X = 0) = \frac{e^{-0.5}\,(0.5)^0}{0!} = \frac{0.607 \times 1}{1} = \underline{\underline{0.607}}$

(iii) $P(X \geq 1) = 1 - P(X < 1)$

$= 1 - [P(X = 0)$

$= 1 - \frac{e^{-0.5}\,(0.5)^0}{0!}$

$= 1 - 0.607$

$= \underline{\underline{0.393}}$

(9) The income of group of 1000 people is normally distributed with mean 15000 and standard deviation Rs. 2000. Find

(i) Number of people with income between 13000 and 17000

(ii) Percentage of people with income less than 19000.

(iii) Proportion of people with income more than 13000

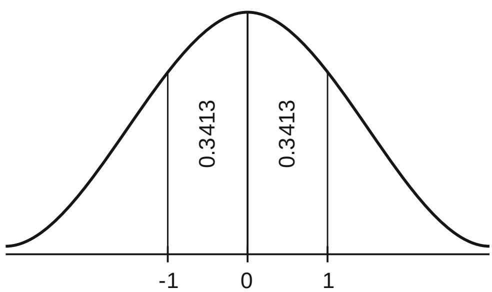
$P(0 < z < 1) = 0.3413$

$P(0 < z < 2) = 0.4772$

Solution:

X: Income of group of people

$X \sim N(\mu = 15000, \sigma = 2000)$



(i) $P(13000 < x < 17000) = P\left(\frac{13000\ \ 15000}{2000} < \frac{x - \mu}{\sigma} < \frac{17000 - 15000}{2000}\right)$

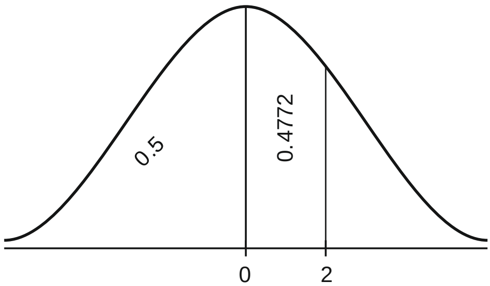
$= P(-1 < Z < 1)$

$= 0.3413 + 0.3413$

$= 0.6826$

Number = N × Prob $= 1000 \times 0.6826$

$= 682.6 = 683$



(ii) $P(X < 19000) = P\left(\frac{x-\mu}{\sigma} < \frac{19000 - 15000}{2000}\right)$

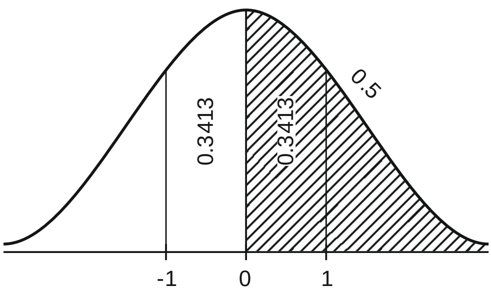
$= P(z < 2)$

$= 0.4772 + 0.5$

$= 0.9772$

Number = N × Prob $= 1000 \times 0.9712$

$= 977.2 = 977$



(iii) $P(X > 13000) = P\left(\frac{x-\mu}{\sigma} > \frac{13000 - 15000}{2000}\right)$

$= P(z > -1)$

$= 0.3413 + 0.5$

$= 0.8413$

(10) The heights of a group of 1000 people is normally distributed with mean 172 cms. and standard deviation 5 cms. Find

(i) Number of people with height greater than 182cms.

(ii) Percentage of people with heights between 162 & 182cms.

(iii) Proportion of people with heights between 162 & 167cms.

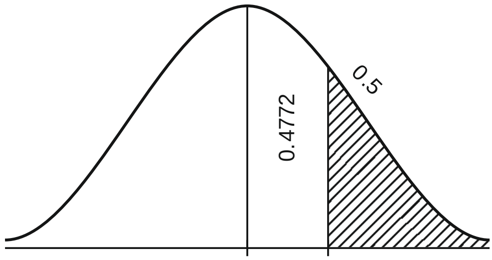
Hint: $P(0 < z < 2) = 0.4772$

$P(0 < z < 1) = 0.3413$

Solution:

X: Height

$X \sim N\ (\mu = 172, \sigma = 5)$



(i) $P(X > 182) = P\left(\frac{x-\mu}{\sigma} > \frac{182-172}{5}\right)$

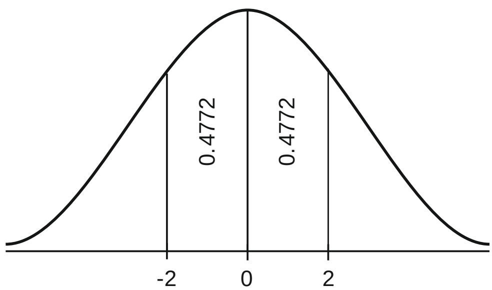
$= P\ (z > 2)$

$= 0.5 - 0.4772$

$= 0.0228$

Number = N × Prob $= 1000 \times 0.0228$

$= 22.8 = 23$



(ii) $P(162 < X < 182) = P\left(\frac{162-172}{5} < \frac{x-\mu}{\sigma} < \frac{182-172}{5}\right)$

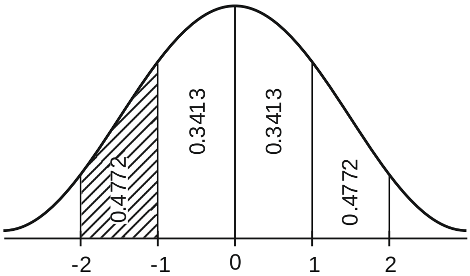
$= P(-2 < z < 2)$

$= 0.4772 + 0.4772$

$= 0.9544$

Number = N × Prob $= 1000 \times 0.9544$

$= 954.4 = 954$



(iii) $P(162 < X < 167) = P\left(\frac{162-172}{5} < \frac{x-\mu}{\sigma} < \frac{167-172}{5}\right)$

$= P(-2 < z < -1)$

$= 0.4772 - 0.3413$

$= 0.1359$

(11) 10000 candidates appeared for a certain examination. The mean scores were 60 with standard deviation 5. Assuming the scores to be normally distributed, find

(i) The minimum marks of top 25% of the students

(ii) The maximum marks of the bottom 25% of the students

Solution:

X: Candidates appeared

$X \sim N(\mu = 60, \sigma = 5)$

To find minimum marks of top 25% of the students, we compute Q_3.

$Q_3 = \mu + 0.675\,\sigma$

$\therefore Q_3 = 60 + 0.675 \times 5 = 63.375 \approx 64$

$\therefore$ the minimum marks of top 25% of the students is 64.

To find maximum marks of bottom 25% of the students, we compute Q_1.

$Q_1 = \mu - 0.675\,\sigma$

$\therefore Q_1 = 60 - 0.675 \times 5 = 56.625 \approx 57$

$\therefore$ the maximum marks of bottom 25% of the students is 57.