Linear Regression

EXERCISE 3.1 [PAGES 41 - 42]

Exercise 3.1 | Q 1 | Page 41

The HRD manager of a company wants to find a measure which he can use to fix the monthly income of persons applying for the job in the production department. As an experimental project, he collected data of 7 persons from that dep artment referring to years of service and their monthly incomes.

Years of service (X)	11	7	9	5	8	6	10
Monthly Income (₹ 1000's)(Y)	10	8	9	5	9	7	11

i. Find the regression equation of income on years of service.
ii. What initial start would you recommend for a person applying for the job after having served in a similar capacity in another company for 13 years?

Solution: (i) Here, X = Years of service,

Y = Income (₹ in 1000's)

$X = x_i$	**$Y = y_i$**	x_i^2	**$x_i y_i$**
11	10	121	110
7	8	49	56
9	6	81	54
5	5	25	25
8	9	64	72
6	7	36	42
10	11	100	110
56	**56**	**476**	**469**

From the table, we have

$n = 7, \sum x_i = 56, \sum y_i = 56, \sum x_i^2 = 476,$

$\sum x_i y_i = 469$

$$\bar{x} = \sum \frac{x_i}{n} = \frac{56}{7} = 8$$

$$\bar{y} = \sum \frac{y_i}{n} = \frac{56}{7} = 8$$

Now, $b_{YX} = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$\therefore b_{YX} = \frac{469 - 7 \times 8 \times 8}{476 - 7 \times (8)^2} = \frac{469 - 448}{476 - 448} = \frac{21}{28} = \frac{3}{4}$$

$\therefore b_{YX} = 0.75$

Also, $a = \bar{y} - b_{YX}\,\bar{x} = 8 - 0.75 \times 8 = 8 - 6 = 2$

∴ The regression equation of income (Y) on years of service (X) is

$Y = a + b_{YX} X$

∴ $Y = 2 + 0.75 X$

(ii) Given, years of service (X) = 13

∴ Substituting X = 13 in regression equation, we get

$Y = 2 + 0.75 (13)$

∴ $Y = 2 + 9.75$

∴ $Y = 11.75$ (₹ in 1000’s)

∴ I would recommend an initial start of ₹ 11,750 for a person who has served in similar capacity in another company for 13 years.

Exercise 3.1 | Q 2 | Page 41

Calculate the regression equations of X on Y and Y on X from the following data:

X	10	12	13	17	18
Y	5	6	7	9	13

Solution:

X = x_i	**Y = y_i**	x_i^2	y_i^2	**x_i y_i**
10	5	100	25	50
12	6	144	36	72
13	7	169	49	91
17	9	289	81	153
18	13	324	169	234
70	**40**	**1026**	**360**	**600**

From the table, we have,

n = 5, $\sum x_i = 70$, $\sum y_i = 40$, $\sum x_i y_i = 600$, $\sum x_i^2 = 1026$,

`sum"y"_"i"^2 = 360

$$\bar{x} = \sum \frac{x_i}{n} = \frac{70}{5} = 14,$$

$$\bar{y} = \sum \frac{y_i}{n} = \frac{40}{5} = 8$$

Now, for regression equation of X on Y

$$b_{XY} = \frac{\sum x_i y_i - n\,\bar{x}\,\bar{y}}{\sum y_i^2 - n\bar{y}^2}$$

$$= \frac{600 - 5 \times 14 \times 8}{360 - 5(8)^2} = \frac{600 - 560}{360 - 320} = \frac{40}{40} = 1$$

Also, $a' = \bar{x} - b_{XY}\,\bar{y} = 14 - 1(8) = 14 - 8 = 6$

∴ The regression equation of X on Y is

$X = a' + b_{XY}Y$

∴ $X = 6 + Y$

Now, for regression equation of Y on X

$$b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\,\bar{x}^2}$$

$$= \frac{600 - 5(14)(8)}{1026 - 5(14)^2} = \frac{600 - 560}{1026 - 980} = \frac{40}{46} = 0.87$$

Also, $a = \bar{y} - b_{YX}\,\bar{x}$

$$= 8 - 0.87 \times 14 = 8 - 12.18 = -4.18$$

∴ The regression equation of Y on X is

$Y = a + b_{YX}\,X$

∴ $Y = - 4.18 + 0.87X$

Exercise 3.1 | Q 3 | Page 41

For a certain bivariate data on 5 pairs of observations given

$\sum x = 20$, $\sum y = 20$, $\sum x^2 = 90$, $\sum y^2 = 90$, $\sum xy = 76$

Calculate:

i. cov (X, Y)
ii. b_{YX} and b_{XY}
iii. r

Solution: Given, $\sum x = 20$, $\sum y = 20$, $\sum x^2 = 90$, $\sum y^2 = 90$, $\sum xy = 76$, n = 5

Now,

$$\bar{x} = \frac{\sum x}{n} = \frac{20}{5} = 4$$

$$\bar{x} = \frac{\sum y}{n} = \frac{20}{5} = 4$$

(i) cov (X, Y) = $\frac{1}{n} \sum xy - \bar{x}\,\bar{y}$

$$= \frac{1}{5} \times 76 - 4 \times 4$$

= 15.2 - 16 = - 0.8

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

$$= \frac{76 - 5 \times 4 \times 4}{90 - 5(4)^2} = \frac{76 - 80}{90 - 80} = \frac{-4}{10} = -0.4$$

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

$= \frac{76 - 5 \times 4 \times 4}{90 - 5(4)^2} = \frac{76 - 80}{90 - 80}$ = - 0.4

(iii) r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{(-0.4)(-0.4)} = \pm\sqrt{0.16} = \pm 0.4$$

Since b_{YX} and b_{XY} both are negative,

r is negative.

∴ r = - 0.4

Exercise 3.1 | Q 4 | Page 41

From the following data estimate y when x =125.

X	120	115	120	125	126	123
Y	13	15	14	13	12	14

Solution: In order to estimate y, we have to find the regression equation of Y on X.

$X = x_i$	$Y = y_i$	x_i^2	$x_i y_i$
120	13	14400	1560
115	15	13225	1725
120	14	14400	1680
125	13	15625	1625
126	12	15876	1512
123	14	15129	1722
729	**81**	**88655**	**9824**

From table, we get

$n = 6, \sum x_i = 729, \sum y_i = 81, \sum x_i^2 = 88655, \sum x_i y_i = 9824$

$$\therefore \bar{x} = \frac{\sum x_i}{n} = \frac{729}{6} = 121.5$$

$$\therefore \bar{x} = \frac{\sum x_i}{n} = \frac{81}{6} = 13.5$$

Now,

$$b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{9824 - 6 \times 121.5 \times 13.5}{88655 - 6(121.5)^2}$$

$$= \frac{9824 - 9841.5}{88655 - 88573.5} = \frac{-17.5}{81.5} = -0.22$$

Also, $a = \bar{y} - b_{YX}\ \bar{x}$

$= 13.5 - (-0.22) \times 121.5$

$= 13.5 + 26.73 = 40.23$

∴ The regression equation of Y on X is

$Y = a + b_{YX} X$

∴ $Y = 40.23 - 0.22 X$

For X = 125,

$Y = 40.23 - 0.22 \times 125 = 40.23 - 27.5 = 12.73$

∴ The estimated value of Y is 12.73 for X = 125.

Exercise 3.1 | Q 5.1 | Page 41

The following table gives the aptitude test scores and productivity indices of 10 workers selected at random.

Aptitude score (X)	60	62	65	70	72	48	53	73	65	82
Productivity Index (Y)	68	60	62	80	85	40	52	62	60	81

Obtain the two regression equations and estimate the productivity index of a worker whose test score is 95.

Solution: Here, X = Aptitude score, Y = Productivity index

X = xi	Y =yi	$x_i - \bar{x}$	$\bar{y}_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
60	68	-5	3	25	9	-15
62	60	-3	-5	9	25	15
65	62	0	-3	0	9	0
70	80	5	15	25	225	75
72	85	7	20	49	400	140
48	40	-17	-25	289	625	425
53	52	-12	-13	144	169	156
73	62	8	-3	64	9	-24
65	60	0	-5	0	25	0
82	81	17	16	289	256	272
650	**650**	-	-	**894**	**1752**	**1044**

From the table, we have

n = 10, $\sum x_i = 650$, $\sum y_i = 650$

$$\therefore \bar{x} = \frac{\sum x_i}{n} = \frac{650}{10} = 65$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{650}{10} = 65$$

Since the mean of X and Y are whole numbers, we will use the formula

$$b_{YX} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{and} \quad b_{XY} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2}$$

From the table, we have

$$\sum(x_i - \bar{x})(y_i - \bar{y}) = 1044, \sum(x_i - \bar{x})^2 = 894, \sum(y_i - \bar{y}) =$$

$$b_{YX} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{1044}{894} = 1.16$$

Now, $a = \bar{y} - b_{YX}\bar{x}$

= 65 - 1.16 × 65 = 65 - 75.4 = - 10.4

∴ The regression equation of productivity index (Y) on Aptitude score (X) is

Y = a + b_{YX} X

∴ Y = - 10.4 + 1.16 X

For X = 95,

Y = - 10.4 + 1.16(95) = - 10.4 + 110.2 = 99.8

∴ The productivity index of worker with a test score of 95 is 99.8.

Exercise 3.1 | Q 5.2 | Page 41

The following table gives the aptitude test scores and productivity indices of 10 workers selected at random.

Aptitude score (X)	60	62	65	70	72	48	53	73	65	82
Productivity Index (Y)	68	60	62	80	85	40	52	62	60	81

Obtain the two regression equations and estimate the test score when the productivity index is 75.

Solution: Here, X = Aptitude score, Y = Productivity index

X = xi	Y =yi	$x_i - \bar{x}$	$\bar{y}_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
60	68	-5	3	25	9	-15
62	60	-3	-5	9	25	15
65	62	0	-3	0	9	0
70	80	5	15	25	225	75
72	85	7	20	49	400	140
48	40	-17	-25	289	625	425
53	52	-12	-13	144	169	156
73	62	8	-3	64	9	-24
65	60	0	-5	0	25	0
82	81	17	16	289	256	272
650	**650**	-	-	**894**	**1752**	**1044**

From the table, we have

n = 10, $\sum x_i$ = 650, $\sum y_i$ = 650

$$\therefore \bar{x} = \frac{\sum x_i}{n} = \frac{650}{10} = 65$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{650}{10} = 65$$

Since the mean of X and Y are whole numbers, we will use the formula

$$b_{YX} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{and} \quad b_{XY} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2}$$

From the table, we have

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = 1044, \sum (x_i - \bar{x})^2 = 894, \sum (y_i - \bar{y}) = 1752$$

$$b_{XY} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{1044}{1752} = 0.59$$

Now, $a' = \bar{x} - b_{XY}\bar{y}$

= 65 - 0.59 × 65 = 65 - 38.35 = 26.65

∴ The regression equation of Aptitude score (X) on productivity index (Y) is

$X = a' + b_{XY} Y$

∴ X = 26.65 + 0.59 Y

For Y = 75,

X = 26.65 + 0.59 × 75 = 26.65 + 44.25 = 70.9

∴ The test score is 70.9 when productivity index is 75

Exercise 3.1 | Q 6 | Page 42

Compute the appropriate regression equation for the following data:

X [Independent Variable]	2	4	5	6	8	11
Y [dependent Variable]	18	12	10	8	7	5

Solution: Since X is independent and Y is dependent variable, we find the regression equation of Y on X.

$X = x_i$	$Y = y_i$	x_i^2	$x_i y_i$
2	18	4	36
4	12	16	48
5	10	25	50
6	8	36	48
8	7	64	56
11	5	121	55
36	**60**	**266**	**293**

From the table, we have,

n = 6, $\sum x_i = 36$, $\sum y_i = 60$, $\sum x_i^2 = 266$, $\sum x_i y_i = 293$

$$\bar{x} = \frac{\sum x_i}{n} \frac{36}{6} = 6$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{60}{6} = 10$$

Now, $b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$= \frac{293 - 6 \times 6 \times 10}{266 - 6(6)^2} = \frac{293 - 360}{266 - 216} = \frac{-67}{50}$$

∴ b_{YX} = - 1.34

Also, $a/ = \bar{y} - b_{YX}\ \bar{x}$

= 10 - (- 1.34)(6) = 10 + 8.04 = 18.04

The regression equation of Y on X is

Y = a + b_{YX} X

∴ Y = 18.04 - 1.34 X

Exercise 3.1 | Q 7 | Page 42

The following are the marks obtained by the students in Economics (X) and Mathematics (Y)

X	59	60	61	62	63
Y	78	82	82	79	81

Find the regression equation of Y on X.

Solution: X = Marks obtained in Economics

Y = Marks obtained in Mathematics

$X = x_i$	$Y = y_i$	x_i^2	$x_i y_i$
59	78	3481	4602
60	82	3600	4920
61	82	3721	5002
62	79	3844	4898
63	81	3969	5103
305	**402**	**18615**	**24525**

From the table, we have

$n = 6, \sum x_i = 305, \sum y_i = 402, \sum x_i^2 = 18615, \sum x_i y_i = 24525$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{305}{5} = 61$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{402}{5} = 80.4$$

Now, $b_{YX} = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$= \frac{24525 - 5 \times 61 \times 80.4}{18615 - 5(61)^2}$$

$$= \frac{24525 - 24522}{18615 - 18605} = \frac{3}{10} = 0.3$$

Also, $a = \bar{y} - b_{YX}\bar{x}$

$= 80.4 - (0.3)(61) = 80.4 - 18.3 = 62.1$

The regression equation of Y on X is

$Y = a + b_{YX} X$

$\therefore Y = 62.1 + 0.3 X$

Exercise 3.1 | Q 8 | Page 42

For the following bivariate data obtain the equations of two regression lines:

X	1	2	3	4	5
Y	5	7	9	11	13

Solution:

$X = x_i$	**$Y = y_i$**	x_i^2	y_i^2	**$x_i y_i$**
1	5	1	25	5
2	7	4	49	14
3	9	9	81	27
4	11	16	121	44
5	13	25	169	65
15	**45**	**55**	**445**	**155**

From the table, we have

$n = 6, \sum x_i = 15, \sum y_i = 45, \sum x_i^2 = 55, \sum y_i^2 = 445, \sum x_i y_i = 155$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{15}{5} = 3$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{45}{5} = 9$$

Now, for regression equation of Y on X,

$$b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{155 - 5 \times 3 \times 9}{55 - 5(3)^2} = \frac{155 - 135}{55 - 45} = \frac{20}{10} = 2$$

Also, $a = \bar{y} - b_{XY}\ \bar{x}$ = 9 - 2(3) = 9 - 6 = 3

The regression analysis of Y on X is

Y = a + b_{YX} X

∴ Y = 3 + 2X

Now, for regression equation of X on Y,

$$b_{XY} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n\bar{y}^2}$$

$$= \frac{155 - 5 \times 3 \times 9}{445 - 5(9)^2} = \frac{155 - 135}{445 - 405} = \frac{20}{40} = 0.5$$

Also, $a\prime = \bar{x} - b_{XY}\ \bar{y}$

= 3 - (0.5)(9) = 3 - 4.5 = - 1.5

The regression equation of X on Y is

X = a' + b_{XY} Y

∴ X = - 1.5 + 0.5Y

Exercise 3.1 | Q 9 | Page 42

From the following data obtain the equation of two regression lines:

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

Solution:

X = x_i	**Y = y_i**	x_i^2	y_i^2	**$x_i y_i$**
6	9	36	81	54
2	11	4	121	22
10	5	100	25	50
4	8	16	64	32
8	7	64	49	56
30	**40**	**220**	**340**	**214**

From the table, we have

n = 6, $\sum x_i = 30$, $\sum y_i = 40$, $\sum x_i^2 = 220$, $\sum y_i^2 = 340$, $\sum x_i y_i =$ 214

$$\bar{x} = \frac{\sum x_i}{n} = \frac{30}{5} = 6$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{40}{5} = 8$$

Now, for regression equation of Y on X,

$$b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{214 - 5 \times 6 \times 8}{220 - 5(6)^2} = \frac{214 - 240}{220 - 180} = \frac{-26}{40} = -0.65$$

Also, $a = \bar{y} - b_{YX}\,\bar{x}$

= 8 - (- 0.65)(6) = 8 + 3.9 = 11.9

The regression equation of Y on X is

$Y = a + b_{YX} X$

∴ Y = 11.9 - 0.65X

Now, for regression equation of X on Y,

$$b_{XY} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum y_i^2 - n\bar{y}^2}$$

$$= \frac{214 - 5 \times 6 \times 8}{340 - 5(8)^2} = \frac{214 - 240}{340 - 320} = \frac{-26}{20} = -1.3$$

Also, $a' = \bar{x} - b_{XY}\,\bar{y}$

= 6 - (- 1.3)8 = 6 + 10.4 = 16.4

The regression equation of Y on X is

$X = a' + b_{XY} Y$

∴ X = 16.4 - 1.3Y

Exercise 3.1 | Q 10 | Page 42

For the following data, find the regression line of Y on X

X	1	2	3
Y	2	1	6

Hence find the most likely value of y when x = 4.

Solution:

X = x_i	Y = y_i	x_i^2	$x_i y_i$
1	2	1	2
2	1	4	2
3	6	9	18
6	**9**	**14**	**22**

From the table, we have

n = 3, $\sum x_i = 6$, $\sum y_i = 9$, $\sum x_i^2 = 14$, $\sum x_i y_i = 22$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{6}{3} = 2$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{9}{3} = 3$$

Now, $b_{YX} = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$= \frac{22 - 3 \times 2 \times 3}{14 - 3(2)^2} = \frac{22 - 18}{14 - 12} = \frac{4}{2} = 2$$

Also, $a = \bar{y} - b_{YX}\,\bar{x}$ = 3 - 2(2) = - 1

The regression equation of Y on X is,

Y = a + b_{YX} X

∴ Y = - 1 + 2X

For X = 4,

Y = - 1 + 2(4) = - 1 + 8 = 7

The most likely value of Y for X = 4 is 7.

Exercise 3.1 | Q 11 | Page 42

From the following data, find the regression equation of Y on X and estimate Y when X = 10.

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

Solution:

$X = x_i$	$Y = y_i$	x_i^2	$x_i y_i$
1	2	1	2
2	4	4	8
3	7	9	21
4	6	16	24
5	5	25	25
6	6	36	36
21	**30**	**91**	**116**

From the table, we have

$n = 6, \sum x_i = 21, \sum y_i = 30, \sum x_i^2 = 91, \sum x_i y_i = 116$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{21}{6} = 3.5$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{30}{6} = 5$$

Now, $b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$= \frac{116 - 6 \times 3.5 \times 5}{91 - 6(3.5)^2} = \frac{116 - 105}{91 - 73.5} = \frac{11}{17.5} = 0.63$$

Also, $a = \bar{y} - b_{YX}\ \bar{x}$

= 5 - 0.63 × 3.5

= 5 - 2.205 = 2.8

The regression equation of Y on X is,

$Y = a + b_{YX} X$

∴ Y = 2.8 + 0.63 X

For X = 10,

Y = 2.8 + 0.63 × 10

= 2.8 + 6.3 = 9.1

∴ The value of Y when X =10 is 9.1

Exercise 3.1 | Q 12 | Page 42

The following sample gives the number of hours of study (X) per day for an examination and marks (Y) obtained by 12 students.

X	3	3	3	4	4	5	5	5	6	6	7	8
Y	45	60	55	60	75	70	80	75	90	80	75	85

Obtain the line of regression of marks on hours of study.

Solution:

$X = x_i$	$Y = y_i$	x_i^2	$x_i y_i$
3	45	9	135

3	60	9	180
3	55	9	165
4	60	16	240
4	75	16	300
5	70	25	350
5	80	25	400
5	75	25	375
6	90	36	540
6	80	36	480
7	75	49	525
8	85	64	680
59	**850**	**319**	**4370**

From the table, we have

n = 12, $\sum x_i = 59$, $\sum y_i = 850$, $\sum x_i^2 = 319$, $\sum x_i y_i = 4370$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{59}{12} = 4.92$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{850}{12} = 5$$

Now, $b_{YX} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$$= \frac{4370 - 12 \times 4.92 \times 70.83}{319 - 12 \times (4.92)^2}$$

$$= \frac{4370 - 4181.80}{319 - 290.48}$$

$$= \frac{188.2}{28.52} = 6.6$$

Also, $a = \bar{y} - b_{YX}\,\bar{x}$

= 70.83 - 6.6 × 4.92

= 70.83 - 32.47

= 38.36

∴ The regression equation of Y on X is,

$Y = a + b_{YX} X$

∴ Y = 38.36 + 6.6 X

EXERCISE 3.2 [PAGES 47 - 48]

Exercise 3.2 | Q 1.1 | Page 47

For bivariate data. $\bar{x} = 53, \bar{y} = 28, b_{YX} = -1.2, b_{XY} = -0.3$

Find Correlation coefficient between X and Y.

Solution:

Here, $\bar{x} = 53, \bar{y} = 28, b_{YX} = -1.2, b_{XY} = -0.3$

$r = \pm\sqrt{b_{XY} \cdot b_{YX}}$

$= \pm\sqrt{(-0.3)(-1.2)}$

$= \pm\sqrt{0.36} = \pm 0.6$

Since b_{YX} and b_{XY} both are – negative,

r is also negative.

∴ r = - 0.6

Exercise 3.2 | Q 1.2 | Page 47

For bivariate data. $\bar{x} = 53, \bar{y} = 28, b_{YX} = -1.2, b_{XY} = -0.3$

Find estimate of Y for X = 50.

Solution:

Here, $\bar{x} = 53, \bar{y} = 28, \mathrm{b_{YX}} = -1.2, \mathrm{b_{XY}} = -0.3$

The regression equation of Y on X is

$$(\mathrm{Y} - \bar{y}) = \mathrm{b_{YX}}(\mathrm{X} - \bar{x})$$

∴ (Y - 28) = (- 1.2)(X - 53)

∴ Y - 28 = - 1.2 X + 63.6

∴ Y = - 1.2 X + 63.6 + 28

∴ Y = - 1.2 X + 91.6

For X = 50

∴ Y = - 1.2(50) + 91.6 = - 60 + 91.6 = 31.6

Exercise 3.2 | Q 1.3 | Page 47

For bivariate data. $\bar{x} = 53, \bar{y} = 28, \mathrm{b_{YX}} = -1.2, \mathrm{b_{XY}} = -0.3$

Find estimate of X for Y = 25.

Solution:

Here, $\bar{x} = 53, \bar{y} = 28, \mathrm{b_{YX}} = -1.2, \mathrm{b_{XY}} = -0.3$

The regression equation of X on Y is

$$(\mathrm{X} - \bar{x}) = \mathrm{b_{XY}}(\mathrm{Y} - \bar{y})$$

∴ (X - 53) = (- 0.3)(Y - 28)

∴ X - 53 = - 0.3 Y + 8.4

∴ X = - 0.3 Y + 8.4 + 53

∴ X = - 0.3 Y + 61.4

For Y = 25

∴ X = - 0.3(25) + 61.4 = - 7.5 + 61.4 = 53.9

Exercise 3.2 | Q 2 | Page 47

From the data of 20 pairs of observations on X and Y, following results are obtained.

$$\bar{x} = 199, \bar{y} = 94,$$

$$\sum (x_i - \bar{x})^2 = 1200, \sum (y_i - \bar{y})^2 = 300,$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = -250$$

Find:

i. The line of regression of Y on X.
ii. The line of regression of X on Y.
iii. Correlation coefficient between X and Y.

Solution:

Given, n = 20, $\bar{x} = 199, \bar{y} = 94,$

$$\sum (x_i - \bar{x})^2 = 1200, \sum (y_i - \bar{y})^2 = 300,$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = -250$$

(i) $\mathrm{b_{YX}} = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \dfrac{-250}{1200} = -\dfrac{5}{24}$

∴ The regression equation of Y on X is

$$(\mathrm{Y} - \bar{y}) = \mathrm{b_{YX}}(\mathrm{X} - \bar{x})$$

∴ (Y - 94) = $-\dfrac{5}{24}$(X - 199)

∴ 24Y - 2256 = - 5X + 995

∴ 5X + 24Y = 3251

(ii) $b_{YX} = \dfrac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(y_i - \bar{y})^2} = \dfrac{-250}{300} = -\dfrac{5}{6}$

∴ The regression equation of X on Y is

$(X - \bar{x}) = b_{XY}(Y - \bar{y})$

∴ (X - 199) = $-\dfrac{5}{6}$ (Y - 94)

∴ 6X - 1194 = - 5Y + 470

∴ 6X + 5Y = 1664

(iii) r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$= \pm\sqrt{\left(-\dfrac{5}{24}\right)\left(-\dfrac{5}{6}\right)} = \pm\sqrt{\dfrac{25}{144}} = \pm\dfrac{5}{12}$

Since b_{XY} and b_{YX} both are negative,

r is also negative.

∴ r = - 5/12

Exercise 3.2 | Q 3 | Page 47

From the data of 7 pairs of observations on X and Y, following results are obtained.

$\sum(x_i - 70) = -35$, $\sum(y_i - 60) = -7$,

$\sum(x_i - 70)^2 = 2989$, $\sum(y_i - 60)^2 = 476$,

$\sum(x_i - 70)(y_i - 60) = 1064$

[Given: $\sqrt{0.7884}$ = 0.8879]

Obtain

i. The line of regression of Y on X.
ii. The line regression of X on Y.
iii. The correlation coefficient between X and Y.

Solution: Given: n =7, $\sum(x_i - 70) = -35$, $\sum(y_i - 60) = -7$,

$\sum(x_i - 70)^2 = 2989$, $\sum(y_i - 60)^2 = 476$,

$\sum(x_i - 70)(y_i - 60) = 1064$

Let $u_i = x_i - 70$ and $v_i = y_i - 60$

$\therefore \sum u_i = -35, \sum v_i = -7$

$$\sum u_i^2 = 2989, \sum v_i^2 = 479$$

$\sum u_i v_i = 1064$

$$\therefore \bar{u} = \frac{\sum u_i}{n} = \frac{-35}{7} = -5$$

$$\therefore \bar{v} = \frac{\sum v_i}{n} = \frac{-7}{7} = -1$$

$$\text{Now, } \sigma_u^2 = \frac{\sum u_i^2}{n} - (\bar{u})^2$$

$$= \frac{2989}{7} - (-5)^2 = 427 - 25 = 402$$

$$\text{and } \sigma_v^2 = \frac{\sum v_i^2}{n} - (\bar{v})^2$$

$$= \frac{476}{7} - (-1)^2 = 68 - 1 = 67$$

$\text{cov}(u, v) = \frac{\sum u_i v_i}{n} - \overline{uv}$

$= \frac{1064}{7} - (-5)(-1) = 152 - 5 = 147$

Since the regression coefficients are independent of change of origin,

$b_{YX} = b_{VU}$ and $b_{XY} = b_{UV}$

$\therefore b_{YX} = b_{VU} = \frac{\text{cov}(u, v)}{\sigma_U^2} = \frac{147}{402} = 0.36$

and $b_{XY} = b_{UV} = \frac{\text{cov}(u, v)}{\sigma_V^2} = \frac{147}{67} = 2.19$

Also, $\bar{x} = \bar{u} + 70 = -5 + 70 = 65$

and $\bar{y} = \bar{v} + 60 = -1 + 60 = 59$

(i) The line of regression of Y on X is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

$\therefore (Y - 59) = (0.36)(X - 65)$

$\therefore Y - 59 = 0.36X - 23.4$

$\therefore Y = 0.36X + 59 - 23.4$

$\therefore Y = 0.36X + 35.6$

(ii) The line of regression of X on Y is

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

$\therefore (X - 65) = (2.19)(Y - 59)$

$\therefore X - 65 = 2.19Y - 129.21$

$\therefore X = 2.19Y + 65 - 129.21$

$\therefore X = 2.19Y - 64.21$

(iii) r = $\pm\sqrt{b_{YX} \cdot b_{XY}}$

$= \pm\sqrt{(0.36)(2.19)}$

$= \pm\sqrt{0.7884} = \pm 0.8879$

Since b_{YX} and b_{XY} both are positive,

r is also positive.

∴ r = 0.8879

Exercise 3.2 | Q 4 | Page 47

You are given the following information about advertising expenditure and sales.

	Advertisement expenditure (₹ in lakh) (X)	Sales (₹ in lakh) (Y)
Arithmetic Mean	10	90
Standard Mean	3	12

Correlation coefficient between X and Y is 0.8

i. Obtain the two regression equations.
ii. What is the likely sales when the advertising budget is ₹ 15 lakh?
iii. What should be the advertising budget if the company wants to attain sales target of ₹ 120 lakh?

Solution:

Given: $\bar{x} = 10, \bar{y} = 90, \sigma_x = 3, \sigma_y = 12$, r = 0.8

$$b_{YX} = r\frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{12}{3} = 0.8 \times 4 = 3.2$$

$$b_{XY} = r\frac{\sigma_x}{\sigma_y} = 0.8 \times \frac{3}{12} = 0.8 \times 0.25 = 0.2$$

(i) The regression equation of Y on X is

$$(\mathrm{Y}-\bar{y})=\mathrm{b}_{\mathrm{YX}}(\mathrm{X}-\bar{x})$$

∴ (Y - 90) = 3.2 (X - 10)

∴ Y - 90 = 3.2 X - 32

∴ Y = 3.2 X - 32 + 90

∴ Y = 3.2 X + 58 .....(i)

The regression equation of X on Y is

$$(\mathrm{X}-\bar{x})=\mathrm{b}_{\mathrm{XY}}(\mathrm{Y}-\bar{y})$$

∴ (X - 10) = 0.2 (Y - 90)

∴ X - 10 = 0.2 Y - 18

∴ X = 0.2 Y - 18 + 10

∴ X = 0.2 Y - 8 .....(ii)

(ii) For X = 15, from equation (i) we get

Y = 3.2 (15) + 58 = 48 + 58 = 106

∴ Likely sales is ₹ 106 lakh when advertising budget is ₹ 15 lakh.

(iii) For Y = 120, from equation (ii) we get

X = 0.2 (120) - 8 = 24 - 8 = 16

∴ To attain sales target of ₹ 120 lakh, advertising budget must be ₹ 16 lakh.

Exercise 3.2 | Q 5.1 | Page 47

Bring out the inconsistency in the following:

$b_{YX} + b_{XY}$ = 1.30 and r = 0.75

Solution: Given, $b_{YX} + b_{XY}$ = 1.30, r = 0.75

Consider, $\dfrac{b_{YX} + b_{XY}}{2} = \dfrac{1.30}{2} = 0.65$

$\therefore \dfrac{b_{YX} + b_{XY}}{2} < r$

But, for consistent data $\left|\dfrac{b_{YX} + b_{XY}}{2}\right| > |r|$

$\therefore$ Given data is inconsistent.

Exercise 3.2 | Q 5.2 | Page 47

Bring out the inconsistency in the following:

$b_{YX} = b_{XY} = 1.50$ and $r = -0.9$

Solution: Given, $b_{YX} = b_{XY} = 1.50$ and $r = -0.9$

Here, the coefficient of regressions is positive and the coefficient of correlation is negative.

But, for consistent data they must have the same signs.

$\therefore$ The given data is inconsistent.

Exercise 3.2 | Q 5.3 | Page 47

Bring out the inconsistency in the following:

$b_{YX} = 1.9$ and $b_{XY} = -0.25$

Solution: Given, $b_{YX} = 1.9$ and $b_{XY} = -0.25$

Here, b_{YX} and b_{XY} have different signs.

But, for consistent data, they must have the same signs.

$\therefore$ The given data is inconsistent.

Exercise 3.2 | Q 5.4 | Page 47

Bring out the inconsistency in the following:

$b_{YX} = 2.6$ and $b_{XY} = 1/2.6$

Solution: Given, $b_{YX} = 2.6$ and $b_{XY} = 1/2.6$

Here, b_{YX} and b_{XY} have the same signs.

Also, $b_{YX} > 1$ and $b_{XY} < 1$

Also, for consistent data, the signs of b_{YX} and b_{XY} are same and $b_{YX} > 1$, $b_{XY} < 1$

Here, b_{YX}. $b_{XY} = 1$

∴ The given data is consistent.

Exercise 3.2 | Q 6 | Page 47

Two samples from bivariate populations have 15 observations each. The sample means of X and Y are 25 and 18 respectively. The corresponding sum of squares of deviations from respective means is 136 and 150. The sum of the product of deviations from respective means is 123. Obtain the equation of the line of regression of X on Y.

Solution:

Given, n = 15, $\bar{x} = 25, \bar{y} = 18,$

$$\sum (x_i - \bar{x})^2 = 136, \sum (y_i - \bar{y})^2 = 150,$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = 123$$

$$\text{Now, } b_{XY} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(y_i - \bar{y})} = \frac{123}{150} = 0.82$$

Also, $a' = \bar{x} - b_{XY}\ \bar{y}$

= 25 - 0.82 × 18 = 25 - 14.76 = 10.24

∴ The regression equation of X on Y is

X = a' + b_{XY} Y

∴ X = 10.24 + 0.82Y

Exercise 3.2 | Q 7 | Page 48

For a certain bivariate data

	X	Y
Mean	25	20
S.D.	4	3

And r = 0.5. Estimate y when x = 10 and estimate x when y = 16

Solution:

Given, $\bar{x} = 25, \bar{y} = 20, \sigma_{\mathrm{X}} = 4, \sigma_{\mathrm{Y}} = 3$, r =0.5

$$\mathrm{b}_{\mathrm{YX}} = \mathrm{r}\frac{\sigma_y}{\sigma_x} = (0.5)\frac{3}{4} = 0.375$$

$$\mathrm{b}_{\mathrm{XY}} = \mathrm{r}\frac{\sigma_y}{\sigma_x} = (0.5)\frac{4}{3} = 0.667$$

The regression equation of Y on X is

$$(\mathrm{Y} - \bar{y}) = \mathrm{b}_{\mathrm{YX}}(\mathrm{X} - \bar{x})$$

(Y - 20) = 0.375 (X - 25)

Y - 20 = - 9.375 + 0.375 X

Y = 10.625 + 0.375 X

For X = 10

Y = 10.625 +0.375 × 10 = 10.625 + 3.75 = 14.375

The regression equation of X on Y is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

(Y - 20) = 0.375 (X - 25)

Y - 20 = - 9.375 + 0.375 X

Y = 10.625 + 0.375 X

For X = 10

Y = 10.625 +0.375 × 10 = 10.625 + 3.75 = 14.375

The regression equation of X on Y is

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

(X - 25) = 0.667(Y - 20)

X - 25 = - 13.34 + 0.667 Y

X = 11.66 + 0.667 Y

For Y = 16,

X = 11.66 + 0.667(16) = 11.66 + 10.672 = 22.332

[**Note:** Answer in the textbook is incorrect.]

Exercise 3.2 | Q 8 | Page 48

Given the following information about the production and demand of a commodity obtain the two regression lines:

	X	Y
Mean	85	90
S.D.	5	6

The coefficient of correlation between X and Y is 0.6. Also estimate the production when demand is 100.

Solution:

Given, $\bar{x} = 85, \bar{y} = 90, \sigma_X = 5, \sigma_Y = 6$, r =0.6

$$b_{YX} = r\frac{\sigma_Y}{\sigma_X} = 0.6 \times \frac{6}{5} = 0.72$$

$$b_{XY} = r\frac{\sigma_X}{\sigma_Y} = 0.6 \times \frac{5}{6} = 0.5$$

The regression equation of Y on X is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

(Y - 90) = 0.72 (X - 85)

Y - 90 = 0.72 X - 61.2

Y = 0.72X - 61.2 + 90

Y = 28.8 + 0.72 X ....(i)

The regression equation of X on Y is

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

(X - 85) = 0.5(Y - 90)

X - 85 = 0.5 Y - 45

X = 0.5 Y - 45 + 85

X = 40 + 0.5Y ....(ii)

For Y = 100, from equation (ii) we get

X = 40 + 0.5(100) = 40 + 50 = 90

∴ The production is 90 when demand is 100.

Exercise 3.2 | Q 9 | Page 48

Given the following data, obtain a linear regression estimate of X for Y = 10, $\bar{x} = 7.6, \bar{y} = 14.8, \sigma_x = 3.2, \sigma_y = 16$ and r = 0.7

Solution:

Given, $\bar{x} = 7.6, \bar{y} = 14.8, \sigma_x = 3.2, \sigma_y = 16$ and r = 0.7

$$b_{XY} = r\frac{\sigma_X}{\sigma_Y} = 0.7 \times \frac{3.2}{16} = 0.14$$

$$a' = \bar{x} - b_{XY}\ \bar{y}$$

= 7.6 - 0.14 × 14.8 = 7.6 - 2.072 = 5.528

The regression equation of X on Y is

X = a' + b_{XY} Y

∴ X = 5.528 + 0.14 Y

For Y = 10

X = 5.528 + 0.14 × 10 = 5.528 + 1.4 = 6.928

Exercise 3.2 | Q 10 | Page 48

An inquiry of 50 families to study the relationship between expenditure on accommodation (₹ x) and expenditure on food and entertainment (₹ y) gave the following results:

∑ x = 8500, ∑ y = 9600, σ_X = 60, σ_Y = 20, r = 0.6

Estimate the expenditure on food and entertainment when expenditure on accommodation is Rs 200.

Solution: X = Expenditure on accommodation.

Y = Expenditure on food and entertainment

Given, ∑ x = 8500, ∑ y = 9600, σ_X = 60, σ_Y = 20, r = 0.6, n = 50

$$\therefore \bar{x} = \frac{\sum x}{\text{n}} = \frac{8500}{50} = 170$$

$$\bar{y} = \frac{\sum y}{\text{n}} = \frac{9600}{50} = 192$$

Now, $\text{b}_{\text{YX}} = \text{r}\frac{\sigma_{\text{Y}}}{\sigma_{\text{X}}} = 0.6 \times \frac{20}{60} = 0.2$

Also, $\text{a} = \bar{y} - \text{b}_{\text{YX}}\ \bar{x}$

= 192 - 0.2 × 170 = 192 - 34 = 158

The regression equation of Y on X is

$Y = a + b_{YX} X$

∴ Y = 158 + 0.2 X

For X = 200,

Y = 158 + 0.2 × 200 = 158 + 40 = 198

∴ The expenditure on food and entertainment is
₹ 198 when expenditure on accommodation is ₹ 200.

[**Note:** Answer in the textbook is incorrect.]

Exercise 3.2 | Q 11.1 | Page 48

The following data about the sales and advertisement expenditure of a firms is given below (in ₹ Crores)

	Sales	Adv. Exp.
Mean	40	6
S.D.	10	1.5

Coefficient of correlation between sales and advertisement expenditure is 0.9.
Estimate the likely sales for a proposed advertisement expenditure of ₹ 10 crores.

Solution: Let X = Sales,

Y = Advertisement expenditure

Given, $\bar{x} = 40, \bar{y} = 6, \sigma_X = 10, \sigma_Y = 1.5$, r = 0.9

$$b_{XY} = r\frac{\sigma_X}{\sigma_Y} = 0.9 \times \frac{10}{1.5} = 6$$

$$b_{YX} = r\frac{\sigma_Y}{\sigma_X} = 0.9 \times \frac{1.5}{10} = 0.135$$

The regression equation of X on Y is

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

∴ (X - 40) = 6(Y - 6)

∴ X - 40 = 6Y - 36

∴ X = 6Y - 36 + 40

∴ X = 6Y + 4

For Y = 10, we get

X = 6(10) + 4 = 60 + 4 = 64

∴ The likely sale is ₹ crores for a proposed advertisement expenditure of ₹ 10 crores.

Exercise 3.2 | Q 11.2 | Page 48

The following data about the sales and advertisement expenditure of a firms is given below (in ₹ Crores)

	Sales	Adv. Exp.
Mean	40	6
S.D.	10	1.5

Coefficient of correlation between sales and advertisement expenditure is 0.9.
What should be the advertisement expenditure if the firm proposes a sales target ₹ 60 crores?

Solution: Let X = Sales,

Y = Advertisement expenditure

Given, $\bar{x} = 40, \bar{y} = 6, \sigma_X = 10, \sigma_Y = 1.5$, r = 0.9

$$b_{XY} = r\frac{\sigma_X}{\sigma_Y} = 0.9 \times \frac{10}{1.5} = 6$$

$$b_{YX} = r\frac{\sigma_Y}{\sigma_X} = 0.9 \times \frac{1.5}{10} = 0.135$$

The regression equation of Y on X is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

∴ (Y - 6) = 0.135(X - 40)

∴ Y - 6 = 0.135X - 5.4

∴ Y = 0.135X - 5.4 + 6

∴ Y = 0.135X + 0.6

For X = 60,

Y = 0.135(60) + 0.6 = 8.1 + 0.6 = 8.7

∴ The advertisement expenditure should be ₹ 8.7 crores if the firm proposes a sales target ₹ 60 crores

Exercise 3.2 | Q 12 | Page 48

For certain bivariate data the following information is available.

	X	Y
Mean	13	17
S.D.	3	2

Correlation coefficient between x and y is 0.6. estimate x when y = 15 and estimate y when x = 10.

Solution:

Given, $\bar{x} = 13, \bar{y} = 1, \sigma_X 3, \sigma_Y = 2,$ r = 0.6

$$b_{YX} = r\frac{\sigma_Y}{\sigma_X} = 0.6 \times \frac{2}{3} = 0.4$$

$$b_{XY} = r\frac{\sigma_X}{\sigma_Y} = 0.6 \times \frac{3}{2} = 0.9$$

The regression equation of X on Y is given by

$$(X - \bar{x}) = b_{XY}(Y - \bar{y})$$

(X - 13) = 0.9 (Y - 17)

X - 13 = 0.9Y - 15.3

X = 0.9Y - 15.3 + 13

X = - 2.3 + 0.9Y ....(i)

For Y = 15, from equation (i) we get

X = - 2.3 + (0.9)(15) = - 2.3 + 13.5 = 11.2

The regression equation of Y on X is given by

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

(Y - 17) = 0.4(X - 13)

Y - 17 = 0.4X - 5.2

Y = 0.4X - 5.2 + 17

Y = 11.8 + 0.4X .....(ii)

For X = 10, from equation (ii) we get

Y = 11.8 + 0.4(10) = 11.8 + 4 = 15.8

EXERCISE 3.3 [PAGES 49 - 50]

Exercise 3.3 | Q 1 | Page 49

From the two regression equations, find r, $\bar{x}$ and $\bar{y}$. 4y = 9x + 15 and 25x = 4y + 17

Solution: Given regression equations are

4y = 9x + 15

i.e., - 9x + 4y = 15 ....(i)

and 25x = 4y + 17

i.e., 25x - 4y = 17 ...(ii)

Adding equations (i) and (ii), we get

$$\begin{array}{l} -9x + 4y = 15 \\ \underline{\ 25x - 4y = 17\ } \\ 16x \quad = 32 \end{array}$$

∴ x = 2

Substituting x = 2 in (i), we get

- 9(2) + 4y = 15

∴ - 18 + 4y = 15

∴ 4y = 33

∴ y = 8.25

Since the point of intersection of two regression lines is

$(\bar{x}, \bar{y}), \bar{x} = 2 \text{ and } \bar{y} = 8.25$

Let 4y = 9x + 15 be the regression equation of Y on X.

∴ The equation becomes Y = $\frac{9}{2}X + \frac{15}{4}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{9}{4} = 2.25$$

Now, the other equation, i.e., 25x = 4y + 17 is the regression equation of X on Y.

∴ The equation becomes X = $\frac{4}{25}Y + \frac{17}{25}$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = \frac{4}{25} = 0.16$$

r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$= \pm\sqrt{0.16 \times 2.25}$

$= \pm\sqrt{0.36} = \pm 0.6$

Since b_{YX} and b_{XY} are positive,

r is also positive.

∴ r = 0.6

∴ $\bar{x} = 2 \text{ and } \bar{y} = 8.25$ and r = 0.6

Exercise 3.3 | Q 2 | Page 49

In a partially destroyed laboratory record of an analysis of regression data, the following data are legible:

Variance of X = 9
Regression equations:
$8x - 10y + 66 = 0$
and $40x - 18y = 214$.
Find on the basis of above information

i. The mean values of X and Y.
ii. Correlation coefficient between X and Y.
iii. Standard deviation of Y.

Solution:

Given, $\sigma_X^2 = 9$

$\therefore \sigma_X = 3$

(i) The two regression equations are

$8x - 10y + 66 = 0$

i.e., $8x - 10y = -66$ ...(i)

and $40x - 18y = 214$ ....(ii)

By 5 × (i) - (ii), we get

$40x - 50y = -330$

$40x - 18y = 214$

(-) (+) (-)

$-32y = -544$

$\therefore y = \dfrac{544}{32} = 17$

Substituting y = 17 in (i), we get

8x - 10 × 17 = - 66

∴ 8x - 170 = - 66

∴ 8x = - 66 + 170

∴ 8x = 104

∴ x = $\frac{104}{8} = 13$

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x}$ = mean value of X = 13, and

$\bar{y}$ = mean value of X = 17.

(ii) Let 8x - 10y + 66 = 0 be the regression equation of Y on X.

∴ The equation becomes 10Y = 8X + 66

i.e., Y = $\frac{8}{10}\mathrm{X} + \frac{66}{10}$

i.e., Y = $\frac{4}{5}\mathrm{X} + \frac{33}{5}$

Comparing it with Y = b_{YX} X + a, we get

$$\mathrm{b_{YX}} = \frac{4}{5}$$

Now, the other equation, i.e., 40x - 18y = 214 is the regression equation of X on Y.

∴ The equation becomes X = $\frac{18}{40}\mathrm{Y} + \frac{214}{40}$

i.e., X = $\frac{9}{20}\mathrm{Y}+\frac{107}{20}$

Comparing it with X = b_{XY} Y + a', we get

$\mathrm{b_{XY}}=\frac{9}{20}$

r = $\pm\sqrt{\mathrm{b_{XY}\cdot b_{YX}}}$

∴ r = $\pm\sqrt{\frac{9}{20}\times\frac{4}{5}}=\pm\sqrt{\frac{9}{25}}=\pm\frac{3}{5}=\pm0.6$

Since b_{YX} and b_{XY} both are positive,

r is also positive.

∴ r = 0.6

(iii) $\mathrm{b_{YX}}=\mathrm{r}\frac{\sigma_\mathrm{Y}}{\sigma_\mathrm{X}}$

$\therefore\ \frac{4}{5}=0.6\times\frac{\sigma_\mathrm{Y}}{3}$

$\therefore\ \frac{4}{5}=\frac{\sigma_\mathrm{Y}}{5}$

$\therefore\ \sigma_\mathrm{Y}=4$

Exercise 3.3 | Q 3 | Page 50

For 50 students of a class, the regression equation of marks in statistics (X) on the marks in Accountancy (Y) is 3y − 5x + 180 = 0.
The mean marks in accountancy is 44 and the variance of marks in statistics is $(9/16)^{th}$ of the variance of marks in accountancy. Find the mean marks in statistics and the correlation coefficient between marks in two subjects.

Solution: Given, n = 50, X = marks in Statistics, Y = marks in Accountancy,

Regression equation of X on Y is

3y - 5x + 180 = 0,

$$\bar{y} = 44,\ \sigma_X^2 = \frac{9}{16}\sigma_Y^2$$

Now, 3y - 5x + 180 = 0 is the regression equation of X on Y.

∴ The equation becomes 5X = 3Y + 180

i.e., X = $\frac{3}{5}Y + \frac{180}{5}$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = \frac{3}{5}, a' = \frac{180}{5} = 36$$

$$a' = \bar{x} - b_{XY}\ \bar{y}$$

∴ 36 = $\bar{x} - \frac{3}{5} \times 44$

∴ 36 = $\bar{x}$ - 26.4

∴ $\bar{x}$ = 36 + 26.4 = 62.4

Also, $\sigma_X^2 = \frac{9}{16}\sigma_Y^2$

$$\therefore \frac{\sigma_X^2}{\sigma_Y^2} = \frac{9}{16}$$

$$\therefore \frac{\sigma_X}{\sigma_Y} = \frac{3}{4}$$

$$b_{XY} = r \times \frac{\sigma_X}{\sigma_Y}$$

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

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

$$\therefore r = \frac{4}{5} = 0.8$$

$\therefore$ Mean marks in statistics $(\bar{x})$ are 62.4 and correlation coefficient (r) between marks in the two subjects is 0.8.

Exercise 3.3 | Q 4 | Page 50

For bivariate data, the regression coefficient of Y on X is 0.4 and the regression coefficient of X on Y is 0.9. Find the value of the variance of Y if the variance of X is 9.

Solution:

Given, $b_{YX} = 0.4$, $b_{XY} = 0.9$, $\sigma_X^2 = 9$

$$\therefore \sigma_X = 3$$

$$r = \pm\sqrt{b_{XY} \cdot b_{YX}}$$

$$= \pm\sqrt{0.9 \times 0.4} = \pm\sqrt{0.36} = \pm 0.6$$

Since b_{YX} and b_{XY} both are positive,

r is also positive.

$\therefore r = 0.6$

Now, $b_{YX} = r \times \frac{\sigma_Y}{\sigma_X}$

$\therefore 0.4 = 0.6 \times \frac{\sigma_Y}{3}$

$\therefore 0.4 = 0.2 \times \sigma_Y$

$\therefore \sigma_Y = \frac{0.4}{0.2} = 2$

$\therefore \sigma_Y^2 = 4$

$\therefore$ The value of variance of Y is 4.

Exercise 3.3 | Q 5 | Page 50

The equations of two regression lines are

$2x + 3y - 6 = 0$

and $2x + 2y - 12 = 0$ Find

i. Correlation coefficient

ii. $\frac{\sigma_X}{\sigma_Y}$

Solution: The given regression equations are
$2x + 3y - 6 = 0$ and $2x + 2y - 12 = 0$

(i) Let $2x + 3y - 6 = 0$ be the regression equation of Y on X

$\therefore$ The equation becomes $3Y = -2X + 6$

i.e., $Y = \frac{-2}{3}X + \frac{6}{3}$

Comparing it with $Y = b_{YX}X + a$, we get

$b_{YX} = -\frac{2}{3}$

Now, the other equation, i.e., $2x + 2y - 12 = 0$ is the regression equation of X on Y.

∴ The equation becomes 2X = –2Y + 12

i.e., X = $-\frac{2}{2}Y + \frac{12}{2}$

Comparing it with X = b_{XY} Y + a' we get

$$b_{XY} = -\frac{2}{2} = -1$$

∴ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{-1 \cdot \left(-\frac{2}{3}\right)} = \pm\sqrt{\frac{2}{3}} = \pm 0.82$$

since b_{XY} and b_{YX} are negative,

r is also negative.

∴ r = - 0.82

(ii) $b_{XY} = r\frac{\sigma_X}{\sigma_Y}$

$$\therefore -1 = -0.82 \times \frac{\sigma_X}{\sigma_Y}$$

$$\therefore \frac{\sigma_X}{\sigma_Y} = \frac{-1}{-0.82}$$

∴ $\frac{\sigma_X}{\sigma_Y}$ = 1.22

Exercise 3.3 | Q 6 | Page 50

For a bivariate data: $\bar{x} = 53, \bar{y} = 28,$ b_{YX} = - 1.5 and b_{XY} = - 0.2.

Estimate Y when X = 50.

Solution:

Given, $\bar{x} = 53, \bar{y} = 28,$ b_{YX} = - 1.5 and b_{XY} = - 0.2.

$\mathbf{a} = \bar{y} - \mathbf{b}_{YX}\,\bar{x}$

∴ a = 28 - (- 1.5)(53) = 28 + 79.5 = 107.5

Now, the regression equation Y on X is

Y = a + b_{YX} X

i.e., Y = 107.5 + (- 1.5)X

i.e., Y = 107.5 - 1.5 X

When X = 50, we get

Y = 107.5 - 1.5 × 50 = 107.5 - 75 = 32.5

Exercise 3.3 | Q 7 | Page 50

The equations of two regression lines are x − 4y = 5 and 16y − x = 64. Find means of X and Y. Also, find correlation coefficient between X and Y.

Solution: Given equations of regression lines are

x - 4y = 5 ...(i)

16y - x = 64

i.e., - x + 16y = 64 ...(ii)

Adding (i) and (ii), we get

$$\begin{array}{r} x - 4y = 5 \\ \underline{-x + 16y = 64} \\ 12y = 69 \end{array}$$

$\therefore$ y = $\frac{69}{12} = 5.75$

Substituting y = 5.75 in (i), we get

x - 4(5.75) = 5

$\therefore$ x - 23 = 5

$\therefore$ x = 5 + 23 = 28

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\therefore \bar{x} = 28 \text{ and } \bar{y} = 5.75$

Let, x - 4y = 5 be the regression equation of X on Y

$\therefore$ The equation becomes X = 4Y + 5

Comparing it with X = b_{XY} Y + a', we get

b_{XY} = 4

Now, the other equation i.e. 16y - x = 64 is regression equation of Y on X

$\therefore$ The equation becomes 16Y = X + 64

i.e., Y = $\frac{1}{16}X + \frac{64}{16}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{1}{16}$$

r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{4 \times \frac{1}{16}} = \pm\sqrt{\frac{1}{4}} = \pm\frac{1}{2} = \pm 0.5$$

Since b_{XY} and b_{YX} both are positive,

r is also positive.

$\therefore$ r = 0.5

Exercise 3.3 | Q 8 | Page 50

In a partially destroyed record, the following data are available: variance of X = 25, Regression equation of Y on X is 5y − x = 22 and regression equation of X on Y is 64x − 45y = 22 Find

i. Mean values of X and Y
ii. Standard deviation of Y
iii. Coefficient of correlation between X and Y.

Solution:

Given, $\sigma_X^2 = 25$

$\therefore \sigma_X = 5$

Regression equation of Y on X is

5y – x = 22

Regression equation of X on Y is

64x - 45y = 22

(i) Consider, the two regression equation

- x + 5y = 22 ....(i)

64x - 45y = 22 ....(ii)

By (i) × + (ii), we get

$-9x + 45y = 198$
$+64x - 45y = 22$
$55x \quad = 220$

$\therefore x = 4$

Substituting x = 4 in (i), we get

$-4 + 5y = 22$

$\therefore 5y = 22 + 4$

$\therefore y = \frac{26}{5} = 5.2$

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x}$ = mean value of X = 4 and

$\bar{y}$ = mean value of Y = 5.2

(ii) To find standard deviation of Y we should first find the coefficient of correlation between X and Y.

Regression equation of Y on X is

$5y - x = 22$

i.e., $5Y = X + 22$

i.e., $Y = \frac{X}{5} + \frac{22}{5}$

Comparing it with $Y = b_{YX} X + a$, we get

$b_{YX} = 1/5$

Now, regression equation of X on Y is

$64x - 45y = 22$

i.e., $64X - 45Y = 22$

i.e., $64X = 45Y + 22$

i.e., X = $\frac{45Y}{64} + \frac{22}{64}$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = \frac{45}{64}$$

r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{\left(\frac{1}{5}\right)\left(\frac{45}{64}\right)} = \pm\sqrt{\frac{9}{64}} = \pm\frac{3}{8}$$

Since b_{YX} and b_{XY} are positive,

r is also positive.

∴ r = $\frac{3}{8}$

Now, $b_{YX} = r\frac{\sigma_Y}{\sigma_X}$

$$\therefore \frac{1}{5} = \frac{3}{8} \times \frac{\sigma_Y}{5}$$

$$\therefore \sigma_Y = \frac{1}{5} \times \frac{8}{3} \times 5$$

∴ σ_Y= Standard deviation of Y = $\frac{8}{3}$

(iii) The correlation coefficient of X and Y is

r = $\frac{8}{3} = 0.375$

Exercise 3.3 | Q 9 | Page 50

If the two regression lines for a bivariate data are 2x = y + 15 (x on y) and 4y = 3x + 25 (y on x), find

i. $\bar{x}$,

ii. $\bar{y}$,

iii. b_{YX}

iv. b_{XY}

v. r [Given $\sqrt{0.375}$ = 0.61]

Solution: Given,

regression equation of X on Y is 2x = y + 15,

i.e., 2x - y = 15 …(i)

regression equation of Y on X is 4y = 3x + 25,

i.e., - 3x + 4y = 25 …(ii)

By (i) × 4 + (ii) we get

$$\begin{array}{r} 8x - 4y = 60 \\ + \ -3x + 4y = 25 \\ \hline 5x \quad = 85 \end{array}$$

∴ x = 17

Substituting the value of x = 17 in (i), we get

2(17) – y = 15

∴ 34 – y = 15

∴ y = 34 – 15 = 19

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

(i) $\bar{x}$ = 17

(ii) $\bar{y}$ = 19

(iii) Regression equation of Y on X is 4y = 3x + 25

i.e., 4Y = 3X + 25

i.e., Y = $\frac{3}{4}X + \frac{25}{4}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{3}{4}$$

(iv) Regression equation of X on Y is 2x = y + 15

i.e., 2X = Y + 15

i.e., X = $\frac{Y}{2} + \frac{15}{2}$

Comparing it with X = b_{XY} Y + a' we get,

$$b_{XY} = \frac{1}{2}$$

(v) r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{\left(\frac{1}{2}\right) \cdot \left(\frac{3}{4}\right)} = \pm\sqrt{0.375} = \pm 0.61$$

Since b_{YX} and b_{XY} both are positive,

r is also positive.

$\therefore$ r = 0.61

Exercise 3.3 | Q 10 | Page 50

The two regression equations are 5x − 6y + 90 = 0 and 15x − 8y − 130 = 0. Find $\bar{x}, \bar{y}$, r.

Solution:

Given, the two regression equations are

5x - 6y + 90 = 0

i.e., 5x - 6y = - 90   ...(i)

and 15x - 8y - 130 = 0

i.e., 15x - 8y = 130   ...(ii)

By (i) × 3 – (ii), we get

15x - 18y = - 270

15x - 8y = 130

-     +   -

- 10y = - 400

$\therefore$ y = 40

Substituting y = 40 in (i), we get

5x - 6(40) = –90

$\therefore$ 5x - 240 = - 90

$\therefore$ 5x = - 90 + 240 = 150

$\therefore$ x = 30

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\therefore \bar{x} = 30$ and $\bar{y} = 40$

Now, let 5x – 6y + 90 = 0 be the regression equation of Y on X.

$\therefore$ The equation becomes 6Y = 5X + 90 i.e., Y = $\frac{5}{6}X + \frac{90}{6}$

Comparing it with Y = b_{YX} X + a, we get

$\therefore b_{YX} = \frac{5}{6}$

Now, other equation 15x – 8y – 130 = 0 be the regression equation of X on Y.

∴ The equation becomes 15X = 8Y + 130

i.e., X = $\frac{8}{15}Y + \frac{130}{15}$

Comparing it with X = b_{XY} Y + a', we get

$\therefore b_{XY} = \frac{8}{15}$

∴ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$= \pm\sqrt{\frac{8}{15} \cdot \frac{5}{6}} = \pm\sqrt{\frac{4}{9}} = \pm\frac{2}{3}$

Since b_{YX} and b_{XY} both are positive,

r is positive.

∴ r = $\frac{2}{3}$

Exercise 3.3 | Q 11 | Page 50

Two lines of regression are 10x + 3y – 62 = 0 and 6x + 5y – 50 = 0. Identify the regression of x on y. Hence find $\bar{x}, \bar{y}$ and r.

Solution:

Given, two lines of regression are

10x + 3y – 62 = 0

i.e., 10x + 3y = 62 …(i)

and 6x + 5y – 50 = 0

i.e., 6x + 5y = 50 …(ii)

By (i) × 5 - (ii) × 3, we get

$$\begin{array}{l} 50x + 15y = 310 \\ 18x + 15y = 150 \\ \underline{- \quad - \qquad -} \\ 32x \quad = 160 \end{array}$$

∴ x = 5

Substituting x = 5 in (i) we get,

10(5) + 3y = 62

∴ 50 + 3y = 62

∴ 3y = 62 - 50 = 12

∴ y = 4

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x} = 5$ and $\bar{y} = 4$

Now,

Let 10x + 3y - 62 = 0 be the regression equation of X on Y.

∴ The equation becomes 10x = –3y + 62

i.e., 10X = –3Y + 62

i.e., X = $-\frac{3}{10}\mathrm{Y}+\frac{62}{10}$

Comparing it with X = b_{XY} Y + a, we get

$$\therefore \mathrm{b_{XY}} = -\frac{3}{10}$$

Now, other equation 6x + 5y – 50 = 0 be the regression equation of Y on X.

∴ The equation becomes 5y = – 6x + 50

i.e., 5Y = – 6X + 50

i.e., Y = $-\frac{6}{5}\mathrm{x}+\frac{50}{5}$

Comparing it with Y = b_{YX} X + a', we get

$$\mathrm{b_{YX}} = -\frac{6}{5}$$

Now, $\mathrm{b_{YX}} \cdot \mathrm{b_{XY}} = \left(-\frac{3}{10}\right)\left(-\frac{6}{5}\right) = \frac{9}{25}$

i.e., b_{XY} . b_{YX} < 1

∴ Assumption of regression equations is true.

∴ r = $\pm\sqrt{\mathrm{b_{XY}} \cdot \mathrm{b_{YX}}} = \pm\sqrt{\frac{9}{25}} = \pm\frac{3}{5}$

Since b_{YX} and b_{XY} both are negative,

r is negative.

$\therefore$ r = $-\frac{3}{5} = -0.6$

Exercise 3.3 | Q 12 | Page 50

For certain X and Y series, which are correlated the two lines of regression are 10y = 3x + 170 and 5x + 70 = 6y. Find the correlation coefficient between them. Find the mean values of X and Y.

Solution: Let 10y = 3x + 170 be the regression equation of Y on X.

$\therefore$ The equation becomes 10y = 3x + 170

i.e., Y = $\frac{3}{10}X + \frac{170}{10}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{3}{10}$$

Now other equation 5x + 70 = 6y be the regression equation of X on Y.

$\therefore$ The equation becomes 5x = 6y – 70

i.e., X = $\frac{6}{5}Y - \frac{70}{5}$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = \frac{6}{5}$$

$\therefore$ r = $\pm\sqrt{b_{XY} \cdot b_{YX}} = \pm\sqrt{\frac{6}{5} \times \frac{3}{10}} = \pm\sqrt{\frac{9}{25}} \pm \frac{3}{5}$

Since b_{YX} and b_{XY} both are positive,

r is positive.

$\therefore r = \frac{3}{5} = 0.6$

Now, two correlated lines of regression are

10y = 3x + 170

i.e., - 3x + 10y = 170 ...(i)

and 5x + 70 = 6y

i.e., 5x - 6y = –70 ...(ii)

By (i) × 5 + (ii) × 3, we get

- 15x + 50y = 850

+ 15x - 18y = - 210

32y = 640

∴ y = 20

Substituting y = 20 in equation (i), we get

- 3x +10(20) = 170

∴ - 3x + 200 = 170

∴ 3x = 200 - 170

∴ x = 10

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x}$ = mean value of X = 10, and

$\bar{y}$ = mean value of Y = 20.

Exercise 3.3 | Q 13 | Page 50

Regression equations of two series are 2x – y – 15 = 0 and 3x – 4y + 25 = 0. Find $\bar{x}, \bar{y}$ and regression coefficients. Also find coefficients of correlation. [Given $\sqrt{0.375}$ = 0.61]

Solution: Given regression equation of two series are

2x - y – 15 = 0

i.e., 2x – y = 15 …(i)

and 3x - 4y + 25 = 0

i.e., 3x – 4y = –25 …(ii)

By (i) × – (ii), we get

8x - 4y = 60
3x - 4y = - 25
\- + +

5x = 85

∴ x = 17

Substituting x = 17 in in equation (i), we have

2(17) - y = 15

∴ 34 - y = 15

∴ y = 34 - 15 = 19

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$$\bar{x} = 17 \text{ and } \bar{y} = 19$$

Now,

Let 2x - y – 15 = 0 be the regression equation of X on Y.

∴ The equation becomes 2x = y + 15

i.e., X = $\frac{Y}{2} + \frac{15}{2}$

Comparing it with X = b_{XY} Y + a, we get

$\therefore b_{XY} = \frac{1}{2}$

Now, other equation 3x - 4y + 25 = 0 be the regression equation of Y on X.

$\therefore$ The equation becomes 4y = 3x + 25

i.e., Y = $\frac{3}{4}X + \frac{25}{4}$

Comparing it with Y = b_{YX} X + a', we get

$\therefore b_{YX} = \frac{3}{4}$

$\therefore$ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{\frac{1}{2} \cdot \frac{3}{4}} = \pm\sqrt{0.375} = \pm 0.61$$

Since b_{YX} and b_{XY} both are positive,

r is positive.

$\therefore$ r = 0.61

Exercise 3.3 | Q 14 | Page 50

The two regression lines between height (X) in inches and weight (Y) in kgs of girls are,
4y − 15x + 500 = 0
and 20x − 3y − 900 = 0

Find the mean height and weight of the group. Also, estimate the weight of a girl whose height is 70 inches.

Solution: Given, X = Height (in inches), Y = weight (in Kg)

The equation of regression are

$4y - 15x + 500 = 0$

i.e., $-15x + 4y = -500$ …(i)

and $20x - 3y - 900 = 0$

i.e., $20x - 3y = 900$ …(ii)

By 3 × (i) + 4 × (ii), we get

$- 45x + 12y = - 1500$
$\underline{+ 80x - 12y = 3600}$
$35x = 2100$
$\therefore x = 60$

Substituting x = 60 in (i), we get

$-15(60) + 4y = -500$

$\therefore 4y = 900 - 500$

$\therefore y = 100$

Since the point of intersection of two regression lines is $\bar{x}, \bar{y}$,

$\bar{x}$ = mean height of the group = 60 inches, and

$\bar{y}$ = mean weight of the group = 100 kg.

Let $4y - 15x + 500 = 0$ be the regression equation of Y on X.

$\therefore$ The equation becomes $4y = 15x - 500$

i.e., $Y = \frac{15}{4}X - \frac{500}{4}$ ...(i)

Comparing it with Y = b_{YX} X + a, we get

$$\therefore b_{YX} = \frac{15}{4}$$

∴ Now, other equation 20x – 3y – 900 = 0 be the regression equation of X on Y

∴The equation becomes 20x – 3y – 900 = 0

i.e., 20x = 3y + 900

X = $\frac{3}{20}Y + \frac{900}{20}$

Comparing it with X = b_{XY} Y + a',

$$\therefore b_{XY} = \frac{3}{20}$$

Now, $b_{YX} \cdot b_{XY} = \frac{15}{4} \cdot \frac{3}{20} = 0.5625$

i.e., $b_{XY} . b_{YX} < 1$

∴ Assumption of regression equations is true.

Now, substituting x = 70 in (i) we get

y = $\frac{15}{4} \times 70 - \frac{500}{4} = \frac{1050 - 500}{4} = \frac{550}{4} = 137.5$

∴ Weight of girl having height 70 inches is 137.5 kg

MISCELLANEOUS EXERCISE 3 [PAGES 51 - 54]

Miscellaneous Exercise 3 | Q 1.01 | Page 51

Choose the correct alternative.

Regression analysis is the theory of

1. Estimation
2. Prediction
3. **Estimation and Prediction**
4. Calculation

Solution: Regression analysis is the theory of **Estimation and Prediction.**

Miscellaneous Exercise 3 | Q 1.02 | Page 51

Choose the correct alternative.

We can estimate the value of one variable with the help of other known variable only if they are

1. **Correlated**
2. Positively correlated
3. Negatively correlated
4. Uncorrelated

Solution: We can estimate the value of one variable with the help of other known variable only if they are **correlated.**

Miscellaneous Exercise 3 | Q 1.03 | Page 51

Choose the correct alternative.

There are ___________ types of regression equations.

1. 4
2. **2**
3. 3
4. 1

Solution: There are **2** types of regression equations.

Miscellaneous Exercise 3 | Q 1.04 | Page 51

Choose the correct alternative.

In the regression equation of Y on X

1. **X is independent and Y is dependent.**
2. Y is independent and X is dependent.
3. Both X and Y are independent.
4. Both X and Y are dependent.

Solution: In the regression equation of Y on X, **X is independent and Y is dependent.**

Miscellaneous Exercise 3 | Q 1.05 | Page 52

Choose the correct alternative.

In the regression equation of X on Y

1. X is independent and Y is dependent.
2. **Y is independent and X is dependent.**
3. Both X and Y are independent.
4. Both X and Y are dependent.

Solution: Y is independent and X is dependent.

Miscellaneous Exercise 3 | Q 1.06 | Page 52

Choose the correct alternative.

b_{XY} is _____

1. Regression coefficient of Y on X
2. **Regression coefficient of X on Y**
3. Correlation coefficient between X and Y
4. Covariance between X and Y

Solution: b_{XY} is **regression coefficient of X on Y.**

Miscellaneous Exercise 3 | Q 1.07 | Page 52

Choose the correct alternative.

b_{YX} is __________.

1. **Regression coefficient of Y on X**
2. Regression coefficient of X on Y
3. Correlation coefficient between X and Y
4. Covariance between X and Y

Solution: b_{YX} is **regression coefficient of Y on X**.

Miscellaneous Exercise 3 | Q 1.08 | Page 52

Choose the correct alternative.

‘r’ is __________.

1. Regression coefficient of Y on X
2. Regression coefficient of X on Y
3. **Correlation coefficient between X and Y**
4. Covariance between X and Y

Solution: ' r ' is **correlation coefficient between X and Y.**

Miscellaneous Exercise 3 | Q 1.09 | Page 52

Choose the correct alternative.

$b_{XY} . b_{YX}$ is __________.

1. v(x)
2. σx
3. **r^2**
4. (σy)2

Solution: $b_{XY} . b_{YX}$ is **r^2**

Miscellaneous Exercise 3 | Q 1.1 | Page 52

Choose the correct alternative.

$b_{YX} > 1$ then b_{XY} is ________

1. > 1
2. **< 1**
3. > 0
4. < 0

Solution: $b_{YX} > 1$ then b_{XY} is **< 1.**

Miscellaneous Exercise 3 | Q 1.11 | Page 52

Choose the correct alternative.

$|b_{xy} + b_{yx}| \geq$ ________

1. |r|
2. **2 |r|**
3. r
4. 2r

Solution: $|b_{xy} + b_{yx}| \geq$ **2 |r|.**

Miscellaneous Exercise 3 | Q 1.12 | Page 52

Choose the correct alternative.

b_{xy} and b_{yx} are ________

1. Independent of change of origin and scale
2. **Independent of change of origin but not of scale**
3. Independent of change of scale but not of origin
4. Affected by change of origin and scale

Solution: b_{xy} and b_{yx} are **independent of change of origin but not of scale.**

Miscellaneous Exercise 3 | Q 1.13 | Page 52

Choose the correct alternative.

If $u = \frac{x - a}{c}$ and $v = \frac{y - b}{d}$ then b_{yx} = ________

Options

$\frac{d}{c}b_{vu}$

$\frac{c}{d}b_{vu}$

$\frac{a}{b}b_{vu}$

$\frac{b}{a}b_{vu}$

Solution:

$$\text{If } u = \frac{x - a}{c} \text{ and } v = \frac{y - b}{d} \text{ then} b_{yx} = \frac{d}{c}b_{vu}$$

Miscellaneous Exercise 3 | Q 1.14 | Page 52

Choose the correct alternative.

If u = $\frac{x - a}{c}$ and $v = \frac{y - b}{d}$ then b_{xy} = _______

Options

$\frac{d}{c} b_{uv}$

$\frac{c}{d} b_{uv}$

$\frac{a}{b} b_{uv}$

$\frac{b}{a} b_{uv}$

Solution:

If u = $\frac{x - a}{c}$ and $v = \frac{y - b}{d}$ then $b_{xy} = \underline{\frac{c}{d} b_{uv}}$

Miscellaneous Exercise 3 | Q 1.15 | Page 52

Choose the correct alternative.

Corr (x, x) = _____

1. 0
2. **1**
3. - 1
4. can't be found

Solution: Corr (x, x) = **1.**

Miscellaneous Exercise 3 | Q 1.16 | Page 52

Choose the correct alternative.

Corr (x, y) = _____

1. corr (x,x)
2. corr (y,y)
3. **corr (y,x)**
4. cov (y,x)

Solution: Corr (x, y) = **corr (y,x).**

Miscellaneous Exercise 3 | Q 1.17 | Page 52

Choose the correct alternative.

$$\text{Corr}\left(\frac{x - a}{c}, \frac{y - b}{d}\right) = -\text{corr}(x, y) \text{ if,}$$

1. **c and d are opposite in sign**
2. c and d are same in sign
3. a and b are opposite in sign
4. a and b are same in sign

Solution: c and d are opposite in sign

Miscellaneous Exercise 3 | Q 1.18 | Page 52

Choose the correct alternative.

Regression equation of X on Y is ____

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

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

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

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

Solution:

Regression equation of X on Y is $x - \bar{x} = b_{xy}(y - \bar{y})$

Miscellaneous Exercise 3 | Q 1.19 | Page 53

Choose the correct alternative.

Regression equation of Y on X is ___

Options

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

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

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

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

Solution:

Regression equation of Y on X is $(y - \bar{y}) = b_{yx}(x - \bar{x})$

Miscellaneous Exercise 3 | Q 1.2 | Page 53

Choose the correct alternative.

b_{yx} = _______

Options

$r\frac{\sigma_x}{\sigma_y}$

$r\frac{\sigma_y}{\sigma_x}$

$\frac{1}{r}\frac{\sigma_y}{\sigma_x}$

$\frac{1}{r}\frac{\sigma_x}{\sigma_y}$

Solution:

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

Miscellaneous Exercise 3 | Q 1.21 | Page 53

Choose the correct alternative.

b_{xy} = ______

Options

$r\frac{\sigma_x}{\sigma_y}$

$r\frac{\sigma_y}{\sigma_x}$

$\frac{1}{r}\frac{\sigma_y}{\sigma_x}$

$\frac{1}{r}\frac{\sigma_x}{\sigma_y}$

Solution:

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

Miscellaneous Exercise 3 | Q 1.22 | Page 53

Choose the correct alternative.

Cov (x, y) = ________

Options

$\sum(x-\bar{x})(y-\bar{y})$

$\dfrac{\sum(x-\bar{x})(y-\bar{y})}{\text{n}}$

$\dfrac{\sum xy}{\text{n}} - \bar{x}\,\bar{y}$

both $\dfrac{\sum(x-\bar{x})(y-\bar{y})}{\text{n}}$ **and** $\dfrac{\sum xy}{\text{n}} - \bar{x}\,\bar{y}$

Solution:

Cov (x, y) = $\underline{\text{both } \dfrac{\sum(x-\bar{x})(y-\bar{y})}{\text{n}} \text{ and } \dfrac{\sum xy}{\text{n}} - \bar{x}\,\bar{y}}$

Miscellaneous Exercise 3 | Q 1.23 | Page 53

Choose the correct alternative.

If bxy < 0 and byx < 0 then 'r' is __________

1. > 0
2. **< 0**
3. > 1
4. not found

Solution: If bxy < 0 and byx < 0 then 'r' is **<u>< 0.</u>**

Miscellaneous Exercise 3 | Q 1.24 | Page 53

Choose the correct alternative.

If equations of regression lines are 3x + 2y − 26 = 0 and 6x + y − 31 = 0 then means of x and y are __________

1. (7, 4)
2. **(4, 7)**
3. (2, 9)
4. (-4, 7)

Solution: If equations of regression lines are 3x + 2y − 26 = 0 and 6x + y − 31 = 0 then means of x and y are
(4, 7).

Miscellaneous Exercise 3 | Q 2.01 | Page 53

Fill in the blank:

If $b_{xy} < 0$ and $b_{yx} < 0$ then 'r' is __________

Solution: If $b_{xy} < 0$ and $b_{yx} < 0$ then 'r' is **negative.**

Miscellaneous Exercise 3 | Q 2.02 | Page 53

Fill in the blank:

Regression equation of Y on X is_________

Solution:

Regression equation of Y on X is $(y - \bar{y}) = b_{YX}(x - \bar{x})$

Miscellaneous Exercise 3 | Q 2.03 | Page 53

Fill in the blank:

Regression equation of X on Y is_________

Solution:

Regression equation of X on Y is $(x - \bar{x}) = b_{XY}(y - \bar{y})$

Miscellaneous Exercise 3 | Q 2.04 | Page 53

Fill in the blank:

There are __________ types of regression equations.

Solution: There are **2** types of regression equations.

Miscellaneous Exercise 3 | Q 2.05 | Page 53

Fill in the blank:

Corr (x, −x) = __________

Solution: Corr (x, −x) = **-1.**

Miscellaneous Exercise 3 | Q 2.06 | Page 53

Fill in the blank:

If $u = \dfrac{x - a}{c}$ and $v = \dfrac{y - b}{d}$ then b_{xy} = ______

Solution:

If $u = \dfrac{x - a}{c}$ and $v = \dfrac{y - b}{d}$ then $b_{xy} = \underline{\dfrac{c}{d} b_{uv}}$

Miscellaneous Exercise 3 | Q 2.07 | Page 53

Fill in the blank:

If $u = \dfrac{x - a}{c}$ and $v = \dfrac{y - b}{d}$ then b_{yx} = ______

Solution:

If $u = \dfrac{x - a}{c}$ and $v = \dfrac{y - b}{d}$ then $b_{yx} = \underline{\dfrac{d}{c} b_{vu}}$

Miscellaneous Exercise 3 | Q 2.08 | Page 53

Fill in the blank:

$|b_{xy} + b_{yx}| \geq$ ______

Solution: $|b_{xy} + b_{yx}| \geq$ **2 |r|**

Miscellaneous Exercise 3 | Q 2.09 | Page 53

Fill in the blank:

If $b_{yx} > 1$ then b_{xy} is _______

Solution: If $b_{yx} > 1$ then b_{xy} is **< 1.**

Miscellaneous Exercise 3 | Q 2.1 | Page 53

Fill in the blank:

$b_{xy} . b_{yx}$ = ________

Solution: $b_{xy} . b_{yx}$ = **$\underline{r^2}$**

Miscellaneous Exercise 3 | Q 3.01 | Page 53

State whether the following statement is True or False.

Corr (x, x) = 1

1. **True**
2. False

Solution: True.

Miscellaneous Exercise 3 | Q 3.02 | Page 53

State whether the following statement is True or False.

Regression equation of X on Y is $(y - \bar{y}) = b_{yx}(x - \bar{x})$

1. True
2. **False**

Solution: False.

Miscellaneous Exercise 3 | Q 3.03 | Page 53

State whether the following statement is True or False.

Regression equation of Y on X is $(y - \bar{y}) = b_{yx}(x - \bar{x})$

1. **True**
2. False

Solution: True.

Miscellaneous Exercise 3 | Q 3.04 | Page 53

State whether the following statement is True or False.

Corr (x, y) = Corr (y, x)

1. **True**

2. False

Solution: True.

Miscellaneous Exercise 3 | Q 3.05 | Page 53

State whether the following statement is True or False.

b_{xy} and b_{yx} are independent of change of origin and scale.

1. True
2. **False**

Solution: False.

Miscellaneous Exercise 3 | Q 3.06 | Page 53

State whether the following statement is True or False.

'r' is regression coefficient of Y on X

1. True
2. **False**

Solution: False.

Miscellaneous Exercise 3 | Q 3.07 | Page 53

State whether the following statement is True or False.

b_{yx} is correlation coefficient between X and Y

1. True
2. **False**

Solution: False.

Miscellaneous Exercise 3 | Q 3.08 | Page 53

State whether the following statement is True or False.

If u = x - a and v = y - b then $b_{xy} = b_{uv}$

1. **True**
2. False

Solution: True.

Miscellaneous Exercise 3 | Q 3.09 | Page 53

State whether the following statement is True or False.

If u = x - a and v = y - b then $r_{xy} = r_{uv}$

1. **True**
2. False

Solution: True.

Miscellaneous Exercise 3 | Q 3.1 | Page 53

State whether the following statement is True or False.

In the regression equation of Y on X, b_{yx} represents slope of the line.

1. **True**
2. False

Solution: True.

Miscellaneous Exercise 3 | Q 4.01 | Page 54

The data obtained on X, the length of time in weeks that a promotional project has been in progress at a small business, and Y, the percentage increase in weekly sales over the period just prior to the beginning of the campaign.

X	1	2	3	4	1	3	1	2	3	4	2	4
Y	10	10	18	20	11	15	12	15	17	19	13	16

Find the equation of the regression line to predict the percentage increase in sales if the campaign has been in progress for 1.5 weeks.

Solution: Here, X = Length of time in weeks,
Y = Percentage increase in weekly sales

X = x_i	**Y = y_i**	x_i^2	**$x_i y_i$**
1	10	1	10
2	10	4	20
3	18	9	54
4	20	16	80
1	11	1	11
3	15	9	45
1	12	1	12
2	15	4	30
3	17	9	51

4	19	16	76
2	13	4	26
4	16	16	64
30	176	90	479

From the table, we have

n = 12, $\sum$ xi = 30, $\sum$ yi = 176, $\sum x_i^2 = 90$, $\sum x_i y_i = 479$

$$\therefore \bar{x} = \frac{\sum x_i}{\mathrm{n}} = \frac{30}{12} = 2.5$$

$$\bar{y} = \frac{\sum y_i}{\mathrm{n}} = \frac{176}{12} = 14.67$$

Now, $\mathrm{b_{YX}} = \frac{\sum \mathrm{x_i y_i} - \mathrm{n\bar{x}\bar{y}}}{\sum \mathrm{x_i^2} - \mathrm{n\bar{x}^2}}$

$$= \frac{479 - 12 \times 2.5 \times 14.67}{90 - 12 \times (2.5)^2}$$

$$= \frac{479 - 440.1}{90 - 75} = \frac{38.9}{15} = 2.59$$

Also,

$\mathbf{a} = \bar{y} - \mathrm{b_{YX}}\ \bar{x}$

= 14.67 - 2.59 × 2.5 = 14.67 - 6.475 = 8.195

∴ a ≈ 8.2

∴ The regression equation of percentage increase in weekly sales (Y) on length of weeks (X) is

Y = a + b_{YX} + X

i.e., Y = 8.2 + 2.59 X

For X = 1.5, we get

Y = 8.2 + 2.59(1.5) = 8.2 + 3.885 = 12.085

∴ Increase in sales is 12.085% if the campaign has been in progress for 1.5 weeks.

Miscellaneous Exercise 3 | Q 4.02 | Page 54

The regression equation of y on x is given by 3x + 2y − 26 = 0. Find b_{yx}.

Solution: Given, regression equation of y on x is

3x + 2y – 26 = 0

∴ The equation becomes 2Y = –3X + 26

i.e., Y = $\frac{-3}{2}\mathrm{X} + \frac{26}{2}$

Comparing it with Y= b_{YX} X + a, we get

$\mathrm{b_{YX}} = \frac{-3}{2}$

Miscellaneous Exercise 3 | Q 4.03 | Page 54

If for bivariate data $\bar{x} = 10, \bar{y} = 12,$ v(x) = 9, σ_y = 4 and r = 0.6

estimate y, when x = 5.

Solution:

Given, $\bar{x} = 10, \bar{y} = 12,$ v(x) = 9, σ_y = 4, r = 0.6

$\therefore \sigma_x = 3$

To estimate y, we should first find the regression equation of Y on X.

$$\therefore b_{YX} = r\frac{\sigma_y}{\sigma_x} = 0.6 \times \frac{4}{3} = 0.8$$

Also, $a = \bar{y} - b_{YX}\,\bar{x}$

= 12 - 0.8(10) = 12 - 8 = 4

The regression equation of Y on X is

$Y = a + b_{YX} X$

$\therefore$ Y = 4 + 0.8 X

For X = 5,

Y = 4 + 0.8 (5) = 4 + 4 = 8

Miscellaneous Exercise 3 | Q 4.04 | Page 54

The equation of the line of regression of y on x is y = 2/9 x and x on y is x = y/2 + 7/6.

Find (i) r, (ii) σ_y^2 if $\sigma_x^2 = 4$

Solution: Given, regression equation of Y on X is

$y = \frac{2}{9}x$

i.e., $Y = \frac{2}{9}X$

Comparing with $Y = b_{YX} X + a$, we get

$$b_{YX} = \frac{2}{9}$$

and regression equation of X on Y is

$$x = \frac{y}{2} + \frac{7}{6}$$

i.e., $X = \frac{1}{2}Y + \frac{7}{6}$

Comparing it with X = bXYY + a', we get

$$b_{XY} = \frac{1}{2}$$

(i) $r = \pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{\frac{1}{2} \cdot \frac{2}{9}} = \pm\sqrt{\frac{1}{9}} = \pm\frac{1}{3}$$

Since b_{YX} and b_{XY} both are positive,

r is positive.

$\therefore r = \frac{1}{3}$

(ii) Given, $\sigma_X^2 = 4$

$\therefore \sigma_X = 2$

we know that, $b_{YX} = r\frac{\sigma_Y}{\sigma_X}$

$$\therefore \sigma_Y = \frac{b_{YX} \times \sigma_X}{r} = \frac{\frac{2}{9} \times 2}{\frac{1}{3}} = \frac{4 \times 3}{9} = \frac{4}{3}$$

$$\therefore \sigma_Y^2 = \frac{16}{9}$$

Miscellaneous Exercise 3 | Q 4.05 | Page 54

Identify the regression equations of x on y and y on x from the following equations, 2x + 3y = 6 and 5x + 7y − 12 = 0

Solution: Given two equations are
2x + 3y = 6 and 5x + 7y – 12 = 0

Let 2x + 3y = 6 be the regression equation of Y on X.

∴ The equation becomes 2X + 3Y = 6

i.e., 3Y = 6 - 2X

i.e., $Y = \frac{-2}{3}X + \frac{6}{3}$

Comparing it with $Y = b_{YX} X + a$, we get

$$b_{YX} = -\frac{2}{3}$$

Now, other equation 5x + 7y – 12 = 0 be the regression equation of X on Y.

∴ The equation becomes 5X + 7Y – 12 = 0

i.e., 5X = - 7Y + 12

$\therefore X = -\frac{7}{5}Y + \frac{12}{5}$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = -\frac{7}{5}$$

Now, $b_{XY} \cdot b_{YX} = -\frac{7}{5} \times \left(-\frac{2}{3}\right) = \frac{14}{15} < 1$

∴ Our assumption of regression equation is true.

∴ 2x + 3y = 6 is the regression equation of Y on X, and 5x + 7y – 12 = 0 is the regression equation of X on Y.

Miscellaneous Exercise 3 | Q 4.06 | Page 54

If for a bivariate data b_{yx} = – 1.2 and b_{xy} = – 0.3 then find r.

Solution: Given, b_{yx} = – 1.2 and b_{xy} = – 0.3

∴ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{(-1.2)(-0.3)} \pm \sqrt{0.36} = \pm 0.6$$

Since b_{XY} and b_{YX} both are negative,

r is negative.

∴ r = - 0.6

Miscellaneous Exercise 3 | Q 4.06 | Page 54

From the two regression equations y = 4x – 5 and 3x = 2y + 5, find $\bar{x}$ and $\bar{y}$.

Solution: Given two regression equations are

y = 4x - 5

i.e., – 4x + y = –5 …(i)

and 3x = 2y + 5

i.e., 3x – 2y = 5   …(ii)

By (i) × 2 + (ii), we get

- 8x + 2y = - 10
+ 3x - 2y = 5
- 5x = - 5

∴ x = 1

Substituting x = 1 in (i)

-4(1) + y = - 5

∴ y = 4 - 5 = - 1

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x} = 1$ and $\bar{y} = -1$

Miscellaneous Exercise 3 | Q 4.07 | Page 54

The equations of the two lines of regression are 3x + 2y − 26 = 0 and 6x + y − 31 = 0
Find

i. Means of X and Y
ii. Correlation coefficient between X and Y
iii. Estimate of Y for X = 2
iv. var (X) if var (Y) = 36

Solution: (i) Given equations of regression are

3x + 2y - 26 = 0

i.e., 3x + 2y = 26   ….(i)

and 6x + y - 31 = 0

i.e., 6x + y = 31   ….(ii)

By (i) - 2 × (ii), we get

3x + 2y = 26

12x + 2y = 62

-      -   -
- 9x = - 36

$\therefore$ x = $\frac{-36}{-9} = 4$

Substituting x = 4 in (ii), we get

6 × 4 + y = 31

$\therefore$ 24 + y = 31

$\therefore$ y = 31 - 24

$\therefore$ y = 7

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x}$ = mean of X = 4, and

$\bar{y}$ = mean of Y = 7

(ii) Let 3x + 2y 26 = 0 be the regression equation of Y on X.

$\therefore$ The equation becomes 2Y = 3X + 26

i.e., Y = $\frac{-3}{2}\mathrm{X} + \frac{26}{2}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{-3}{2}$$

Now, the other equation 6x + y - 31 = 0 is the regression equation of X on Y.

$\therefore$ The equation becomes 6X = - Y + 31 i.e., X = $\frac{-1}{6}Y + \frac{31}{6}$

Comparing it with X = b_{XY} Y+ a', we get

$b_{XY} = \frac{-1}{6}$

∴ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$= \pm\sqrt{\frac{-1}{6} \times \frac{-3}{2}} = \pm\sqrt{\frac{1}{4}} = \pm\frac{1}{2} = \pm 0.5$

Since the values of b_{XY} and b_{YX} are negative,

r is also negative.

∴ r = - 0.5

(iii) The regression equation of Y on X is

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

For X = 2, we get

Y = $\frac{-3}{2} \times 2 + \frac{26}{2} = -3 + 13 = 10$

(iv) Given, Var (Y) = 36, i.e., σ_Y^2 = 36

∴ $\sigma_Y = 6$

Since $b_{XY} = r \times \frac{\sigma_X}{\sigma_Y}$

$$\frac{-1}{6} = -0.5 \times \frac{\sigma_X}{6}$$

$$\therefore \sigma_X = \frac{-6}{-6 \times 0.5} = 2$$

$\therefore \sigma_X^2$ = Var(X) = 4

Miscellaneous Exercise 3 | Q 4.08 | Page 54

Find the line of regression of X on Y for the following data:

n = 8, $\sum(x_i - \bar{x})^2 = 36, \sum(y_i - \bar{y})^2 = 44, \sum(x_i - \bar{x})(y_i - \bar{y}) = 24$

Solution:

Given, n = 8, $\sum(x_i - \bar{x})^2 = 36$

$\sum(y_i - \bar{y})^2 = 44, \sum(x_i - \bar{x})(y_i - \bar{y}) = 24$

$$\therefore \text{b}_{\text{XY}} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(y_i - \bar{y})^2} = \frac{24}{44} = \frac{6}{11}$$

Now, the regression equation of X on Y is

$(\text{X} - \bar{x}) = \text{b}_{\text{XY}}(\text{Y} - \bar{y})$

i.e., $(\text{X} - \bar{x}) = \frac{6}{11}(\text{Y} - \bar{y})$

Miscellaneous Exercise 3 | Q 4.09 | Page 54

Find the equation of line of regression of Y on X for the following data:

n = 8, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 120, \bar{x} = 20, \bar{y} = 36, \sigma_x = 2, \sigma_y = 3$

Solution:

Given, n = 8, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 120,$

$\bar{x} = 20, \bar{y} = 36, \sigma_x = 2, \sigma_y = 3$

∴ Var (X) = $\sigma_X^2 = 4$

Since Var (X) = $\frac{\sum(x_i - \bar{x})}{n}$,

4 = $\frac{(\sum(x_i - \bar{x}))^2}{8}$

∴ $\sum(x_i - \bar{x})^2$ = 32

Now, $b_{YX} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{120}{32} = 3.75$

∴ The regression equation of Y on X is

$(Y - \bar{y}) = b_{YX}(X - \bar{x})$

∴ (Y - 36) = 3.75 (X - 20)

∴ Y - 36 = 3.75X - 75

∴ Y = 3.75X - 75 + 36

∴ Y = 3.75 X - 39

Miscellaneous Exercise 3 | Q 4.1 | Page 54

The following results were obtained from records of age (X) and systolic blood pressure (Y) of a group of 10 men.

	X	Y
Mean	50	140
Variance	150	165

and $\sum(x_i - \bar{x})(y_i - \bar{y}) = 1120$

Find the prediction of blood pressure of a man of age 40 years.

Solution: Given, X = Age, Y = Systolic blood pressure,

n = 10, $\bar{x} = 50, \bar{y} = 140,$

$\sigma_X^2 = 150,\ \sigma_Y^2 = 165\ \text{and}$

$\sum(x_i - \bar{x})(y_i - \bar{y}) = 1120$

Since Var(X) = $\dfrac{\sum(x_i - \bar{x})^2}{n}$,

$$\sigma_x^2 = \frac{\sum(x_i - \bar{x})^2}{n}$$

$$\therefore 150 = \frac{\sum(x_i - \bar{x})^2}{10}$$

$$\therefore \sum(x_i - \bar{x})^2 = 1500$$

Now, $b_{YX} = \dfrac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \dfrac{1120}{1500} = 0.7$

∴ The regression equaiton of systolic blood pressure of the men (Y) on their age (X) is

$$(Y - \bar{y}) = b_{YX}(X - \bar{x})$$

∴ (Y - 140) = 0.7(X - 50)

∴ Y - 140 = 0.7X - 35

∴ Y = 0.7X - 35 + 140

∴ Y = 0.7X + 105

For X = 40,

Y = 0.7(40) + 105 = 28 + 105 = 133

∴ The man of age 40 years has systolic blood pressure 133.

Miscellaneous Exercise 3 | Q 4.11 | Page 54

The equations of two regression lines are 10x − 4y = 80 and 10y − 9x = − 40 Find:

i. $\bar{x}$ and $\bar{y}$

ii. b_{YX} and b_{XY}

iii. If var (Y) = 36, obtain var (X)

iv. r

Solution: (i) Given equations of regression are

10x - 4y = 80

i.e., 5x - 2y = 40 ….(i)

and 10y - 9x = - 40

i.e., - 9x + 10y = - 40 ….(ii)

By 5 × (i) + (ii), we get

25x - 10y = 200

<u>- 9x + 10y = - 40</u>

16x = 160

∴ x = 10

Substituting x = 10 in (i), we get

5(10) - 2y = 40

∴ 50 - 2y = 40

∴ - 2y = 40 - 50

∴ - 2y = - 10

∴ y = 5

Since the point of intersection of two regression lines is $(\bar{x}, \bar{y})$,

$\bar{x} = 10$ and $\bar{y} = 5$

(ii) Let 10y - 9x = - 40 be the regression equation of Y on X.

∴ The equation becomes 10Y = 9X - 40

i.e., Y = $\frac{9}{10}X - \frac{40}{10}$

Comparing it with Y = b_{YX} X + a, we get

$$b_{YX} = \frac{9}{10} = 0.9$$

Now, the other equation 10x □ 4y = 80 be the regression equation of X on Y.

∴ The equation becomes 10X = 4Y + 80

i.e., X = $\frac{4}{10}Y + \frac{80}{10}$

i.e., X = $\frac{2}{5}Y + 8$

Comparing it with X = b_{XY} Y + a', we get

$$b_{XY} = \frac{2}{5} = 0.4$$

(iii) Given, Var (Y) = 36, i.e., σ_Y^2 = 36

∴ σ_Y = 6

Since $b_{XY} = r \times \frac{\sigma_X}{\sigma_Y}$

$$\frac{2}{5} = 0.6 \times \frac{\sigma_X}{6}$$

$$\therefore \frac{2}{5} = 0.1 \times \sigma_X$$

$$\therefore \frac{2}{5 \times 0.1} = \sigma_X$$

$\therefore \sigma_X = 4$

$\therefore \sigma_X^2 = 16$ i.e., Var(X) = 16

(iv) r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{\frac{2}{5} \times \frac{9}{10}} \pm \sqrt{\frac{9}{25}} = \pm\frac{3}{5} = \pm 0.6$$

Since b_{YX} and b_{XY} are positive,

r is also positive.

$\therefore$ r = 0.6

Miscellaneous Exercise 3 | Q 4.12 | Page 54

If b_{YX} = − 0.6 and b_{XY} = − 0.216, then find correlation coefficient between X and Y. Comment on it.

Solution: Given, b_{YX} = − 0.6, b_{XY} = − 0.216

$\therefore$ r = $\pm\sqrt{b_{XY} \cdot b_{YX}}$

$$= \pm\sqrt{-0.216 \cdot (-0.6)} = \pm\sqrt{0.1296}$$

$\therefore$ r = ± 0.36

Since b_{XY} and b_{YX} are negative,

r is also negative.

$\therefore$ r = - 0.36

$\therefore$ X and Y negatively correlated.