3. Linear regression

Exercise: 3.1

1. 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 the department Referring to years of service and their monthly incomes.

Years of service (x)	11	7	9	5	8	6	10
Monthly income (Rs.1000's) (y)	10	8	6	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 similar capacity in another company for 13 Years?
Solution:
Let x = years of service. Y = monthly income (in ₹ 1000's)

x	y	Xy	X^2	
11	10	110	121	
7	8	56	49	
9	6	54	81	
5	5	25	25	
8	9	72	64	
6	7	42	36	
10	11	110	100	
$\sum x = 56$	$\sum y = 56$	$\sum xy = 469$	$\sum x^2 = 476$	N = 7

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

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

Regression coefficient of y on x:

$$byx = \frac{\frac{\sum xy}{n} - \bar{x}.\bar{y}}{\frac{\sum x^2}{n} - \bar{x}^2}$$

$$= \frac{\frac{469}{7} - 8 \times 8}{\frac{479}{7} - (8)^2}$$

$$= \frac{67 - 64}{68 - 64} = \frac{3}{4} = 0.75$$

i) Regression equation of income (y) on years of Service (x):

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

∴ y – 8 = 0.75 (x - 8)
∴ y – 8 = 0.75x – 6
∴ y = 0.75x - 6 + 8
∴ y =2 + 0.75x

ii) Estimate of initial start (y) when years of service
x = 13:
y =2 + 0.75x
Putting x = 13 we get
Y =2 + 0.75 × 13
∴ y =2 + 9.75
∴ y = 11.75
Hence, the initial start, the person will get is ₹ 11.75 × 1000 = ₹ 11750.

2. Calculate the regression equation 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:
We prepare the following table showing the calculation:

$x = x_i$	$y = y_i$	$(x_i - \bar{x})$ $\bar{x} = 14$	$(y_i - \bar{y})$ $\bar{y} = 8$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - y)^2$
10	5	-4	-3	12	16	9
12	6	-2	-2	4	4	4
13	7	-1	-1	1	1	1
17	9	3	1	3	9	1
18	13	4	5	20	16	25
$\sum x = 70$	$\sum y = 40$	$\sum(x_i - \bar{x}) = 0$	$\sum(y_i - \bar{y}) = 0$	$\sum(x_i - \bar{x})(y_i - \bar{y}) = 40$	$\sum(x_i - \bar{x})^2 = 46$	$\sum(y_i - y)^2 = 40$

Here, n = 5

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

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

Regression equation of x and y:

$x = a' + b_{xy} . y$

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

Putting the value, we get

$b_{xy} = \frac{40}{40} = 1$

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

$Putting\ \bar{x} = 14, b_{xy}$

$\rightarrow = 1, \bar{y} = 8, we\ get$

a' = 14 – 1(18) = 14 – 8 = 6.

$Putting\ a' = 6\ and\ b_{xy}$

= 1, we get the regression equation on x on y as

X = 6 + 1 (y) = x = y + 6

Regression equation of y on x.

$y = a + b_{xy} . x$

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

$= \frac{40}{46} = 0.87$

$\therefore a = \bar{y} - b_{xy} . \bar{x}$

$putting\ \bar{y} = 8, b_{xy}$

$\rightarrow = 0.87\ and\ \bar{x} = 14, we\ get$

a = 8 – 0.87 (14)
= 8 – 12.18
= - 4 .18
$Putting\ a = -4.18\ and\ b_{xy}$
= 0.87, we get the regression
Equation of y on x as
Y = -4.18 + 0.87x = y = 0.87x – 4.18
Hence, the aggression equation of x on y is x = y + 6
And the aggression equation of y on x is y = 0.87x – 4.18.

3. for a certain bivariate data on 5 pairs of observations given.
$\boldsymbol{\sum x = 20, \sum y = 20, \sum x^2}$
$\boldsymbol{\rightarrow = 90, \sum y^2 = 90, \sum xy = 76}$

Calculate
i) cov(x,y)

$\boldsymbol{ii)\ b_{yx}\ and\ b_{xy}}$

iii) r.

Solution:

$Given: \sum x = 20, \sum y = 20, \sum x^2$
$\rightarrow = 90, \sum y^2 = 90, \sum xy = 76, n = 5.$
$\therefore\ \overline{x} = \frac{\sum x}{n} = \frac{20}{5} = 4;$
$\rightarrow\ \overline{y} = \frac{\sum y}{n} = \frac{20}{5} = 4$

i) cov(x,y):
$cov\,(x\,,y) = \frac{1}{n}\sum xy - \overline{x}\,\overline{y}$
$= \frac{76}{5} - (4 \times 4) = 15.2\ - 16$
∴ cov (x , y) = -0.8

$ii)\ b_{yx}\ and\ b_{xy}:$
$b_{xy} = \frac{\sum xy - n\overline{x}\,\overline{y}}{\sum x^2 - n\overline{x}2}$
$= \frac{76 - 5(4 \times 4)}{90 - 5\,(4)^2} = \frac{76 - 80}{90 - 80}$
$= \frac{-4}{10} = \ -0.4$

$\therefore\ b_{yx} = -0.4$

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

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

$$= \frac{-4}{10} = -0.4$$

$\therefore\ b_{xy} = -0.4$

iii) r .

$$r = \pm\sqrt{b_{yx}\,.\,b_{xy}}$$

$$= -\sqrt{b_{xy}\,.\,b_{xy}}$$

→ $\ldots.(\therefore\ b_{yx}\ and\ b_{xy}\ are\ negative)$

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

$$= -\sqrt{0.16}$$

= - 0.4

∴ r = - 0.4

4. From the following data estimate y when x = 125

x	120	115	120	125	126	123
y	13	15	14	13	12	14

Solution:

We have to estimate Y when X = 125. So we obtain the regression
Equation of Y on X.
The value of aggression coefficient are independent
of change of origin. Therefore we take the new variable.
U =(x – a) and u = (y –b), taking a = 120 and b = 14
We prepare the following table for calculation:

X = x	Y = y	U = (x –a) a = 120
120	13	0
115	15	-5
120	14	0
125	13	5
126	12	6

123	14	3
$\sum$x = 729	$\sum$y = 81	14 -5 $\sum$u = 9

(Table continue here)

U = (y – b) B = 14	uv	U2
-1	0	0
1	-5	25
0	0	0
-1	-5	25
-2	-12	36
0	0	9
1 – 4 $\sum$u = -3	$\sum$uv = -22	$\sum u^2 = 95$

Here, n = 6

$\overline{x} = \frac{\sum x}{n} = \frac{729}{6} = 121.5;$

→ $\overline{y} = \frac{\sum y}{n} = \frac{81}{6} = 13.5$

$\overline{u} = \frac{\sum u}{n} = \frac{9}{6} = 1.5;$

→ $\overline{v} = \frac{\sum v}{n} = \frac{-3}{6} = 0.5$

Regression equation of Y on x.

$y = a + b_{yx}.x$

$b_{yx} = b_{vu} = \frac{\sum uv - n\overline{u}\,\overline{v}}{\sum u^2 - n\,\overline{u^2}}$

$= \frac{-22 - 6\,[1.5 \times (-0.5)]}{95 - 6(1.5)^2}$

$= \frac{-22 + 4.5}{95 - 13.5}$

$= \frac{-17.5}{81.5}$

= - 0.21

$a = \overline{y} - b_{yx}.\overline{x}$

$Putting\ \overline{y} = 13.5, b_{yx}$
$\rightarrow = -0.21, \overline{x} = 121.5, we\ get$
a = 13.5 – (- 0.21 × 121.5)
= 13.5 + 25.515 = 39.015
$Putting\ a\ =\ 39.015\ and\ \ b_{yx}$
= -0.21, we get the regression
Equation on Y on Y as follows:
Y = 39.015 – 0.21x
Estimation of Y when X = 125
Putting X = 125 in the equation y = 39.015 – 0.21x, we get
Y = 39.015 – 0.21 (125)
= 39.015 – 26.25
= 12.765

5. 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:
i) The productivity index of a worker whose test score is 95

ii) The test score when productivity index is 75.

Solution:
We prepare the following table for calculation:

Aptitude score $\overline{x}$	Productivity index $\overline{y}$	$(x - \overline{x})\overline{x} = 65$
60	68	-5
62	60	-3
65	62	0
70	80	5
72	85	7
48	40	-17
53	52	-12
73	62	8
65	60	0

82	81	17
$\sum x = 650$	$\sum y = 650$	$\sum(x-\overline{x}) = 0$

(Table continue here)

$(y-\overline{y})\overline{y} = 65$	$(x-\overline{x})$ × $(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	
3	-15	25	9	
-5	15	9	25	
-3	0	0	9	
15	75	25	225	
20	140	49	400	
-25	425	289	625	
-13	156	144	169	
-3	-24	64	9	
-5	0	0	25	
16	272	289	256	
$\sum(y-\overline{y}) = 0$	1083 -39 $\sum(x-\overline{x})(y-\overline{y}) = 1044$	$\sum(x-\overline{x}2 = 894$	$\sum(y-\overline{y})^2 = 1752$	N = 10

$\overline{x} = \frac{\sum x}{n} = \frac{650}{10} = 65;$

$\rightarrow \overline{y} = \frac{\sum y}{n} = \frac{650}{10} = 65$

Regression equation of Y on Y:

$y = a + b_{yx}.x$

$b_{yx} = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})2}$

$\rightarrow = \frac{1044}{894} = 1.1678$

$a = \overline{y} - b_{yx}.x$

$Putting\ \overline{y} = 65, b_{yx} = 1.1678, we\ get$

a = 65 – 1.1678 (65) = 65 – 75.907 = - 10.907

Therefore the aggression equation of Y on X is

y = -10.907 + 1.1678x .... (1)

Regression equation of X on Y:

$y = a' + b_{xy}.y$

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

$\rightarrow = \frac{1044}{1752} = 0.5959$

$a' = \bar{x} - b_{xy} \,.\bar{y}$

$putting\ \bar{x} = 65, b_{xy}$

= 0.5959, we get

a' = 65 – 0.5959 × 65 = 65 – 38.73 = 26.27

Therefore, the aggression equation of X on Y is

X = 26.27 + 0.5959y ... (2)

i) Estimation of productivity index (y) when X = 95

putting x = 95 in equation (1), we get

y = - 10.907 + 1.1678(95)

∴ y = - 10.907 + 110.941

∴ y = 100.034

ii) Estimation of aptitude score (x) when y= 75

Putting y = 75 in equation (2), we get

X = 26.27 + 0.5959 (75)

∴ x = 26.27 + 44.6925

∴ x = 70.9625

6. 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:

The appropriate regression equation for the given data is of Y on X.

We prepare the following table for calculation:

$x_i = x$	$y_i = y$	Xy	X^2
2	18	36	4
4	12	48	16
5	10	50	25
6	8	48	36
8	7	56	64
11	5	55	121
$\sum x = 36$	$\sum y = 60$	$\sum xy = 293$	$\sum x^2 = 266$

Here, n = 6

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

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

Regression equation of Y on X:

$y = a + b_{yx} .x$

$now, b_{yx} = \frac{\sum xy - n\bar{x}\,\bar{y}}{\sum x^2 - n\,\overline{\overline{x^2}}}$

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

$\rightarrow = \frac{293 - 360}{266 - 216}$

$= \frac{-67}{50} = -1.34$

$\therefore\ b_{yx} = -1.34$

$a = \bar{y} - b_{yx} .\bar{x}$

$Putting\ \bar{y} = 10,$

$b_{yx} = -1.34\ and$

$\bar{x} = 6, we\ get$

a = 10 – (-1.34) (6)

= 10 + 8.04

= 18.04

$Putting\ a = 18.04\ and\ b_{yx}$

= -1.34 in y

$\rightarrow = a + b_{yx}\ .\ x, we\ get$

The aggression equation of Y on X as follows:

y = 18.04 - 1.34x

y = - 1.34x + 18.04

7. The following are the marks obtained by the students
In economic (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 in economic, Y = marks in mathematics
We prepare the following table for calculation:

x	y	$(x-\bar{x})$ $\bar{x}=61$	$(y-\bar{y})$ $\bar{y}=80.4$	$(x-\bar{x})$ $(y-\bar{y})$	$(x-\overline{x)^2}$
59	78	-2	-2.4	4.8	4
60	82	-1	1.6	-1.6	1
61	82	0	1.6	0	0
62	79	1	-1.4	-1.4	1
63	81	2	0.6	1.2	4
$\sum x = 305$	$\sum y = 402$	$\sum(x-\bar{x})=0$	$\sum(y-\bar{y})=0$	6.0 – 3.0 $\sum(x-\bar{x})$ $(y-\bar{y})=3$	$\sum(x-\overline{x)^2}$ 10

Here, n =5

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

$\rightarrow\ \bar{y}=\frac{\sum y}{n}=\frac{405}{5}=80.4$

Regression equation of Y on X:

$y=a+\ b_{yx}\,.\,x$

$b_{yx}=\frac{\sum(x-\bar{x})\,(y-\bar{y})}{\sum(x-\overline{x)^2}}$

$\rightarrow=\frac{3}{10}=0.3$

$a=\bar{y}-b_{yx}\,.\bar{x}$

$Putting\ \bar{y}=80.4, b_{yx}$

$\rightarrow=0.3\ and\ \bar{x}=61, we\ get$

a = 80.4 – 0.3(61)

= 80.4 – 18.3 = 62.1

∴ a = 62.1

$Putting\ a=62.1\ and\ b_{yx}$

$\rightarrow=0.3\ in\ y=a+b_{yx}\,.x, we\ get\ the$

Regression equation of y on x as follows:

Y = 62.1 + 0.3x

= y = 0.3x + 62.1

8. 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:
We prepare the following table for calculation:

x	**y**	**xy**	x^2	y^2
1	5	5	1	25
2	7	14	4	49
3	9	27	9	81
4	11	44	16	121
5	15	65	25	169
$\sum x = 15$	$\sum y = 45$	$\sum xy = 155$	$\sum x^2 = 55$	$\sum y^2 = 445$

Here, n = 5

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

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

Regression equation of Y on X:

$y = a + b_{yx} . x$

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

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

$\rightarrow = \frac{155 - 135}{55 - 45}$

$= \frac{20}{10} = 2$

$\therefore\ b_{yx} = 2$

$a = \bar{y} - b_{yx} . \bar{x}$

$Putting\ \bar{y} = 9, b_{yx}$

$\rightarrow = 2, \bar{x} = 3, we\ get$

a = 9 -2 (3) = 9 – 6 = 3

$Putting\ a = 3\ and\ b_{yx}$

$\rightarrow = 2\ in\ y = a + b_{yx} .$

$x, we\ get\ the$

Regression equation Y on X as follows:

Y = 3 + 2x

= y = 2x + 3

Regression equation of X on Y:

$x = a' + b_{xy} . y$

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

$= \frac{155 - 5(3 \times 9)}{445 - 5(9)^2}$

$\rightarrow = \frac{155 - 135}{445 - 405}$

$= \frac{20}{40} = \frac{1}{2} = 0.5$

$\therefore\ b_{yx} = 0.5$

$a = \bar{x} - \ b_{yx}\,.\bar{y}$

$Putting\ \bar{x} = 3, b_{yx}$

$\rightarrow = 0.5, and\ \bar{y} = 9,$

We get

a' = 3 -0.5 (9) = 3 – 4.5 = -1.5

∴ a' = - 1.5

$Putting\ a' = -1.5\ and\ b_{xy}\ in\ x$

$\rightarrow = a' + \ b_{yx}\,.y, we\ get\ the$

Regression equation of X on Y as follows:

X = -1.5 + 0.5y

= x = 0.5y – 1.5

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

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

Solution:

We prepare the following table for calculation:

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

Here n = 5

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

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

Regression equation of Y on y:

$y = a + b_{yx} .x$

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

→ $= \frac{-26}{40} = 0.65$

$\therefore\ b_{yx} = -0.65$

$a = \bar{y} - b_{yx} .\bar{x}$

$Putting\ \bar{y} = 8, b_{yx}$

→ $= -0.65\ and\ \bar{x} = 6, we\ get$

a = 8 – (-0.65) 6 = 8+ 3.9 = 11.9

∴ a = 11.9

Hence, regression equation of Y on x is

Y = 11.9 – 0.65x

∴ y = - 0.65x + 11.9

Regression equation of Y on y

$x = a' + b_{xy} .y$

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

→ $= \frac{-26}{20} = -1.3$

$\therefore\ b_{xy} = -1.3$

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

$Putting\ \bar{x} = 6, b_{xy}$

→ $= -1.3\ and\ \bar{y} = 8, we\ get$

a’ = 6 –(-1.3) 8

∴ a’ = 6+ 10.4 ∴ 16.4

Hence the regression equation X on Y is

X = 16.4 – 1.3y

∴ x = - 1.3y + 16.4

10. 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:
We prepare the following table for calculation:

x	y	$(x-\overline{x})$ $\overline{x}=2$	$(y-\overline{y})$ $\overline{y}=3$	$(x-\overline{x})$ $(y-\overline{y})$	$(x-\overline{x})^2$
1	2	-1	-1	1	1
2	1	0	-2	0	0
3	6	1	3	3	1
$\sum x=6$	$\sum y=9$	$\sum(x-\overline{x})=0$	$\sum(y-\overline{y})=0$	$\sum(x-\overline{x})(y-\overline{y})=4$	$\sum(x-\overline{x})^2=2$

Here n = 3

$\overline{x}=\frac{\sum x}{n}=\frac{6}{3}=2;$

$\rightarrow\ \overline{y}=\frac{\sum y}{n}=\frac{9}{3}=3$

Regression line of Y on X:

$y=a+\ b_{yx}\ .x$

$b_{yx}=\frac{\sum(x-\overline{x})\ (y-\overline{y})}{\sum(x-\overline{x})^2}$

$\rightarrow=\frac{4}{2}=2$

$\therefore\ b_{yx}=2$

$a=\ \overline{y}-b_{xy}.\overline{x}$

$Putting\ \overline{y}=3, b_{yx}$

$\rightarrow=\ 2, \overline{x}\ =2, we\ get$

a = 3 -2(2) =3 -4 = -1

∴ a = -1

Hence, regression line of Y on x is

Y = -1 + 2x

∴ y = 2x -1

Mostly likely value of Y when x =4:

Putting x =4 in y = 2x – 1, we get

Y =2 (4) -1

∴ y =8 -1

∴ y = 7

11. 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:
We prepare the following table for calculation:

x	y	xy	x^2
1	2	2	1
2	4	8	4
3	7	21	9
4	6	24	16
5	5	25	25
6	6	36	36
$\sum x = 21$	$\sum y = 30$	$\sum xy = 116$	$\sum x^2 = 91$

Here n = 6

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

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

Regression equation of Y on X:

$y = a + b_{yx} . x$

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

$= \frac{116 - 6(3.5 \times 5)}{91 - 6(3.5)^2}$

$\rightarrow = \frac{116 - 105}{91 - 73.5}$

$= \frac{11}{17.5} = 0.63$

$\therefore \ b_{yx} = 0.63$

$a = \bar{y} - b_{yx} . \bar{x}$

$Putting\ \bar{y} = 5, b_{yx}$

$\rightarrow = 0.63, \bar{x} = 3.5,$

We get

a = 5 - 0.63 (3.5)

= 5 - 2.205 = 2.795

∴ a = 2.795 ≈ 2.8

Hence, regression line of Y on x is

Y = 2.8 + 0.63x

∴ y = 0.63x + 2.8

Estimate of y when x = 10:
Put x =10 in y = 0.63x + 2.8
∴ y = 0.63 (10) + 2.8 ∴ y = 6.3 + 2.8
∴ y = 9.1

12. 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:
We take new variables u = (x –a), a =5 and v = (y –b)
b =70 to make the calculation easier:
We prepare the following table for calculation:

No .of hours of study x	Grades obtained y	U = (x-a) a =5	V = (y – b) b = 70	uv	u^2
3	45	-2	-25	50	4
3	60	-2	-10	20	4
3	55	-2	-15	30	4
4	60	-1	-10	10	1
4	75	-1	5	-5	1
5	70	0	0	0	0
5	80	0	10	0	0
5	75	0	5	0	0
6	90	1	20	20	1
6	80	1	10	10	1
7	75	2	5	10	4
8	85	3	15	45	9
∑x =59	∑y = 850	7-8 ∑u = -1	70 – 60 ∑v = 10	195 – 5 ∑uv = 190	$\sum u^2 = 29$

Here n = 12

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

→ $\bar{y} = \frac{\sum y}{n} = \frac{850}{12} = 70.83$

$\overline{u} = \frac{\sum u}{n} = \frac{-1}{12} = -0.083;$

$\rightarrow \ \overline{v} = \frac{\sum v}{n}$

$= \frac{10}{12} = 0.83$

Regression line of marks (Y) on hours of study (X):

$y = a + \ b_{yx}\,.x$

$\therefore \ b_{yx} = \ b_{vu}$

$\rightarrow = \frac{\sum uv - (\overline{u}\,\overline{v})}{\sum u^2 - n\overline{u^2}}$

$= \frac{190 - 12\ (-0.083 \times 0.83)}{29 - 12\ (\,-0.083)^{\,2}}$

$= \frac{190 + 0.069}{29 - 0.083}$

$\rightarrow = \frac{190.069}{28.917} = 6.57$

$b_{yx} = 6.57\ \approx 6.6$

$a = \ \overline{y} - \ b_{yx}\,.\overline{x}$

$Putting\ \overline{y} = 70.83, b_{yx}$

$\rightarrow = 6.6\,, and\ \overline{x}$

= 4.92, we get

$a = 70.83, b_{yx} = 6.6\ and\ \overline{x}$

= 4.92, we get

a = 70.83 – 6.6(4.92)

∴ a = 70.83 – 32.47

∴ a = 38.36

Hence, the regression line of marks (Y) on hours of study (X) is

Y = 38.36 + 6.6x

Y = 6.6x + 38.36

Exercise: 3.2

1. for a bivariate data:

$\boldsymbol{\overline{x} = 53, \overline{y} = 28, b_{yx}}$

$\rightarrow \boldsymbol{= -1.2\ and\ b_{xy} = -0.3.}$

Find:

i) Correlation coefficient between X and Y:

ii) Estimate of Y for X = 50.

iii) Estimate of X for Y = 25.

Solution:

$Given: \overline{x} = 53, \overline{y} = 28, b_{yx}$
$\rightarrow = -1.2\ and\ b_{xy} = -0.3.$

i) Correlation coefficient between X and Y:

$r = \pm\sqrt{b_{yx} \,.\, b_{xy}}$
$= \pm\sqrt{(-1-2)\,(-0.3)}$
$= \pm\sqrt{0.36}$
∴ r = - 0.6

ii) Estimate of Y for X = 50.
Regression equation Y on X is:
$y = a + b_{yx}\ .x$
$b_{yx} = -1.2$
$a = \overline{y} - b_{yx}.\overline{x}$
= 28 – (- 1.2) 53
= 28 + 63.6
= 91.6
∴ y = 91.6 – 1.2x
= y = - 1.2x + 91.6
Put x = 50
∴ Y = - 1.2(50) + 91.6
∴ y = - 60 + 91.6 ∴ y = 31.6

iii) Estimate of X for Y = 25.
Regression equation of X on Y is
$x = a' + b_{xy}.y$
$b_{xy} = -0.3$
$a' = \overline{x} - b_{xy}\,(\overline{y})$
= 53 – (- 0.3) (28)
= 53 + 8.4
= 61.4
∴ x = 61.4 – 0.3y
X = - 0.3y + 61.4
Put y = 25,
∴ X = - 0.3(25) + 61.4

$\therefore$ X = - 75 + 61.4
$\therefore$ x =53.9

2. From the data of 20 pairs of observation on X and Y, Following results are obtained:

$\overline{x} = 199, \overline{y} = 94$
$\rightarrow \sum(x_i - x)^2 = 1200$
$\sum(y_i - \overline{y})2 = 300,$
$\rightarrow \sum(x_i - \overline{x})(y_i - \overline{y})$
$= -250.$

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:

$Given: \overline{x} = 199, \overline{y} = 94$
$\rightarrow \sum(x_i - x)^2 = 1200$

$\sum(y_i - \overline{y})2 = 300,$
$\rightarrow \sum(x_i - \overline{x})(y_i - \overline{y}) = -250.$

i) The line of regression of Y on X.

$b_{yx} = \frac{\sum(x - \overline{x})(y - \overline{y})}{\sum(x - \overline{x})^2}$
$= \frac{-250}{1200} = -\frac{5}{24}$
$y = a + b_{yx} . x$
$a = \overline{y} - b_{yx} . \overline{x}$
$= 94 - \left(-\frac{5}{24}\right) 199$
$= 94 + \frac{995}{24} =$
$\rightarrow \frac{2256 + 995}{24}$
$= \frac{3251}{24}$
$\therefore$ Line of aggression of Y on X is
$y = \frac{3251}{24} - \frac{5}{24} x$

∴ 24y = 3251 – 5x
∴ 5x + 24y = 3251

ii) The line of regression of X on Y.

$$b_{yx} = \frac{\sum(x-\overline{x})\,(y-\overline{y})}{\sum(y_i-\overline{y})^2}$$

$$= \frac{-250}{300} = -\frac{5}{6}$$

$$y = a' + b_{yx}\,.y$$

$$a' = \overline{x} - b_{yx}\,.\overline{y}$$

$$= 199 - \left(-\frac{5}{6}\right)(94)$$

$$= 199 + \frac{470}{6}$$

$$= \frac{1194+470}{6} = \frac{1664}{6}$$

∴ Line of regression of X on Y is

$$x = \frac{1664}{6} - \frac{5}{6}y$$

∴ 6x = 1664 – 5y
∴ 6x + 5y = 1664

iii) Correlation coefficient between X and Y.

$$b_{yx} = -\frac{5}{24},$$

$$b_{xy} = -\frac{5}{6}$$

$$Now, r = \pm\sqrt{b_{yx}\,.\,b_{xy}}$$

$$= \pm\sqrt{\left(-\frac{5}{24}\right)\left(-\frac{5}{6}\right)}$$

$$= \pm\sqrt{\frac{25}{144}}$$

$$\therefore r = -\frac{5}{12}$$

3. From the data of 7 pairs of observation on X and Y Following results are obtained:

$$\boldsymbol{\sum(x_i - 70) = -35,}$$

$$\boldsymbol{\sum(y_i - 60) = -7,}$$

$\sum(\boldsymbol{x_i} - \mathbf{70})^2 = \mathbf{2989},$
$\rightarrow \sum(\boldsymbol{y_i} - \ \mathbf{60})^2 = \mathbf{476},$

$\sum(\boldsymbol{x_i} - \mathbf{70}) \sum(\boldsymbol{y_i} - \mathbf{60}) = \mathbf{1064}.,$
$[\boldsymbol{Given}: \sqrt{\mathbf{0.7884}} = \mathbf{0.8879}]$

Obtain:
i) The line of regression of Y on X.

ii) The line of aggression of X on Y.

iii) The correlation coefficient between X and Y.

Solution:
Given: n = 7
$\rightarrow \sum(x_i - 70) = -35$

$\sum(y_i - 60) = -7,$
$\rightarrow \ \sum(x_i - 70)^2 = 2989$

$\sum(y_i - \ 60)^2 = 476$
$\sum(x_i - 70)$
$\rightarrow \sum(y_i - 60) = 1064$
Regression coefficient are independence of change of Origin.
$Therefore, let\ u_i$
$\rightarrow = \ x_i - 70, v_i = \ y_i - 60$
$\therefore \ \sum u_i = -35,$
$\rightarrow \sum v_i = \ -7, \sum ui^2 = 2989,$
$\rightarrow \ \sum vi^2 = 476,$

$\sum u_i v_i = 1064$
Now,
$\overline{u} = \frac{\sum u_i}{n} = \frac{-35}{7} = \ -5$
$\overline{v} = \frac{\sum v_i}{n} = \frac{-7}{7} = \ -1$
$Now, \overline{x} = \ \overline{u} + 70$
$\rightarrow : \overline{y} = \ \overline{v} + 60$

= -5 + 70	= - 1 + 60
= 65	= 59

i) The line of regression of Y on X.

$y = a + b_{yx} .x$

$b_{yx} = b_{vu} = \frac{\sum u_i v_i - n\,\overline{uv}}{\sum u_{i\,2} - n\overline{u^2}}$

$= \frac{1064 - 7\,(-5)(-1)}{2989 - 7(-5)^2}$

$\rightarrow = \frac{1064 - 35}{2989 - 175}$

$= \frac{1029}{2814} = 0.3657$

$\therefore\ b_{yx} = 0.3657\ \approx 0.37$

$a = \overline{y} - b_{yx} .\overline{x}$

= 59 – 0.37(65)

= 59 – 24.05

= 34.95

= 35

∴ line of regression of Y on X is

Y = 35 + 0.37x

Y = 0.37x + 35

ii) The line of aggression of X on Y.

$x = a' + b_{xy} .y$

$b_{xy} = b_{uv}$

$\rightarrow = \frac{\sum u_i v_i - n\,\overline{uv}}{\sum u_{i\,2} - n\overline{u^2}}$

$= \frac{1064 - 7\,(-5)(-1)}{476 - 7(-1)^2}$

$\rightarrow = \frac{1064 - 35}{476 - 7}$

$= \frac{1029}{469} = 2.19$

$\therefore\ b_{xy} = 2.19$

$a' = \overline{x} - b_{xy} .y$

= 65 – 2.19(59)

= 65 – 129.21 = - 64.21

∴ line of aggression of X on Y is

∴ x = - 64.21 + 2.19y

∴ x = 2.19y – 64.21y

iii) The correlation coefficient between X and Y.

$b_{yx} = 0.37, b_{xy} = 2.19$

$Now, r = \pm\sqrt{b_{yx} . b_{xy}}$

$\therefore r = \pm\sqrt{0.37 \times 2.19}$

$\therefore r = \pm\sqrt{0.8103}$

$\therefore r = 0.90$

4. You are given the following information about Advertising expenditure and sales:

	Advertising expenditure (₹ in lakh)(X)	Sales (₹ in lakh) (Y)
Arithmetic mean	10	90
Standard deviation	3	12

Correlation coefficient between X and Y is 0.8

i) Obtain the two regression equation

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:

$Given: \overline{x} = 10, \overline{y} = 90,$

$\rightarrow \sigma x = 3, \sigma y = 12, r = 0.8$

i) Regression equation of Y on X:

$y = a + b_{yx} . x$

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

$\therefore\ b_{yx} = 0.8 \times \frac{12}{3} = 3.2$

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

= 90 – 3.2 (10)

= 90 – 32

= 58

Now,

$y = a + b_{yx} . x$

∴ y = 58 + 3.2x

= y = 3.2x + 58

Regression equation of X on Y:
$x = a' + b_{xy} . y$
$b_{xy} = r . \frac{\sigma_x}{\sigma_y}$
$\therefore\ b_{xy} = 0.8 \times \frac{3}{12} = 0.2$
$a' = \overline{x} - b_{xy} . \overline{y}$
= 10 – 0.2 (90)
= 10 – 18
= - 8
Now,
$x = a' + b_{yx} . y$
∴ x = - 8 + 0.2y
= x = 0.2y – 8

ii) Likely sales (Y) when X =15:
Putting x = 15 in y = 58 + 3.2x we get
Y = 58 + 3.2 (15)
∴ y = 58 + 48.0 ∴ y = 106
Hence, likely sales is ₹ 106 lakh when advertising
Budget is ₹ 15 lakh.

iii) Estimation of advertising budget (X) when Y = 120:
Putting y =120 in x = - 8 + 0.2y, we get
X = - 8 + 0.2 (120)
∴ x = - 8 + 24
∴ x = 16
Hence, The advertising budget is ₹ 16 lakh to attain
Sales target of ₹ 120 lakh.

5. Bring out inconsistency if any, in the following:

$\boldsymbol{i)\ b_{yx} +\ b_{xy}}$
= 1.30 and r = 0.75

$\boldsymbol{ii)\ b_{yx} +\ b_{xy}}$
= 150 and r = 0.9

$\boldsymbol{iii)\ b_{yx} = 1.9\ and\ b_{xy} = -0.25}$

$\boldsymbol{iv)\ b_{yx} = 2.6\ and\ b_{xy} = \frac{1}{2.6}}$

Solution:

$\boldsymbol{i)\ b_{yx} + b_{xy} = 1.30\ and\ r = 0.75}$

$\frac{b_{yx} + b_{xy}}{2} =$

$\rightarrow \frac{1.30}{2} = 0.65\ and\ r = 0.75$

$\therefore \frac{b_{yx} + b_{xy}}{2} < r$

Hence, the data is inconsistent.

$\boldsymbol{ii)\ b_{yx} + b_{xy}}$

$\rightarrow \boldsymbol{= 1.50\ and\ r = 0.9}$

The Sign of r is not similar to the signs of

$b_{yx}\ and\ b_{xy}.$

Hence, the data is inconsistence

$\boldsymbol{iii)\ b_{yx} = 1.9\ and\ b_{xy} = -0.25}$

$The\ sign\ of\ b_{yx}\ and$

$\rightarrow b_{xy}\ are\ not\ similar.$

Hence, the data is inconsistence.

$\boldsymbol{iv)\ b_{yx} = 2.6\ and\ b_{xy} = \frac{1}{2.6}}$

$b_{yx} > 1\ and\ b_{xy}$

1 and both are of similar signs.

Hence, the data is inconsistence.

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

Solution:

Given:

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

$\rightarrow \sum(x - \overline{x})^2 = 136.$

$\sum(y - \overline{y})^2 = 148,$

$\rightarrow \sum(x - \overline{x})^2\ (y - \overline{y})^2$

$= 122.$

Regression coefficient of X on Y:

Equation of line of aggression of X on Y.

$x = a' + b_{xy} . y$

$b_{yx} = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sum(x-\overline{x})^2}$

$= \frac{122}{148} = 0.8243 \approx 0.82$

$a' = \overline{x} - b_{xy} . \overline{y}$

= 25 – 0.82 (18)

= 25 – 14.76

= 10.24

∴ Equation of line of regression of X on Y is

X = 10.24 + 0.82y

X = 0.82y + 10.24

7. 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:

$\overline{x} = 25, \overline{y} = 20, \sigma_x = 4,$

$\rightarrow \sigma_y = 3, r = 0.5$

Estimation of Y when X = 10:

Regression equation of Y on X us

$y = a + b_{yx} . x$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.5 \times \frac{3}{4} = \frac{3}{8}$

$a = \overline{y} - b_{yx} . \overline{x}$

$= 20 - \frac{3}{8} \times 25$

$= 20 - \frac{75}{8}$

$= \frac{160 - 75}{8} = \frac{85}{8}$

$\therefore y = \frac{85}{8} + \frac{3}{8}x$

Putting x = 10,

$$y = \frac{85}{8} + \frac{3}{8} \times 10$$
$$= \frac{85+30}{8} \frac{115}{8} = 14.375$$

Estimation of X when Y= 16:

Regression equation of X on Y is

$x = a' + b_{xy} .y$

$$b_{xy} = r . \frac{\sigma_y}{\sigma_x}$$
$$= 0.5 \times \frac{4}{3}$$
$$= \frac{2}{3}$$

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

$$= 25 - \frac{2}{3} (20)$$
$$= 25 - \frac{40}{3}$$
$$\rightarrow = \frac{75-40}{3} = \frac{35}{3}$$
$$\therefore x = \frac{35}{3} + \frac{2}{3} y$$

Putting y = 16

$$x = \frac{35}{3} + \frac{2}{3} \times 16$$
$$= \frac{35}{3} + \frac{32}{3} = \frac{35+32}{3}$$
$$= \frac{67}{3}$$
$$= 22.33$$

8. Given the following information about the production

And the demand of a commodity obtain the two regression lines:

	Production X	Demand Y
Mean	85	90
S.D.	5	6

Coefficient of correlation between X and Y is 0.6. Also, Estimate the production when demand is 100.

Solution:

$Given: \overline{x} = 85, \overline{y} = 90,$
$\rightarrow \sigma_x = 5, \sigma_y = 6, r = 0.6$
Estimation of production (X) when demand (Y) = 100.
Regression equation of X on Y.
$x = a' + b_{xy} . y$
$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$
$= 0.6 \times \frac{5}{6} = 0.5$
$\therefore b_{yx} = 0.5$
$a' = \overline{x} - b_{xy} . \overline{y}$
$\therefore a' = 85 - 0.5\ (90)$
= 85 – 45
= 40
Hence, regression equation of X on X is
X = 40 + 0.5y
X = 0.5y + 40
Put y = 100, we get
X = 0.5 (100) + 40
∴ x = 50+ 40
∴ x = 90
Hence, production (X) will be 90 when demand (Y) is 100.
Regression line of Y on Y.
$y = a + b_{yx} . x$
$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$
$0.6 \times \frac{6}{5}$
= 0.72
$a = \overline{y} - b_{yx} . (\overline{x})$
∴ a = 90 – 0.72 (85)
= 90 – 61.2
= 28.8
Hence, the regression line of Y on X is
Y = 28.8 + 0.72x
Y = 0.72x + 28.8

9. Given the following data, obtain linear regression

Estimate of X for Y = 10
$\boldsymbol{\overline{x} = 7.6, \overline{y} = 14.8,}$
$\boldsymbol{\rightarrow \sigma_x = 3.2,}$
$\boldsymbol{\rightarrow \sigma_y = 16, r = 0.7}$

Solution:

$Given: \overline{x} = 7.6, \overline{y} = 14.8,$

$\rightarrow \sigma_x = 3.2,$

$\rightarrow \sigma_y = 16, r = 0.7$

Linear regression estimate of X for Y = 10

Regression equation of X on Y:

$x = a' + b_{yx} . y$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.7 \times \frac{3.2}{16}$

0.7 × 0.2 = 0.14

$\therefore \ b_{xy} = 0.14$

$a' = \overline{x} - b_{xy} . \overline{y}$

∴ a = 7.6 – 0.14(14.8)

= 7.6 – 2.072

= 5.528

Hence, regression equation of X on Y is

X = 5.528 + 0.14y

X = 0.14y + 5.528

Putting y =10, we get

X = 0.14(10) + 5.528

= 1.4 + 5.528 = 6.928

∴ x = 6.928

Hence, linear regression estimate of x is 6.928 for y = 10.

10. 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:

$\boldsymbol{\sum x = 8500, \sum y = 9600,}$

$\boldsymbol{\rightarrow \sigma_x = 60, \sigma_y = 20, r = 0.6}$

Estimate the expenditure on food and entertainment
When expenditure on accommodation is ₹ 200.

Solution:

Here, X = Expenditure on accommodation (in ₹)

Y = expenditure on food and entertainment (in ₹)

Given:

$\sum x = 8500, \sum y = 9600,$

$\rightarrow \sigma_x = 60, \sigma_y = 20, r = 0.6$

N = 50

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

$\rightarrow = \frac{8500}{50} =$ ₹ 170

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

Estimation of y when x = 200:

Regression equation of Y on X:

$y = a + b_{yx} . x$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.6 \times \frac{20}{60} = 0.2$

$\therefore\ b_{yx} = 0.2$

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

= 192 – 0.2 (170)

= 192 – 34

= 158

∴ a = 158

Hence, Regression equation of Y on X is

Y = 158 + 02x

Y = 0.2x + 158

Putting x = 200, we get

Y = 0.2 × 200 + 158

= 40 + 158

= 198

∴ y = ₹ 198

Hence, the estimated expenditure on food and entertainments

Will be ₹ 198 when expenditure on accommodation is ₹ 200.

11. The following data about the sales and advertisement expenditure of a firm is given below (in ₹ crores).

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

Correlation of coefficient between sales and expenditure = 0.9

i) estimate the likely sales for a proposed advertisement expenditure of ₹ 10 crores.

ii) What should be the advertisement expenditure if the firm propose a sale target ₹ 60 crores.

Solution:

Here, x = sales, Y = advertisement expenditure.

$Given: \overline{x} = 40, \overline{y} = 6,$

$\rightarrow \sigma_x = 10, \sigma_y = 1.5, r = 0.9$

i) Estimation of likely sales (X) for Y = 10:

Regression equation of X on Y:

$x = a' + b_{xy} . y$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.9 \times \frac{10}{1.5}$

= 6

$\therefore\ b_{xy} = 6$

$a' = \overline{x} - b_{xy} . \overline{y}$

= 40 – 6(6)

= 40 – 36 = 4

∴ a =4

Hence, the regression equation of X on Y is

X = 4 + 6y

X = 6y + 4

Putting y =10 we get

X = 6 × 10 + 4

∴ x = 60 + 4

∴ x = 64

Hence, the estimated sales is ₹ 64 crores when advertisement Expenditure (Y) is ₹ 10 crores.

12. For a certain bivariate data the following information are available.

	X	Y
A.M.	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:

$\overline{x} = 13, \overline{y} = 17,$

→ $\sigma_x = 3, \sigma_y = 2, r = 0.6$

Estimation of x when y = 15:

Regression equation of x on Y is

$x = a' + b_{xy} . y$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.6 \times \frac{3}{2} = 0.9$

$\therefore \; b_{xy} = 0.9$

$a' = \overline{x} - b_{xy} . \overline{y}$

= 13 – 0.9 (17)

= 13 – 15.3

$\therefore a' = \; -2.3$

Hence, the regression equation of X on Y is

X = - 2.3 + 0.9y

X = 0.9y – 2.3

Putting y = 15, we get

X =0.9 × 15 – 2.3

= 13.5 – 2.3 = 11.2

∴ x = 11.2

Hence, the estimate value of x is 11.2, when y =15.

Estimation of Y when X = 10:

Regression equation of Y on X is:

$y = a + \; b_{yx} . x$

$Now, b_{yx} = r . \frac{\sigma_y}{\sigma_x}$

$= 0.6 \times \frac{2}{3} = 0.4$

$\therefore \; b_{yx} = 0.4$

$a = \; \overline{y} - \; b_{yx} . \overline{x}$

= 17 – 0.4 (13)

= 17 – 5.2 = 11.8

∴ a =11.8

Hence, the aggression equation of Y on X is

Y =11.8 + 0.4x

Y = 0.4x + 11.8

Putting x = 10, we get

Y = 0.4(10) + 11.8

= 4 + 11.8 = 15.8

∴ y = 15.8

Hence, the estimated value of y is 15.8, when x =10.

Exercise 3.3

1. from the two regression equation.
$Find\ r, \overline{x}\ and\ \overline{y}$
4y = 9x + 15 and 25x = 4y + 17

Solution:
To find r:
Let 4y = 9x + 15 be the equation of line regression of Y on X.
$\therefore y = \frac{9}{4}x + \frac{15}{4}$
$\therefore\ b_{yx} = \frac{9}{4}$
Then the other equation 25x = 4y + 17 is the equation
Of line regression of X on Y.
$\therefore x = \frac{4}{25}y + \frac{17}{25}$
$\therefore\ b_{yx} = \frac{4}{25}$
$Now, r^2 = b_{yx}.b_{xy}$
$\therefore r = \pm\sqrt{b_{yx}\,.\,b_{xy}}$
$\rightarrow = \pm\sqrt{\frac{9}{4}\times\frac{4}{25}}$
$= \pm\sqrt{\frac{9}{25}}$
$\rightarrow = \pm\sqrt{0.36} = \pm 0.6$
$\therefore$ r = 0.6
$To\ find\ \overline{x}, \overline{y}$:
$9x - 4y + 15 = 0\ \ \ldots.(1)$
$25x - 4y - 17 = 0\ \ \ldots.(2)$
Subtracting equation (2) from equation (1)

$$\begin{array}{l} 9x - 4y + 15 = 0 \\ 25x - 4y - 17 = 0 \\ -\quad\ +\quad\ + \\ \hline -16x \qquad\quad + 32 = 0 \end{array}$$

$\therefore$ 32 = 16x
$\therefore\ x = \frac{32}{16} = 2$

$\therefore\ \overline{x} = 2$

Putting x = 2 in the equation (1),
9(2) – 4y + 15 = 0
∴ 18 – 4y + 15 = 0
∴ 33 = 4y

$\therefore y = \frac{33}{4} = 8.25$

$\therefore\ \overline{y} = 8.25$

Hence,
$r = 0.6, \overline{x} = 2, \overline{y} = 8.25$

2. In a partially destroyed laboratory record of an Analysis of regression data, the following data are
Legible: Regression equation: 8x – 10y + 66 = 0 and 40x – 18y = 214
Find on the basic 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: 8x – 10y + 66 = 0, 40x – 18y = 214,
$\sigma_x^2 = 9 \quad \therefore\ \sigma_x = 3$

i) The mean values of X and Y:

$8x - 10y = -66 \quad \ldots(1)$

$40x - 18y = 214 \quad \ldots.(2)$

$- \quad + \quad -$

$- \quad 32y = -544$

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

Put y = 12 in equation (1)
∴ 8x – 10(17) = -66
∴ 8x = - 66+ 170
∴ 8x = 104

$\therefore x = \frac{104}{8} = 13$

$Hence, \overline{x} = 13\ and\ \overline{y} = 17.$

ii) Correlation coefficient between X and Y.

Let 8x – 10y + 66 = 0 be the regression equation of Y on X.

$\therefore$ 10y = 8x + 66

$\therefore\ y = \frac{8}{10}x + \frac{66}{10}$

$\therefore\ b_{yx} = \frac{8}{10}$

And the other equation 40x – 18y = 214 be the regression

Equation of X and Y.

$\therefore$ 40x = 18y + 214

$\therefore x = \frac{18}{40}y + \frac{214}{40}$

$\therefore\ b_{xy} = \frac{18}{40}$

Now,

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

$= \pm\sqrt{\frac{8}{10} \times \frac{18}{40}} =$

$\rightarrow \pm\sqrt{\frac{9}{25}} = \quad \pm\ \frac{3}{5}$

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

iii) Standard deviation of Y.

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

$\therefore\ \frac{8}{10} = \frac{3}{5} \times \frac{\sigma_y}{3}$

$\therefore\ \frac{8 \times 5 \times 3}{10 \times 3} = \sigma_y$

$\therefore\ \sigma_y = 4$

3. For 50 students of a class, the regression 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 variances of marks in

statistics is $\left(\frac{9}{16}\right)^{th}$

of the variences of marks in accountancy

Find the mean marks in statistics and the correlation

Coefficient between marks in the two subjects.

Solution:

Let X, marks in statistics, Y = marks in accountancy
Given: Regression equation of X on Y is
3y – 5x + 180 = 0
∴ 5x = 3y + 180

$\therefore x = \frac{3}{5}y + 36$

$\therefore \ b_{xy} = \frac{3}{5}$

$\overline{y} = 44, \sigma_x^2 = \frac{9}{16}\ \sigma_x^2$

$\rightarrow \therefore \ \sigma_x = \frac{3}{4}\sigma_y$

Mean marks in statistics:
$\overline{y} = 44$
∴ Put y = 44 in 3y – 5x + 180 = 0
∴ 3(44) – 5x + 180 = 0
∴ 5x = 180 + 132

$x = \frac{312}{5} = 62.4$

$\overline{x} = 62.4\, marks.$

Correlation coefficient between X and Y.

$we\ have, b_{xy} = \frac{3}{5}$

$Now, b_{xy} = r.\frac{\sigma_x}{\sigma_y}$

$\rightarrow \therefore \ \frac{3}{5} = r.\frac{\frac{3}{4}\sigma_y}{\sigma_y}$

$\rightarrow \therefore \ \frac{3}{5} = r\ .\ \ \frac{3}{4}$

$\therefore \ r = \frac{3}{5} \times \frac{4}{3} = \frac{4}{5} = 0.8$

4. For a 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 variances of Y if variance of X is 9.

Solution:

$Given: b_{yx} = 0.4, b_{xy} = 0.9,$

$\rightarrow r\ = ?\,, \sigma_x^2 = 9, \sigma_y^2 = ?$

$r = \ \pm\sqrt{b_{yx}\,.\,b_{xy}}\ =$

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

= 0.6
Variance of Y:

$Now, b_{yx} = 0.4$

$\therefore r\,.\dfrac{\sigma_y}{\sigma_x} = 0.4$

$\therefore 0.6 \times \dfrac{\sigma_y}{3} = 0.4$

$\therefore 0.2\,\sigma_y = 0.4$

$\therefore\ \sigma_y = \dfrac{0.4}{0.2} = 2$

$\therefore\ \sigma_y^2 = (2)^2 = 4$

Hence, the variance of Y is 4.

5. The equation of two regression lines are 2x + 3y – 6 = 0 and 2x + 2y – 12 = 0. Find: (i) Correlation coefficient

$(\boldsymbol{ii})\ \dfrac{\boldsymbol{\sigma_x}}{\boldsymbol{\sigma_y}}$

Solution:

Let the equation 2x + 3y – 6 = 0 be the regression
Equation of Y on X.

∴ 3y = -2x + 6

$\therefore y = -\dfrac{2}{3}x + 6$

$\therefore\ b_{yx} = -\dfrac{2}{3}$

Another equation 2x + 2y -12 = 0 be the regression
Equation of X on Y.

∴ 2x = - 2y + 12

$\therefore = -\dfrac{2}{2}y + \dfrac{12}{2}$

$\rightarrow \therefore x = -y + 6$

$\therefore\ b_{xy} - 1$

i) Correlation coefficient:

$r = \pm\sqrt{b_{yx}\,.b_{xy}}$

$\rightarrow = \pm\sqrt{\left(-\dfrac{2}{3}\right) \times (-1)}$

$= \pm\sqrt{\dfrac{2}{3}}$

$= \pm\ 0.82$

$\therefore r = -0.82$

$ii)\ \frac{\sigma_x}{\sigma_y}:$

$We\ know, b_{xy} = r\ .\frac{\sigma_x}{\sigma_y}$

$\therefore\ -1 = -0.82 \times \frac{\sigma_x}{\sigma_y}$

$\therefore (-1) \times \left(\frac{1}{-0.82}\right) = \frac{\sigma_x}{\sigma_y}$

$\therefore\ \frac{\sigma_x}{\sigma_y} = \frac{1}{0.82}$

$\therefore\ \frac{\sigma_x}{\sigma_y} = 1.22$

6. for a bivariate data:

$\boldsymbol{\bar{x} = 53, \bar{y} = 28,}$

$\rightarrow \boldsymbol{b_{yx} = -1.5, and\ b_{xy} = -0.2}$

Estimate Y when X = 50.

Solution:

$\bar{x} = 53, \bar{y} = 28,$

$\rightarrow b_{yx} = -1.5, and\ b_{xy} = -0.2$

Using regression equation of Y on X, we estimate Y

When X = 50.

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

∴ y – 28 = - 1.5 (x – 53)

∴ y = - 15x + 79.5 + 28

∴ y = - 1.5x + 107.5

Estimate of Y when X = 50:

Put x =50 in Y = -1.5x + 107.5

∴ y = - 1.5 × 50 + 107.5

∴ y = - 75 + 107.5

∴ y = 32.5

7. The equation 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: Two regression lines: x – 4y = 5, 16y – x = 64

Means of X and Y:

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

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

$= 5..75$

Put $y = 5.75$ in $x - 4y = 5$

$\therefore x - 4(5.75) = 5$

$\therefore x = 5 + 23$

$\therefore x = 28$

Hence,

$\bar{x} = 28, \ \bar{y} = 5.75$

Correlation of coefficient

$X - 4y = 5$

$\therefore x = 4y + 5$

$\therefore \ b_{xy} = 4$

$16y - x = 64$

$\therefore 16y = x + 64$

$\therefore y = \frac{1}{16}x + 4$

$\therefore \ b_{yx} = \frac{1}{16}$

$Now, r = \pm\sqrt{b_{yx}.\ b_{xy}}$

$\therefore r = \pm\sqrt{\frac{1}{16} \times 4}$

$\therefore r = \frac{1}{2} = 0.5$

8. 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) Coefficient of correlation between X and Y.

iii) Standard deviation of Y if

$\boldsymbol{\sigma_x = 5}$

Solution:

Given: Var(x)=25
Regression equation of Y on X: 5y – x =22
Regression equation of X ON y: 64x – 45y = 22

(i) Mean Values of X and Y:
5y –x =22 = - x + 5y = 22 ……(1)
64x – 45y = 22 …. (2)
Form equation (1) x = 5y -22
Putting x = 5y – 22 in equation (2), we get
64 (5y -22) – 45y = 22
∴ 320y – 1480 – 45y = 22
∴ 275y = = 22 + 1408
∴ 275y = 1430

$$\therefore\ y\ =\ \frac{1430}{275}$$

$$\rightarrow \therefore y = 5.2$$

$$\therefore\ \overline{y} = 5.2$$

Putting y = 5.2 in x = 5y – 22, we get
x = 5(5.2) – 22
∴ x = 26 – 22
∴ x = 4

$$\therefore \overline{x}\ = 4$$

$$Hence, \overline{x}\ = 4,\ \overline{y}\ = 5.2$$

ii) Regression equation of Y on X is 5y – x = 22
∴ 5y = x +22

$$\therefore y = \frac{x}{5} + \frac{22}{5}$$

$$\therefore\ b_{yx} = coefficient\ of\ x$$

$$\rightarrow = \frac{1}{5}$$

Regression equation of x on y is 64x – 45y = 22
∴ 64x = 45y + 22

$$\therefore x\ = \frac{45y}{64} + \frac{22}{64}$$

$$\therefore\ b_{yx} = coefficient\ of\ y\ \ = \frac{45}{64}$$

$$Now, r\ = \pm\sqrt{b_{yx}.\ b_{xy}}$$

$$=\ \pm\ \sqrt{\frac{1}{5} \times \frac{45}{64}} = \pm\sqrt{\frac{9}{64}}$$

$$=\ \pm\ \frac{3}{8}$$

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

$\boldsymbol{iii)\ b_{yx} = \frac{1}{5}},$

$\rightarrow\ r = \frac{3}{8}, \sigma_x = 5$

$Now, b_{yx} = r\,.\frac{\sigma_y}{\sigma_x}$

$\therefore\ \frac{1}{5} = \frac{3}{8}\,.\frac{\sigma_y}{\sigma_x}$

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

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

$\therefore\ \sigma_y = \frac{5 \times 8}{5 \times 3} = \frac{8}{3}$

Hence, standard deviation

$\rightarrow of\ Y\ is\ \frac{8}{3}$

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

$\boldsymbol{i)\ \overline{x}\ , ii\ )\ \overline{y}, iii)\ b_{yx}}\ ,$

$\rightarrow \boldsymbol{iv)\ b_{xy}},$

$\rightarrow \boldsymbol{v)\ r\ \ [Given: \sqrt{0.375} = 0.61]}$

Solution:

Given: regression equation of X on Y: 2x = y + 15
Regression equation of Y on X: 4y = 3x + 25

i) and ii)

2x = y + 15
∴ 2x – y = 15…. (1)
4y = 3y + 25
∴ - 3x + 4y = 25… (2)
From, equation (1), y = 2x – 15
Putting y = 2x – 15 in equation (2), we get
- 3x + 4 (2x - 15) = 25
∴ - 3x + 8x – 60 = 25
∴ 5x = 25 + 60

$\therefore x = \frac{85}{5} = 17$

$\therefore \bar{x} = 17$

Putting x = 17 in y = 2x – 15, we get

Y = 2(17) – 15

= 34 – 15 = 19

$\therefore \bar{y} = 19$

Hence,

$\bar{x} = 17, \bar{y} = 19$

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

$\therefore y = \frac{3}{4}x + \frac{25}{4}$

$\therefore b_{yx} = coefficient\ of\ x$

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

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

$\therefore x = \frac{1}{2}y + \frac{15}{2}$

$\therefore b_{xy}\ corfficient\ of\ y$

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

$v)\ r = \pm\sqrt{b_{yx}.\ b_{xy}}$

$Putting, b_{xy} = \frac{3}{4}$

$\rightarrow and\ b_{xy} = \frac{1}{2}$

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

$\therefore r = \pm\sqrt{\frac{3}{8}}$

$\therefore r = \pm\sqrt{0.375}$

∴ r = 0.61

10. The two regression equation are 5x – 6y + 90 = 0 and 15x -8y – 130 = 0
$\boldsymbol{Find\ \bar{x}, \bar{y}, r.}$

Solution:

Given: Two regression equation

5x – 6y + 90 = 0.... (1)
15x -8y – 130 = 0.... (2)
Multiply equation (1) by 3 and subtracting equation
From it, we get

$$\begin{array}{l} 5x - 6y + 90 = 0 \\ 15x - 8y - 130 = 0 \\ - \quad + \quad + \\ \hline \quad\quad -10y + 400 = 0 \end{array}$$

$\therefore$ 400 = 10y

$\therefore y = \frac{400}{10} = 40$

→ $\therefore \bar{y} = 40$

Putting y = 40 in equation (1) we get,
5x – 6(40) +90= 0
$\therefore$ 5x – 240 + 90 =0
$\therefore$ 5x – 150 = 0
$\therefore$ 5x = 150

$\therefore x = \frac{150}{5} = 30$

→ $\therefore \bar{x} = 30$

Let the equation of Y on X be 5x – 6y + 90 = 0
$\therefore$ 6y = 5x + 90

$\therefore y = \frac{5}{6}x + 15$

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

And the equation of aggression of X on Y be
15x – 8y – 130 = 0
$\therefore$ 15x = 8y + 130

$\therefore x = \frac{8}{15}y + \frac{130}{15}$

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

$Now, r = \pm\sqrt{b_{yx} . b_{xy}}$

$= \pm\sqrt{\frac{5}{6} \times \frac{8}{15}}$

$= \pm\sqrt{\frac{8}{18}} = +\sqrt{\frac{4}{9}}$

$= \pm \frac{2}{3}$

$\therefore r = \frac{2}{3}$

11. Two lines of aggression are 10x + 3y – 62 = 0 and 6x + 5y – 50 = 0. Identify the aggression of x on Y Hence.

$\boldsymbol{\bar{x}, \bar{y}, and\ r}.$

Solution:

Given: Two lines of aggression

10x + 3y – 62 = 0; 6x + 5y – 50 = 0

∴ 10x = - 3y + 62

$\therefore x = -\frac{3}{10}y + 6.2$

$\therefore\ b_{xy} = -\frac{3}{10}$

And the regression equation of Y on X be

6x + 5y – 50 = 0

∴ 5y = - 6x + 50

$\therefore\ y = -\frac{6}{5}x + 10$

$\therefore\ b_{yx} = -\frac{6}{5}$

$Now, b_{xy}\ .\ b_{yx}$

$\rightarrow = \left(-\frac{3}{10}\right) \times \left(-\frac{6}{5}\right)$

$\rightarrow = \frac{18}{50}\ which\ is\ less\ than\ 1.$

Hence, assumption regarding regression equation holds

True,

∴ Regression equation of X on Y is 10x + 3y - 62 = 0

To find

$\bar{x}, \bar{y}$

10x + 3y - 62 = 0

6x + 5y – 50 = 0

Multiply equation (1) by 5 and equation (2) by 3.

Then subtracting equation (2) from equation (1), we get

$$\begin{array}{l} 50x + 15y - 310 = 0 \\ 18x + 15y - 150 = 0 \\ -\quad\ -\qquad\ + \\ \hline 32x \qquad\quad - 160 = 0 \end{array}$$

∴ 32x = 160

$\therefore x = \frac{160}{32} = 5$

$\therefore \bar{x} = 5$

Putting x = 5 in equation (1), we get

$10 \times 5 + 3y - 62 = 0$

∴ 50 + 3y – 62 = 0

∴ 3y – 12 = 0

∴ 3y = 12

$\therefore y = \frac{12}{3} = 4$

$\therefore \bar{y} = 4$

$Hence, \bar{x} = 5\ \bar{y} = 4$

To find r:

$b_{xy} = -\frac{3}{10}, b_{yx} = -\frac{6}{5}$

$Now, r = \pm\sqrt{b_{yx} . b_{xy}}$

$= \pm\sqrt{\left(-\frac{3}{10}\right)\left(-\frac{6}{5}\right)}$

$= \pm\sqrt{\frac{18}{50}} = \pm\sqrt{0.36}$

∴ r = - 0.6

12. for certain x and y series, which are correlated the Two lines of aggression are 10y = 3x + 170 and 5x + 70 = 6y. Find the correlation coefficient Between them. Find the mean value of x and y.

Solution:

Given: Two lines of regression

10y = 3x + 170; 5x + 70 = 6y

Correlation coefficient:

Let the regression equation of Y on X be

10y = 3x + 170

$\therefore y = \frac{3}{10}x + 17$

$\therefore b_{yx} = \frac{3}{10}$

And the regression equation of X on Y be

5x + 70 = 6y

∴ 5x 6y – 70

$\therefore x = \frac{6}{5}y - 14$

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

$Now, r = \pm\sqrt{b_{yx} . b_{xy}}$

$= \pm\sqrt{\frac{3}{10} \times \frac{6}{5}}$

$= \pm\sqrt{\frac{18}{50}}$

$= \pm\sqrt{0.36}$

$\therefore$ r = 0.6

Mean values of X and Y:

10y = 3x + 170

$\therefore$ 3x – 10y + 170 = 0 .... (1)

5x + 70 = 6y

$\therefore$ 5x – 6y + 70 = 0 ... (2)

Multiply equation (1) by 5 and multiply equation (2) by 3

Then subtracting equation (2) from equation (1), we get

$$\begin{array}{r} 15x - 50y + 850 = 0 \\ 15x - 18y + 210 = 0 \\ - \quad + \quad - \\ \hline -32y + 640 = 0 \end{array}$$

$\therefore$ 640 = 32y

$\therefore y = \frac{640}{32} = 20$

$\therefore\ \bar{y} = 20$

Putting y =20 in equation (1)we get

3x – 10 (20)+ 170 = 0

$\therefore$ 3x – 200 + 170 = 0

$\therefore$ 3x – 30 = 0 $\quad\therefore$ 3x = 30

$\therefore x = \frac{30}{3} = 10$

$\rightarrow \therefore\ \bar{x} = 10$

$Hence, \bar{x} = 10, \bar{y} = 20$

13. Regression equation of two series are 2x – y – 15 = 0 and 3x – 4y + 25 = 0.
$Find\ \bar{x}, \bar{y},$

and regression coefficient, also find
Coefficient of correlation.

$\left[Given: \sqrt{0.375 = } \ \ 0.61\right]$

Solution:

Given: regression equation 2x – y – 15 = 0 and 3x – 4y + 25 = 0.

$To\ Find\ \overline{x}, \overline{y},$

2x – y – 15 = 0

3x – 4y + 25 = 0

From equation (1) y = 2x – 15

Putting y = 2x – 15 in the equation (2), we get

3x – 4 (2x - 15) + 25 = 0

∴ 3x – 8x + 60 + 25 = 0

∴ - 5x + 85 = 0

∴ 85 = 5x

$\therefore x = \frac{85}{5} = 17$

$\therefore\ \overline{x} = 17$

Now, Putting x = 17 in y = 2x – 15, we get

Y = 2(17) - 15

= 34 – 15 = 19

$\therefore\ \overline{y} = 19$

$Hence, \overline{x} = 17, \overline{y} = 19$

Regression coefficient:

Let the regression equation of X on T be.

2x – y – 15 = 0

∴ 2x = y + 15

$\therefore x = \frac{1}{2}y + \frac{15}{2}$

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

And the regression equation of Y on X be

3x – 4y + 25 = 0

∴ 4y = 3x + 25

$\therefore y = \frac{3}{4}x + \frac{25}{4}$

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

$Now, b_{yx}\ .b_{xy}$

$\rightarrow = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$

Which is less han 1. Hence the

Assumption regarding regression equation holds true.

$b_{xy} = \frac{1}{2}\ and\ b_{yx} = \frac{3}{4}$

Coefficient correlation:

$b_{xy} = \frac{1}{2}, b_{yx} = \frac{3}{4}$

$Now, r = \pm\sqrt{b_{yx} . b_{xy}}$

$= \pm\sqrt{\frac{3}{4} \times \frac{1}{2}} = \pm\sqrt{\frac{3}{8}}$

$= \pm 0.375$

∴ r = 0.61

14. 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 mean height and weight of the group, also,
Estimate weight of a girls whose height is 70 inches.

Solution:

Given: 4y – 15x + 500 = 0, 20x – 3y - 900 = 0

Mean height (x) and mean weight (y):

-15x + 4y + 500 = 0 .... (1)

20x – 3y – 900 = 0 ...(2)

Multiplying equation (1) by 3 and equation (2) by 4 and then adding them,

-45x + 12y + 1500 = 0

80x – 12y – 3600 = 0

∴ 35x -2100 = 0

∴ 35x = 2100

$\therefore \bar{x} = \frac{2100}{35} = 60$

Put x =60 in equation (1), we get

4y -15 (60) + 500 = 0

∴ 4y – 900 + 500 = 0

∴ 4y = 400

$\therefore \bar{y} = \frac{400}{40} = 100$

Hence,

mean height $\bar{x} =$

60 inches and mean weight

$\bar{y} = 100kg$

Estimate of weight (y) of a girl when x = 70:

We use the regression equation of Y on X

From the given equation let 4y – 15x + 500 = 0 be the

Regression equation of Y on X.

∴ 4y = 15x – 500

$$\therefore y = \frac{15}{4}x - \frac{500}{4}$$

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

Now, regression equation of X on Y is

20x – 3y – 900 = 0

∴ 20x = 3y + 900

$$\therefore x = \frac{3}{20}y + 45$$

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

Now,

$$b_{yx}, b_{xy} = \frac{15}{4} \times \frac{3}{20} = \frac{45}{80}$$

which is less than 1. hence

Assumption regarding regression equation is true

$$y = \frac{15}{4}x - \frac{500}{4}$$

Put x = 70 in the equation

$$\therefore y = \frac{15}{4}(70) - \frac{500}{4}$$

$$\therefore y = \frac{1050 - 500}{4}$$

$$\rightarrow = \frac{550}{4} = 137.5$$

Hence, the estimate weight of a girl is 137.5 kg