Q. 2 a In the estimation of regression equations in respect of two correlated variables X and Y the following results were obtained.

$\bar{X} = 90$, $\bar{Y} = 70$, $n = 10$, $\Sigma x^2 = 6360$

$\Sigma y^2 = 2860$, $\Sigma xy = 3900$, where

$x = X - \bar{X}$, $y = Y - \bar{Y}$. i.e x and y are the deviations of X and Y from their respective means. Obtain the two regression equations.

Sol^n We know that

the regression coefficient of Y on X is

$$b_{yx} = \frac{cov(x,y)}{\sigma_x^2}$$

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

$$= \frac{\Sigma xy}{\Sigma x^2} = \frac{3900}{6360} = .6132$$

IIly the regression coefficient of X on Y is

$$b_{xy} = \frac{cov(x,y)}{\sigma_y^2} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(Y-\bar{Y})^2}$$

$$= \frac{\Sigma xy}{\Sigma y^2} = \frac{3900}{2860} = 1.3636$$

The regression equation of y on x is

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

$$\Rightarrow y - 70 = .6132(x - 90)$$

$$\Rightarrow y = .6132x - 55.188 + 70$$

$$\Rightarrow y = .6132x + 14.812$$

The regression equation of x on y is.

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

$$\Rightarrow x - 90 = 1.3636(y - 70)$$

$$\Rightarrow x = 1.3636y - 5.452.$$

Q Given the bivariate data :

x :	2	4	5	6	8	11
y :	18	12	10	8	7	5

(i) Fit the regression line of y on x and estimate y when x = 10.

(ii) Fit the regression line of x on y and estimate x if y = 8.5

(iii) Calculate Karl Pearson's coefficient of correlation.

(iv) Interpret the regression coefficient.

Sol^n

x	$(x-\bar{x})$	$(x-\bar{x})^2$	y	$(y-\bar{y})$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
2	-4	16	18	8	64	-32
4	-2	4	12	2	4	-4
5	-1	1	10	0	0	0
6	0	0	8	-2	4	0
8	2	4	7	-3	9	-6
11	5	25	5	-5	25	-25
		50			106	-67

$$\bar{x} = \frac{\Sigma x}{n} = \frac{36}{6} = 6$$

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

now

The regression coefficient of y on x is

$$b_{yx} = \frac{cov(x,y)}{\sigma_x^2}$$

$$= \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

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

$$= -1.34$$

||ly regression coefficient of x on y is

$$b_{xy} = \frac{cov(x,y)}{\sigma_y^2}$$

$$= \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(y-\bar{y})^2} = \frac{-67}{106}$$

$$= -0.63$$

(i) Now the regression line of y on x is

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

$$\Rightarrow y - 10 = -1.34(x - 6)$$

$$\Rightarrow y = 10 - 1.34x + 8.04$$

$$\Rightarrow y = -1.34x + 18.04$$

when $x = 10$, the value of y is

$$y = -1.34(10) + 18.04$$

$$= 4.64$$

(ii) the regression line of x on y is

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

$$\Rightarrow x - 6 = -.63(y - 10)$$

$$\Rightarrow x = -.63y + 12.3$$

when $y = 8.5$ then

$$x = -.63(8.5) + 12.3$$

$$= 6.945$$

(iii) We know that the karl pearson correlation coefficient is

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

$$= \pm\sqrt{(-1.34)(-0.63)}$$

$$= -.92$$

(iv) $\because$ $b_{yx} = -1.34$, which shows that when the independent variable increases by 1 unit, the dependent variable y decreases by 1.34 units.

IIly when $b_{xy} = -.63$, it means that corresponding to one unit increase in y there is 0.63 unit decrease in x.

---

in y there is 0.63 unit decrease in x.

Example 5. The following data are given :

		x	y
Arithmetic mean	:	*36*	*85*
Standard deviation	:	*11*	*8*

Correlation coefficient between x and y = 0.66

(i) Find the two regression equations.

(ii) Estimate the value of x when y = 75.

Solution : Given $\bar{x} = 36$, $\bar{y} = 85$, $\sigma_x = 11$, $\sigma_y = 8$, r or $r_{xy} = 0.66$

(i) The regression equation of y or x is:

$$y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x}) \qquad \left[\because\ b_{yx} = r\frac{\sigma_y}{\sigma_x}\right]$$

$$\Rightarrow\ y - 85 = 0.66 \times \frac{8}{11}(x - 36)$$

$$\Rightarrow\ y = 0.48(x - 36) + 85$$

$$\Rightarrow \quad y = 0.48\,x - 17.28 + 85$$
$$\Rightarrow \quad \mathbf{y = 0.48\,x + 67.72}$$

Again, the regression equation of x on y is :

$$x - \bar{x} = r.\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$\Rightarrow \quad x - 36 = 0.66 \times \frac{11}{8}(y - 85)$$
$$\Rightarrow \quad x - 36 = 0.91\,(y - 85)$$
$$\Rightarrow \quad x = 0.91\,y - 77.35 + 36$$
$$\Rightarrow \quad \mathbf{x = 0.91\,y - 41.35}$$

(*ii*) To estimate the value of x corresponding to a given value of y we must use the regression equation of x or y.

$\therefore$ When $\quad y = 75, \; x = 0.91 \times 75 - 41.35$

$\quad = 68.25 - 41.35 = \mathbf{26.9.}$

Example 6. In trying to evaluate the effectiveness in its advertising campaign a firm com-

$\Rightarrow X = 1.3030\,Y - 3.432$

Example 9. *Given that the variance of $X = 9$. The regression equations are :*

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

Find the following :

(i) *Average values of X and Y,*

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

(iii) *Standard deviation of Y.*

Solution : (i) The regression equations are

$$8X - 10Y = -66 \qquad \text{...(1)}$$

$$40X - 18Y = 214 \qquad \text{...(2)}$$

$\because (\bar{X}, \bar{Y})$ is the point of intersection of (1) and (2), hence this point must satisfy (1) and (2) together. Thus

$$8\bar{X} - 10\bar{Y} = -66 \qquad \text{...(3)}$$

$$40\bar{X} - 18\bar{Y} = 214 \qquad \text{...(4)}$$

$(3) \times 5 \Rightarrow 40\bar{X} - 50\bar{Y} = -330$

Again from (4) $\quad 40\bar{X} - 18\bar{Y} = 214$

Substracting $\quad -32\bar{Y} = -544$

$\Rightarrow \bar{Y} = 17$

Putting the value of $\bar{Y}$ in (3) we get

$$\bar{X} = 13$$

(*ii*) We observe that (1) is the regression equation of Y on X and (2) is the regression equation of X on Y [See Note (*ii*) below.]

$$(1) \Rightarrow \; Y = \frac{4}{5}X + \frac{33}{5}, \quad \therefore b_{YX} = \frac{4}{5}$$

$$(2) \Rightarrow \; X = \frac{9}{20}Y + \frac{107}{20}, \quad \therefore b_{XY} = \frac{9}{20}$$

Since the regression coefficients are positive, r will also be positive.

$$\therefore \qquad r = +\sqrt{b_{YX} \times b_{XY}} = \sqrt{\frac{4}{5} \times \frac{9}{20}} = \frac{3}{5} = \mathbf{0.6}$$

(*iii*) Given, variance of $X = \sigma_X^2 = 9$

$$\therefore \qquad \sigma_X = 3$$

We know that,

$$b_{YX} = r \cdot \frac{\sigma_Y}{\sigma_X} \Rightarrow \frac{4}{5} = \frac{3}{5} \times \frac{\sigma_Y}{3} \Rightarrow \sigma_Y = \mathbf{4}$$

The required yield when rainfall is 2 inches is 317.960 kg.

Example 11. The following data are given regarding expenditures on advertising and sales of a particular farm:

	Advertisement Expenditure (X) (₹ in lakh)	*Sales (Y)* (₹ in lakh)
Mean	*10*	*90*
S.D.	*3*	*12*

Correlation coefficient $r_{XY} = 0.8$.

(i) Obtain the regression equation of Y on X.

(ii) Estimate the advertisement expenditure required to attain a sales target of ₹ 120 lakhs.

Solution: (*i*) *The regression equation of Y on X is given by :*

$$Y - \bar{Y} = b_{YX}(X - \bar{X}) \quad \text{...(1)}$$

where
$$b_{YX} = r \cdot \frac{\sigma_Y}{\sigma_X} = 0.8 \times \frac{12}{3} = 3.2$$

$$\bar{X} = 10, \ \bar{Y} = 90$$

$\therefore$ (1) $\Rightarrow$ $Y - 90 = 3.2(X - 10)$

$\Rightarrow$ $Y - 90 = 3.2X - 32$

$\Rightarrow$ $Y = 3.2X - 32 + 90$

$\Rightarrow$ $Y = 3.2X + 58$, which is the regression equation of Y on X.

(*ii*) *To estimate advertisement expenditure for a given sales target, we have to find the regression equation of X on Y.* The regression equation of X on Y is :

$$X - \bar{X} = r \cdot \frac{\sigma_X}{\sigma_Y}(Y - \bar{Y})$$

$\Rightarrow$ $X - 10 = 0.8 \times \frac{3}{12}(Y - 90)$

$\Rightarrow$ $X - 10 = 0.2(Y - 90)$

$\Rightarrow$ $X - 10 = 0.2Y - 18$

$\Rightarrow$ $X = 0.2Y - 8$

When $Y = 120$, $X = 0.2 \times 120 - 8 = 24 - 8 = 16$

Hence the required advertisement expenditure is **₹ 16 lakhs.**

Example 12. The two regression lines obtained from certain data were Y = X + 5 and 16X =

Example 13. Given the following two lines of regression:

$$5x + 7y - 22 = 0 \text{ and } 6x + 2y - 20 = 0.$$

Find the mean values of x and y. If the variance of y is 15, find the standard deviation of x.

Solution: We know that the two lines of regression intersect at the point $(\bar{x}, \bar{y})$. Therefore, this point satisfies both the regression equations. Hence, we have

$$5\bar{x} + 7\bar{y} = 22, \text{ and } 6\bar{x} + 2\bar{y} = 20$$

Solving these equations simultaneously for $\bar{x}$ and $\bar{y}$, we get $\bar{x} = 3, \bar{y} = 1.$

Following the criterion $b_{yx} \times b_{xy} \le 1$, we observe that $5x + 7y - 22 = 0$ is the regression equation of y on x and $6x + 2y - 20 = 0$ is the regression equation of x on y.

Now, $5x + 7y - 22 = 0$

$\Rightarrow \quad 7y = -5x + 22$

$\Rightarrow \quad y = \left(-\frac{5}{7}\right)x + \frac{22}{7}$

$\Rightarrow \quad b_{yx} = -\frac{5}{7}$

And $6x + 2y - 20 = 0$

$\Rightarrow \quad 6x = -2y + 20$

$\Rightarrow \quad 6 = \left(-\frac{1}{3}\right)y + \frac{10}{3}$

$\Rightarrow \quad b_{xy} = -\frac{1}{3}.$

Since the regression coefficients are negative, the correlation coefficient will also be negative

$$r_{xy} = -\sqrt{b_{yx} \times b_{xy}}$$

$$= -\sqrt{\left(-\frac{5}{7}\right) \times \left(-\frac{1}{3}\right)} = -\sqrt{\frac{5}{21}}$$

Given that $\sigma_y^2 = 15$

$\therefore \quad \sigma_y = \sqrt{15}$

We have

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

$\Rightarrow \quad = \sqrt{\frac{5}{21}} \times \frac{\sigma_x}{\sqrt{15}}$

$\Rightarrow \quad \frac{1}{3} = \sqrt{\frac{5}{21 \times 15}} \times \sigma_x$

$\Rightarrow \quad \frac{1}{3} = \sqrt{\frac{1}{63}} \times \sigma_x$

$\Rightarrow \quad \sigma_x = \frac{1}{3} \times \sqrt{63} = \sqrt{\frac{63}{9}} = \sqrt{7}.$