Regression Analysis :-

Regression line of y on x

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

Regression line of x on y

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

b_{yx}, b_{xy} are called regression coefficients.

$$b_{xy} = r \frac{\sigma x}{\sigma y}, \quad b_{yx} = r \frac{\sigma y}{\sigma x}$$

$$\therefore b_{xy} \times b_{yx} = r^2 \qquad \therefore r = \pm\sqrt{b_{xy} \times b_{yx}}$$

If both b_{xy}, b_{yx} are +ve then r is +ve
& if b_{xy}, b_{yx} are −ve then r is −ve.

1. The equations of the regression lines between two variables are expressed as $2x - 3y = 0$ and $4y - 5x - 8 = 0$. Find $\bar{x}$ and $\bar{y}$, the regression coefficients and the correlation coefficient between x and y.

Soln:- The two regression equations are

$$2x - 3y = 0 \quad \text{—— (1)}$$

$$4y - 5x - 8 = 0 \quad \text{—— (2)}$$

The point $(\bar{x}, \bar{y})$ lies at the intersection of both the lines of regression equations.

$$\begin{aligned} 2\bar{x} - 3\bar{y} &= 0 \\ 4\bar{y} - 5\bar{x} - 8 &= 0 \end{aligned} \Rightarrow \begin{aligned} \bar{x} &= -3.429 \\ \bar{y} &= -2.286 \end{aligned}$$

$$2x-3y=0 \Rightarrow y=\frac{2}{3}x \Rightarrow b_{yx}=\frac{2}{3}$$

$$4y-5x=8 \Rightarrow x=\frac{4}{5}y-\frac{8}{5} \Rightarrow \text{coeff. of } y = b_{xy}=\frac{4}{5}$$

$\therefore$ Coefficient of correlation $r(x,y)=\sqrt{b_{xy}\times b_{yx}}$

$$=\sqrt{\frac{4}{5}\times\frac{2}{3}}$$

$\therefore \quad r(x,y)=0.73.$ [Range of r ; $-1\le r\le 1$]

Ex2: The two regression lines obtained from certain data were:

$y=x+5$

$16x=9y-94$

Find the variance of x, if the variance of y is 16. Also find the covariance betn x and y.

→ $y=\underbrace{1x}_{\text{coeff. of } x = b_{yx}}+5$ —— ①

$16x=9y-94 \Rightarrow x=\underbrace{\frac{9}{16}}_{=\text{coeff. of } y = b_{xy}}y-\frac{47}{8}$

$b_{yx}=1$ From eq^n ①

$b_{xy}=\frac{9}{16}$ From eq^n ②

$\therefore \quad r=\sqrt{b_{yx}\, b_{xy}}=\frac{3}{4}$

Variance $=\sigma^2$

variance of $x=\sigma_x^2=?$

Given variance of $y=\sigma_y^2=16 \Rightarrow \sigma_y=4$

$\therefore \quad b_{xy}=r\frac{\sigma_x}{\sigma_y} \Rightarrow \frac{9}{16}=\left(\frac{3}{4}\right)\frac{\sigma_x}{4}$

$\therefore \quad \sigma_x=3 \qquad \therefore$ Variance of $x=9$

$r=\frac{\text{cov}(x,y)}{\sigma_x\sigma_y} \Rightarrow \text{cov}(x,y)=r\,\sigma_x\,\sigma_y=\frac{3}{4}\times 3\times 4=9$

Ex 3: Find the mean value of random variable x and y and the correlation coefficient between them, when the two lines of regression are given by

$$5x+7y-22=0$$
$$6x+2y-20=0$$

If the variance of y is 15, find the std. deviation of x.

Soln:- Point $(\bar{x},\bar{y})$ satisfies the given eqns.

$$\therefore \quad 5\bar{x}+7\bar{y}-22=0 \quad \Rightarrow \quad \bar{x}=3, \ \bar{y}=1$$
$$6\bar{x}+2\bar{y}-20=0$$

$$5x+7y=22 \Rightarrow y=\frac{22}{7}-\frac{5}{7}x \Rightarrow b_{yx}=-\frac{5}{7}$$

$$6x+2y=20 \Rightarrow x=-\frac{1}{3}y+\frac{10}{3} \Rightarrow b_{xy}=-\frac{1}{3}$$

$$r=-\sqrt{b_{yx}\,b_{xy}}=-\sqrt{\frac{10}{42}}=-\sqrt{\frac{5}{21}}$$

b_{xy}, b_{yx} are negative $\therefore r$ is negative.

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

$$b_{xy}=r\frac{\sigma_x}{\sigma_y} \Rightarrow \sigma_x=2.646$$

$\therefore$ Std. deviation of x is 2.646.

Ex 4:- Given that the variance of $x=9$ & the regression equations are

$$8x-10y+66=0$$
$$40x-18y=214$$

Find (a) Mean values of x & y

(b) coeff. of correlation betn x & y

(c) std deviation of y.

→ $(\bar{x}, \bar{y})$ satisfies eq^{n}s

$$8\bar{x} - 10\bar{y} = -66$$
$$40\bar{x} - 18\bar{y} = 214$$

$\therefore \quad \bar{x} = 13, \ \bar{y} = 17.$

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

$b_{yx} = \frac{4}{5}$

$40x = 18y + 214 \Rightarrow \quad x = \frac{18}{40}y + \frac{214}{48}$

$\Rightarrow \quad b_{xy} = \frac{9}{20}$

→ $\therefore \ r = \sqrt{b_{yx}\, b_{xy}} = \frac{6}{10}$

$\sigma_x^2 = 9 \Rightarrow \sigma_x = 3$

$b_{yx} = r \, \frac{\sigma_y}{\sigma_x} \Rightarrow \boxed{\sigma_y = 4}$