Department of Statistics

Class: FY M.Sc (Data Science) **Subject: Statistics**

Practical : III Correlation and Regression **Date:**

Q.1 Compute correlation coefficient between X & Y for the following data & interprer the value.

X=65,45,50,60,40

Y= 70,35,60,50,40

Q.2 For a bivariate data $n=50$, $\Sigma x=20$, $\Sigma y=25$, $\Sigma x^2=85$, $\Sigma y^2=90$, $\Sigma xy=75$. Find the equation of line of regression y on x.

Q.3. The equation of two regression lines. Obtain in a regression analysis are $3x+12y=19$

$3y+9x=46$.

Obtain i) correlation coefficient between x & y

ii) Mean of x & y.

38

Practical - 03

seen

Page No.
Date

1) → Given -

$X = 65, 45, 50, 60, 40$
$Y = 70, 35, 60, 50, 40$

Soln -

X	Y	X.Y	X^2	Y^2
65	70	4550	4225	4900
45	35	1575	2025	1225
50	60	3000	2500	3600
60	50	3000	3600	2500
40	40	1600	1600	1600
$\Sigma X =$ 260	$\Sigma Y =$ 255	$\Sigma XY = 13725$	13950	13825

$$\bar{X} = \frac{\Sigma x}{n} = \frac{260}{5} = 52$$

$$\bar{Y} = \frac{\Sigma Y}{n} = \frac{255}{5} = 51$$

$$Cov(X,Y) = \frac{\Sigma x_i y_i}{n} - \bar{X}\bar{Y}$$

$$= \frac{13725}{5} - (52)(51)$$

$$= 2745 - 2625$$

$$Cov(X,Y) = 93$$

$$\sigma x = \sqrt{\frac{\Sigma x_i^2}{n} - \bar{X}^2}$$

$$= \sqrt{\frac{13950}{5} - (52)^2}$$

$$6x = 9.27$$

$$6y = \sqrt{\frac{\Sigma y_i^2 - y^2}{n}} = \sqrt{\frac{13825}{5} - (51)^2}$$

$$6y = 12.80$$

$$r = \frac{Cov(X,Y)}{6x\ 6y}$$

$$= \frac{93}{9.27 \times 12.80}$$

$$r = 0.7838$$

As $r = 0.7838$ their is a high positive corrlation between ~~mo~~ X and Y

Q.2) → Given :-

$n = 50$, $\Sigma x = 20$, $\Sigma y = 25$, $\Sigma x^2 = 85$

$\Sigma y^2 = 90$, $\Sigma xy = 75$

To find eq^n of line regression

$$\cancel{Y - \bar{y} = bxy\ (x - \bar{x})}$$

$$bxy = \frac{Cov(X,Y)}{6^2 x}$$

$$\bar{X} = \frac{\Sigma x}{n} = \frac{20}{50} = 0.4$$

$$\bar{Y} = \frac{\Sigma y}{n} = \frac{25}{50} = 0.5$$

$$Cov(X,Y) = \frac{\Sigma xy}{n} - \bar{X}\bar{Y}$$

$$= \frac{75}{50} - (0.4)(0.5)$$

Page No
Date

$Cov(XY) = 1.3$

$\sigma^2 x = \frac{\Sigma x^2}{n} = (0.4)^2 = 0.16$

$byx = \frac{Cov(XY)}{\sigma^2 x}$

$= \frac{1.3}{0.16}$

$byx = 8.1250$

Eq^n of line regression y on x is 8.1250

Given - $3x + 12y = 19$
$3y + 9x = 46$

Sol^n -
Let us assume that
$3x + 12y = 19$ is eq^n of line of regression x on y and
$9x + 3y = 45$ is eq^n of line of regression y on x.

$\Rightarrow$ $x = -4x + \frac{19}{3}$

$bxy = -4$

$\Rightarrow$ $Y = -3x + \frac{46}{3}$

$byx = -3$

$r^2 = bxy \cdot byx$
$= -4 \times -3$
$= 12$

$\therefore$ r^2 value is lies between -1 to 1
hence assumption is wrong.

Page No
Date

Now we assume that
$3x + 12y = 19$ is regression line ~~on~~ y on x &
$9x + 3y = 45$ is regression line x on y

$$\Rightarrow \quad y = \frac{-3}{12}x + \frac{19}{12}$$

$$b_{yx} = \frac{-3}{12} = \frac{-1}{4}$$

$$\Rightarrow \quad x = \frac{-3y}{9} + \frac{46}{9}$$

$$b_{xy} = \frac{-1}{3}$$

$$r^2 = b_{xy} \cdot b_{yx}$$

$$= \frac{-1}{3} \cdot \frac{-1}{4} = \frac{1}{12}$$

$$r = \sqrt{0.8333}$$

$$r = -0.9$$

Sign of regression coeff^n be always be same as sign of corr^n coeff^n.

ii) Since both line of regression passes to point $\bar{x}, \bar{y}$
then

$$3\bar{x} + 12\bar{y} = 19 \quad \text{———} \quad \times 3$$

$$9\bar{x} + 3\bar{y} = 46$$

$$\cancel{9\bar{x}} + 36\bar{y} = 57$$

$$-\cancel{9\bar{x}} - 3\bar{y} = -46$$

$$33\bar{y} = 11$$

$$\bar{y} = \frac{1}{3} = 0.33$$

$$3\bar{x} + 12\left(\frac{1}{3}\right) = 19$$

$$\bar{x} = \frac{15}{3}$$

$$\boxed{\bar{x} = 5}$$

~~Mean of $\bar{x} = 5$~~
~~Mean of $\bar{y} = 0.33$~~