

(An Autonomous Institution)



#### **DEPARTMENT OF MATHEMATICS**

(34)

### REGRESSION:

Regression is the estimation or prediction of unknown values of one variable from the known values of another variable. It actually measures the nature and extent to which the variables are correlated.

## Lines of Regression:

If the variables in a bivariate distribution are selated we will find that the points in the scattered diagram will cluster around some curve called the curve of regression. If the curve is a straight line, it is called the line of regression.

Equations of lines of regression:

The lines of regression of X on Y is,

$$x - \overline{x} = \gamma \cdot \frac{\sigma_{X}}{\sigma_{\overline{Y}}} (y - \overline{y})$$

The lines of regression of Y on X is,

$$y-\bar{y}=x\cdot\underline{\sigma_{X}}(x-\bar{x})$$

Regression Coefficients:

Regression Coefficient of X on Y is,

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



(An Autonomous Institution)



### PROBLEMS:

- 1) From the following data, find
  - (i) the two regression lines
  - (ii) the Coefficient of Correlation between the manks in Economics and Statistics.
  - (iii) the most likely marks in Statistics when marks in Economics are 30

		-	-	y mesons			-	-		1
Marks in	25	28	35	32	31	36	29	38	34	32
Economics	~ >	20	00	٥٠	0.	00	1	00		
Marks in	1,3	46	49	41	36	72	71	30	33	39
Statistics	40	70	Τ,	1.	00	3 ^	01	30		,

### Solution:

	1					7
Х	У	$\begin{array}{c} X - \overline{X} \\ = X - 3\lambda \end{array}$	Y- <del>Y</del> = Y-38	$(x-\bar{x})^{a}$	(y-y)2	$(x-\overline{x})(y-\overline{y})$
25	43	-7	5	49	25	- 35
28	46	-4	8	16	64	-32
35	49	3	11	9	121	33
32	41	0	3	O	9	б
31	36	-1	- 2	1	4	2
36	32	4	-6	16	36	- 24
29	31	-3	-7	9	49	21
38	30	6	-8	36	64	- 48
34	33	2	-5	4	25	-10
<i>3</i> 2	39	0	-	0	1	0
320	380	0	0	140	398	- 93



(An Autonomous Institution)



#### **DEPARTMENT OF MATHEMATICS**

$$\bar{X} = \underbrace{\Sigma X}_{10} = \underbrace{320}_{10} = 32$$

$$\bar{y} = \frac{Zy}{n} = \frac{380}{10} = 38$$

$$b_{yx} = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sum (x - \overline{x})^2} = \frac{-93}{140} = -0.6643$$

Coefficient of regression of X on Y is,

$$b_{XY} = \underbrace{\frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sum (Y - \overline{Y})^2}} = \underbrace{-93}_{398} = -0.2337.$$

Equation of the line of regression of x on y is,

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

$$\chi - 3\lambda = -0.2337 (y - 38)$$

$$x = -0.2337y + 0.2337 \times 38 + 32$$

$$x = -0.2337y + 40.8806$$

Equation of the line of regression of Y on X is,

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

$$y-38 = -0.6643 (x-32)$$

$$y = -0.6643x + 0.6643x32 + 38$$

$$y = -0.6643x + 0.6643 \times 32 + 38$$

$$y = -0.6643x + 59.2576$$

(") Coefficient of Correlation,

$$\gamma^2 = b_{yx} \times b_{xy}$$



(An Autonomous Institution)



37) 40

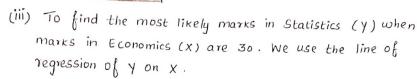
#### **DEPARTMENT OF MATHEMATICS**

$$\frac{1}{3}(x+y), 0 \le x \le 1, 0 \le y \le a.$$

$$\gamma^2 = 0.155a$$

$$\gamma = \pm \sqrt{0.155a}$$

$$\gamma = \pm 0.394$$



$$y = -0.6643 \times + 59.2575$$
Put  $x = 30$ ,
$$y = -0.6643 (30) + 59.2575$$

$$y = 39.3286$$

$$y = 39$$

2) The two lines of regression are 8x - 10y + 66 = 0

$$40 \times - 18 y - 214 = 0$$

The variance of X is 9. Find (i) the mean values of X and Y (ii) Correlation Coefficient between X and Y. Solution:

Given: 
$$8x - 10y = -66$$
  
 $40x - 18y = 214$ 

Since both the lines of regression passes through the mean Values  $\bar{x}$  and  $\bar{y}$ ,

$$8\bar{x} - 10\bar{y} = -66 \rightarrow 0$$

$$40\overline{2} - 18\overline{9} = 214 \longrightarrow 2$$

