6 ← std d

(13)
(1)

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

Lines of Regression

Regression is the study of the relationship between the variables.

Regression shows a relationship between the average values of two variables.

Thus regression is very helpful in estimating and predicting the average value of one variable for a given value of the other variable.

Regression Line which shows the average value of one variable x for a given value of other variable y.

Regression Equation - The best average value if one variable associated with the given value of the other variable may also be estimated or predicted by means of an equation.

Types of variables

Independent Variable - The variable which influences the value or is used for prediction is called "Independent variable."

Dependent Variable - The variable whose value is influenced or is to be predicted is called "Dependent variable"

X on Y → for given value of y find value of X

(Dependent var, Independent var)

Types of Regressions:

1) Simple Regression: Study of only two variable at a time.

2) Multiple Regression: Studying more than two variables at a time.

3) Linear Regression : If the regression curve is a straight line.

4) Non-linear Regression: If the curve or regression is not a straight line.

Lines of Regression

I] Line of regression of x on y

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

II] Line of regression of y on x

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

$\bar{x}$, $\bar{y}$ - means of x and y series

σ_x, σ_y - Std deviation of X and Y series

r - correlation coefficient between x and y.

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

$$r\frac{\sigma x}{\sigma y} = b_{xy} = \frac{N\Sigma xy - \Sigma x \cdot \Sigma y}{N\Sigma y^2 - (\Sigma y)^2}$$

$$r\cdot\frac{\sigma y}{\sigma x} = b_{yx} = \frac{N\Sigma xy - \Sigma x \cdot \Sigma y}{N\Sigma x^2 - (\Sigma x)^2}$$

So,

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

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

PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

Example 1,

Find the equation of regression lines for the following data

x	1	2	3	4	5	6	7	8	9
y	9	8	10	12	11	13	14	16	15

Solution

X	Y	X^2	Y^2	XY
1	9	1	81	9
2	8	4	64	16
3	10	9	100	30
4	12	16	144	48
5	11	25	121	55
6	13	36	169	78
7	14	49	196	98
8	16	64	256	128
9	15	81	225	135
45	108	285	1356	597
Σx	Σy	Σx^2	Σy^2	$\Sigma x \cdot y$

Regression Equation of x on y:

$$X - \bar{X} = b_{xy}(Y - \bar{Y})$$

$$b_{xy} = \frac{N\Sigma xy - \Sigma x \cdot \Sigma y}{N \cdot \Sigma y^2 - (\Sigma y)^2}$$

~~$b_{xy} = 9 \quad N = 9$~~

$N = 9$

$\bar{X} = \frac{45}{9} = 5$

$\bar{Y} = \frac{108}{9} = 12$

$\Sigma x = 45$
$\Sigma y = 108$
$\Sigma x^2 = 285$
$\Sigma y^2 = 1356$
$\Sigma xy = 597$

PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

Semester : VI Subject :CSC601 Data Analytics and Visualization Academic Year: 2023- 2024

$$b_{xy} = \frac{N \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{N \cdot \Sigma y^2 - (\Sigma y)^2}$$

$$= \frac{9 \cdot (597) - 45 \cdot (108)}{9 \cdot 1356 - (108)^2}$$

$$= \frac{5373 - 4860}{12204 - 11664}$$

$$= \frac{513}{540} = 0.95$$

$$\boxed{b_{xy} = 0.95}$$

$$X - \bar{X} = b_{xy}(Y - \bar{Y})$$

$$X - 5 = 0.95(Y - 12)$$

$$X - 5 = 0.95Y - 11.4$$

$$X = 0.95Y - 11.4 + 5$$

$$\boxed{X = 0.95Y - 6.4} \quad \text{---------- (1)}$$

eqn of X on Y

The Regression Equation of Y on x

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

$$b_{yx} = \frac{N \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{N \cdot \Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{9 \cdot (597) - (45)(108)}{9(285) - (45)^2}$$

$$= \frac{5373 - 4860}{2565 - 2025}$$

$$= \frac{513}{540} = 0.95$$

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

$$y - 12 = 0.95(X - 5)$$
$$y - 12 = 0.95X - 4.75$$
$$y = 0.95X - 4.75 + 12$$
$$\boxed{y = 0.95x + 7.25} \quad \text{---------- (2)}$$

eqn of y on x