ENGINEERING MATHEMATICS

UNIT-3 : *Statistical Techniques – I*

Lec-11

Today's Target

- ***Regression Lines***
- (i) ***Line of Regression of y on x***
- (ii) ***Line of Regression of x on y***
- PYQ
- DPP

Regression Analysis

- Regression means functional relationship between two or more related variables.
- If two varibales are correlated then there exists a relation between them.
- Regression Analysis is a process of estimating relationship between dependent variable and independent variables.

Linear Regression

When points in scatter diagram are concentrated around a straight line then it is Linear Regression otherwise non − linear regression.

There are two lines of Regression

(i) Line of Regression of y on x

(ii) Line of Regression of x on y

(1) Line of Regression of y on x

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

$$b_{yx} = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

OR

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

Where

b_{yx} = Regression coefficient of y on x

Note: Regression lines are always passes through the point $(\bar{x}, \bar{y})$

Equation of a line passing through a given point (x_1, y_1) and given slope (m)

$$y - y_1 = m(x - x_1)$$

$$2x + 4y = 11$$

$$4y = -2x + 11$$

$$y = -\frac{1}{2}x + \frac{11}{4}$$

$$m = -\frac{1}{2}$$

$$b_{yx} = -\frac{1}{2}$$

2) Line of Regression of x on y

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

$$b_{xy} = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2}$$

OR

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

Where

b_{xy} = Regression coefficient of x on y

Note: Regression lines are always passes through the point $(\bar{x}, \bar{y})$

Regression line

$2x + 4y = 11$

$2x = -4y + 11$

$x = -2y + \frac{11}{2}$

$b_{xy} = -2$

Relation between regression Coefficient and corelation Coefficient

We know that

$$b_{yx} = \frac{r\,\sigma_y}{\sigma_x} \quad \text{—— ①}$$

$$b_{xy} = \frac{r\,\sigma_x}{\sigma_y} \quad \text{—— ②}$$

Multiply ① and ②

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

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

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

r is the Geometric mean of b_{yx} and b_{xy}

Q.1 *If the regression coefficients are 0.8 and 0.2, what would be the value of coefficient of correlation?*

AKTU-2021

Given

$b_{yx} = 0.8$

$b_{xy} = 0.2$

We know that

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

$$r = \sqrt{0.8 \times 0.2}$$

$$r = \sqrt{0.16}$$

$$\boxed{r = 0.4}$$

Q. 2 *Find the coefficient of correlation and obtain the equation to the lines of regression for the following data:*

AKTU-2017

x:	6	2	10	4	8
y:	9	11	5	8	7

x	y	x^2	y^2	xy
6	9	36	81	54
2	11	4	121	22
10	5	100	25	50
4	8	16	64	32
8	7	64	49	56
$\Sigma x = 30$	$\Sigma y = 40$	$\Sigma x^2 = 220$	$\Sigma y^2 = 340$	$\Sigma xy = 214$

Regression coefficient

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

$$b_{yx} = \frac{5 \times 214 - 30 \times 40}{5 \times 220 - (30)^2}$$

$$b_{yx} = \frac{-130}{200}$$

$$b_{yx} = -0.65$$

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

$$b_{xy} = \frac{-130}{100}$$

$$b_{xy} = -1.3$$

We know that

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

$$r = \sqrt{0.65 \times 1.3}$$

$$r = -0.9192$$

coefficient of correlation

Mean

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

$$= \frac{30}{5}^{6}$$

$$\boxed{\bar{x} = 6}$$

$$\bar{y} = \frac{\Sigma y}{n}$$

$$\bar{y} = \frac{40}{5}^{8}$$

$$\boxed{\bar{y} = 8}$$

Line of regression of y on x

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

$$y - 8 = -0.65(x - 6)$$

$$y - 8 = -0.65x + 3.90$$

$$\boxed{y = -0.65x + 11.9}$$

Line of regression of x on y

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

$$x - 6 = -1.3(y - 8)$$

$$x - 6 = -1.3y + 10.4$$

$$\boxed{x = -1.3y + 16.4}$$

Q.3 *The following table gives age (x) in years of cars and annual maintenance cost (y) in hundred rupes:*

AKTU-2021

X:	1	~~2~~ 3	5	7	9
Y:	15	18	21	23	22

Estimate the maintenance cost for a 4 year old car after finding the regression equation.

x	y	x^2	xy
1	15	1	15
3	18	9	54
5	21	25	105
7	23	49	161
9	22	81	198
$\Sigma x = 25$	$\Sigma y = 99$	$\Sigma x^2 = 165$	$\Sigma xy = 533$

Regression coefficient

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

$$b_{yx} = \frac{5 \times 533 - 99 \times 25}{5 \times 165 - (25)^2}$$

$$b_{yx} = 0.95$$

Mean

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

$$= \frac{25}{5}$$

$$\bar{x} = 5$$

$$\bar{y} = \frac{\Sigma y}{n}$$

$$\bar{y} = \frac{99}{5}$$

$$\bar{y} = 19.8$$

Regression line of y on x

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

$$y - 19.8 = 0.95(x - 5)$$

$$y - 19.8 = 0.95x - 4.75$$

$$y = 0.95x + 15.05$$

Maintenance cost

$$x = 4$$

$$y = 0.95 \times 4 + 15.05$$

$$y = 18.85$$

$$y = 1885 \text{ Rs}$$

Q.4 *The regression equations calculated from a given set of observations for two random variables are*

AKTU-2016

$$x = -0.4y + 6.4 \text{ and } y = -0.6x + 4.6.$$

Calculate (*i*) $\bar{x}$ (*ii*) $\bar{y}$ (*ii*) r

Regression lines

$$x = -0.4y + 6.4 \quad \text{—— (1)}$$

$$y = -0.6x + 4.6 \quad \text{—— (2)}$$

From (1)

$$b_{xy} = -0.4$$

From (2)

$$b_{yx} = -0.6$$

We know that

$$r = \sqrt{0.6 \times 0.4}$$

$$\boxed{r = 0.48989}$$

coeffcuent of correlation

Since Regression lines always passing through $(\bar{x}, \bar{y})$

From ①

$$\bar{x} = -0.4\bar{y} + 6.4$$

$$\bar{x} + 0.4\bar{y} = 6.4 \quad \text{— ③}$$

From ②

$$\bar{y} = -0.6\bar{x} + 4.6$$

$$0.6\bar{x} + \bar{y} = 4.6 \quad \text{— ④}$$

Solve ③ and ④

$$\bar{x} = 6$$

$$\bar{y} = 1$$

AKTU-2018

Q.5:– *In a partially destroyed laboratory record of an analysis of a correlation data, the following results only are legible:*

Variance of $x = 9$

Regression equations: $8x - 10y + 66 = 0, 40x - 18y = 214.$

What were (a) the mean values of x and y (b) the standard deviation of y and the ceofficient of correlation between x and y ?

Given

$\sigma_x^2 = 9$

$\sigma_x = 3$

Regression lines

$8x - 10y = -66$ — (1)

$40x - 18y = 214$ — (2)

Regression lines always passing through $(\bar{x}, \bar{y})$

From (1) and (2)

$8\bar{x} - 10\bar{y} = -66$ — (3)

$40\bar{x} - 18\bar{y} = 214$ — (4)

Solve ③ and ④

$\bar{x} = 13$

$\bar{y} = 17$

From ① and ②

$8x - 10y = -66$

$-10y = -8x - 66$

$y = 0.8x + 6.6$

$b_{yx} = 0.8$

$40x - 18y = 214$

$40x = 18y + 214$

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

$x = 0.45y + 214$

$b_{xy} = 0.45$

We know that

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

$r = \sqrt{0.8 \times 0.45}$

$r = 0.6$

Also

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

$$0.8 = \frac{0.6\,\sigma_y}{3}$$

$$\sigma_y = \frac{0.8 \times 3}{0.6_2}$$

$$\sigma_y = \frac{8}{2}$$

$$\boxed{\sigma_y = 4}$$

DPP- 11

Topic : *Regression Lines*

Q. 1 *From the data given, find the equation of lines of regression of x on y and y on x. Also, calculate the correlation co − efficient.*

x:	2	4	6	8	10
y:	5	7	9	8	11

Q. 2 *For two random variables, x and y with the same mean, the two regression equations are $y = ax + b$ and $x = \alpha y + \beta$. Show that $\frac{b}{\beta} = \frac{1-a}{1-\alpha}$.*

Find also the common mean. (AKTU – 2017)

$\bar{x} = \bar{y} = m$

Q. 3 *The following data regarding the heights (y) and weights (x) of 100 college students are given:* (AKTU – 2018)

$$\sum x = 15000, \quad \sum x^2 = 2272500, \quad \sum y = 6800, \quad \sum y^2 = 463025 \text{ and } \sum xy = 1022250$$

Find the equation of regression line of height on weight.

Q. 4 *The equations of two regression lines, obtained in a correlation analysis of 60 observations are:*

$$5x = 6y + 24 \text{ and } 1000y = 768x - 3608.$$

What is the correlation coefficient? Show that the ratio of coefficient of variablity of x to that of y is $\frac{5}{24}$. What is the ratio of x and y. (AKTU-2019)

GATEWAY CLASSES

AKTU Full Courses (Paid)

Download Gateway Classes Application

From Google Play store

All Subjects

Link in Description

Thank You

GATEWAY CLASSES