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

Examples of Coefficients of Regression

Example 1

The regression lines of a sample are $x + 6y = 6$ and $3x + 2y = 10$.

Find i) sample means $\bar{X}$ and $\bar{Y}$

ii) the coefficient of correlation between x and y

iii) Also find the value of y at $X = 12$.

Solution

$$x + 6y = 6 \quad \text{— (1)}$$
$$3x + 2y = 10 \quad \text{— (2)}$$

i) Sample means $\bar{x}$ & $\bar{y}$

Solving eqn (1) & (2) we get $\bar{x} = 3$ & $\bar{y} = 0.5$

ii) To find the coefficient of correlation we need to first find

Line x on $y \rightarrow b_{xy}$

Line y on $x \rightarrow b_{yx}$

For eqn ① $x + 6y = 6$

$x = 6 - 6y$

$y = -6$ (coefficient of y)

$b_{xy} = -6$

For eqn ① $x + 6y = 6$

$y = \frac{6}{6} - \frac{x}{6}$

$x = -\frac{1}{6}$ (coefficient of x)

$b_{yx} = -\frac{1}{6}$

$b_{yx} = -0.167$

For eqn ② $3x + 2y = 10$

Calculate b_{xy}

$x = \frac{10}{3} - \frac{2y}{3}$

$y = -\frac{2}{3}$

$b_{xy} = -\frac{2}{3} = -0.667$

Calculate b_{yx}

$y = \frac{10}{2} - \frac{3x}{2}$

$x = -\frac{3}{2}$

$b_{yx} = -1.5$

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

Find the value < 1 out of b_{xy} & b_{yx} calculated for eqn ①
Don't consider the sign

$$b_{yx} = -\frac{1}{6}$$ for eqn ①

Now for eqn ② we need to take value of b_{xy}

$$b_{xy} = -\frac{2}{3}$$ for eqn ②

So Line x on y = eqn ② Line y on x = eqn ①

$$r = \pm\sqrt{b_{yx} \cdot b_{xy}}$$

$$r = \pm\sqrt{\left(-\frac{1}{6}\right)\cdot\left(-\frac{2}{3}\right)} = \frac{1}{3}$$

$$r = -\frac{1}{3}$$ as b_{xy} & b_{yx} both are negative r is also negative

So, the coefficient of correlation is $r = -\frac{1}{3}$

iii) Estimate y when $x = 12$

y on $x = 12$
eqn for y on x is eqn (1)
so put $x = 12$ in eqn (1)

$$x + 6y = 6$$
$$12 + 6y = 6$$
$$6y = \frac{6 - 12}{6}$$
$$\boxed{y = -1} \quad \text{when } x = 12$$

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 2 [Dec 2023] [10 marks]

From the following results, obtain the two regression equations and estimate the yield when the rainfall is 29 cm and the rainfall, when the yield is 600 kg:

	Yield in kg	Rainfall in cm
Mean	508.4	26.7
SD	36.8	4.6

The coefficient of correlation between yield and rainfall is 0.52.

Solution

Let x be the rainfall in cm and
y be the yield in kg

$\bar{X} = 26.7 \quad \sigma_x = 4.6$

$\bar{Y} = 508.4 \quad \sigma_y = 36.8$

$r = 0.52$

The regression coefficients are

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.52\,\frac{36.8}{4.6} = 4.16$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = 0.52\,\frac{4.6}{36.8} = 0.065$$

Now, the regression line of x on y is

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

$$x - 26.7 = 0.065(y - 508.4)$$

$$\boxed{x = 0.065y - 6.346} \quad \text{---- (1) x on y}$$

The regression line of y on x is

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

$$y - 508.4 = 4.16(x - 26.7)$$

$$\boxed{y = 4.16x + 397.328} \quad \text{— (2) y on x}$$

Now, when the rainfall $x = 29$ cm, estimated yield y is

$$y = 4.16(29) + 397.328$$

$$\boxed{y = 517.968 \text{ kg}}$$

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

When the yield y is 600 kg, estimated rainfall X is

$$X = 0.065(600) - 6.346$$

$$X = 32.654 \text{ cm}$$

Example 3. [Dec 23] [10 M]

The following data given the experience of machine operators and their performance rating as given by the number of good parts turned out per 100 pieces.

Operators	1	2	3	4	5	6
Performance rating (x)	23	43	53	63	73	83
Experience (y)	5	6	7	8	9	10

Calculate the regression line of performance rating on experience and also estimate the probable performance if an operator has 11 years of experience.

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

Solution $n = 6$

x	y	y^2	xy
23	5	25	115
43	6	36	258
53	7	49	371
63	8	64	504
73	9	81	657
83	10	100	830
338 Σx	45 Σy	355 Σy^2	2735 Σxy

The regression coefficient of x on y is

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

$$\boxed{b_{xy} = 11.429}$$

Now $\bar{x} = \frac{\Sigma x}{n} = 56.33$

$\bar{y} = \frac{\Sigma y}{n} = 7.5$

So the equation of regression line of X on Y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$x - 56.33 = 11.429\,(y - 7.5)$$
$$x = 11.429(y) - 29.3875$$

When the experience is 11 years of an operator, estimated performance is

$x = 96.33$

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

Example4 :

The number of bacterial cells (y) per unit volume in a cultural at different hours (x) is given below:

x	0	1	2	3	4	5	6	7	8	9
y	43	46	82	98	123	167	199	213	245	272

Fit lines of regression of y on x. and x on y. Also, estimate the number of bacterial cells after 15 hours.

Solution :- n = 10

x	y	x^2	xy	y^2
0	43	0	0	1849
1	46	1	46	2116
2	82	4	164	6724
3	98	9	294	9604
4	123	16	492	15129
5	167	25	835	27889
6	199	36	1194	39601
7	213	49	1491	45369
8	245	64	1960	60025
9	272	81	2448	73984
$\Sigma x = 45$	$\Sigma y = 1488$	$\Sigma x^2 = 285$	$\Sigma xy = 8924$	$\Sigma y^2 = 282290$

Here, $\bar{x} = \frac{\Sigma x}{n} = 4.5$

$\bar{y} = \frac{\Sigma y}{n} = 148.8$

The regression coefficients are

$$b_{xy} = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma y^2 - (\Sigma y)^2} = 0.0366$$

$$b_{yx} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x^2)} = 27.000601$$

The regression line of y on x is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$
$$y - 148.8 = 27.000\,61(x - 4.5)$$

$$\boxed{y = 27.00061x + 21.2726} \quad \text{---} (1)$$

The regression line of x on y is

$$x - \bar{x} = b_{xy}(y - \bar{y})$$
$$x - 4.5 = 0.0366(y - 148.8)$$

$$\boxed{x = 0.0366y - 0.9461} \quad \text{---} (2)$$

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

Thus at X = 15 hours

$y = 27.0061(15) + 27.2726$

$y = 432.3641$