AI1110: Assignment 2

SADINENI ABHINAY - CS21BTECH11055

ICSE class 12 paper 2018

I. Question 20(a)

Find the line of regression of y on x from the following table.

x	1	2	3	4	5
y	7	6	5	4	3

Hence, estimate the y value when x=6.

Solution.

Given observations $\begin{pmatrix}x1\\y1\end{pmatrix},\begin{pmatrix}x2\\y2\end{pmatrix}.....,\begin{pmatrix}xn\\yn\end{pmatrix}$ best fit a straight line to it

$$Y = a_0 + a_1 X \tag{1}$$

The residual error

$$E_i = (y_i - (a_0 + a_1 x_i)) \tag{2}$$

Sum of squares of errors should be minimum

$$S_r = \sum_{i=1}^{n} E_i^2 \tag{3}$$

$$\frac{\partial S_r}{\partial a_0} = 0, \frac{\partial S_r}{\partial a_1} = 0, \tag{4}$$

we will get two equations after solving

$$na_0 + a_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \tag{5}$$

$$a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \tag{6}$$

Finally we will get

$$a_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{7}$$

$$a_0 = \left(\frac{\sum_{i=1}^{n} y_i}{n}\right) - a_1\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \tag{8}$$

For this problem

$$\begin{pmatrix}x\\y\end{pmatrix} : \mathbf{A}\begin{pmatrix}1\\7\end{pmatrix}, \mathbf{B}\begin{pmatrix}2\\6\end{pmatrix}, \mathbf{C}\begin{pmatrix}3\\5\end{pmatrix}, \mathbf{D}\begin{pmatrix}4\\4\end{pmatrix}, \mathbf{E}\begin{pmatrix}5\\3\end{pmatrix} \tag{9}$$

x	y	xy	x^2
1	7	7	1
2	6	12	4
3	5	15	9
4	4	16	16
5	3	15	25
$\sum x = 15$	$\sum y = 25$	$\sum xy = 65$	$\sum x^2 = 55$

Mean values and coefficent a_0, a_1 :

$$\bar{x} = 3, \bar{y} = 5, a_1 = -1, a_0 = 8 \tag{10}$$

The line of regression:

$$Y = 8 - X \tag{11}$$

$$X + Y = 8 \tag{12}$$

When $x = 6$ then y must be 2 from the line of regression.

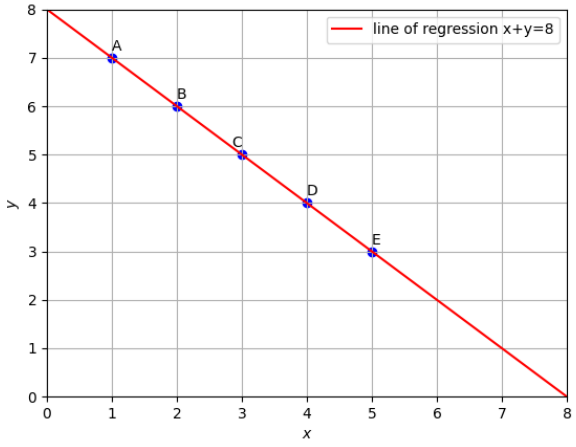

Fig. 0: plot of all points