Assignment 2

Challa Akshay Santoshi-CS21BTECH11012

ICSE 12 2019

Question: 20 (b)

Find the equation of the regression line of y on x, if the observations (x, y) are as follows:
(1, 4), (2, 8), (3, 2), (4, 12), (5, 10), (6, 14), (7, 16), (8, 6), (9, 18)
Also find the estimated value of y when x = 14.

Solution:

Let the regression line of Y on X be given by

$$y = A + Bx \tag{0.0.1}$$

Here A is the Y-intercept and B is the slope of the line.
Using vectors, it can also be represented as:

$$\begin{pmatrix} B & -1 \end{pmatrix} \mathbf{X} = -A \tag{0.0.2}$$

From the definition of line regression, the value of B can be found from a set of data using following relation:

$$B = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} \tag{0.0.3}$$

So, equation (0.0.2) can be written as:

$$\begin{pmatrix} \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} & -1 \end{pmatrix} \mathbf{X} = -A \tag{0.0.4}$$

To find the value of A, we can substitute the point $\begin{pmatrix} \frac{\Sigma x}{n} \\ \frac{\Sigma y}{n} \end{pmatrix}$ in the equation (0.0.3) and solve for it.
Solving for A, we can simplify it as:

$$\begin{pmatrix} \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} & -1 \end{pmatrix} \begin{pmatrix} \frac{\Sigma x}{n} \\ \frac{\Sigma y}{n} \end{pmatrix} = -A \tag{0.0.5}$$

Substituting this value of -A back in equation (0.0.4), we get the regression line of y on x as

$$\begin{pmatrix} \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} & -1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} & -1 \end{pmatrix} \begin{pmatrix} \frac{\Sigma x}{n} \\ \frac{\Sigma y}{n} \end{pmatrix} \tag{0.0.6}$$

From the data given in question,

$$\Sigma x = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 \tag{0.0.7}$$
$$\Sigma y = 4 + 8 + 2 + 12 + 10 + 14 + 16 + 6 + 18 \tag{0.0.8}$$
$$\Sigma xy = 4 + 16 + 6 + 48 + 50 + 84 + 112 + 48 + 162 \tag{0.0.9}$$
$$\Sigma x^2 = 1 + 4 + 9 + 16 + 25 + 36 + 49 + 64 + 81 \tag{0.0.10}$$
$$\tag{0.0.11}$$

From these above equations we get,

$$\Sigma x = 45 \tag{0.0.12}$$
$$\Sigma y = 90 \tag{0.0.13}$$
$$\Sigma xy = 530 \tag{0.0.14}$$
$$\Sigma x^2 = 285 \tag{0.0.15}$$
$$n = 9 \tag{0.0.16}$$

Substituting these values in equation (0.0.6) will give

$$\begin{pmatrix} \frac{530 - \frac{45 \times 90}{9}}{285 - \frac{(45)^2}{9}} & -1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} \frac{530 - \frac{45 \times 90}{9}}{285 - \frac{(45)^2}{9}} & -1 \end{pmatrix} \begin{pmatrix} \frac{45}{9} \\ \frac{90}{9} \end{pmatrix} \tag{0.0.17}$$

Simplifying we get

$$\begin{pmatrix} \frac{80}{60} & -1 \end{pmatrix} \mathbf{X} = \begin{pmatrix} \frac{80}{60} & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 10 \end{pmatrix} \tag{0.0.18}$$
$$\implies \begin{pmatrix} \frac{4}{3} & -1 \end{pmatrix} \mathbf{X} = \frac{-10}{3} \tag{0.0.19}$$

Equation (0.0.19) gives us the linear equation

$$\frac{4}{3}x - y = \frac{-10}{3} \tag{0.0.20}$$

Simplifying we get,

$$4x - 3y + 10 = 0 \tag{0.0.21}$$

Equation (0.0.21) is the required equation of the regression line of y on x for the data given.

The second part of the question asks us to find the estimated value of y when x = 14.
Using equation (0.0.19), we can find y as follows:

$$\begin{pmatrix} \frac{4}{3} & -1 \end{pmatrix} \begin{pmatrix} 14 \\ y \end{pmatrix} = \frac{-10}{3} \tag{0.0.22}$$

$$\implies \frac{56}{3} - y = \frac{-10}{3} \tag{0.0.23}$$

$$\implies y = \frac{66}{3} \tag{0.0.24}$$

$$\implies y = 22 \tag{0.0.25}$$

Therefore, the value of y is 22 when x is 14.

Figure given below represents the regression line.

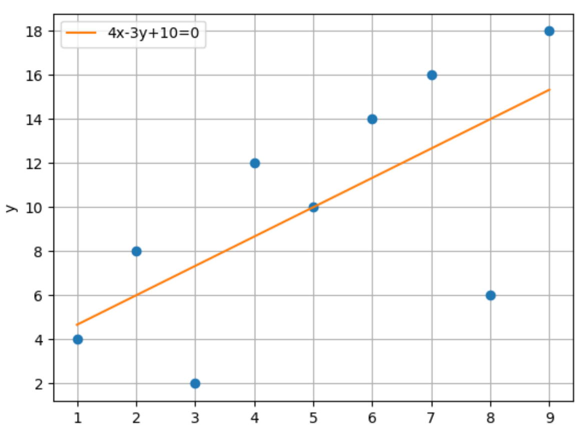

Fig. 0: Graph showing the Regression Line