

Answer on Question #78960 – Math – Statistics and Probability

Question

Given the data,

x 1 2 3 4 5 6 7 8 9

y 9 8 10 12 11 13 14 16 15

(a) Calculate the coefficient of correlation?

(b) Obtain the line of regression.

(c) Estimate the value of y which should correspond to x = 6.2

Solution

(a) Coefficient of correlation is given by

$$r = \frac{\sum_i [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

$$\bar{x} = \frac{1}{9}(1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9) = 5$$

$$\bar{y} = \frac{1}{9}(9 + 8 + 10 + 12 + 11 + 13 + 14 + 16 + 15) = 12$$

$$x_i - \bar{x} : -4; -3; -2; -1; 0; 1; 2; 3; 4$$

$$y_i - \bar{y} : -3; -4; -2; 0; -1; 1; 2; 4; 3$$

$$\sum_i (x_i - \bar{x})^2 = 16 + 9 + 4 + 1 + 0 + 1 + 4 + 9 + 16 = 60$$

$$\sum_i (y_i - \bar{y})^2 = 9 + 16 + 4 + 0 + 1 + 1 + 4 + 16 + 9 = 60$$

$$\sum_i [(x_i - \bar{x})(y_i - \bar{y})] = 12 + 12 + 4 + 0 + 0 + 1 + 4 + 12 + 12 = 57$$

$$\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2} = \sqrt{60 * 60} = 60$$

$$r = \frac{\sum_i [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} = \frac{57}{60} = 0.95$$

Answer (a): $r=0.95$.

(b) In general:

$$y = a + bx,$$

Where

$$b = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2}$$

$$\sum x_i = n\bar{x} = 9 * 5 = 45, \quad \sum y_i = n\bar{y} = 9 * 12 = 108$$

$$\sum x_i^2 = 1 + 4 + 9 + 16 + 25 + 36 + 49 + 64 + 81 = 285$$

$$\sum x_i y_i = 9 + 16 + 30 + 48 + 55 + 78 + 98 + 128 + 135 = 597$$

So

$$b = \frac{9 * 597 - 45 * 108}{9 * 285 - 45 * 45} = \frac{5373 - 4860}{2565 - 2025} = \frac{513}{540} = 0.95$$

a can be found from:

$$\bar{y} = a + b\bar{x}$$
$$a = 12 - 0.95 * 5 = 12 - 4.75 = 7.25$$

Answer (b): $y = 7.25 + 0.95x$.

$$(c) \ y(6.2) = 7.25 + 0.95 * 6.2 = 13.14$$

Answer (c): $y(6.2) = 13.14$.