In order to find the correlation coefficient between two variables X and Y from 20 pairs of observations, the following calculations were made: $\Sigma x = 15$, $\Sigma y = -6$, $\Sigma x y = 50$, $\Sigma x^2 = 61$ and $\Sigma y^2 = 90$. Calculate the correlation coefficient, and the slope of the regression line of Y on X.

Solution

Correlation coefficient is calculated as follows:

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$
$$r_{xy} = \frac{20 \cdot 50 - 15 \cdot (-6)}{\sqrt{20 \cdot 61 - 15^2} \cdot \sqrt{20 \cdot 90 - 6^2}} = 0.823$$

The slope of the regression line of Y on X is calculated as follows:

$$a = r \frac{s_x}{s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{20 \cdot 50 - 15 \cdot -6}{20 \cdot 61 - 15^2} = 1.095$$