Answer on Question #62810 – Math – Statistics and Probability

Question

The data of an experiment to apply sufficient quantities of fertilizer to optimize vegetation growth (grass yield) and avoid excessive application that could lead to runoff and nutrient enrichment of a nearby lake is shown in the data below. Determine the regression of y on x.

y	25	50	75	100	125	150	175	200	225	250
x	84	9	90	154	148	169	206	244	212	248

Solution

By definition, the equation for a regression line is

$$\tilde{y} = a + b \cdot x,$$

where the slope is

$$b = \frac{n(\sum_i xy) - (\sum_i x)(\sum_i y)}{n(\sum_i x^2) - (\sum_i x)^2}$$

and the y-intercept is

$$a = \frac{(\sum_i y)(\sum_i x^2) - (\sum_i x)(\sum_i xy)}{n(\sum_i x^2) - (\sum_i x)^2}.$$

By simple calculation we obtain the following table

$\Sigma_i x$	$\Sigma_i y$	$\Sigma_i xy$	$(\Sigma_i x)^2$	$\Sigma_i x^2$	n
1564	1375	263100	2446096	297838	10

Now, substituting these values into the formula

$$b = \frac{480500}{532284} \approx 0.9027, \qquad a = \frac{-1961150}{532284} \approx -3.6844.$$

Thus, the regression equation of y on x is

$$\tilde{y} = -3.6844 + 0.9027 \cdot x$$

Answer:

$$\tilde{y} = -3.6844 + 0.9027 \cdot x.$$