

To prove $\beta_1 = \sum(X_i - \bar{X})(Y_i - \bar{Y}) / \sum(X_i - \bar{X})^2$, we can start with the ordinary least squares (OLS) regression model, which is:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

where Y_i is the dependent variable, X_i is the independent variable, β_0 and β_1 are the intercept and slope coefficients, respectively, ϵ_i is the error term, and i is the observation index.

The OLS method estimates the coefficients β_0 and β_1 by minimizing the sum of squared errors (SSE) between the observed values of Y_i and the predicted values of $\beta_0 + \beta_1 X_i$:

$$SSE = \sum_{i=1}^n \epsilon_i^2$$

To find the values of β_0 and β_1 that minimize SSE, we take the partial derivatives of SSE with respect to β_0 and β_1 and set them to zero:

$$\partial SSE / \partial \beta_0 = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\partial SSE / \partial \beta_1 = -2 \sum_{i=1}^n X_i (Y_i - \beta_0 - \beta_1 X_i) = 0$$

Solving these equations for β_0 and β_1 , we get:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

$$\beta_1 = \sum(X_i - \bar{X})(Y_i - \bar{Y}) / \sum(X_i - \bar{X})^2$$

where \bar{X} is the mean of the independent variable X_i and \bar{Y} is the mean of the dependent variable Y_i .

Therefore, we have proved that $\beta_1 = \sum(X_i - \bar{X})(Y_i - \bar{Y}) / \sum(X_i - \bar{X})^2$ is the expression for the slope coefficient β_1 in the OLS regression model.