

UNIVERSITY OF TEXAS AT AUSTIN

HW Assignment 2Regression.

Please, provide your **complete solutions** to the following problems. Final answers only, even if correct will earn zero points for those problems.

In all the problems below, you want to perform a simple linear regression with X being the explanatory and Y the response random variable, i.e., your aim is to fit the following model:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

with errors ε independent from X and normal with mean zero and a common standard deviation σ .

Problem 2.1. ($5 \times 2 = 10$ points)

- (i) The parameter β_1 can be interpreted as the mean increase in the response variable Y per unit increase in the explanatory variable X . *True or false?*
- (ii) The parameter β_0 is the mean of the response variable Y . *True or false?*
- (iii) The coefficient of determination R^2 can be interpreted as the proportion of variation in Y that is explained by the linear model. *True or false?*
- (iv) The coefficient of determination R^2 is defined as the ratio of the residual sum of squares to the total sum of squares. *True or false?*
- (v) $\sqrt{\frac{RSS}{n-1}}$ is the appropriate estimate of the standard deviation of the error σ . *True or false?*

Problem 2.2. (10 points) For a data set consisting of 10 observations of the pair (X, Y) , you are given, in our usual notation,

$$\bar{x} = 8, \quad \bar{y} = 10, \quad \sum_{i=1}^{10} x_i^2 = 4000, \quad \sum_{i=1}^{10} x_i y_i = 5000.$$

Determine the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ fitted from the above data.

Problem 2.3. (10 points) For a data set consisting of 20 observations of the pair (X, Y) , you are given, in our usual notation,

$$\sum_{i=1}^{20} x_i = 200, \quad \sum_{i=1}^{20} y_i = 300, \quad \sum_{i=1}^{20} x_i^2 = 3000, \quad \sum_{i=1}^{20} y_i^2 = 4600, \quad \sum_{i=1}^{20} x_i y_i = 3200.$$

Determine the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ fitted from the above data.

Problem 2.4. (10 points) For a data set consisting of observations of the pair (X, Y) , you are given, in our usual notation,

$$\bar{x} = 4, \quad \bar{y} = 3, \quad \sum (x_i - \bar{x})^2 = 12, \quad \sum (y_i - \bar{y})^2 = 1.25, \quad \sum (x_i - \bar{x})(y_i - \bar{y}) = 3.$$

Determine the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ fitted from the above data.

Problem 2.5. (5 points) For a data set consisting of 25 observations of the pair (X, Y) , you are given, in our usual notation,

$$\bar{x} = 5, \quad \bar{y} = 3, \quad \sum (x_i)^2 = 5000, \quad \sum (y_i)^2 = 1000, \quad \sum x_i y_i = 450.$$

The residual sum of squares is 300. Find the coefficient of determination R^2 .

Problem 2.6. (5 points) *Source: An old CAS exam from 1995.*

You fit a simple linear regression model with dependent variable values $y_i = i$ for $i = 1, \dots, 5$. You determine that the estimate of the variance of the error term is $s^2 = 1$. What is the coefficient of determination?