

METU Department of Economics

Spring 2025

ECON 206

T.A.: Pınar Özpala

Problem Set VIII

Content: Simple Linear Regression III

Question 1

How do we make sure that our model's predictive performance is good?

Question 2

We have the regression line for the retail sales per households which depends disposable income. $\hat{y}_{n+1} = 1.922 + 0.3815x_{n+1}$. Also we got the information that

$n = 22$, $\bar{x} = 10799$, $\sum (x_i - \bar{x})^2 = 34110178$, $s_e^2 = 21806$.

Find 95% confidence interval for next year's retail sales when disposable income is 12000\$.

Question 3

A sample of 25 blue-collar employees at a production plant was taken. Each employee was asked to assess his or her own job satisfaction(x) on a scale from 1 to 10. In addition, the numbers of days absent(y) from work during the last year were found for these employees. The sample regression line $\hat{y} = 12.6 - 1.2x$ was estimated by least squares for this data. Also found were

$\bar{x} = 6$, $\sum_{i=1}^{25} (x_i - \bar{x})^2 = 130$, $s_e^2 = 80.6$

A particular employee has a job satisfaction level 4. Find a %90 interval for the number of days this employee would absent from work in a year.

Question 4

Suppose you estimate the consumption function:

$$Y_i = \alpha X_i + u_{1i}$$

And the saving function

$$Z_i = \beta X_i + u_{2i}$$

Where Y = consumption, Z = savings, X = income, and $X = Y + Z$, that is income is equal to consumption plus savings.

- What is the relationship, if any, between $\hat{\alpha}$ and $\hat{\beta}$?
- Will the SSR be the same for the two models?
- Can you compare the R^2 terms of the two models?

Question 5

A researcher wishes to estimate the parameters of the following model:

Model 1: $Y_t = \beta_0 + \beta_1 X_t + u_t$

However, her assistant confuses the dependent and independent variables and estimates the following model:

Model 2: $X_t = \alpha_0 + \alpha_1 Y_t + v_t$, and reports

$$\hat{X}_t = -1.44 + 0.709Y_t$$

$$(-1.55) \quad (30.68)$$

$$R^2_2 = 0.986, \text{SSR}_2 = 35.76, T = 15$$

- Using the results for Model 2, estimate β_1 and its standard error.
- Find the coefficient of determination for the first model.
- Test if β_1 is not significantly different from zero.

Question 6

Let x denote the biochemical oxygen demand (BOD) mass loading and y denote BOD mass removal. A linear regression model is formed

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

And the following information is given about the model:

$$n = 14, \sum x_i = 517, \sum y_i = 346, \sum x_i^2 = 39,095, \sum x_i y_i = 25,825$$

- Obtain the OLS estimators.
- Predict the value of *BOD* mass removal for a single observation made when BOD mass loading is 35.
- Calculate the SSE and then a point estimate of σ .
- What portion of observed variation in removal can be explained by the approximate linear relationship between two variables?