Erasmus School of Economics

MOOC Econometrics

Test Exercise 1

Notes:

- See website for how to submit your answers and how feedback is organized.
- This exercise uses the datafile `TestExer1` and requires a computer.
- The dataset `TestExer1` is available on the website.

Goals and skills being used:

- Get hands-on experience with performing simple regressions.
- Get feeling for consequences of violations of regression assumptions.
- Obtain some experience with how to diagnose that an assumption is violated.

Questions

This exercise considers an example of data that do not satisfy all the standard assumptions of simple regression. In the considered case, one particular observation lies far off from the others, that is, it is an outlier. This violates assumptions A3 and A4, which state that all error terms ε_i are drawn from one and the same distribution with mean zero and fixed variance σ^2. The dataset contains twenty weekly observations on sales and advertising of a department store. The question of interest lies in estimating the effect of advertising on sales. One of the weeks was special, as the store was also open in the evenings during this week, but this aspect will first be ignored in the analysis.

(a) Make the scatter diagram with sales on the vertical axis and advertising on the horizontal axis. What do you expect to find if you would fit a regression line to these data?

(b) Estimate the coefficients a and b in the simple regression model with sales as dependent variable and advertising as explanatory factor. Also compute the standard error and t-value of b. Is b significantly different from 0?

(c) Compute the residuals and draw a histogram of these residuals. What conclusion do you draw from this histogram?

(d) Apparently, the regression result of part (b) is not satisfactory. Once you realize that the large residual corresponds to the week with opening hours during the evening, how would you proceed to get a more satisfactory regression model?

(e) Delete this special week from the sample and use the remaining 19 weeks to estimate the coefficients a and b in the simple regression model with sales as dependent variable and advertising as explanatory factor. Also compute the standard error and t-value of b. Is b significantly different from 0?

(f) Discuss the differences between your findings in parts (b) and (e). Describe in words what you have learned from these results.

Erasmus