Universität des Saarlandes
Statistik und Ökonometrie
PD Dr. Stefan Klößner
M.Sc. Sandra Baar

UNIVERSITÄT DES SAARLANDES

2nd Tutorial to
Econometric Methods and Applications WS 2017/18

Exercise 8 [2.5%]

Consider the following simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + u_i.$$

Explain and/or discuss the following notions:

(a) model assumptions

(b) regressand, regressor, and error term

(c) depdendent and independent variable

(d) regression line, intercept, slope

(e) regression coefficient and estimated regression coefficient

(f) ordinary least squares

(g) What is the relation between the conditional expectation of Y and the regression line if the assumption $E(u_i|X_i) = 0$ is fulfilled?

Exercise 9 [2,5%]

Regressing average weekly salaries (AWS, in Euro) on workers' age (in years) yielded the following result:

$$\widehat{AWS} = 696.7 + 9.6 \cdot AGE, \quad R^2 = 0.023, \quad SER = 624.1.$$

(a) Interpret the values 696.7 and 9.6.

(b) What is the unit of SER (Euro? Years? No unit?)

(c) Interpret the coefficient of determination.

(d) What estimate does the model produce for 25 and 45 year old workers, respectively?

(e) The average age of the workers in this study was 41.6 years. What was the average weekly salary in this study?

Exercise 10 [2,5%]

(a) For the quantities TSS $= \sum_{i=1}^{n}(Y_i - \overline{Y})^2$ (total sum of squares), ESS $= \sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2$ (explained sum of squares), and SSR $= \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ (sum of squared residuals), show the relation TSS $=$ ESS $+$ SSR. How are these quantities related to the coefficient of determination?

(b) Assume that in a regression, we find $\widehat{\beta}_1 = 0$. What does this imply for the coefficient of determination?

Exercise 11 [2,5%]

To study the effect of time pressure on results of exams, 400 students were randomly given either 90 or 120 minutes of time for tackling the exercises. The scores Y_i achieved by the students ($0 \leq Y_i \leq 100$) were then regressed on the amount of time X_i the students could spend ($X_i \in \{90, 120\}$) using the simple linear regression model $Y_i = \beta_0 + \beta_1 \cdot X_i + u_i$. The result of the regression was: $\widehat{Y}_i = 49 + 0.24 \cdot X_i$.

(a) Calculate the expected scores of students who had 90 (120) minutes at their disposal.

(b) How would the results change if the two treatments (90 min, 120 min) were coded by a dummy variable taking the values 0 and 1?

Exercise 12 [2.5%]

Consider the simple linear regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$.

(a) Assume that the intercept β_0 is unnecessary, i.e. $\beta_0 = 0$ such that the true data generating process is $Y_i = \beta_1 X_i + u_i$.

(i) Show that the OLS estimator for β_1 given $\beta_0 = 0$ is given by $\tilde{\beta}_1 = \frac{\overline{XY}}{\overline{X^2}}$.

(ii) Show that $\tilde{\beta}_1 = \frac{\overline{XY}}{\overline{X^2}}$ can be written as $\beta_1 + \frac{1}{n}\sum_{i=1}^{n}\frac{X_i}{\overline{X^2}}u_i$.

(iii) Show that, for homoskedastic errors, the variance of $\tilde{\beta}_1$ is equal to $\frac{\sigma_u^2}{n\overline{X^2}}$. Compare this to $\frac{\sigma_u^2}{n s_X^2}$, the variance of the 'standard' estimator $\hat{\beta}_1 = \frac{s_{XY}}{s_X^2}$.

(b) [facultative, additional 2.5%]
Conduct an analogous analysis of the case of a spurious regressor, i.e. for the model $Y_i = \beta_0 + u_i$.