ECON 3150/4150: Seminars spring semester 2012.

23 January, 2012.

The seminars start in week 5. The last seminar is week 17. There are no seminars in week 8, 14 and 15.

Exercises to seminar 1

The two first exercises are computer exercises where you can use Stata. The third is an exercise in the use of "OLS algebra"

1. Download *ConsInc.dta* from the course web page. This file contains annual observations for the period 1960 to 2006. There are two variables, dubbed I and C. The data is computer generated (artificial data) and we interpret I as real disposable income (in million kroner) and C as real private consumption expenditure (in million kroner).

   (a) Estimate two consumption functions with the use of OLS: One for the sample period 1960-1984 and another for the sample period 1985-2006. Take care to include an intercept in the regression. Show that the estimated regression coefficient (with three decimal points) is 0.897 on the first sample and 0.893 on the second sample. Does this mean that the regression lines based on the two samples are identical?

   (b) Estimate a third regression for the combined sample 1960-2006 which gives an estimated regression coefficient of 0.805. Can you give an intuitive explanation for why both the two sub-sample estimates were higher?

   (c) Estimate the regression that has I on the left hand side and C as the regressor (this is often called the "reverse regression"). Do this for the full sample 1960-2006. Can you, with reference to OLS algebra explain why is this estimate becomes different from both 0.805 and 1/0.805?

2. In this question, use *ConsInc.dta* and the full sample 1960-2006.

   (a) Calculate the means $\bar{C}$ and $\bar{I}$ of private consumption and disposable income. Re-do the regression in question 1b. Use the estimated equation and the mean of income ($\bar{I}$), to calculate "the fitted mean" of C which we can denote $\widehat{\bar{C}}$. Compare $\widehat{\bar{C}}$ to $\bar{C}$. What do you find? Can you give an interpretation in terms of a scatter-plot and a regression line?

   (b) Calculate the two variables $c = C - \bar{C}$ and $i = I - \bar{I}$. Estimate two linear equations between c and i with the use of OLS: One with , and another without an intercept. What do you find? Can you explain you findings?

(c) Divide C and I by 1000 so that the the unit of measurement becomes billion kroner instead of million kroner. How does this affect the OLS estimate of the intercept and of the regression coefficient?

(d) Calculate the two variables $LC = \ln(C)$ and $LI = \ln(I)$. Regress LC on LI. What is the economic interpretation of the regression coefficient in this case?

3. Consider the regression model

(1) $$y_i = \beta_1 + \beta_2 x_i + e_i,\ i = 1, 2, ..., n$$

where we assume that the random errors e_i $(i = 1, 2, ..., n)$ have properties that are in accordance with the assumptions on page 47 in HGL (an example of "classical assumptions"). In particular, we have that $Var(e_i) = \sigma^2$ for all i.

(a) Show that the OLS estimator of β_2, which we denote $\hat{\beta}_2$, can be obtained by applying OLS on the re-parameterized model

$$y_i = \alpha + \beta_2(x_i - \bar{x}) + e_i,\ i = 1, 2, ..., n$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(b) Assume that the regressor is a deterministic variable. Find expressions for $var(\hat{\alpha})$, $var(\hat{\beta}_2)$, and $var(\hat{\beta}_1)$ with the aid of the notation $var(e_i) = \sigma^2$ and using the assumptions of the model.

(c) With reference to the lectures we have that $cov(\hat{\alpha}, \hat{\beta}_2) = 0$, but what is $cov(\hat{\beta}_1,\hat{\beta}_2)$?