hw03: Linear Regression

*Your name**

04 07, 2017

Questions in this problem set are taken from *Kleiber, C., & Zeileis, A. (2008). Applied econometrics with R. Springer Science & Business Media, pp. 91-92 (questions 2 - 5).*

Question 2

Estimate a hedonic regression for the HousePrices data taken from Anglin and Gen¸cay (1996), which contain prices of houses sold in the city of Windsor, Canada, during July, August, and September 1987. These data are also used in the textbook by Verbeek (2004).

a)

Fit a multiple linear regression model to the logarithm of the price, using all remaining variables as regressors. Experiment with models containing lot size, number of bathrooms, number of bedrooms, and stories in logarithms and in levels, respectively. Which model do you prefer?

```
# Write here your codes
```

b)

What is the expected price of a two-story house of 4,700 sq. ft. with three bedrooms, two bathrooms, a driveway, no recreational room, a full finished basement, without gas heating or air conditioning, and two-car garage, that is not located in a preferred area? Report also a prediction interval.

```
# Write here your codes
```

c)

In order to determine whether the logarithmic transformation of the dependent variable is really appropriate, a Box-Cox transformation might be helpful. Use the function `boxcox()` from the package **MASS**. What do you conclude?

```
# Write here your codes
```

Question 3

Consider the `PSID1982` data from Cornwell and Rupert (1988) and discussed further in Baltagi (2002).

*affiliation

a)

Regress the logarithm of the wage on all available regressors plus experience squared.

```
# Write here your codes
```

b)

Does gender interact with education and/or experience?

Question 4

Section 3.5 considered two competing models for US consumption utilizing an encompassing test. Different approaches to comparing nonnested models are the J test suggested by Davidson and MacKinnon (1981) and the Cox test. Both are available in the package **lmtest** in the functions `jtest()` and `coxtest()`. For the methodological background, we refer to Greene (2003, Chapter 8) and Davidson and MacKinnon (2004).

a)

Test `cons_lm1` vs. `cons_lm2` using the J test.

```
# Write here your codes
```

b)

Test `cons_lm1` vs. `cons_lm2` using the Cox test.

```
# Write here your codes
```

Do all tests lead to similar conclusions?

Question 5

Use the `PSID1982` data and consider the following two nonnested models (compare with Baltagi 2002, p. 230):

$$\begin{aligned} M_1 : log(wage) = \beta_0 + \beta_1 education + \beta_2 experience + \\ \beta_3 experience^2 + \beta 4 weeks + \beta_5 married + \\ \beta_6 gender + \beta_7 ethnicity + \beta_8 union + \epsilon \end{aligned}$$

$$\begin{aligned} M_2 : log(wage) = \beta_0 + \beta_1 education + \beta_2 experience \\ + \beta_3 experience^2 + \beta 4 weeks + \beta_5 occupation \\ + \beta_6 south + \beta_7 smsa + \beta_8 industry + \nu \end{aligned}$$

a)

Compute the J tests for M_1 vs. M_2 and M_2 vs. M_1, respectively.

```
# Write here your codes
```

b)

Both M_1 and M_2 can be artificially nested within a larger model. Use an F test for M_1 versus this augmented model. Repeat for M_2 versus the augmented model. What do you conclude?

```
# Write here your codes
```