

Homework 2

ECON 4848

Professor Martins-Filho

Due date: 2.11.2011 in class

1. Using the data 401K.dta available on the class website write a Stata `do` file to answer question C2.1 on p. 64 of your textbook. In addition answer the following questions:

1. Under the assumption that $E(\text{prate}|\text{mrate}) = \beta_0 + \beta_1\text{mrate}$, $V(\text{prate}|\text{mrate}) = \sigma^2$, that $\text{prate}_i|\text{mrate}_i \sim N(\beta_0 + \beta_1\text{mrate}, \sigma^2)$ and that we have a random sample, obtain the estimator we labeled s^2 for σ^2 .
2. Suppose that you knew the true value of σ^2 to be the 260. Find the value z^* that satisfies $P\left(\left|\frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}}}\right| < z^*\right) = 0.95$.
3. Produce a graph with the estimated regression and the observed data.
4. If we label the predicted value of `prate` by $\widehat{\text{prate}}$. What is the correlation between `prate` and $\widehat{\text{prate}}$? Is this what we called the regression R^2 ?
5. Suppose we multiply `prate` and `mrate` by different constants, say a and b . What is the impact of this transformation on R^2 and the estimated coefficients? What if the constants are equal? Explain algebraically.

2. Using the data WAGE2.dta available on the class website write a Stata `do` file to answer question C2.4 on p. 65 of your textbook. In addition answer the following questions:

1. Obtain the regression residuals for the model described in item C2.4 (iii). Produce a scatter plot of the squared residuals against the regressor. Does this graph give intuitive support for the assumption of homocedasticity? Explain carefully.

2. Test the hypothesis that IQ does not impact earnings in the model described in C2.4 (ii) at 5 percent (under the assumption of conditional normality of the regressand).
3. The data `anscombe73.dta` contains two variables that are used as regressors `reg1` and `reg2` and four variables that are used as regressands `rsd1 – 3` and `rgs4`. Write a `do` code that performs estimation of the following linear regressions:

1. regress `rsd1` on `reg1`
2. regress `rsd2` on `reg1`
3. regress `rsd3` on `reg1`
4. regress `rgs4` on `reg2`

In each situation produce a scatter plot with the observations and the fitted regression. Do you believe that the data for each of the four regressions considered comes from the same stochastic process? What do you learn from this exercise?

4. Do questions 2.6, 2.11 in your textbook pp. 62,64.