Homework 1 for Econometrics I

(due October 30, 2019)

Fall 2019

Instructor: Jihai Yu

TA: Zeyang Chen and Yuqing Wang

1. Consider the linear regression model $y = X\beta + \epsilon$ where the $n \times K$ matrix X contains a constant term, $rank(X) = K$, and $\epsilon \sim N(0, \sigma^2 I_n)$

(1) To test $R\beta = q$ where R is $J \times K$ constant matrix, we use F-test statistics. Show that it is F distributed with degrees of freedom J and $n - K$.

(Hint: If w_1 and w_2 are two independent χ^2 variables with, respectively, degrees of freedom p and q, then, $\frac{w_1/p}{w_2/q}$ is an F with p and q degrees of freedom.)

(2) Write down the form of F-test when we are interested in $H_0 : \beta_2 = ... = \beta_K = 0$.

(3) Express the F-test in (2) in terms of the coefficient of determination R^2.

2. The file MROZ contains data[1] on the labor participation and the wage level for married women, where we have $n = 753$ observations. **We need only first 428 observations.** We are going to examine how the log(wage) changes if there is a change in some explanatory variables. Following are the observations we have

inlf: infl= 1 denotes that the observed woman was in labor force

lwage: log(wage)

exper: past years of labor market experience

expersq: square of exper

educ: years of education

age: age of the married woman

kidslt6: number of children less than 6 years old

kidsge6: number of kids between 6 and 18.

The estimation equation is

$$\log(wage) = \beta_0 + \beta_1 \text{exper} + \beta_2 \text{exper}^2 + \beta_3 educ + \beta_4 age + \beta_5 kidslt6 + \beta_6 kidsge6 + \epsilon.$$

(1) Obtain the OLSE by using a computer (e.g. Stata).

(2) Is $\hat{\beta}_2$ significant at 5% significance level? How about the 10% level?

[1] https://www.msu.edu/~ec/faculty/wooldridge/book2.htm

(3) Construct an interval for β_1 such such this interval will cover the true value of β_1 with 95% probability.

(4) Test $H_0 : \beta_1 = ... = \beta_6 = 0$ under 1% significance level.

3. For the MLE of θ_0 in the exponential model:

$$F(u;\theta_0) = \begin{cases} 0 & \text{if } u < 0 \\ 1 - e^{-u/\theta_0} & \text{if } u \geq 0 \end{cases}$$

(1) Given the observations $u_1, ..., u_n$, what is the MLE?

(2) Is the MLE unbiased? What is the variance of the MLE?

(Hint: use $E(u_i) = \theta_0$ and $Var(u_i) = \theta_0^2$ for the exponential distribution)

(3) Find the information matrix. Is the MLE you obtain in (1) efficient relative to other unbiased estimators?

4. For the linear regression model $y = X_1\beta_1 + X_2\beta_2 + \epsilon$ where X_1 is $n \times k_1$, X_2 is $n \times k_2$, and elements of ϵ are i.i.d. $(0, \sigma^2)$.

The null hypothesis is $H_0 : \beta_1 = 0$. Assuming that we know σ^2 so that we only need to estimate β. Also, X_1 and X_2 are of full rank.

(1) Derive the Wald test statisitc.

(2) Derive the LM test statistic.