Utrecht University
Utrecht University School of Economics

Midterm exam Econometrics (Wisb377)

Tuesday, 3 November 2020, 13:00-15:30 CET

For those who have permission from the Board of Examiners for an extension of time, they can hand the answer sheet in on 16:20 CET ultimately.

- Please do not post copies on the Internet.

Remarks:
- This entrance test consists of 4 questions (12 sub-questions) on 5 numbered pages (included front page).
- Write your name and registration number on each page of your exam.
- You should write down your answers in full, as if it were an on campus exam.
- Please make pictures of all pages of your answers (with a camera, copy the pictures in **ONE Word document**, and upload the document on Blackboard.
- You should upload the Word document **on 15:35 CET ultimately** (or mail it to W.H.J.Hassink@uu.nl).
- Please do not post copies of this exam on the Internet.

By taking the midterm exam of Econometrics wisb377 of November 3rd 2020, you automatically confirm that

"By taking this exam I declare that I will formulate the answers myself, without the help of others, and without using unauthorized tools, and take the exam according to its instructions. Violation of these rules is regarded as fraud / plagiarism."

Questions

In the questions below – unless otherwise stated – we consider the linear regression equation

$$y_i = \boldsymbol{\beta}'\mathbf{x}_i + u_i \qquad i=1,\dots,n$$

for which y is the dependent variable. $\boldsymbol{\beta}$ is a (k+1) dimensional column vector. $\mathbf{x}$ is the (k+1) dimensional column vector of explanatory variables, and u is an error term. Subscript i refers to the i-th individual. n is the number of observations of the data set.

Question 1 (Chapter 7, 8)

a) We apply a Z-test by applying Theorem 8.1 on page 363 in which the test statistic

$$\hat{\boldsymbol{\beta}} \mid \mathbf{X} \sim Normal(\boldsymbol{\beta}, \sigma_u^2(\mathbf{X}'\mathbf{X})^{-1})$$

is derived. Suppose that the six assumptions of Theorem 8.1 are valid. With respect to the following regression equation we would like to test whether *Educ* has a statistical significant effect on $\log(Wage)$

$$\log(Wage_i) = \beta_0 + \beta_1 Educ_i + \beta_2 Exper_i + u_i \qquad i=1,\dots,n$$

Questions:

1. Starting with the aforementioned statistic of $\hat{\boldsymbol{\beta}} \mid \mathbf{X}$, which additional proposition(s) of the Normal distribution are further required to test whether *Educ* has a significant effect on $\log(Wage)$. How would the test statistic be formulated?

2. Consider the six assumptions of Theorem 8.1. What is the disadvantage of this testing procedure? In other words, which assumptions of the list of assumptions are too strong? Please carefully motivate your answer.

b) For

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \mid \mathbf{X} \sim Normal(\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, 4\mathbf{I}_3)$$

please formulate the corresponding Chi-square statistic.

c) Next, we examine the Wald test

$$U = \frac{(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})/q}{\sigma_u^2} \mid \mathbf{X}$$

$$V = (n-k-1)\cdot\frac{\hat{\sigma}_u^2}{\sigma_u^2}$$

U and V are stochastically independent.

$$F(1, n-k-1) = \left(t_{(n-k-1)}\right)^2$$

Could you please explain how the information of above can be used to test $H0: \beta_1 = 0$ versus $H1: H0 \text{ not true}$ for the equation

$$\log(Wage_i) = \beta_0 + \beta_1 Educ_i + \beta_2 Exper_i + u_i \qquad i=1,\ldots,n$$

Question 2 (Chapter 9, 10)

We can make use of the Central Limit Theorem (CLT) to develop a statistical testing procedure. Consider the (k+1)-dimensional random vectors $\mathbf{x}_1,\ldots,\mathbf{x}_n$ which are identically and independently distributed. In the proof of Theorem 9.6, the Cramér-Wold device is applied. Two questions on the following expression:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{x}_i u_i \xrightarrow{d} Normal(\mathbf{0}, \sigma_u^2\mathbf{C}) \Leftrightarrow \frac{1}{\sqrt{n}}\sigma_u^{-1}\boldsymbol{\lambda}'\mathbf{C}^{-\frac{1}{2}}\sum_{i=1}^{n}\mathbf{x}_i u_i \xrightarrow{d} Normal(0,1)$$

$$\forall \boldsymbol{\lambda} \in \mathbb{R}^{k+1}, \text{ expect for } \boldsymbol{\lambda} = \mathbf{0}$$

a) Why is it useful to apply the Cramér-Wold device in this proof?

b) Could you please further explain this expression?

c) Do we need to assume Normality of the error term u of the regression equation when relying on the Central Limit Theorem? Could you please explain your answer in great detail?

d) Is it necessary to assume Normality of the error term u to calculate the robust standard errors of the estimated OLS estimator? Please motivate your answer.

Question 3 (Chapter 11)

a) Let's assume a general covariance structure $Var(\mathbf{u} \mid \mathbf{X}) = \boldsymbol{\Psi}$ and let's suppose that $\boldsymbol{\Psi}$ is known. Please explain why it is important to make a comparison between the estimates of the OLS estimator $\hat{\boldsymbol{\beta}}_{\text{OLS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and the GLS estimator $\hat{\boldsymbol{\beta}}_{\text{GLS}} = (\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Psi}^{-1}\mathbf{y}$
Please, carefully explain your answer.

b) Consider the second-order moving average model (the MA(2) model)

$$u_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} \qquad t = 3,\ldots,T$$

for which the error term e_t is i.i.d. (identically and independently distributed), with expected value zero and constant variance: $Ee_t = 0$ and $Var(e_t) = \sigma_e^2$.

<u>Question:</u> Could you please derive the $(T-2)\text{x}(T-2)$ covariance matrix of the error terms?

Question 4 (Chapter 13)

In this question, we are interested in the linear regression equation

$$y_{it} = \alpha_i + \beta_1 x_{it} + u_{it} \qquad i=1,\ldots,n;\ t=1,\ldots,T$$

a) Show that the within estimator and the first-differences estimator of β_1 are equal for T=2.

b) When do you favor pooled OLS and when Random effects?

c) Please carefully explain why for the first-difference model it is assumed that the error term u_{it} follows a unit root process.

< End of the exam >