UNIVERSITY OF WARWICK

May Examinations 2018/19

Advanced Econometric Theory

Time Allowed: 3 Hours

Answer ALL questions in each section.

Use a separate answer booklet for each section.

Read carefully the instructions on the answer book provided and make sure that the particulars required are entered on each answer book.

Section A: Answer ALL questions

---

1. (a) Define Convergence Almost Surely. **(3 marks)**

(b) Consider the linear regression model, $y_i = \mathbf{x}_i'\beta_0 + \epsilon_i$, where $\mathbf{x}_i$ is an L-dimensional vector of endogenous covariates, i.e., $E[\mathbf{x}_i\epsilon_i] \neq 0$. Suppose that there is a set of instruments $\mathbf{z}_i \in \mathbb{R}^{K\times 1}$ that is available to the researcher to achieve identification of β_0. State and explain the two conditions that the set of instruments $\mathbf{z}_i$ needs to satisfy to identify β_0. **(5 marks)**

---

2. In order to study the effect of education on wages in a certain population, an i.i.d. sample of n individuals are chosen and, for each individual, the following variables are recorded:

$$X_i = \text{number of years of education of individual } i.$$
$$Y_i = \text{wage in dollars of individual } i.$$

Assume that $Var(X) > 0$, $Var(Y) > 0$, and that $E\left[(X - E(X))^2 (Y - E(Y))^2\right] < \infty$. The researcher is interested in conducting large sample inference about the correlation coefficient between wages and years of education:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

(Question 2 continued overleaf)

In particular, she is interested in testing a hypothesis about the sign of $\rho_{X,Y}$. To conduct this test, the researcher construct a statistic based on the sample covariance, given by

$$\widehat{cov}(X,Y) = n^{-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)\left(Y_i - \bar{Y}_n\right)$$

where $\bar{X}_n$ and $\bar{Y}_n$ are the sample averages of $\{X_i\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$, respectively. Complete the following points.

(a) Derive the asymptotic distribution of $\widehat{cov}(X,Y)$.
Hint: In order to have a non-degerate distribution, the statistic should be properly normalized. You may find useful deriving the asymptotic distribution of:

$$n^{-1}\sum_{i=1}^{n}\left(X_i - E(X)\right)\left(Y_i - E(Y)\right)$$

(properly normalized). **(8 marks)**

(b) Provide a consistent estimator of the asymptotic **variance** of $\widehat{cov}(X,Y)$. To answer this question you do not need to formally prove consistency, just sketch the arguments and indicate which results are being used. **(5 marks)**

(c) Prove that $\rho_{X,Y} \leq 0$ if and only if $cov(X,Y) \leq 0$. **(2 marks)**

(d) Using the finding in previous points, propose a procedure to conduct the following hypothesis test:

$$\begin{aligned} H_0 &: \rho_{X,Y} \leq 0 \\ H_1 &: \rho_{X,Y} > 0 \end{aligned} \tag{1}$$

For the answer to be valid, the proposed test (i) should be consistent, (ii) have an asymptotic size of $\alpha = 5\%$. In particular, you should:

(i) Describe the test procedure. **(5 marks)**
(ii) Sketch an argument to show that the proposed test is consistent. **(4 marks)**
(iii) Sketch an argument to show that the proposed test has an asymptotic size of $\alpha = 5\%$. **(4 marks)**

(e) Suppose that the null hypothesis of the test in Equation (1) is rejected. Is this sufficient evidence to conclude that an increase in education will cause an increase in expected wages? Justify your answer. **(2 marks)**

(Continued overleaf)

3. Consider the following sampling process. We select a number n of individuals from a population, and for each individual, we collect $Y_{1,i}$ and $Y_{2,i}$. We assume that:

$$\{(Y_{1,i}, Y_{2,i})\}_{i=1}^{n} \quad \text{are independent and} \quad \begin{bmatrix} Y_{1,i} \\ Y_{2,i} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_i \\ \mu_i \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \right)$$

where the unknown parameters are $(\{\mu_i\}_{i=1}^{n}, \sigma^2) \in \Theta_n = \mathbb{R}^n \times \mathbb{R}_{++}$. As the notation indicates, the parameter space Θ_n depends on the sample size. Our objective in this problem is to estimate σ^2. We are not interested in the parameters $\{\mu_i\}_{i=1}^{n}$, i.e., from our perspective $\{\mu_i\}_{i=1}^{n}$ is a vector of nuisance parameters.

(a) Explain why $\{(Y_{1,i}, Y_{2,i})\}_{i=1}^{n}$ is NOT necessarily an i.i.d. sequence of random variables. **(2 marks)**

(b) Consider the following MLE estimator of $(\{\mu_i\}_{i=1}^{n}, \sigma^2)$

$$\hat{\mu}_i = \frac{Y_{1,i} + Y_{2,i}}{2}, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(Y_{1,i} - Y_{2,i})^2}{n},$$

Prove the following results.

Hint: For the following points, you may find useful that $Y_{1i} \perp Y_{2i}$ when $(Y_{1,i}, Y_{2,i})$ is distributed bivariate normal and $cov(Y_{1,i}, Y_{2,i}) = 0$.

(i) Show that the MLE estimator of μ_i is inconsistent. **(5 marks)**

(ii) Show that the MLE estimator of σ^2 is inconsistent. **(5 marks)**

Remark: The situation in this excercise is known as incidental parameters problem. The size of the vector of nuisance parameters $\{\mu_i\}_{i=1}^{n}$ grows with the sample size n (i.e., they are incidental) and this negatively affects the properties of the parameter of interest.

Section B: Answer ALL questions

4. Consider the following model for an observed binary variable y_{it} such that:

$$\Pr(y_{it} = 1 | y_{i(t-1)}, y_{i(t-2)}, .., y_{i1}, x_{iT}, ..., x_{i1}) = \Lambda(x_{it}^{'}\beta + \alpha_i) \qquad (i = 1, ..., N; t = 1, .., T)$$

where Λ is the logistic cumulative distribution function (cdf) $[\Lambda(z) = \exp(z)/[1 + \exp(z)]]$.

(Question 4 continued overleaf)

(a) Derive the likelihood function for this model and explain how you can obtain a consistent estimator of β, without making any assumptions about the form of the distribution of $\alpha_i|x_{i1}, x_{i2}, ..., x_{iT}$. All necessary assumptions need to be stated and linked to the derivations. **(4 marks)**

(b) Verify that the above method discussed in (a) works for the case of $T = 2$. **(4 marks)**

(c) Compare and contrast the method with the method that relies on the assumption regarding the distribution of $\alpha_i|x_{i1}, x_{i2}, ..., x_{iT}$. **(5 marks)**

---

5. Consider the following dynamic probit model:

$$\Pr(y_{it} = 1|y_{i(t-1)}, y_{i(t-2)}, .., y_{i1}) = \Phi(\gamma y_{i(t-1)} + x_{it}^{'}\beta + \alpha_i) \qquad (i = 1, ..., N; t = 1, .., T)$$

where Φ is the standard normal cdf.

(a) What are the 'initial conditions' problems in these type of models? **(3 marks)**

(b) Explain how you would estimate the parameters of the above model addressing the initial conditions problem. Discussion of one method is enough. However, all the necessary assumptions need to be discussed to motivate the method. **(9 marks)**

Section C: Answer ALL questions

---

6. Provide a statement of the Wold Representation Theorem and discuss its meaning and limitations.

(a) Sketch out a proof of the Theorem for univariate stochastic processes. **(6 marks)**

Consider a purely nondeterministic, stationary VARMA(p,q) process $Y_t = (y_{1,t}' \; y_{2,t}')'$

$$\begin{pmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ \Psi_{21}(L) & \Psi_{22}(L) \end{pmatrix} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}. \qquad (2)$$

(b) Assume Y_t to be covariance-stationary. Does $y_{1,t}$ have an independent ARMA representation (not involving $y_{2,t}$)? Provide a proof of your statement. **(6 marks)**

7. In a VAR (possibly with infinite lag order) involving current and past values of a stochastic vector Y_t, the innovations are assumed to be linear combinations of the 'true shocks' of the system and, moreover, the spaces spanned by the innovations and the true 'true shocks' are assumed to coincide, i.e.

$$\nu_t = \Theta \varepsilon_t \tag{3}$$

where Θ is non singular, ν_t are the VAR residuals and ε_t are the 'true shocks'.

(a) Show that necessary condition for Equation 3 to hold is that the number of variables in the VAR equal the number of shocks. **(3 marks)**

(b) Show that condition in Equation 3 is equivalent to requiring that ε_t can be linearly determined from current and lagged values of Y_t:

$$\varepsilon_t = Proj(\varepsilon_t | Y_t, Y_{t-1}, Y_{t-2}, \dots) \ .$$

(5 marks)

Recall that $\nu_t = Proj(Y_t | Y_{t-1}, Y_{t-2}, \dots)$.

(c) How does this condition relate to the condition of invertibility of the moving average? **(5 marks)**

(End)