STATISTICS AND ECONOMETRICS

JUNE 11, 2004 (10:00-13:00)

All questions carry equal marks. Credit will be given for partial answers to questions. The marks allocated to each part of each question are indicated in parentheses.

1. Consider the following regression model

$$y = X\beta + u$$

where y is an $N \times 1$ vector; X is an $N \times k$ matrix of exogenous variables; β is a $k \times 1$ vector of parameters; and u is an $N \times 1$ vector of error terms.

a) (2) Clearly state the necessary assumptions of the classical linear regression model.

b) (6) Derive the OLS estimator for β and its covariance matrix. State and prove the properties of the OLS estimator for β under these assumptions.

c) (12) What are the effects on the OLS estimator for β and its covariance matrix of :

i) Autocorrelation

ii) Heteroskedasticity

iii) Endogeneity of the $X's$.

For each of (i)-(iii) state explicitly what is meant by the above term and provide a proof of your answer.

2. Suppose one is interested in estimating the parameters of the following Cobb Douglas production function

$$q_i = \pi + \alpha \ln k_i + \theta \ln l_i + e_i$$

where q is the log of output; and k and l denote the level of capital and labor inputs. This production function was estimated over a sample of 1000 firms and the following estimates were obtained

$$q_i = .01 + .33 * \ln k_i + .58 * \ln l_i$$

where the standard error for the estimate of α is .10; the standard error for the estimate of θ is .25; and the covariance between the two parameters is .125:

a) (2) Interpret the estimated coefficients of α and θ.

b) (4) Test whether $\alpha = 0$ at the 95 percent level.

c) (4) Test whether $\theta = 1$ at the 95 percent level.

d) (5) Test whether $\alpha = \theta$ at the 95 percent level.

e) (5) Test whether there are constant returns to scale at the 95 percent level.

3. Suppose we have the following model

$$y_i = \beta + \beta_1 x_i + u_i$$

where one suspects that the correlation between x and u is non-zero. Furthermore, suppose we have a series of j variables $z_{1i}, ...z_{ji}$ which are all correlated with the regressor but not the error term.

a) (3) Derive a consistent estimator for β using z_{1i} as an instrument for x.

b) (5) Derive the covariance matrix for this estimator.

c) (5) Show that a more efficient estimate for β can be obtained by using all j instruments.

d) (4) Under what conditions will the estimate in question b) differ from that in question c) ?

e) (3) Describe, and provide, a test of the null hypothesis that $H_0 : E(x_i u_i) = 0$.

4. Consider a random variable X which is uniformly distributed between a and b. where $b > a$.

a) (5) Is $f_X(x) = (b-a)^{-1}$ for $a \leq x \leq b$, and 0 elsewhere, an admissible specification of the probability density function of this random variable over this interval? Why?

b) (3) For this specification of the probability density function, derive the distribution function $F_X(x) = \int_{-\infty}^{x} f_X(z)dz$ of X.

c) (5) Calculate $E(X)$ and $E(X^2)$, and hence $var(X)$, where var stands for the 'variance' of a random variable.

d) (7) When X_1 and X_2 are two independent drawings from a uniform distribution with $a = 0$ and $b = 1$ and $\overline{X} = (X_1 + X_2)/2$, sketch the density function of $\overline{X}$.

5. A sample of n random variables $\{X_i\}$ is independently distributed $N[\mu_X, \sigma_x^2]$.

a) (1) What is $E(X_i)$ for each i?

b) (5) Let $\overline{X} = \sum_{i=1}^n X_i$ denote the sample mean of the variables. Show that the sample mean is an unbiased estimator of the population mean? How does this finding help to estimate μ_X when it is unknown?

c) (8)Let $\vartheta^2 = \sum_{i=1}^n (X_i - \overline{X})^2/n-1$ be the sample variance of X around its sample mean. Using the result that $E(X_i - \mu_X)^2 = \sigma_x^2$, show that $E(\overline{X} - \mu_X)^2 = \sigma_x^2/n$.

d) (6) Finally, show that $E(\vartheta^2) = \sigma_x^2$