29 October 2002

Economics 573
Fall 2002
Examination 2

1. Show that

$$E\left(X_n^2\right) \to 0 \quad \Rightarrow X_n \overset{P}{\to} 0 \qquad \textbf{(1)}$$

Hint:

$$E\left(X_n^2\right) = \int_{-\infty}^{\infty} x^2 dF_n(x) = \int_{x \mid x^2 \geq \varepsilon^2} x^2 dF_n(x) + \int_{x \mid x^2 < \varepsilon^2} x^2 dF_n(x) \qquad \textbf{(2)}$$

2. The classical linear regression model can be written in a variety of forms. We can write it in matrix notation as follows

$$
\begin{aligned}
&I \quad y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\
&II \quad E(\boldsymbol{\varepsilon} \mid X) = 0 \\
&III \quad E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \mid X) = \boldsymbol{\sigma}^2 I \qquad (3) \\
&IV \quad X \textit{ is a nonstochastic matrix of rank } k \\
&V \quad \boldsymbol{\varepsilon} \sim N(0; \boldsymbol{\Sigma} = \boldsymbol{\sigma}^2 I)
\end{aligned}
$$

We sometimes make an additional assumption in order to discuss large sample properties of the least squares estimator.

$$
\begin{aligned}
&IV' \quad \text{X is a nonstochastic matrix of rank k} \\
&\quad \textit{and} \; \lim_{n \to \infty} \frac{X'X}{n} = Q \qquad (4) \\
&\quad \text{Q exists as a finite nonsingular matrix}
\end{aligned}
$$

Briefly discuss the meaning of each assumption.

3. Consider the standard model from problem 2. Define the estimated residuals and fitted values of y as follows

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
y &= X\hat{\beta} + e \qquad \textbf{(5)} \\
\hat{y} &= X\hat{\beta}
\end{aligned}
$$

where $\hat{\beta} = (X'X)^{-1}X'y$. Also define $s^2 = \dfrac{e'e}{n-k}$.

a. Show that $e'e = y'M_X y = \varepsilon' M_X \varepsilon$ where $M_X = (I - X(X'X)^{-1}X')$.

b. Show that $E(e'e) = (n-k)\sigma^2$.

c. Show that s^2 is unbiased.

d. Show that $\frac{(n-k)s^2}{\sigma^2} \sim \chi^2(tr M_X)$. You make invoke theorems without proving them.

4. Given the assumptions of the classical model we can show that

$$\frac{1}{\sqrt{n}} X' \boldsymbol{\varepsilon} \overset{d}{\rightarrow} N(0, \sigma^2 Q) \tag{9}$$

Show that the asymptotic distribution of $\hat{\beta} = (X'X)^{-1} X'y$ is given by

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) &\overset{d}{\rightarrow} N(0, \sigma^2 Q^{-1}) \\ Q^{-1} &= \lim_{n \to \infty} \left(\frac{X'X}{n} \right)^{-1} \end{aligned} \tag{11}$$

5. Consider the following model

$$y_t = e^{\beta_1 + \beta_2 t} L_t^{\beta_3} K_t^{\beta_4} M_t^{\beta_5} \cdot \varepsilon_t$$

a. How would you propose estimating the model using linear regression?

b. Explain a way to test the hypothesis that $\beta_3 + \beta_4 + \beta_5 = 1$ running one regression. Set up the form of the test statistic and show how it is distributed.

c. Describe what multicollinearity is, two possible ways to detect it, and two possible ways to handle it.

d. Explain why the variance of $\hat{\beta}_i$ increases as x_i is more highly correlated with the other variables in the model.

6. Consider the classical linear regression model as outlined in problem 2.

   a. Show that the OLS estimator will be consistent.

   b. Show that the OLS estimator may be biased if the X matrix is stochastic and possibly correlated with ε.

   c. Define the instrumental variables estimator for the model in question 2 if the X matrix is stochastic and possibly correlated with ε. Discuss the conditions that instruments must satisfy and show that this estimator is consistent

7. Consider the following output from a regression relating costs in the electric utility industry to output levels and input prices. The variables are all normalized with respect to the price of fuel to insure homogeneity of degree one in input prices. The variables are defined as follows:

LNPC3	The log of cost normalized by the price of fuel
LNP13	The price of labor normalized by the price of fuel
LNP23	The price of capital normalized by the price of fuel
LNKWH	The size of the plant in kilowatt hours

OLS ESTIMATION
145 OBSERVATIONS   DEPENDENT VARIABLE = LNPC3

R-SQUARE = 0.9316   R-SQUARE ADJUSTED = 0.9301
VARIANCE OF THE ESTIMATE-SIGMA**2 = 0.15348
STANDARD ERROR OF THE ESTIMATE-SIGMA = 0.39176
SUM OF SQUARED ERRORS-SSE= 21.640
MEAN OF DEPENDENT VARIABLE = -1.4842
LOG OF THE LIKELIHOOD FUNCTION(IF DEPVAR LOG) = 147.370

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO	PARTIAL CORR. COEFFICIENT
LNP13	0.59291	0.20457	2.8983	0.2371
LNP23	-0.73811E-02	0.19074	-0.38698E-01	-0.0033
LNKWH	0.72069	0.17436E-01	41.334	0.9611
CONSTANT	-4.6908	0.88487	-5.3011	-0.4077

a. Does size affect cost? How would you test this fact? How many degrees of freedom are there for this problem?

b. What happens to costs as labor prices rise? How would you test whether this coefficient was greater than 0.5?

c. How does capital price seem to affect cost? Is this effect as you would postulate? Is this effect significant? How might you test it.

d. How would you test if the coefficient on the price of labor was the same as the coefficient on the price of capital? What information for constructing such a test is available in the above printout? What is missing?

8. Consider the model in question 2.

   a. What are sufficient statistics and why are they useful?

   b. How does the Neyman-Fisher factorization theorem help us in finding sufficient statistics?

   c. Discuss how one would show that SSE and $\hat{\beta}$ are jointly sufficient statistics for the classical model. You need not actually do so, but provide enough information that it is clear you know how.

   d. How do we know that $\hat{\beta}$ and s^2 are efficient for the classical model?

   e. What are two possible ways to show that an estimator is asymptotically efficient?

9. Consider the following model

$$\begin{aligned} y_t &= \beta x_t + \varepsilon_t, \quad t = 1,2,\ldots,n \\ E(\varepsilon_t \mid x_t) &= 0, \quad \forall t \\ Var(\varepsilon_t \mid x_t) &= \sigma^2, \quad \forall t \\ E(\varepsilon_t \varepsilon_s) &= 0, \quad t \neq s \\ E(\varepsilon_t x_s) &= 0, \quad \forall t, s \end{aligned} \tag{16}$$

a. Find the BLUE estimator of β by direct computation.

b. What is the variance of this estimator?

10. Consider the model in question 9. The least squares estimator is given by

$$\hat{\beta}_n = \frac{\sum_{t=1}^{n} y_t x_t}{\sum_{t=1}^{n} x_t^2}$$

a. Show that

$$\left(\hat{\beta}_n - \beta\right) = \frac{\sum_{t=1}^{n} x_t \varepsilon_t}{\sum_{t=1}^{n} x_t^2}$$

b. Show that

$$\left(\sum_{t=1}^{n} x_t^2\right)^{\frac{1}{2}} \left(\frac{\hat{\beta}_n - \beta}{\sigma}\right) = \frac{\sum_{t=1}^{n} x_t \varepsilon_t}{\left(\sigma^2 \sum_{t=1}^{n} x_t^2\right)^{\frac{1}{2}}} = \frac{\sum_{t=1}^{n} w_t}{\left(\sum_{t=1}^{n} Var(w_t)\right)^{\frac{1}{2}}}$$

where $w_t = x_t \varepsilon_t$.

For your reading enjoyment, the Lindberg-Feller Central Limit Theorem is provided below.

> **Central Limit Theorem (Lindberg-Feller):** Let $X_1, X_2, ..., X_n$ be a sequence of independent random variables. Suppose for every t, X_t has finite mean μ_t and finite variance σ_t^2; Let F_t be the distribution function of X_t. Define C_n as follows:

$$C_n = \left(\sum_{t=1}^{n} \sigma_t^2 \right)^{\frac{1}{2}} \qquad \textbf{(20)}$$

If effect, C_n^2 is the variance of the sum of the X_t. We then consider a condition on "expected value" of squared difference between X_t and its expected value

$$\text{If } \lim_{n \to \infty} \frac{1}{C_n^2} \sum_{t=1}^{n} \int_{|\omega - \mu_t| > \epsilon C_n} (\omega - \mu_t)^2 \, dF_t(\omega) = 0 \quad \forall \ \epsilon > 0$$

$$\text{then} \qquad \textbf{(22)}$$

$$Z_n = \sum_{t=1}^{n} \frac{x_t - \mu_t}{C_n} \sim N(0, 1)$$

where ω is a variable of integration.

c. Assuming the appropriate conditions of the central limit theorem hold (you need not verify them), show that

$$\left(\sum_{t=1}^{n} x_t^2\right)^{\frac{1}{2}}\left(\frac{\hat{\beta}_n - \beta}{\sigma}\right) \overset{d}{\to} N(0,\ 1)$$

d. Assuming the result in part c, argue that

$$\hat{\beta}_n \overset{a}{\sim} N\left(\beta,\ \sigma^2\left(\sum_{t=1}^{n} x_t^2\right)^{-1}\right)$$