PhD/MA. Econometrics Examination
January 2008

Total Time: 8 hours

MA students are required to answer from A and B.
PhD students are required to answer from A, B, and C.

PART A
(Answer TWO questions from Part A)

1. a) Given that p.d.f. of y is $f(y) = 2y \quad 0 \le y \le 1$ and $X = 300Y - 60$, find $E(X)$.
b) The cumulative distribution function for the random variable Y is given by $F(y) = y/5 \quad 0 \le y \le 5$. Find the p.d.f. of y.
c) Prove that $V(Y) = EY^2 - (EY)^2$

2. Consider the following cumulative distribution function (CDF) $F(x) = 1 - e^{-2x}, 0 < x < \infty$, and consider $Y = e^X$

a) Derive the probability density function $f(x)$ from the above CDF
b) Find the mean and variance of X.
c) Using the method of transformations to derive the probability density function of $y, f(y)$. Also give the range of y.

3. a) For the model $y_t = \beta + e_t$, derive the OLS estimator for β using the method of least squares.
b) Derive the estimator for β using the method of maximum likelihood method.
c) Discuss the meaning of BLUE and prove that the least squares estimator has this property.

PART B
(Answer TWO questions from Part B)

4. Given the following linear model in matrix notation

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{U}$$

where the error vector suffers from a heteroscedasticity problem.

a. Using the method of decomposition (e.g., **P** matrix) applied to the variance-covariance matrix, derive the GLS estimator and its variance-covariance matrix.
b. Present the decomposed matrix **P** for the following heteroscedastic form:
$V(u_t) = \exp(k_0 + k_1 Z_t)$

c. What is the benefit of having an exponential form in the variance function? Is it essential? Explain.

d. Assuming that the error vector **U** is white noise and that the **P** matrix is just a set of instruments (e.g., **Z**, a matrix of independent variables), present the GLS estimator of this instrumented model. (Instrumentation is performed to remove the endogeneity in the some of the right hand side variables.)

5. Given the following simultaneous model:

$$Y_t = C_t + I_t + G_t$$
$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + u_t$$
$$I_t = \beta_0 + \beta_1 R_t + \beta_2 Y_{t-1} + v_t$$

where $\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1, \beta_2$ are coefficients. Y(t) = income, C(t) = consumption, I(t) = investment, G(t) = government expenditure, R(t) = interest rate, Y(t-1) = lagged income.

a. Present the endogenous and predetermined variables that are in the model.
b. Use the order and rank condition to see if the equations in the above model are identified.
c. Present the step-by-step method to estimate the equations that are identified.
d. Discuss the generated regressor bias problem.
e. Simplify the above model: ignore the Investment equation, drop Y(t-1)) from the consumption function, drop intercept from the consumption function). Now, for this simple specification, demonstrate the least squares bias for the MPC coefficient "alpha_one".

6. Consider a linear model:

LnHogProduction(t) = a0 + a1*LnPriceHog(t) + a2*LnPriceChicken(t) + a3*LnIncome(t) + u(t)

Present the testing procedures for the following (each separately):
a. HO: a1 = 0
b. HO: a1 = 0, a2 = 0
c. HO: a1=a2;
d. HO: a1 = a2 and a3 = 1

(Present all the necessary steps with null, alternate, formulas, critical table value graph, restriction matrix etc…)

7. Given the following autoregressive model: y(t) = a0 + a1 * y(t-1) + e(t) where e(t) is a white noise error process.

a. Derive the unconditional mean and variance of y (t). That is, E(y(t))=?, and V(y(t)) = ?
b. Repeat (a) for conditional case.
c. How would you test for an Arch process of order two?
d. What is the difference between Arch and Garch?

e. Assuming an ARCH(1) process for the conditional variance function:

h(t) = b0 + b1 * (e(t-1))^2

set up a likelihood function.

PART C
(Answer TWO questions from Part C)

8. Consider the following set of reduced form equations (VAR):
Consider the following set of reduced form equations (VAR):

$$I_{t1}^{US} = \pi_{11} + \pi_{12} I_{t-1,1}^{US} + \pi_{13} I_{t-1,2}^{MEX} + u_t$$
$$I_{t2}^{MEX} = \pi_{21} + \pi_{22} I_{t-1,1}^{US} + \pi_{23} I_{t-1,2}^{MEX} + v_t$$

(Where I's represent inflation for Mexico and US.)

a. Show that the variance-covariance matrix, $Var - Cov(U) = \Sigma \otimes I$
b. Present the FGLS method of estimating this model. (Show your work in details)
c. How would you test the Null: Mexican inflation does not Granger-cause the US inflation? (Set up the null and alternate, present test statistics etc.)
d. Can you demonstrate that under the condition of zero covariance between the two errors, the GLS and OLS estimators will give identical estimators?

9. a. Given the following censored model:

$$Y_t^* = \beta X_t + u_t \quad where \quad Y_t = Y_t^* \; if \; Y_t^* > 0 \quad else \; Y_t = 0 \; if \; Y_t^* \leq 0$$

Set up the log likelihood function for this model.

b. Consider the following research problem in New Mexico housing loan market: study of the factors affecting the credit rating. Assume that there are 1000 individuals (four different races: White, Black, Hispanic, other; three neighborhoods: North Valley, North East, and South Valley), who applied for the loans. First, they were ranked based on their credit rating (1 poor, to 4, excellent).

i. Which model would you use, and why?
ii. Set up the log-likelihood function
iii. After controlling for factors like income, education, and age, it was hypothesized that the credit ranking itself was racially biased. You as an economist were hired to prove or disprove this hypothesis. How would you test and what test would you use?

c. Now, assume that you have information on loan amount given out to applicants. Your hypothesis is that the loan disbursement is racially biased. But, the information on loans is not available above $250,000. How would your set up your econometrics model?

i. Show step by step methods to test the underlying hypothesis (set up the null and alternate hypothesis, present the testing formula, explain which critical table would you use, and present the degrees of freedom etc…)
ii. How would you handle/test the claim that the discrimination was geographically motivated? That is, people from the South valley were "redlined (provided smaller amount)".

10. Given the following nonlinear consumption function:

Consumption(t) = a + b*Income(t)^c + u(t)

Where a, b, and c are unknown parameters.

a. Show the Newton-Raphson method of estimating it using a distribution of your choice.
b. The model above was estimated using the maximum likelihood method and the estimates are presented below

Con(t) = 2.3 + .15 * Inc(t) ^.87

Se(a) = 1.1, Se(b) = .03, Se(c) = .23 , Cov(a,b) = .001, Cov(a,c) = -.002, Cov(b,c) = -.001
Average Consumption = 230.2 Billion$, Average Income = 300.4 Billion$
LnL = -123.54

a. Calculate the marginal propensity to consume (MPC)
b. Calculate income elasticity.
c. Calculate the standard error of the MPC estimate derived in a.
d. Using the Wald test, test the hypothesis that MPC = .5

11 a. Given the following Moving Average model:

$$y_t = e_t + \theta_1 e_{t-1}$$

Present the autocorrelation functions (a.c.f.) up to lag 3.

b. Why is it called a short memory model?
c. Assuming $\theta_1 = .67$, draw the a.c.f. and p.a.c.f. graphs.
d. How would you estimate this model?
e. Explain the concept of cointegration and error correction model.