PhD/MA Econometrics Examination
January 2011

Total Time: 8 hours

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

PART A
(ANSWER ANY TWO)

Q. 1 (a) For $g(y) = ky(1-y) \quad 0 \le y \le 1$ find k so that $g(y)$ is a probability density function (p.d.f).

(b) The cumulative distribution function for random variable *Y* is given by $F(y) = y/3 \quad 0 \le y \le 3$. Find the p.d.f. for *Y*, $f(y)$. Graph $f(y)$ and discuss some of the characteristics of this p.d.f.

Q. 2 The random variables Y_i, $i = 1,...,n$ are independently and identically distributed with probability density function (p.d.f.):

$$f(y) = \frac{1}{\beta} e^{-y/\beta} \quad y > 0$$

(a) Give (i) the joint p.d.f. (ii) the likelihood function and (iii) the loglikelihood function of the random variables Y_i, $i = 1,...,n$.

(b) Find the maximum likelihood estimator of the parameter β.

Q. 3

(a) Derive the least squares estimator of β, $\hat{\beta}$, for the model $y_i = \beta + u_i$. Assume a sample size of *n*.

(b) Show that residuals sum to zero, that is, show that $\sum_{i=1}^{n} \hat{u}_i = 0$.

(c) Find the mean and variance of $\hat{\beta}$.

PART B
(ANSWER ANY TWO)

Q. 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 of the beta-vector 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. Using a simple model of the form y(t) = a0 + b0*x(t) + u(t), present a weighted least squares method based on the variance form as stipulated in "b".
e. Present the log-likelihood function for this multiplicative model (combining expressions in "b" and "c")

Q. 5 Given the following ARCH model:

$$y_t = \phi_1 y_{t-1} + e_t \qquad e_t = u_t \sqrt{\alpha_0 + \alpha_1 e_{t-1}^2}$$

where u_t is standard normal.

a. Derive the unconditional variance $V(e_t)$.
b. Derive the conditional variance $V(e_t|e_{t-1})$.
c. How would you test for an ARCH(2) process? Show all the steps.
d. Set up a log-likelihood function for this model.

Q. 6. Given the following simple national income account model:

$$Y(t) = C(t) + I(t) + G(t) \quad (1)$$
$$C(t) = a + b\,Y(t) + c\,C(t-1) + U(t) \quad (2)$$
$$I(t) = d + e\,R(t) + f\,Y(t-1) + V(t) \quad (3)$$

$Y(t)$ = GDP, $C(t)$ = consumption, $G(t)$ = government expenditure, $C(t-1)$ = one period lagged consumption, $Y(t-1)$ = lagged GDP, $U(t)$ and $V(t)$ = error terms, a, b, c, d, e, and f are coefficients.

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 two-step method (2SLS) to estimate the consumption function. Discuss the estimation method for the investment equation.
d. Discuss briefly the generated bias problem.
e. Make the above equation simple.

$$Y(t) = C(t) + I(t) \quad (1)$$
$$C(t) = b\,Y(t) + u(t) \quad (2)$$

f. Prove that the OLS estimator of b is not consistent.

PART C
(ANSWER ANY TWO)

Q. 7. Consider the following research problem in New Mexico housing loan market. 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. They were ranked based on their credit rating (1 poor, to 4, excellent). Other socio-economic information was also available: age, income ($ per month), and the number of schooling years. It was hypothesized that the loan amount was a function of the race of the borrower. *You as an economist were hired to prove or disprove the hypothesis that the minority borrowers were given smaller loan amount.* But the problem is that the bank did not want to reveal the loan amount for those who borrowed less than 5K. This information was recorded as missing. [Hint: the dependent variable (y or y*) is in $.]

a. Should you use the OLS method? Why or why not?
b. Set up the appropriate model by spelling out all the steps. (latent equation, coefficients, variables etc.)
c. Present the log-likelihood function.
d. Show as to how you would test the hypothesis (using the Wald or the Likelihood Ratio) about the link between the race and the magnitude of the loan amount as stipulated in the question above.
e. Set up the log likelihood function for a case when the data would be censored from below as well as from above also. That is, the amount more than >150K would also be recorded as missing.

Q. 8. Given the following Moving Average model:

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

a. Present the autocorrelation functions (a.c.f.) up to lag 3.
b. Why is it called a short term memory model?
c. Assuming $\theta_1 = .67$, draw the a.c.f. and p.a.c.f. graphs.

Q. 9. Consider the following non-linear model describing the hourly video viral data showing the link between the internet youtube discussion hits (y) on a social media blog about the news event of birds dropping from the sky. X is the hourly indicator 1st hour, 2nd, hour, 3rd hour etc…

$$y_t = \frac{b_0}{[1 + e^{b_1(x_t - b_2)}]} + u_t$$

Where the subscript t stands for time, Y is the internet , and x is the news outlet coverage. For simplicity, assuming that every hour the news outlet covering the

a. Assuming b_0 = 70.; b_1 = -1.34 ; b_2 = 0. Plot Y versus X.
b. Assuming b_0 = 72; b_1 = -1.34 ; b_2 = 18. Plot Y versus X. What is the difference between a and b?
c. Present all the necessary steps to estimate this model using the Newton Raphson method.
d. Also, discuss and present the expression to formulate the variance-covariance matrix of the estimated coefficients.