PhD/MA. Econometrics Examination
January 14, 2009

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 any two from part A)

Q. 1

The random variables Y_i, $i = 1,...,n$ are independently and identically distributed with p.d.f.

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

(a) Find the cumulative distribution function (c.d.f.) for y.

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

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

Q. 2

A recent graduate has submitted his application to the World Bank for a position in the Young Professional Program. He knows the Bank hires 4% of its applicants. Only some of the applicants receive an interview. It is known that among all the applicants hired 98% receive interviews, the other 2% being related to Division Chiefs! Furthermore, among all applicants not hired only 1% are interviewed.

(a) Define events and organize information in the text above into probability statements.

(b) State Bayes' theorem.

(c) Our recent graduate was just called for an interview and asks you to tell him the probability that he will be hired. Based on your calculations using Bayes' theorem, what is your answer?

Q. 3

As you walk into your econometrics core exam, a friend bets you $10 that she will outscore you on the exam. Let X be a random variable denoting your winnings. X can take on the values 10, -10, or 0 (you tie on the exam). You know that the probability density function of X, p(x) depends on whether she studies for the exam or not. Let Y=0 if she studied and Y=1 if she did not. Consider the following joint distribution table

			Y	
		0	1	$p(x)$
X	-10	0.18		
	0	0		0.3
	10		0.45	
	$p(y)$		0.75	

(a) Fill in the missing elements in the table.
(b) Given your result in (a), compute E(X) and E(Y).
(c) Would you take the bet? Why?

Q. 4

The following regression equation is estimated:

$$\hat{sleep} = 3{,}638.25 - .148\ totwrk - 11.13\ educ + 2.20\ age$$
$$(112.28) \quad (.017) \qquad (5.88) \qquad (1.45)$$
$$n = 706,\ R^2 = .113$$

where sleep = minutes slept per week, totwrk = minutes worked per week, and educ (= education) and age are measured in years. The standard errors are reported in the parenthesis and the sample size is 706.

(a) Test whether educ or age are individually significant at the 5% level against a two-sided alternative. Show your work.

(b) Dropping educ and age from the equation gives:

$$\hat{sleep} = 3{,}586.38 - .151\ totwrk$$
$$(38.91) \quad (.017)$$
$$n = 706,\ R^2 = .103$$

Test whether educ and age are jointly significant in the original equation at the 5% level.

(c) Suppose that the sleep equation contains heteroskedasticity. What does this mean about the tests computed in parts (a) and (b)?

PART B
(Answer any two from part B)

Q. 5

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.)

Q. 6

Given the following ARCH model:

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

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 steps.
d. Set up a log-likelihood function for this model.

Q. 7

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

$$y_{t1} = \alpha_{11} + \alpha_{12} y_{t-1,1} + \alpha_{13} y_{t-1,2} + u_t$$
$$y_{t2} = \alpha_{21} + \alpha_{22} y_{t-1,1} + \alpha_{23} y_{t-1,2} + v_t$$

(Where y(t1) = US inflation, y(t2) = Canadian inflation)

A. Show that the variance-covariance matric, $Var - Cov(U) = \Sigma \otimes I$
B. Present the FGLS method of estimating this model.
C. Your research buddy Mark insists upon stacking the two equations (i.e., US and CANADA data) and estimating it as one regression. Under what conditions and/or assumptions would you allow him to do so?
D. How would you test the Null: Canadian inflation does not Granger-cause the US inflation? (Set up the null and alternate, present test statistics etc.)

Q. 8

Consider the following stock-adjustment model:

LnY(t)* = a0 + a1 LnX(t) + U(t) *(1)*

where Y(t)* is desired inventory of cars in b$; X(t) is the oil price ($ per gallon)

The stock-adjustment mechanism is as follows:

[Y(t)/Y(t-1)] = [Y(t)*/Y(t-1)]^*d* 0<*d*<1 *(2)*

(Hints: take the log of the equation *2* to make it operational.)

a. Combine 1 and 2 to come up with a Koyck type distributed lag model and discuss the estimation method. What is the interpretation of *d*?

b. Given the following distributed lag type model for the annual car inventory data (1980 – 2004):
LnY(t) = B0 + B1 LnY(t-1) + B2 LnX(t) + V(t)
Where Y(t) = Car Sales $ billion; X(t) = Advertisement Expenses ($).
Estimates: b0 = 1.5, b1 = .52, b2 = .32

c. Calculate the long-run multiplier effect of advertisement on car sales.
d. Calculate the half life value and interpret the result.
c. Calculate the impact multipliers at lag 0, 1, and 2.

PART C
(Answer any two from part C)

Q. 9

Quantity_Beef(t) = a0 + a1*Price_Beef(t) + a2*Income(t) + u(t)
Quantity_Beef(t) = b0 + b1*Price_Beef(t) + b2*Price_Chicken + b3*MadCowNews(t) + v(t)

a. Present the above model in a vector notation – Y, Z etc..
b. Perform the general identification test for this model.

Writing the above model generically in vector notation – Y, Z etc., derive the following:

c. Var-Cov(U), where U is the stacked vector of errors, u and v.
d. Present the step-by-step method of 3SLS.
e. Why is it called a GLS method? Show that the 2 SLS estimator is inconsistent and biased (Hint: You can do this in a generic notation).

Q. 10

Consider the following research problem in New Mexico housing loan market. Assume that there are 1000 individuals who applied for the loans. The dependent variable is the credit rating (1 poor, to 4, excellent).

Variables of interest are: *HISPANIC*, *BLACK*, *INCOME*, *MALE*, *EDUCATION*, *AGE*

a. To assess the impact of various factors on the credit rating, which model would you use, and why?
b. Set up the model in a regression equation format (e.g., spell out **equation with variables and coefficients**,
c. Write out the necessary steps to derive the log-likelihood function
d. Discuss the estimation algorithm of your choice (e.g., numerical optimization of your choice.)

d. 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)

Influence of race: *HISPANIC and BLACK* jointly

Q . 11

Given the following nonlinear consumption function:

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

Where a, b, and c are unknown parameters.

a. You may express this function generically --con(t) = f(a,b,c,Inc(t)) + u(t) and show how you might estimate using the method of non-linear least squares.

OR

b. Show the Newton-Raphson method of estimating it using a distribution of your choice.

The model above was estimated using the maximum likelihood method and the estimates are presented below

Con(t) = 2.3 + .75 * 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$, Average Income = 320.4 Billion$
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

Q. 12

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.
d. How would you estimate this model? Econometrics II: 510