UCLA Spring 99 **Quantitative Methods**

Core Exam

The exam consists of three parts. For each part answer two of the three questions. Use separate blue books for the three parts.

Part I

Choose two out of the following three questions.

1. Let the random variable X have a binomial distribution with $Pr(X = 1) = 1/4$. Conditional on $X = x$, the random variable Y has an exponential distribution with mean $\mu \cdot (x + 1)$.

   (a) Find the conditional probability that $X = 1$ given $Y = y$.

   (b) Calculate the mean and variance of Y (marginal, not conditional on X).

   (c) Calculate the Cramer-Rao bound for μ.

   (d) Find, if possible, an unbiased estimator for μ with variance equal to the Cramer-Rao bound.

   (e) Let $(x_1, y_1), \ldots, (x_N, y_N)$ be a random sample from this distribution. Find a one-dimensional sufficient statistic for μ.

2. Let $X_1, X_2, \ldots, X_N$ be independent random variables with normal distributions with mean μ and variance σ^2.

   (a) Find the moment generating function for a random variable with a normal distribution with mean μ and variance σ^2.

   (b) Find the maximum likelihood estimator for σ^2.

   (c) Find an unbiased estimator for σ^2.

   (d) Show that the difference between the mvue and mle converges to zero in probability as the sample size goes to infinity.

3. Let X be a random variable with probability density function

$$f_X(x; \mu) = \frac{1}{\mu} \exp(-x/\mu),$$

   for $x > 0$ and zero elsewhere.

(a) Calculate the mean and variance of X.

(b) Calculate the mean and variance of X conditional on $X < 8$.

(c) Let $x_1, x_2, \ldots, x_N$ be a random sample from this distribution, with $N = 20$, $\sum x = 95$, and $\sum x^2 = 590$. Calculate the maximum likelihood estimate.

(d) Estimate the probability that $X < 8$ and its asymptotic variance.

(e) Test the hypothesis that $\mu = 4$ at the 10% level using a likelihood ratio test.

(f) Test the same hypothesis using a Lagrange multiplier (score) test.

Part II

Choose two out of the following three questions.

1. Suppose you have n independent observations $\{(y_i, X_i)\}_{i=1}^n$, where y_i is a scalar random variable, and X_i is a $(1 \times k)$ random vector. The observations can be grouped into G subgroups, not necessarily of the same size, for each one of which,

$$y_{ig}| \, X_i \sim N\left(X_i\beta, \sigma_g^2\right)$$

where $g = 1, ..., G$.

(a) Compute the MLE's of β and σ_g^2.

(b) Construct the Likelihood Ratio test statistic for groupwise heteroskedasticity, that is, for testing the hypothesis that $\sigma_1^2 = ... = \sigma_G^2$. Make sure that you state the asymptotic distribution of the test statistic.

2. For each one of the following claims show whether they are true or false.

(a) The R^2 of a regression does not change if we add to the dependent variable a constant and/or if multiply the dependent variable by a constant.

(b) The OLS estimator is consistent when the dependent variable in a regression is subject to classical measurement error.

(c) The OLS estimator is consistent when the independent variables in a regression are subject to classical measurement error.

NOTE: For claims (b) and (c) it suffices to consider a bivariate regression model:

$$y_i^* = \alpha + \beta x_i^* + \varepsilon_i$$

where instead of observing y^* or x^* in (b) and (c) respectively, we observe:

$$\begin{aligned} y_i &= y_i^* + v_i \\ x_i &= x_i^* + v_i \end{aligned}$$

where $E(\varepsilon_i) = E(v_i) = E(\varepsilon_i x_i^*) = E(v_i x_i^*) = E(v_i y_i^*) = E(v_i \varepsilon_i) = 0$.

3. Suppose you have n independent observations $\{(y_i, x_i)\}_{i=1}^{n}$, where y_i and x_i are scalar random variables. The density of y_i conditional on x_i is:

$$\frac{(\beta + x_i)^{-\rho}}{\Gamma(\rho)} y_i^{\rho-1} e^{-y_i/(\beta+x_i)}$$

Thus, y_i is Gamma distributed. Describe 3 ways for estimating β and ρ. Recall that for a Gamma distribution with density

$$\frac{\lambda^{\rho}}{\Gamma(\rho)} y^{\rho-1} e^{-\lambda y}$$

the mean is ρ/λ and the variance is ρ/λ^2.

Part III

Choose two out of the following three questions.

1. Consider the following two-equation model relating schooling (s_i) and log earnings (y_i) to an exogenous variable x_i, measuring distance (in hundreds of miles) from the nearest state university. (All variables are assumed to be recorded in deviations from means.)

$$\begin{aligned} s_i &= \beta x_i + \epsilon_i \\ y_i &= \delta s_i + u_i \end{aligned}$$

Assume that (ϵ_i, u_i) is jointly normally distributed. The reduced form is

$$\begin{pmatrix} s_i & y_i \end{pmatrix} = x_i \Pi + v_i',$$

where

$$E(v_i) = 0, \qquad E(v_i v_i') = \Omega$$

The following least-squares estimates are obtained from a sample of data:

$$\hat{\Pi} = \begin{pmatrix} -2 & -.18 \end{pmatrix}, \qquad \hat{\Omega} = \begin{pmatrix} 2.0 & .78 \\ .78 & .6242 \end{pmatrix}.$$

(a) Examine the identification of both equations (directly, or by using order and rank conditions).

(b) Obtain the indirect least squares estimate of δ. If you were to use x_i as an instrument and estimate δ using the IV estimator, what would your result be? How would the standard errors of the ILS and IV estimators compare?

(c) You are told that for individual $(n+1)$, $x_{n+1} = 1$ and $s_{n+1} = .5$. Provide a forecast for y_{n+1}.

Note: For the bivariate normal distribution:

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \right),$$

the marginal distributions are given by

$$\begin{aligned} x &\sim N(\mu_x, \sigma_x^2), \\ y &\sim N(\mu_y, \sigma_y^2). \end{aligned}$$

and the conditional distribution of y given x is

$$y|x \sim N(\alpha + \beta x, \sigma_y^2(1-\rho^2)).$$

where $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ is the correlation between x and y, and

$$\begin{aligned} \alpha &= \mu_y - \beta\mu_x \\ \beta &= \frac{\sigma_{xy}}{\sigma_x^2} \end{aligned}$$

2. Consider a linear model with a single regressor in mean-deviated form:

$$y_i = x_i\beta + u_i, \qquad i = 1, \ldots, n.$$

   (a) What are the properties of OLS with $E(x_i u_i) \neq 0$?

   (b) A researcher proposes the following test of orthogonality:

      i. Calculate the OLS residual $\hat{u}_i = y_i - x_i b$.

      ii. Define $\lambda = \sum_{i=1}^{n} x_i \hat{u}_i$.

      iii. If $\lambda \approx 0$, do not reject $H_0 : E(x_i u_i) = 0$. Otherwise, reject H_0.

   What will happen if the researcher uses this test?

   (c) The researcher decides to try to find an instrumental variable and use IV to estimate β. Consider the following instrument: use the computer to randomly generate

$$z_i \overset{\text{i.i.d.}}{\sim} N(0,1).$$

   Will the IV estimator based on z_i be consistent? Justify your answer by taking the probability limit of the IV estimator.

3. An individual decides to migrate if his/her lifetime expected utility from staying at the current location, u_{ip}, is less than or equal to his/her lifetime expected utility from migrating, u_{im}, minus the migration cost c_i. Assume that

$$u_{ip} = x'_{ip}\beta_p + \epsilon_{ip}, \tag{1}$$

$$u_{im} = x'_{im}\beta_m + \epsilon_{im}, \tag{2}$$

$$c_i = z_i\gamma + u_i. \tag{3}$$

where x_{ip}, x_{im}, and z_i are observed variables. Suppose x_{ip} and x_{im} include the individual's education, experience, age, race, and gender, and local unemployment rates and average wages in the different locations. Variables in z_i include whether the individual is self-employed and whether he/she has recently changed industry of employment. The model disturbances satisfy

$$\begin{pmatrix} \epsilon_{ip} \\ \epsilon_{im} \\ u_i \end{pmatrix} | x_{ip}, x_{im}, z_i \overset{\text{i.i.d.}}{\sim} N(0, \Sigma),$$

where Σ is a symmetric positive definite matrix. We observe x_{ip}, x_{im}, z_i, and an indicator for whether or not the individual chose to migrate.

(a) Construct a probit model for the migration decision of individual i.

(b) Write down the log-likelihood function for a sample of n individuals (assuming independence across observations).

(c) Describe as precisely as possible the maximum likelihood estimator for this model. Are all the coefficients in equations 1, 2, and 3, consistently estimable?

(d) What are the asymptotic properties of the maximum likelihood estimator in part (c)? (You do not need to provide formal proofs, but justify your reasoning.)