ECONOMETRICS - QUALIFIER

July 2007

ANSWER ONLY 5 QUESTIONS.

1. Consider a probit model

$$y_i = \begin{cases} 0 & if \quad y_i^* = x_i\beta + \varepsilon_i < 0, \\ 1 & if \quad y_i^* = x_i\beta + \varepsilon_i \geq 0. \end{cases}$$

   where ε_i is $N(0,1)$. The log-likelihood function is given by

$$\ln L(\beta) = \sum_{i=1}^{n} y_i \ln \Phi(x_i\beta) + \sum_{i=1}^{n} (1 - y_i) \ln [1 - \Phi(x_i\beta)],$$

   where $\Phi(\cdot)$ is the cdf of a standard normal random variable.

   (a) Show that $Prob([y_i = 1|x_i] = \Phi(x_i\beta)$.

   (b) Discuss how to estimate β by the maximum likelihood method. You need to derive the first-order condition for estimating β.

   (c) Suppose that $y_i = 1$ if individual i owns a car, 0 otherwise, $(x_i = 1, income_i, age_i)$. If the income of an individual goes up by \$1500, what is the change of probability in owning a car?

   (d) Discuss how you can test the null hypothesis that an individual's age does not affect the probablity of owning a car. You need to give the test statistics, state the asymptotic null distribution of your test statistic, and discus under what condition you will reject the null hypothesis.

2. Consider a Cobb-Douglas production function

$$y_i = Cx_{i1}^{\beta_1}x_{i2}^{\beta_2} \quad i = 1, \ldots, n$$

   where x_{i1} is firm i's labor input, and x_{i2} is firm i's capital input.

   (a) Discuss how one can obtain a linear regression model by taking logarithm of the variables.

   (b) Discuss how one can test the restriction of $\beta_1 + \beta_2 = 1$. You need to give a test statitic, state its null distribution, etc.

   (c) Why one might be interested in testing the above null hypothesis? Is there an economic interpretation of the above null hypothesis?

   (d) Discuss how you can test the above null hypothesis against the one-sided alternative: H_1: $\beta_1 + \beta_2 < 1$.

3. Let $X_1, X_2, \ldots$ be an independent and identically distributed sequence of random variables such that $\mathbf{E}X_1^8 < \infty$, and let $\bar{X}_n = \sum_{i=1}^n X_i/n$ be the sample mean. Show that the sample kurtosis

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right)^2}$$

converges in probability to the population kurtosis

$$\frac{\mathbf{E}(X_1 - \mu)^4}{\left(\mathbf{E}(X_i - \mu)^2\right)^2},$$

where $\mu = \mathbf{E}X_1$ and $\mathbf{E}(X_1 - \mu)^2 > 0$. Make your derivation as rigorously as possible and provide every detail.

4. Consider the regression model

$$y_i = \alpha x_i + u_i$$

where (x_i, u_i) is iid with $\mathbf{E}(u_i|x_i) = 0$ and $\mathbf{E}(u_i^2|x_i) = \sigma^2 x_i^2$. We assume that $\mathbf{E}x_i^4 < \infty$. Let $\hat{\alpha}$ be the OLS estimator from the regression of y_i on x_i.

(a) Show that $\hat{\alpha}$ is consistent.F

(b) Find the limiting distribution of $\sqrt{n}(\hat{\alpha} - \alpha)$.

Now let $\tilde{\alpha}$ be the GLS estimator which "corrects" the conditional heterogeneity in the errors. Clearly, $\tilde{\alpha}$ can be obtained from the OLS method from the transfromed regression

$$\frac{y_i}{|x_i|} = \alpha\frac{x_i}{|x_i|} + \frac{u_i}{|x_i|}$$

(a) Find the limiting distribution of $\sqrt{n}(\tilde{\alpha} - \alpha)$.

(b) Compare the asymptotic variances of $\hat{\alpha}$ and $\tilde{\alpha}$.

5. Consider a latent variable regression equation

$$y_t{}^* = X_t\beta + u_t', u_t, \quad \sim i.i.d.N(0, \sigma^2)$$

where y_t^* is not fully observed. Suppose that we have observations y_t

$$y_t = \begin{cases} y_t^* & if \quad y_t^* > \gamma \\ 0 & if \quad y_t^* \geq \gamma \end{cases}.$$

where γ is an unknown constant.

(a) Consider the OLS estimator $\widehat{\beta}$ from a regression of y_t on X_t. Show whether $\widehat{\beta}$ is biased or not and present an intuition behind your result.

(b) Consider the OLS estimator $\widetilde{\beta}$ from a regression of positive y_t on X_t(excluding samples with $y_t = 0$). Show whether $\widetilde{\beta}$ is biased or not and present an intuition behind your result.

(c) Specify the log likelihood function in terms of $\gamma, \alpha = \beta/\sigma$and $\tau = 1/\sigma$and derive the first-order conditions.

(d) Present the decomposition of the marginal effect of a regressor X_{tk} on $E(y_t)$ and give a proper interpretation to each term.

6. The maximum likelihood estimation with a known density function $f(x_t; \theta)$ can be interpreted as

$$\max_{\theta} \sum_{t=1}^{n} \ln(p_t), \quad \textit{subject to} \quad p_t = f(x_t \colon \theta)$$

(a) What modification does the maximum empirical likelihood (MEL) estimation introduce into this interpretation?

(b) What is the difference between $MELE$ and $MME/GMME$?

(c) We wish to estimate two parameters θ_1 and θ_2 by using a random sample $y_t, t = 1, 2, \ldots, n$, which satisfies the moment conditions

$$E(y_t - \theta) = 0$$
$$E[(y_t - \theta_1)^2 - \theta_2] = 0.$$

(d) Derive the MME and $MELE$ of θ_1 and θ_2.