Introduction to Econometrics

Final Examination

Fall 2005

Please answer all of the questions and show your work. If you think a question is ambiguous, clearly state how you interpret it before providing an answer.

1. (30 points) A latent variable y^* is related to a set of covariates X by

$$y^* = X\beta + \varepsilon,$$

where ε is independently, normally distributed with mean 0 and variance σ_ε^2. Recall that the probability density function (p.d.f.) of ε in this case is

$$f(\varepsilon) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp(-\frac{1}{2\sigma_\varepsilon^2}\varepsilon^2).$$

In all of the cases below, assume that you have access to a random sample of N observations, $\{y_i, X_i\}_{i=1}^N$.

(a) Say that

$$y_i = \begin{cases} 1 & \text{if} \quad y_i^* > 0 \\ 0 & \text{if} \quad y_i^* \leq 0 \end{cases}.$$

Write down the log likelihood function for the sample. Determine whether β and σ_ε^2 are separately identified.

(b) Say that

$$y_i = \begin{cases} y_i^* & \text{if} \quad y_i^* > 0 \\ 0 & \text{if} \quad y_i^* \leq 0 \end{cases}.$$

Write down the log likelihood function for the sample in this case. Are β and σ_ε^2 separately identified in this case? Why?

(c) Say that

$$y_i = y_i^*.$$

Define the maximum likelihood estimator of β for this case.

2. (30 points) Consider the panel data regression model

$$y_{it} = \beta x_{it} + \varepsilon_{it}$$

where

$$\varepsilon_{it} = \alpha_t m_i + u_{it},$$

and where u_{it} is i.i.d. with mean 0 and variance σ_u^2, $\alpha_1 = 1$. You have access to a random sample of observations from the population for N individuals and 2 time periods, $\{y_{it}, x_{it}\}_{i=1,\ldots N;t=1,2.}$. Assume that x_{it} is a scalar for simplicity, and that all variables are measured as deviations from their means.

(a) Assume that m_i is i.i.d. with mean 0 and variance σ_m^2 for all i. Derive consistent estimators for β, σ_u^2, σ_m^2, and α_2, if they exist.

(b) Now assume that the distribution of m_i conditional on x_{i1}, x_{i2} is unknown. In particular, it is not (necessarily) the case that $E(m_i|x_{i1}, x_{i2}) = 0$.

i. If you know that $\alpha_2 = 1$, explain how you would obtain a consistent estimator of β, assuming that $x_{i2} \neq x_{i1}$ for all i.

ii. If α_2 is unknown, explain whether you could obtain a consistent estimator of β.

3. (20 points) A population of unemployed job searchers has completed spell lengths given by a negative exponential distribution, that is, the p.d.f. and c.d.f. are

$$\begin{aligned} f(t) &= \alpha \exp(-\alpha t),\ t > 0,\ \alpha > 0. \\ F(t) &= 1 - \exp(-\alpha t). \end{aligned}$$

A random sample of N individuals are drawn from this population. For unemployment spells that lasted less than 3 months, the exact duration of the spell was not recorded. For those lasting more than 3 months, t is measured exactly.

(a) Write the log likelihood for this sample as a function of α. Derive the maximum likelihood estimator of α.

(b) If you *exclude* the censored observations (i.e., those for which $t < 3$), the sample mean of unemployment spells is 5. Using this information, is it possible to provide a bound on the maximum likelihood *estimate* you would obtain from the maximization of the log likelihood you derived above? If so, what is it?

4. (20 points) A cross-sectional population regression model is given by

$$y_i = \beta x_i + \varepsilon_i,$$

where y_i and x_i have mean 0 in the population, and ε_i is independently and identically distributed (i.i.d.) with mean 0 and variance σ_ε^2. You have access to a random sample of observations on $\{y_i, x_i^*, z_i\}_{i=1}^N$. The variable x_i^* is a "noisy" measure of x_i, that is,

$$x_i^* = x_i + u_i,$$

where u_i is i.i.d. with mean 0 and variance σ_u^2. In the population, the following moment properties hold:

$$\begin{aligned} E(zx) &\neq 0 \\ E(zu) &= 0 \\ E(z\varepsilon) &= 0. \end{aligned}$$

Determine whether it is possible to obtain a consistent estimator of β given access to the sample information. If not, why not? If so, what is it?