UC Berkeley **Spring Semester 2003**
Economics 241A **Guido Imbens**

Midterm Exam

Monday March 17$^{\text{th}}$, 4.00-5.30pm

All three questions will be weighted equally. Within each question the different parts are weighted equally. Good luck!

1. Suppose you have data on unemployment durations T_i (in weeks) and covariates X_i for N individuals. Assume that the hazard function is

$$h(t|x) = \exp(\beta' x),$$

and that all durations are independent.

(a) Suppose you observe the complete durations. What is the log likelihood function?

The density is

$$f_{T|X}(t|x;\beta) = e^{x'\beta} \exp(-te^{x'\beta}),$$

and the log likelihood function is therefore

$$L(\beta) = \sum_{i=1}^{N} x_i'\beta - t_i \exp(x_i'\beta).$$

(b) Suppose you obtain the data in the following way. On March 1st you go to an unemployment registry and record information about the next 100 individuals registring as unemployed, including their covariates and the date they show up (not enough people show up on the first day). You follow these people till they find a job, or till March 31st, whichever comes first. What is the log likelihood function in this case?

Let c_i be the censoring time (the length of time between the date of registring and March 31st), and d_i be an indicator that is equal to one for failure times and zero for censoring times. Let $y_i = \min(c_i, t_i)$ be the minimum of the failure time and the censoring time. Let $S(t|x;\beta)$ be the survivor function. Then the likelihood function is

$$\mathcal{L}(\beta) = \prod_{i=1}^{N} S(y_i|x_i;\beta) \cdot h(y_i|x_i;\beta)^{d_i}$$

$$= \prod_{i=1}^{N} \exp(-y_i e^{x_i'\beta}) \cdot \left(e^{x_i'\beta}\right)^{d_i},$$

and the log likelihood function is

$$L(\beta) = \sum_{i=1}^{N} d_i x_i'\beta - y_i \exp(x_i'\beta).$$

(c) Suppose instead you go to the unemployment registry and randomly sample 100 people currently unemployed. You record their information, including how long they have been unemployed already. Two weeks later you return and record for each of the 100 individuals whether and when they found a job. What is the log likelihood function in this case?

Let d_i again indicate whether we observe a completed spell. If $d_i = 1$ we observe the length of time already unemployed prior to the start of the observation period, s_i, and the complete duration t_i. If $d_i = 0$ we observe the length of the initial period of unemployment, s_i, and the censoring time, $c_i = s_i + 14$ (in days). Let $y_i = \min(t_i, c_i)$ be the minimum of the censoring and failure time. Then the likelihood function is

$$\mathcal{L}(\beta) = \prod_{i=1}^{N} \frac{S(y_i|x_i;\beta)}{S(s_i|x_i;\beta)} \cdot h(y_i|x_i;\beta)^{d_i} = \prod_{i=1}^{N} \exp(-(y_i - s_i)e^{x_i'\beta}) \cdot (\exp(x_i'beta))^{d_i},$$

with the corresponding log likelihood function

$$L(\beta) = \sum_{i=1}^{N} d_i \cdot x_i'\beta - (y_i - s_i) \cdot \exp(x_i'\beta).$$

2. Suppose that given covariates X_{i1} (age), X_{i2} (education) and X_{i3} (earnings) the probability of commuting by car ($Y_i = 1$) rather than by public transportation ($Y_i = 0$) is

$$Pr(Y_i = 1|X_{i1}, X_{i2}, X_{i3}) = \frac{\exp(\beta_0 + \beta_1 \cdot X_{i1} + \beta_2 \cdot X_{i2} + \beta_3 \cdot X_{i3}}{1 + \exp(\beta_0 + \beta_1 \cdot X_{i1} + \beta_2 \cdot X_{i2} + \beta_3 \cdot X_{i3}}.$$

(a) What is the likelihood function given data on commuting behavior and the three covariates from a random sample?

Let $x_i'\beta = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i3}$. The likelihood function is

$$\mathcal{L}(\beta) = \prod_{i=1}^{N} \left(\frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)} \right)^{y_i} \cdot \left(\frac{1}{1 + \exp(x_i'\beta)} \right)^{1-y_i}.$$

(b) Suppose we wish to test the null hypothesis that $\beta_3 = 0$. Describe how a likelihood ratio test would work.

What is the distribution of the test statistics under the null hypothesis?

First estimate the model by maximum likelihood imposing the restriction. This leads to a restricted estimator $\hat{\beta}_r$, with $\hat{\beta}_{r3} = 0$. Then estimate the model without imposing the restriction, leading to the unrestricted maximum likelihood estimator $\hat{\beta}_u$. The likelihood ratio test statistic is

$$2 \cdot \left(\log \mathcal{L}(\hat{\beta}_u) - \log \mathcal{L}(\hat{\beta}_r) \right).$$

Under the null hypothesis this statistic is distributed as a chi-square random variable with one degree of freedom.

(c) Describe how a Hausman test would work in this case. What is the distribution of the test statistics under the null hypothesis?

For a Hausman test we estimate, using the information matrix, the variance of the restricted and unrestricted maximum likelihood estimators $\hat{\beta}_r$ and $\hat{\beta}_u$. Let these variances be denoted by V_r and V_u. The Hausman test statistic is then

$$H = (\hat{\beta}_u - \hat{\beta}_r)'(V_u - V_r)^{-g}(\hat{\beta}_u - \hat{\beta}_r).$$

The degrees of freedom will depend on the rank of $V_u - V_r$. In this case the rank will be at most one (and in general for this case will be equal to one). In the implementation one has to be careful to take the generalized inverse of the variance since $V_u - V_r$ can be singular.

3. Consider estimating θ by the Generalized Method of Moments, using the moment restrictions

$$\psi(y, \theta) = \begin{pmatrix} y - \theta \\ y^2 - \theta^2 - \theta \end{pmatrix}.$$

(a) Describe how you would estimate θ efficiently using GMM methods

First get a consistent (but possibly inefficient) estimate of θ. Here one possibility is to use only the first moment and estimate θ as $\hat{\theta}_{init} = \sum_{i=1}^N y_i/N$. This also corresponds to minimizing the quadratic form

$$Q_C(\theta) = \left(\frac{1}{N}\sum_{i=1}^N \psi(y_i, \theta)\right)' C \left(\frac{1}{N}\sum_{i=1}^N \psi(y_i, \theta)\right),$$

with the weight matrix

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

An alternative is to use the identity matrix as the initial weight matrix. Given an initial estimator of this type, estimate the optimal weight matrix as

$$\hat{\Delta} = \frac{1}{N}\sum_{i=1}^N \psi(y_i, \hat{\theta}_{init}) \cdot \psi(y_i, \hat{\theta}_{init})'.$$

Finally, minimize $Q_{\hat{\Delta}^{-1}}(\theta)$ to get the efficient GMM estimator.

(b) Describe how you would estimate θ using Empirical Likelihood methods.

One approach is to solve

$$\min_{\theta,\pi} \sum_{i=1}^N \ln(\pi_i),$$

subject to the restrictions

$$\sum_{i=1}^N \pi_i = 1, \qquad \text{and} \quad \sum_{i=1}^N \pi\psi(y_i, \theta) = 0.$$

This optimization problem can also be written as the solution to

$$\sum_{i=1}^N \rho(y_i, \theta, t_1, t_2) = 0,$$

where

$$\rho(y,\theta,t_1,t_2) = \begin{pmatrix} (y-\theta)/(1+t_1\cdot(y-\theta)+t_2\cdot(y^2-\theta-\theta^2)) \\ (y^2-\theta^2-\theta)/(1+t_1\cdot(y-\theta)+t_2\cdot(y^2-\theta-\theta^2)) \\ (-t_1-t_2-2\theta t_2)/(1+t_1\cdot(y-\theta)+t_2\cdot(y^2-\theta-\theta^2)) \end{pmatrix}$$

(c) What is the large sample variance of $\hat{\theta}_{gmm}$?
Let

$$\Delta = \mathbb{E}[\psi(Y,\theta)\psi(Y,\theta)'],$$

and

$$\Gamma = \frac{\partial \psi}{\partial \theta}(Y,\theta) = \begin{pmatrix} -1 \\ -1-2\theta \end{pmatrix}.$$

Then the asymptotic variance of $\sqrt{N}(\hat{\theta}_{gmm}-\theta)$ is $(\Gamma'\Delta^{-1}\Gamma)^{-1}$.

(d) Suppose Y has a Poisson distribution with mean θ. How does the variance of the GMM estimator simplify?
In that case the score function is

$$\mathcal{S}(y:\theta) = \frac{y}{\lambda} - 1.$$

This is proportional to the first moment in the GMM estimator. Hence the GMM estimator using only the first moment is equivalent to maximum likelihood. The GMM estimator using both moments cannot be less efficient, and since it also cannot be more efficient than ml, it must have the same variance. The ml estimator if $\bar{Y}$, with asymptotic variance θ, so that must be the variance of the GMM estimator in that case. You can also get there by simplifying the full variance expression from the previous part.