Field Examination: Econometrics

Department of Economics
University of California, Berkeley

August 2012

Instructions: You have 180 minutes to answer **THREE out of the following four questions.** Please make your answers elegant, that is, clear, concise, and, above all, correct. Good luck!

1. Let $(X_i)_{i=1}^n$ be a sample of real-valued i.i.d. copies of $X \sim P_0$. An (applied) econometrician considers the following statistical model $\{P(\cdot;\theta) \colon \theta \in \Theta\}$ where $P(\cdot;\theta)$ is a probability measure over X and Θ is the parameter space. However, the model is misspecified, i.e., $P_0 \notin \{P(\cdot;\theta) \colon \theta \in \Theta\}$.

Suppose Θ is compact subset of some Euclidean space. Suppose $P(\cdot;\theta)$ has a pdf, $p(\cdot;\theta)$ (with respect to Lebesgue measure), and $\theta \mapsto \log(p(\cdot;\theta))$ is twice continuously differentiable, and $|\nabla_\theta \log(p(\cdot;\theta))| \leq G(\cdot)$ and $|\nabla_{\theta\theta} \log(p(\cdot;\theta))| \leq G(\cdot)$ a.s., with $E_{P_0}[||G(X)||] < \infty$. [1]

Let

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} \hat{Q}_n(\theta)$$

where $\hat{Q}_n(\theta) \equiv n^{-1}\sum_{i=1}^n \log(p(X_i;\theta))$.

(Throughout your calculations below, feel free to add more assumptions if you think you need them; however, adding “unnecessary” assumptions will be penalized.)

(a) Show that $\theta^* \in \arg\max_{\theta\in\Theta} E_{P_0}[\log(p(X;\theta))]$ exists.

(b) Henceforth, assume θ^* is unique. Show that $\hat{\theta}_n = \theta^* + o_{P_0}(1)$.

(c) Henceforth, assume θ^* is in the interior of Θ. Is $\hat{\theta}_n$ root-n asymptotically normal, i.e., $\sqrt{n}(\hat{\theta}_n - \theta^*)$ is asymptotically normal? What is the asymptotic variance?

(d) Show whether the information equality holds or not.

(e) Show that $n(\hat{Q}_n(\hat{\theta}_n) - \hat{Q}_n(\theta^*))$ equals (up to $o_{P_0}(1)$ terms) $n(\hat{\theta}_n - \theta^*)' M (\hat{\theta}_n - \theta^*)$. What is M?

(f) According to your calculations in the previous item; is $n(\hat{Q}_n(\hat{\theta}_n) - \hat{Q}_n(\theta^*))$ chi square distributed? Why or why not?

(g) How would your answer to the previous item change if the model is not misspecified, i.e., θ^* is such that $P_{\theta^*} = P_0$? Please explain.

(h) Consider the following null hypothesis: $A\theta^* = 0$. Can you construct a Wald statistic (i.e., $\hat{W}_n = (A\hat{\theta}_n)'\mathbf{W}_n A\hat{\theta}_n$ where $\mathbf{W}_n$ is chosen by you) such that is asymptotically chi-square distributed (under the null)?

[1] $\nabla_x f$ denotes the derivative of f with respect to x; $\nabla_{xy} f$ denotes the second derivative of f with respect to x and y; etc. E_P means the expectation with respect to probability P.

2. For a sample of N i.i.d. observations on a scalar dependent variable y_i and p-dimensional vector of (non-constant) regressors x_i, the conditional median of y_i given $x_i = x_0$, denoted $m_0 \equiv m(x_0)$, is any value that satisfies

$$\begin{aligned} \Pr\{y_i &\leq m_0 \mid x_i = x_0\} \geq 1/2, \\ \Pr\{y_i &\geq m_0 \mid x_i = x_0\} \geq 1/2. \end{aligned}$$

Assume the random vector $z_i \equiv (y_i, x_i')'$ is jointly continuously distributed, with conditional density $\phi(y|x)$ of y_i given $x_i = x$ and marginal density $f(x)$ of x_i which are both smooth and well-behaved (e.g., permit interchange of limits and expectations and the usual series expansions for nonparametric estimation).

(a) Give conditions on the densities above which ensure that the parameter $m_0 \equiv m(x_0)$ is uniquely determined as the solution to the conditional extremum problem

$$m_0 \equiv \arg\min_{b \in R} E[(|y_i - b| - |y_i|) \mid x_i = x_0].$$

(b) A kernel estimator of m_0 can be defined to minimize a kernel-weighed average of absolute deviations of differences $y_i - b$ over b; that is,

$$\begin{aligned} \hat{m} &\equiv \arg\min_{b \in R} S_n(b), \\ S_n(b) &\equiv \frac{1}{Nh^p} \sum_{i=1}^{N} K\left(\frac{x_0 - x_i}{h}\right) \cdot |y_i - b|, \end{aligned}$$

where the kernel function $K(\cdot)$ satisfies standard regularity conditions – it integrates to one, is nonnegative, symmetric about zero, bounded, and smooth – and the nonrandom bandwidth sequence $h = h_N$ has $h \to 0$ as $N \to \infty$. Find conditions on h_N which ensure consistency of $\hat{m}$ for m_0, and under these conditions (and the conditions used in part (a) above), give an argument for consistency of $\hat{m}$ for using analogous arguments based on those for consistency of the LAD estimator of an unconditional median and for consistency of kernel estimators of density and regression functions.

(c) Suppose you have established the following approximate first-order condition for the minimization problem defining $\hat{m}$:

$$\begin{aligned} \hat{\Psi}_N(\hat{m}) &\equiv \frac{1}{Nh^p} \sum_{i=1}^{N} K\left(\frac{x_0 - x_i}{h}\right) \cdot sgn\{y_i - \hat{m}\} \\ &= o_p\left(\frac{1}{\sqrt{Nh^p}}\right), \end{aligned}$$

where

$$sgn\{u\} \equiv 1\{u \geq 0\} - 1\{u \leq 0\}.$$

Suppose you have also established a "stochastic equicontinuity" result that

$$\hat{\Psi}_N(\hat{m}) - \hat{\Psi}_N(m_0) - [\Lambda(\hat{m}) - \Lambda(m_0)] = o_p\left(\frac{1}{\sqrt{Nh^p}}\right),$$

where

$$\Lambda(m) = \lim_{N \to \infty} E[\hat{\Psi}_N(m)].$$

Use these results to derive an explicit form for the asymptotic (normal) distribution of $\hat{m}$. You need not verify the conditions of the limit theorems you use; instead, use existing results on the asymptotics of quantile estimation and of kernel density and regression estimators wherever possible.

3. Suppose $\left\{(y_t, x_t)' : 1 \leq t \leq T\right\}$ is an observed time series generated by the cointegrated system

$$y_t = \beta x_t + u_t,$$

where

$$\begin{pmatrix} u_t \\ \Delta x_t \end{pmatrix} \sim i.i.d.\ \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

with initial condition $x_0 = 0$. A researcher wants to estimate the scalar parameter β.

(a) Characterize the limiting distribution (after appropriate centering and rescaling) of $\hat{\beta}$, the OLS estimator of β.

(b) As an alternative estimator of β, consider $\tilde{\beta}(q) = \left(\sum_{t=q+1}^{T} x_{t-q} x_t\right)^{-1} \left(\sum_{t=q+1}^{T} x_{t-q} y_t\right)$, where $q \geq 1$ is a fixed integer. Characterize the limiting distribution (after appropriate centering and rescaling) of $\tilde{\beta}(q)$. Is $\tilde{\beta}(q)$ asymptotically equivalent to $\hat{\beta}$?

(c) Does the answer to (b) change if $q = q_T$ is a function of T satisfying $q_T \to \infty$ and $q_T/T \to 0$ (as $T \to \infty$)? If so, how does the answer change?

4. Suppose $\{y_t : 1 \leq t \leq T\}$ is an observed time series generated by the model

$$y_t = \mu + u_t, \qquad u_t = \rho u_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T,$$

where $u_0 = 0$ and the $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$, while $\mu \in \mathbb{R}$ and $\rho \in [0,1]$ are (possibly) unknown parameters.

(a) Find the log likelihood function $\mathcal{L}(\mu, \rho)$ and, for $m \in \mathbb{R}$, derive $\hat{\rho}(m) = \arg\max_\rho \mathcal{L}(m, \rho)$, the maximum likelihood estimator of ρ when μ is assumed to equal m.

Suppose $|\rho| < 1$.

(b) Find the limiting distribution (after appropriate centering and rescaling) of the "oracle" estimator $\hat{\rho}(\mu)$ and find a value of α such that $\hat{\rho}(\mu)$ asymptotically equivalent to $\hat{\rho}(\hat{\mu})$ provided $\hat{\mu}$ satisfies

$$\hat{\mu} = \mu + o_p(T^\alpha).$$

(c) Does $\hat{\mu} = (T-1)^{-1} \sum_{t=2}^{T} y_{t-1}$ satisfy the condition derived in (b)? If not, determine whether $\hat{\rho}(\hat{\mu})$ is asymptotically equivalent to $\hat{\rho}(\mu)$.

(d) Do the answers to (b)-(c) change if $\rho = 1$? If so, how do the answers change?