Field Exam 2018: Econometrics

August 8, 2018

Please read carefully

You have to:

- Answer **3** of the following 5 questions. Your selection of **3** questions must satisfy the following constraints:
  - At least 1 of the 3 chosen questions has to be chosen from Part I.
  - and at least 1 of the 3 chosen questions has to be chosen from Part II.

All questions and all subsections of each question have equal weight. No books, notes, tables, or calculating devices are permitted. You have **180** minutes to answer all three questions.

Please make your answers elegant, that is, clear, concise, and, above all, correct. Good luck!

Part I

Please pick **at least 1 out of 3** questions from this section.

Question I.1

Assume a dependent variable y_i satisfies the mean single-index restriction

$$E[y_i|x_i] \equiv g(x_i) = G(x_i'\beta_0)$$

for some $\beta_0 \in \mathbb{R}^p$ and some unknown, smooth function $G(\cdot)$. Also assume that the p-vector x_i is jointly continuously distributed with density function which is positive everywhere and known up to an unknown location parameter μ_0, i.e.,

$$f_X(x; \mu_0) = f(x - \mu_0)$$

for f known. Given a random sample of size N on y_i and x_i, consider the following procedure: first estimate μ_0 by maximum likelihood using only the observations on x_i, then use the ML estimator $\hat{\mu}$ to compute a feasible version of Stoker's "scaled coefficient" estimator

$$\begin{aligned}\hat{\delta} &= \frac{-1}{N}\sum_{i=1}^{N} y_i \cdot \frac{\partial f(x_i - \hat{\mu})}{\partial x} \\ &\equiv \frac{-1}{N}\sum_{i=1}^{N} y_i \cdot s(x_i - \hat{\mu}),\end{aligned}$$

where $s(u)$ is the "score function" $s(u) \equiv \partial f(u)/\partial u$.

(a) Derive the probability limit of $\hat{\delta}$ assuming all functions are sufficiently differentiable, all needed moments exist, and that terms like $g(x)f(x-\mu_0)$ all equal zero on the boundary (i.e., when $||x_i|| \to \infty$).

(b) Derive the form of the asymptotic distribution of $\hat{\delta}$, taking into account the asymptotic distribution of the first-step ML estimator $\hat{\mu}$. Your expression will involve certain Hessian and Jacobian matrices, which you should provide in explicit algebraic form.

Question I.2

Let $\{W_i, X_i, Y_i\}_{i=1}^{N}$ be a random sample from a distribution F_0 such that

$$Y = X'\beta_0 + h_0(W) + U, \;\; \mathbb{E}[U|\, W, X] = 0.$$

Here $W \in \mathbb{W} = \{w_1, \ldots, w_J\}$ is discretely-valued with $\rho_j = \Pr(W_i = w_j) > 0$ for all $j = 1, \ldots, J$. If D is the $J \times 1$ vector of dummy variables with $D_j = 1$ and $D_k = 0$ for $k \neq j$ if $W = w_j$, then $h_0(W) = D'\delta_0$ for some $J \times 1$ vector of coefficients, δ_0. Note that X does not include a constant term. Further define the conditional mean vector and variance matrix

$$e_0(w) = \mathbb{E}[X|\, W = w] \text{ and } v_0(w) = \mathbb{V}(X|\, W = w).$$

You may additionally assume that the conditional variance of Y given W and X is constant (i.e., $\mathbb{V}(U|\, W, X) = \sigma^2$).

<u>HINT</u>: Due to the discreteness of W we have, for example, that $\mathbb{E}[v_0(W)] = \sum_{j=1}^{J} \rho_j v_0(w_j)$.

1. Consider the OLS fit of Y_i onto X_i and $D_{1i}, \ldots, D_{Ji}$. Sketch are argument justifying the claim that

$$\sqrt{N}\left(\hat{\beta}_{\text{OLS}} - \beta_0\right) \xrightarrow{D} \mathcal{N}\left(0, \sigma^2 \mathbb{E}[v_0(W)]^{-1}\right). \tag{1}$$

   For finite N does the OLS fit always exist? Why or why not?

2. Consider the OLS of Y_i onto a constant and X_i in the subsample of units with $D_{ji} = 1$ (i.e., units with $W_i = w_j$). Let $\hat{b}_j$ be the coefficient on X_i. Show that

$$\sqrt{N}\left(\hat{b}_j - \beta_0\right) \xrightarrow{D} \mathcal{N}\left(0, \frac{\sigma^2}{\rho_j} v_0(w_j)^{-1}\right)$$

and also argue that $\sqrt{N}\left(\hat{b}_j - \beta_0\right)$ and $\sqrt{N}\left(\hat{b}_k - \beta_0\right)$ are asymptotically uncorrelated for $j \neq k$.

3. For this question additionally assume that $\dim(X_i) = 1$. Let

$$\Sigma = \sigma^2 diag\left\{\rho_1 v_0(w_1), \ldots, \rho_J v_0(w_J)\right\}^{-1},$$

$\hat{\mathbf{b}} = \left(\hat{b}_1, \ldots, \hat{b}_J\right)'$ and ι_J be a $J \times 1$ vector of ones. Consider the (infeasible) minimum distance estimate of β_0:

$$\hat{\beta}_{\text{OMD}} = \arg\min_{\beta \in \mathbb{R}} \left(\hat{\mathbf{b}} - \iota_J \beta\right)' \Sigma^{-1} \left(\hat{\mathbf{b}} - \iota_J \beta\right).$$

Show that

$$\sqrt{N}\left(\hat{\beta}_{\text{OMD}} - \beta_0\right) \xrightarrow{D} \mathcal{N}\left(0, \sigma^2 \mathbb{E}\left[v_0(W)\right]^{-1}\right).$$

4. Now assume that the partially linear model is mis-specified; specifically that

$$\hat{b}_j \xrightarrow{p} b_0(w_j).$$

Propose a test with a χ^2_{J-1} null distribution for

$$H_0 : b_0(w_1) = \cdots = b_0(w_J) = \beta_0$$

against the alternative that at least one of these equalities fails. Justify your test (2 - 4 sentences).

5. Propose an estimate of the average coefficient $\beta_0 = \mathbb{E}\left[b_0(W)\right]$. Sketch its large sample properties.

6. Provide a synopsis of the insight you have gleaned from answering the questions above (10 sentences).

Question I.3

Consider the problem of estimating the square of a pdf,

$$\psi(P) = \int (p(z))^2 dz$$

where P is the true probability measure generating the IID data $(Z_1, ..., Z_n)$ and p is the corresponding pdf. We assume $p \in L^2 \cap L^\infty$ and that for some $C < \infty$ and $\rho \geq 0$,

$$|p(x+t) - p(x) - p'(x)t| \leq Ct^\rho, \ \forall x, t.$$

The following estimator that has been considered in the literature

$$\hat{\psi}_{h,n} \equiv n^{-1} \sum_{i=1}^{n} \hat{p}_h(Z_i)$$

where $z \mapsto \hat{p}_h(z) = (nh)^{-1} \sum_{i=1}^{n} \varrho((Z_i - z)/h)$ where ϱ is a smooth symmetric at 0 pdf with all moments finite.

1. Show that $\hat{\psi}_{h,n}$ is of the form[1]

$$\psi_h(P_n) = \int (\kappa_h \star P_n)(z) P_n(dz)$$

with $\kappa_h(.) = h^{-1}\kappa(./h)$ and P_n is the empirical distribution. Characterize κ in terms of ϱ.

2. Show that for each $h > 0$,

$$|\psi_h(P) - \psi(P)| = O(h^\rho).$$

3. Show that

$$\begin{aligned}\sqrt{n}(\psi_h(P_n) - \psi_h(P)) =& n^{-1/2}\sum_{i=1}^{n} 2\{(\kappa_h \star P)(Z_i) - E_P[(\kappa_h \star P)(Z)]\} \\ &+ O_P\left(\frac{\kappa(0)}{h\sqrt{n}} + (\sqrt{nh})^{-1}\right)\end{aligned}$$

4. (i) Suppose $\rho > 1$. Argue that $\sqrt{n}(\psi_{h(n)}(P_n) - \psi(P)) \Rightarrow N(0, V)$ for some $V > 0$ and some $(h(n))_n$. Characterize the rate of $(h(n))_n$. **Hint:** Throughout feel free to assume that $n^{-1/2}\sum_{i=1}^{n}\{(\kappa_{h(n)} \star P)(Z_i) - E_P[(\kappa_{h(n)} \star P)(Z)]\} \Rightarrow N(0, V')$ for some V'.

   (ii) For $\rho \leq 1$ is it possible to obtain the same result? What fails? Please explain your answer. **Hint:** Throughout feel free to assume that $n^{-1/2}\sum_{i=1}^{n}\{(\kappa_{h(n)} \star P)(Z_i) - E_P[(\kappa_{h(n)} \star P)(Z)]\} \Rightarrow N(0, V')$ for some V'.

Another seemingly different estimator is the "leave-one-out" estimator

$$\hat{\psi}^{LOO}_{h,n} \equiv n^{-2}\sum_{i \neq j} h^{-1}\varrho((Z_i - Z_j)/h)$$

5. Show that this estimator is also of the form $\int(\kappa_h \star P_n)(z)P_n(dz)$ and that in this case $\kappa(0) = 0$.

6. (i) Using the previous results find the cutoff value of ρ such that for all values higher than it the estimator is root-n Normal.

   (ii) Is the value you found in part (i) higher or lower than 1 (the cutoff value for the original estimator given in part 4(i))? Explain your answer.

Part II

Please pick **at least 1 out of 2** questions from this section.

[1] The symbol $\star$ is a convolution operator: $x \mapsto (f \star g)(x) = \int f(x-t)g(t)dt$ or $x \mapsto (f \star P)(x) = \int f(x-t)P(dt)$.

Question II.1

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

$$y_t = \mu + u_t, \qquad t = 1, \ldots, T,$$

where $u_t \sim i.i.d.\ \mathcal{N}(0,1)$, while μ is the parameter of interest.

As an estimator of μ, consider

$$\hat{\mu} = \frac{\sum_{t=1}^{T} z_t y_t}{\sum_{t=1}^{T} z_t},$$

where $\{z_t : 1 \leq t \leq T\}$ is some observed time series, which is independent of $\{u_t : 1 \leq t \leq T\}$.

(a) Suppose $z_t = t$. Find the limiting distribution (after appropriate centering and rescaling) of $\hat{\mu}$.

(b) Suppose $z_t = \varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$. Find the limiting distribution of $\hat{\mu}$.

(c) Suppose $z_t = \sum_{s=1}^{t} \varepsilon_s$, where $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$. Find the limiting distribution of $\hat{\mu}$.

(d) Rank the estimators from (a)-(c) in terms of (asymptotic) efficiency.

Question II.2

Suppose $\{(y_t, x_t, z_t)' : 1 \leq t \leq T\}$ is an observed time series generated by the model

$$y_t = \beta x_t + \gamma z_t^2 + u_t,$$

where β and γ are unknown parameters and

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

with initial condition $x_0 = z_0 = 0$.

It can be shown that

$$\begin{pmatrix} T^{-1/2} x_{\lfloor T \cdot \rfloor} \\ T^{-1/2} z_{\lfloor T \cdot \rfloor} \\ T^{-1} \sum_{t=1}^{T} x_t u_t \\ T^{-3/2} \sum_{t=1}^{T} z_t^2 u_t \end{pmatrix} \to_d \begin{pmatrix} B_x(\cdot) \\ B_z(\cdot) \\ \int_0^1 B_x(r) dB_y(r) \\ \int_0^1 B_z(r)^2 dB_y(r) \end{pmatrix},$$

where B_x, B_z, and B_y are independent Wiener processes.

(a) Derive an expression for $\hat{\beta}_{ML}$, the maximum likelihood estimator of β.

(b) Find the limiting distribution (after appropriate centering and rescaling) of $\hat{\beta}_{ML}$.

As an alternative estimator of β, consider

$$\tilde{\beta} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2}$$

(c) Find the limiting distribution (after appropriate centering and rescaling) of $\tilde{\beta}$.

(d) Rank the estimators $\hat{\beta}_{ML}$ and $\tilde{\beta}$ in terms of (asymptotic) efficiency.