Econometrics

Instruction: The econometrics exam consists of two parts: ECON 501 and ECON 510. You may use scratch paper, which will be provided by the proctor as needed. However, you shall not submit your scratch paper. All answers should be written in the space provided below each question in a clear and legible manner. Please show your work!

Print Your Assigned ID:

1. Consider the probability spaces $\big((0,1), \mathcal{B}(0,1), \lambda\big)$, where $\mathcal{B}(0,1)$ is the Borel σ algebra over the interval $(0,1)$, and λ is the Lebesgue measure. Let $X : (0,1) \to \mathbb{R}$ be a random variable defined by $X(\omega) = \exp(-\omega)$.

   (a) Provide at least two ways of completely characterizing the probability law $P_X := \lambda \circ X^{-1}$ on $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ algebra over $\mathbb{R}$.

(b) Find a nonnegative function f on $\mathbb{R}$ such that

$$\int_{(0,1)} g\big(X(\omega)\big)\lambda(d\omega) = \int_{\mathbb{R}} g(x)f(x)\mu(dx)$$

for any nonnegative function g, where μ is the Lebesgue measure on $(\mathbb{R}, \mathcal{B})$.

2. Suppose that the joint probability density function of (X, Y) is given by $f(x, y) = c \cdot x \cdot \mathbb{I}(0 < x < y < 1)$ for some constant $c > 0$.

   (a) Compute $\mathbb{P}(Y < X^2 + 0.5)$.

(b) Compute $\mathrm{Var}(X)$.

(c) Characterize the probability distribution of

$$Z := \mathbb{E}\Big(\mathbb{E}\big[\mathbb{E}\{Y \mid \mathbb{I}(X \leq 0.5)\} \mid \mathbb{I}(X \leq 0.5), X^2\big] \;\Big|\; X\Big).$$

(d) What is the correlation coefficient between $Y - Z$ and $\mathbb{I}(X > 0.5)$, where Z is defined in part (c)?

3. Suppose that X has a probability density function f on $\mathbb{R}$ such that $f(x) > 0$ for all $x \in \mathbb{R}$, and $\int_{-\infty}^{\infty} x^2 f(x)dx < \infty$. For $y \in \mathbb{R}$, define $g(y) := \int_{-\infty}^{\infty} |x - y| f(x)dx$.

   (a) Derive a characterization of the value y^* that minimizes the function g by using the probability density function f.

(b) Let $x^* := \int_{-\infty}^{\infty} x f(x) dx$. What is the largest possible distance between x^* and y^*?

4. Suppose that X and Y are random variables with $\mathbb{E}\{\exp(tX+sY)\} = \exp\{(t^2+s^2)/2\}$ for all $t, s \in \mathbb{R}$.

   (a) Obtain the conditional density function of Y given $2X + Y$.

(b) Obtain $\mathbb{E}(Y \mid 2X + Y < 0)$.

5. Suppose that $X_1, X_2, \cdots$ are independent random variables, and that the probability density function of X_n is given by $f_n(x) = \{1 + \cos(2\pi n x)\}\mathbb{I}(0 \le x \le 1)$. Compute

$$\sum_{n=501}^{600} \mathbb{P}(0.5 \le X_n < 0.7 \text{ and } 0.2 < X_{n+3} \le 0.3).$$

You may provide an approximation with a proof that justifies it.

6. State whether each of the following statements is true or false, and explain why.

   (a) If $X_n \xrightarrow{p} 0$, then X_n is approximately equal to 0 when n is large, and therefore we must have $\mathbb{P}(X_n \leq 0) \to 1$.

   (b) The p-value, significance level, size of a test are all interchangeable terms, which indicate the chance of making a Type I error.

(c) Suppose that we are interested in some parameter θ_0, for which an asymptotically normal estimator $\hat{\theta}$ is available. Suppose that an estimate $\hat{\theta} = 0.10$ is reported with the standard error 0.025. Based on this information, it is correct to say that θ_0 is approximately between 0.05 and 0.15 with probability 95%.

(d) Suppose that we have an i.i.d. sample $\{X_1, X_2, \cdots, X_n\}$, where $X_1 \sim N(\mu_0, \sigma_0^2)$. If we knew σ_0, then we would be able to estimate μ_0 more accurately than we could without knowing σ_0.

Pennsylvania State University, Department of Economics, Patrik Guggenberger, August 2024

Answer all five questions on this part. You can write on the front and back side of each page.

Comprehensive Exam

1. Suppose $Y_i = X_i'\beta + U_i$ for $i = 1, 2, ..., n$, $\{(X_i, U_i) : i \geq 1\}$ are iid with $E(U_i \mid X_i) = 0$ a.s. and $E(U_i^2 \mid X_i) < \infty$ a.s.

   a) Derive the asymptotic distribution of the least squares estimator.

   b) Provide a consistent estimator for the asymptotic variance of the least squares estimator and prove consistency.

   Precisely state any additional assumptions you may need.

2. Under Assumptions EE3 and CF-NS (both stated below) we derived the limiting distribution of the GMM estimator $\widehat{\theta}_n$ with nonsmooth stochastic criterion function, namely

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \rightarrow_d N(0, (\Gamma'\Gamma)^{-1}\Gamma' V_0 \Gamma(\Gamma'\Gamma)^{-1}).$$

a) Provide estimators for V_0 and Γ and discuss their consistency under appropriate conditions.

Consider the example of Quantile Regression, $(Y_i, X_i')'$ i.i.d.

$$Y_i = X_i'\theta_0 + U_i,$$

where the τ-quantile of U_i conditional on X_i equals 0 for some $\tau \in (0, 1)$.

b) Derive what V_0 and Γ are in this case.

Setup and Assumptions for the GMM case with nonsmooth stochastic criterion function:

Let $Q_n(\theta) = ||\overline{g}_n(\theta)||$, where $\overline{g}_n(\theta) = n^{-1}\sum_{i=1}^n g(W_i, \theta)$, and $g(\theta) = Eg(W_i, \theta)$.

Assumption CF-NS: (i) *θ_0 is in the interior of Θ.*

(ii) *$g(\theta)$ is differentiable at θ_0 with* $\Gamma = (\partial/\partial\theta')g(\theta_0)$ of full rank $d \leq k$.

(iii) $g(\theta_0) = 0$.

(iii) $\sqrt{n}\overline{g}_n(\theta_0) \rightarrow_d N(0, V_0)$.

(iv) For every sequence of positive constants $\{\delta_n\}_{n\geq 1}$ that converges to zero,

$$\sup_{\theta\in\Theta, ||\theta-\theta_0||<\delta_n} \sqrt{n}\,||\overline{g}_n(\theta) - g(\theta) - \overline{g}_n(\theta_0)|| \rightarrow_p 0.$$

Assumption EE3: (i) $Q_n(\widehat{\theta}_n) = \inf_{\theta\in\Theta} Q_n(\theta) + o_p(n^{-1/2})$ and (ii) $\widehat{\theta}_n \rightarrow_p \theta_0$.

3. In the Heckit model (Heckman, 1979), using the same notation/assumptions as in class ($y^* = x'\theta + \varepsilon$, $y = dy^*$, $d = 1(x'\pi_1 + z'\pi_2 \geq -\eta$, joint normality of (ε, η)...)

a) show that

$$E(y|x, z, d = 1) = x'\theta + \sigma_{\varepsilon\eta}\lambda(x'\pi_1 + z'\pi_2),$$

where $\lambda(s) = \phi(s)/\Phi(s)$.

b) Identify π_1 and π_2 from a probit regression of d onto x and z.

c) Specify conditions under which one can identify θ and $\sigma_{\varepsilon\eta}$.

4. Consider the model defined by moment inequalities/equalities

$$E_F m_j(W_i, \theta_0) \geq 0 \text{ for } j = 1, ..., p \text{ and}$$
$$E_F m_j(W_i, \theta_0) = 0 \text{ for } j = p+1, ..., p+v,$$

where the $m_j(\cdot,\cdot)$ are known functions, θ_0 is the true unknown value of the parameter of interest θ, and F is the distribution of the data W_i.

Define $\gamma_1 = (\gamma_{1,1}, ..., \gamma_{1,p})' \in R^p_+$, where

$$\gamma_{1,j} = \sigma^{-1}_{F,j}(\theta) E_F m_j(W_i, \theta) \text{ for } j = 1, ..., p$$

and $\sigma_{F,j}(\theta) = Var_F^{1/2}(m_j(W_i, \theta))$. Let

$$\gamma_2 = \Omega = \Omega(\theta, F) = Corr_F(m(W_i, \theta)),$$

where the latter denotes the $k \times k$ correlation matrix of $m(W_i, \theta)$, and $\gamma_3 = (F, \theta)$. Define

$$T_n(\theta) = S(n^{1/2}\overline{m}_n(\theta), \widehat{\Sigma}_n(\theta))$$

where $\widehat{\Sigma}_n(\theta)$ denotes an estimator of the asymptotic variance matrix, $\Sigma(\theta)$, of $n^{1/2}\overline{m}_n(\theta)$, namely

$$\widehat{\Sigma}_n(\theta) = n^{-1} \sum_{i=1}^{n} (m(W_i, \theta) - \overline{m}_n(\theta))(m(W_i, \theta) - \overline{m}_n(\theta))'.$$

Besides certain continuity conditions, assume

(i) $S((m_1, m_2), \Sigma)$ is non-increasing in m_1, for all $m_1 \in R^p$, $m_2 \in R^v$, and variance matrices $\Sigma \in R^{k \times k}$ and

(ii) $S(m, \Sigma) = S(\Delta m, \Delta \Sigma \Delta)$ for all $m \in R^k$, $\Sigma \in R^{k \times k}$, and pd diagonal $\Delta \in R^{k \times k}$.

We showed in class (under some technical conditions) that under drifting sequences $\{\gamma_{n,h} = (\gamma_{n,h,1}, \gamma_{n,h,2}, \gamma_{n,h,3}) : n \geq 1\}$ of parameters $(\gamma_1, \gamma_2, \gamma_3)$ such that

$$n^{1/2}\gamma_{n,h,1} \to h_1 \in [0, \infty]^p, \ \gamma_{n,h,2} \to h_2$$

we have

$$T_n(\theta_{n,h}) \ \to_d S(h_2^{1/2} Z + (h_1, 0_v), h_2) \sim J_h,$$

where $Z \sim N(0, I_k)$ and $h = (h_1, h_2)$. Denote by $c_h(1-\alpha)$ the $1-\alpha$-quantile of J_h.

The objective is to test the null hypothesis $H_0 : \theta = \theta_0$ at nominal size α versus a two-sided alternative using $T_n(\theta_0)$ as the test statistic.

While h_1 cannot be consistently estimated one can construct an asymptotically valid $(1-\beta)$-confidence region $CR_{h_1,n}(1-\beta)$ for h_1 for some $\beta > 0$.

a) Discuss how to construct $CR_{h_1,n}(1-\beta)$.

Assume one uses the critical value

$$cv(1-\alpha) = \sup_{\overline{h}_1 \in CR_{h_1,n}(1-\beta)} c_{(\overline{h}_1, \widehat{h}_{2,n})}(1-\alpha)$$

for the test of the null hypothesis $H_0 : \theta = \theta_0$ for a consistent estimator $\widehat{h}_{2,n}$ of h_2.

b) Provide a candidate for $\widehat{h}_{2,n}$.

c) Is the asymptotic size of the test that rejects if $T_n(\theta_0) > cv(1-\alpha)$ equal to α? Bigger than α? Smaller than α?

d) Is there a modification of the approach above where one rejects the null if $T_n(\theta_0) > cv(1-\delta)$ for an appropriate (nonrandom) $\delta > 0$ (that may depend on α and β) that guarantees that the asymptotic size of the resulting test is bounded by α? If yes, specify δ and prove your claim.

e) If the answer in d) is yes, is there anything that can be said about the power merits of the resulting test relative to the worst-case plug-in approach?

5. Let X_i ~iid with unknown mean $\beta \in R$ and variance 1, $i = 1, ..., n$. Assume we are interested in testing

$$H_0 : \beta = 0 \text{ vs } H_1 : \beta \neq 0.$$

based on the statistic $n^{1/2}\overline{X}_n$, where $\overline{X}_n = \sum_{i=1}^{n} X_i/n$. Regarding a critical value, two resampling schemes are suggested.

a) Let X^*_{ib} ~iid$\widehat{F}_n$, for $i = 1, ..., n$ and $b = 1, ..., B$ (for some user chosen B, say 1000) where $\widehat{F}_n$ denotes the EDF of the data and define $\overline{X}^*_{nb} = \sum_{i=1}^{n} X^*_{ib}/n$. Take as the critical value the $1 - \alpha$-quantile of $n^{1/2}\overline{X}^*_{nb}$, $b = 1, ..., B$.

b) Choose $s = n^{1/2}$. Consider all the different subsets S_j, $j = 1, ..., \binom{n}{s}$, of $\{X_1, ..., X_n\}$ with s elements and for each subset S_j calculate the sample average $\overline{X}_{sj} = \sum_{i \in S_j} X_i/s$. Take as the critical value the $1 - \alpha$-quantile of $s^{1/2}\overline{X}_{sj}$, $j = 1, ..., \binom{n}{s}$.

What can you say about the asymptotic null rejection probability of the two procedures? Consistency? Provide justifications for your answers. Would you recommend any of the two procedures? Explain.