Econometrics Preliminary Exam
Agricultural and Resource Economics, UC Davis

June 2023

There are **four** questions. You must answer all questions. Each question receives equal weight. Within each question, each part will receive equal weight in grading. You have 20 minutes to read the exam and then four hours to complete the exam.

I. Probability and Statistics

(a) International students who would like to work in industry jobs in the U.S. need to apply for an H1B work visa. In recent years, the number of H1B applications always exceeded the annual quota (65,000+20,000) so an H1B visa lottery was triggered every April to determine which applications get work visas. The lottery has two steps. In the first step, all applications join a lottery that randomly draws 65,000 out of all applicants to grant an H1B visa. In the second step, unselected applicants with an advanced degree (master degree or higher) join another lottery that randomly draws an extra 20,000 H1B visas.

i. Suppose in one year there are a total of 500,000 applicants, among which 300,000 have an advanced degree. Assume that each applicant only submits one application. What is the probability of winning the H1B lottery if an applicant does NOT have an advanced degree? Suppose the probability of winning the two-step H1B lottery if an applicant has an advanced degree is a in that year. If one applicant wins the lottery, what is the probability that this applicant has an advanced degree?

ii. Applicants who graduated from STEM majors (including Economics) can join the H1B lottery for a maximum of three consecutive years. Suppose we have a STEM applicant who will keep applying until either winning the H1B lottery or hitting the three-year maximum. Suppose the probability of this applicant winning the H1B lottery in year t conditioning on applying is $p_t \in (0,1)$, for $t = 1, 2, 3$. What is the probability of the applicant obtaining the H1B visa?

(b) A random variable X has density

$$f(x) = a(1 - x^2),\ x \in [-1, 1]$$

where a is a constant.

i. What is the value of the constant a?

ii. Calculate $\mathbb{P}[X > 0]$ and then $\mathbb{E}[X|X > 0]$.

iii. Suppose a random variable Y is continuous. Let $F(.)$ and $f(.)$ be the cumulative distribution function and the probability density function of Y, respectively. Let c be a constant satisfying that $\mathbb{P}[Y > c] > 0$ and $f(c) > 0$. Show that $\mathbb{E}[Y|Y > c]$ increases with c. [Hint: $\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)dt\right) = f(x, b(x)) \cdot \frac{d}{dx}b(x) - f(x, a(x)) \cdot \frac{d}{dx}a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)dt.$]

(c) Consider scalar random variables X and Y, where $X \in \{0, 1\}$ and $\mathbb{P}(X = 1) = p \in (0, 1)$. Denote $\mathbb{E}[Y|X = 0] = \mu_0$ and $\mathbb{E}[Y|X = 1] = \mu_1$.

i. Derive $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

ii. Derive $cov(X, Y)/var(X)$, where $cov(X, Y)$ is the covariance between X and Y and $var(X)$ is the variance of Y. [Hint: $cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.]

(d) Consider scalar continuous random variables $X_1, X_2, ..., X_n$ that are independent and identically distributed following an exponential distribution with density

$$f(x) = \frac{1}{\lambda}\exp(-\frac{x}{\lambda}), \quad x > 0, \quad \lambda > 0.$$

Recall that a random variable following the exponential distribution has mean λ and variance λ^2.

i. Derive the MLE estimator of λ and call it $\hat{\lambda}$.

ii. Is $\hat{\lambda}$ an unbiased estimator of λ? Is $(\hat{\lambda})^2$ an unbiased estimator of λ^2? Is $\hat{\lambda}$ a consistent estimator of λ? Is $(\hat{\lambda})^2$ a consistent estimator of λ^2? Why or why not?

iii. What is the asymptotic distribution of $\sqrt{n}\left((\hat{\lambda})^2 - \lambda^2\right)$ as $n \to \infty$? What is the asymptotic distribution of $\frac{\hat{\lambda}-\lambda}{\hat{\lambda}/\sqrt{n}}$ as $n \to \infty$? Show your steps.

II. **Linear Regression**

Consider the model $y = X\beta + e$, where y and e are $n \times 1$ vectors, X is an $n \times k$ matrix, $\mathbb{E}[X'e] = 0$, $\mathbb{E}[ee'|X] = \sigma^2 I_n$, and I_n is an $n \times n$ identity matrix. Consider two estimators:

$$\hat{\beta} = \left(X'X\right)^{-1} X'y$$
$$\bar{\beta} = \left(X'WX\right)^{-1} X'Wy,$$

where W is a known positive definite matrix.

(a) Is $\hat{\beta}$ unbiased for β? If so, prove it. If not, state additional conditions you require for unbiasedness and prove unbiasedness under those conditions.

(b) Consider the following estimator for the variance of $\hat{\beta}$ conditional on X:

$$\hat{V} = \left(X'X\right)^{-1} \left(X'\hat{D}X\right) \left(X'X\right)^{-1},$$

where $\hat{D}$ is a diagonal matrix with the i^{th} diagonal element equal to $\hat{e}_i^2$ where $\hat{e}_i$ is the i^{th} element of $\hat{e} = y - X\hat{\beta}$. Is $\hat{V}$ unbiased for the variance of $\hat{\beta}$ conditional on X? If so, prove it. If not, propose an unbiased estimator and prove it is unbiased.

(c) Following on from (b), would you recommend using the estimator $\hat{V}$. Explain in words why or why not.

(d) Is $\hat{\beta}$ consistent for β? If so, prove it. If not, state additional conditions you require for consistency and prove consistency under those conditions.

(e) Following on from (d), find the asymptotic distribution of $\sqrt{n}(\hat{\beta} - \beta)$ as $n \to \infty$. You may carry over any assumptions you made in (d) to show consistency. State any additional assumptions you require.

(f) Following on from (e), write down a test statistic for testing the null hypothesis $H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$. State the asymptotic null distribution of your statistic. No proof required.

(g) Under the conditions you imposed for unbiasedness in (a), is $\bar{\beta}$ unbiased for β? If so, prove it. If not, either (i) state additional conditions you require for unbiasedness and prove unbiasedness under those conditions, or (ii) explain why no such conditions exist.

(h) Under the conditions you imposed for unbiasedness in (a), which estimator is more efficient, $\bar{\beta}$ or $\hat{\beta}$? Give a mathematical proof. If you need any additional conditions to answer this question, state them.

III. **Estimation and Testing in the Linear Model**

(a) Consider the model $y_i = \beta x_i + e_i$, where $x_i > 0$ is scalar, $\mathbb{E}[e_i] = 0$, and $\mathbb{E}[x_i e_i] = 0$. You have an *iid* random sample of size n. Consider the estimators:

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$
$$\bar{\beta} = \frac{\sum_{i=1}^n x_i^3 y_i}{\sum_{i=1}^n x_i^4}$$

i. Is $\bar{\beta}$ consistent for β? If so, prove it. If not, state additional conditions you require for consistency and prove consistency under those conditions.

ii. Assume $\mathbb{E}(e_i^2|x_i) = \sigma^2 x_i^{-2}$, where σ^2 is an unknown constant. Describe how you would construct a bootstrap confidence interval for β. Be specific, including about which estimator you would use and why.

iii. Propose a test statistic for the null hypothesis $H_0 : \mathbb{E}(x_i^3 e_i) = 0$ vs $H_1 : \mathbb{E}(x_i^3 e_i) \neq 0$. Derive the asymptotic null distribution of your test statistic. State any additional assumptions you need.

(b) Suppose we have i.i.d. observations $\{(Y_i, X_i, Z_i) : i = 1, \ldots, n\}$. Consider the model $Y_i = X_i'\beta_0 + e_i$ with $\mathbb{E}[Z_i e_i] = 0$. Let $\hat{\beta}_{2sls}$ be the 2SLS estimator. Define $\boldsymbol{Q}_{ZX} = \mathbb{E}[Z_i X_i']$, $\boldsymbol{Q}_{ZZ} = \mathbb{E}[Z_i Z_i']$, $\boldsymbol{Q}_{XZ} = \mathbb{E}[X_i Z_i']$, $\hat{\boldsymbol{Q}}_{ZX} = \frac{1}{n}\sum_{i=1}^n Z_i X_i'$, $\hat{\boldsymbol{Q}}_{ZZ} = \frac{1}{n}\sum_{i=1}^n Z_i Z_i'$, and $\hat{\boldsymbol{Q}}_{XZ} = \frac{1}{n}\sum_{i=1}^n X_i Z_i'$.

i. Provide a sufficient condition for β_0 to be identified. You do not need to prove anything for this question.

ii. Assume that the parameter is identified. Show that $\hat{\beta}_{2sls}$ is a consistent estimator for β_0. In your proof, add assumptions if necessary and clarify how you use the law of large numbers and the central limit theorem.

iii. Assume we have a consistent estimator $\hat{\boldsymbol{V}}_\beta$ for the asymptotic variance of $\hat{\beta}_{2sls}$. Let b be any vector and consider the null hypothesis of $\beta_0 = b$. Using $\hat{\beta}_{2sls}$ and $\hat{\boldsymbol{V}}_\beta$, explain the meaning of "the Type I error for the Wald test converges to α as $n \to \infty$." In your answer, define the test statistic and the critical value. Also, clarify the definition of the Type I error.

IV. **M-estimation/Panel Data**

(a) Suppose we have i.i.d. observations $\{(Y_i, X_i) : i = 1, \ldots, n\}$.

i. Consider the M-estimator

$$\hat{\theta} \in \arg\min_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \rho(Y_i, X_i, \theta)$$

where ρ is a known function and Θ is a parameter space. Suppose there is θ_0 such that

$$\theta_0 \in \arg\min_{\theta \in \Theta} \mathbb{E}[\rho(Y, X, \theta)].$$

Suppose $\theta \mapsto \rho(Y, X, \theta)$ is twice differentiable, $\mathbb{E}[\| \frac{\partial}{\partial \theta} \rho(Y, X, \theta)|_{\theta = \theta_0} \|^2] < \infty$, $\mathbb{E}[\frac{\partial^2}{\partial \theta \partial \theta'} \rho(Y, X, \theta)|_{\theta = \theta_0}]$ is finite and invertible, and

$$\sqrt{n}(\hat{\theta} - \theta_0) + \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta \partial \theta'} \rho(Y_i, X_i, \theta)|_{\theta = \theta_0} \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \rho(Y_i, X_i, \theta)|_{\theta = \theta_0} \to_p 0.$$

This equation is called the asymptotic linear representation for $\hat{\theta}$. Show that $\sqrt{n}(\hat{\theta} - \theta_0)$ converges to a normal distribution. In your proof, add assumptions if necessary and clarify how you use the law of large numbers and the central limit theorem.

ii. Assume there are a known function m and a parameter value θ_0 such that $\mathbb{E}[Y_i \mid X_i] = m(X_i, \theta_0)$. Assume $\theta \mapsto m(x, \theta)$ is continuously differentiable for every x. Let $\hat{\theta}$ be the nonlinear least squares estimator based on the regression model $m(x, \theta)$, and let $\hat{V}$ be an estimator for $\mathrm{var}[\hat{\theta}]$. Suppose we are interested in $\mathbb{E}[Y_i \mid X_i = x]$ for some x. Write down the 95% confidence interval for $\mathbb{E}[Y_i \mid X_i = x]$. You do not need to prove anything for this question.

iii. Assume there is a parameter value θ_0 such that $Pr(Y_i = y \mid X_i) = \frac{1}{y!} \exp(y X_i' \theta_0) \exp(-\exp(X_i' \theta_0))$ for every non-negative integer y. We use the maximum likelihood estimation to estimate θ_0. To guarantee the parameter is identified, which matrix do you assume is positive or negative definite? You do not need to prove anything for this question.

(b) Consider a panel dataset $\{(Y_{it}, X_{it}, Z_i) : i = 1, \ldots, N, \ t = 1, \ldots, T\}$ with $\dim(Y_{it}) = \dim(Z_i) = 1$.[1]

i. Consider the model $Y_{it} = X_{it}' \beta_0 + u_i + \varepsilon_{it}$. Assume $\mathbb{E}[X_{is} \varepsilon_{it}] = 0$ for every $s, t = 1, \ldots, T$. Propose a consistent estimator for β_0 when $N \to \infty$ and T is fixed. Provide a sufficient condition for its consistency. Prove it is sufficient.

[1] For all of the following questions, you have to define the estimator formally. If you are using a nonlinear least squares or maximum likelihood estimator, then you have to define the sample objective function. If you are using a generalized method-of-moments estimator, then you have to define the moment function and specify a weight matrix. You do not need to use the asymptotically optimal weight matrix.

ii. (Anderson-Hsiao meet Hausman-Taylor.) Consider $T = 4$. Consider the model $Y_{it} = \alpha_0 Y_{it-1} + \beta_0 Z_i + u_i + \varepsilon_{it}$. Assume that

$$\mathbb{E}\left[Y_{it-1}\varepsilon_{it+s}\right] = 0 \text{ for every } s \geq 0 \text{ and every } t$$

and that

$$\mathbb{E}[Z_i(u_i + \varepsilon_{it})] = 0 \text{ for every } t.$$

Based on these assumptions, we can derive

$$\mathbb{E}\left[Y_{i1}((Y_{i3} - Y_{i2}) - \alpha_0(Y_{i2} - Y_{i1}))\right] = 0, \tag{1}$$

$$\mathbb{E}\left[Y_{i2}((Y_{i4} - Y_{i3}) - \alpha_0(Y_{i3} - Y_{i2}))\right] = 0, \tag{2}$$

and

$$\mathbb{E}\left[Z_i\left(\frac{Y_{i4} + Y_{i3} + Y_{i2}}{3} - \alpha_0\frac{Y_{i3} + Y_{i2} + Y_{i1}}{3} - \beta_0 Z_i\right)\right] = 0. \tag{3}$$

Assume $\mathbb{E}[Y_{i1}(Y_{i2} - Y_{i1})]\mathbb{E}[Y_{i2}(Y_{i3} - Y_{i2})] \neq 0$ and $\mathbb{E}[Z_i^2] > 0$. Using all the conditions in (1)-(3), propose a consistent estimator for α_0 and β_0. Show that the parameters are identified. In your proof, add assumptions if necessary.

ARE/ECN 240B Reference Sheet

Functions Define $\phi(u) = \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)$ and $\Lambda(u) = \exp(u)/(1 + \exp(u))$. Then

$$\frac{d}{du}\phi(u) = -u\phi(u), \ \frac{d}{du}\Lambda(u) = \Lambda(u)(1-\Lambda(u)), \ \frac{d}{du}\log(u) = \frac{1}{u}, \text{ and } \frac{d}{du}\exp(u) = \exp(u).$$

First-order condition Consider a convex and differentiable real-valued function f on a convex open set $\mathcal{D}$ in $\mathbb{R}^k$. Then, x is a minimizer of f on $\mathcal{D}$ if and only if the derivative of f is zero at x.

Second-order condition Consider a real-valued function f on a convex set in a finite-dimensional Euclidean space. Suppose f is strictly convex. The number of minimizers of f is 0 or 1.

Theorem 22.2 Uniform Law of Large Numbers Assume

(a) (Y_i, X_i) are i.i.d.,
(b) $\rho(Y, X, \theta)$ is continuous in $\theta \in \Theta$ with probability one,
(c) there is a function $(x, y) \mapsto G(y, x)$ such that $|\rho(Y, X, \theta)| \leq G(Y, X)$ where $\mathbb{E}[G(Y, X)] < \infty$, and
(d) Θ is compact.

Then $\sup_{\theta\in\Theta} |S_n(\theta) - S(\theta)| \to_p 0$.

Theorem 22.3 Consider a parameter $\theta \in \Theta$ and a function $\rho(Y, X, \theta)$ of (Y, X, θ). Define $\hat{\theta} = \arg\min_{\theta\in\Theta} \frac{1}{n} \sum_{i=1}^n \rho(Y, X, \theta)$. Suppose

(a) (Y_i, X_i) are i.i.d.,
(b) $\rho(Y, X, \theta)$ is continuous in $\theta \in \Theta$ with probability one,
(c) there is a function $(x, y) \mapsto G(y, x)$ such that $|\rho(Y, X, \theta)| \leq G(Y, X)$ where $\mathbb{E}[G(Y, X)] < \infty$,
(d) Θ is compact, and
(e) θ_0 is uniquely minimizes $\mathbb{E}[\rho(Y, X, \theta)]$.

Then $\hat{\theta}$ is a consistent estimator for θ_0.

Theorem 22.4 Consider a parameter $\theta \in \Theta$ and a function $\rho(Y, X, \theta)$ of (Y, X, θ). Define $\hat{\theta} = \arg\min_{\theta\in\Theta} \frac{1}{n} \sum_{i=1}^n \rho(Y, X, \theta)$. Suppose $\theta \mapsto \rho(Y, X, \theta)$ is differentiable with $\psi(Y, X, \theta) = \frac{\partial}{\partial\theta}\rho(Y, X, \theta)$. Assume the conditions of Theorem 22.3 hold, plus

(a) $\mathbb{E}\|\psi(Y, X, \theta_0)\|^2 < \infty$,
(b) $\mathbf{Q}$ is positive definite,
(c) there is a neighborhood $\mathcal{N}$ of θ_0 such that $\mathbf{Q}(\theta)$ is continuous in $\theta \in \mathcal{N}$,
(d) $\theta \mapsto \psi(Y, X, \theta)$ is Lipschitz-continuous over $\theta \in \mathcal{N}$ (i.e., there is a function $(x, y) \mapsto B(y, x)$ such that $\|\psi(Y, X, \theta_1) - \psi(Y, X, \theta_2)\| \leq B(Y, X)\|\theta_1 - \theta_2\|$ and that $\mathbb{E}[B(Y, X)^2] < \infty$),
(e) θ_0 is in the interior of Θ.

Then $\sqrt{n}(\hat{\theta} - \theta_0) \to_d N(0, \boldsymbol{Q}^{-1}\Omega\boldsymbol{Q}^{-1})$ as $n \to \infty$.