Econometrics Preliminary Exam

Agricultural and Resource Economics, UC Davis

August 14, 2018

There are **FOUR** questions. Answer each part of each question. All questions are weighted equally. 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) Consider (X, Y) with joint p.d.f. $f_{X,Y}(x,y) = \begin{cases} x+y & 0 \leqslant x \leqslant 1, \quad 0 \leqslant y \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$

(i) Obtain $f_X(x)$, the marginal density of X.

(ii) Obtain $E[X]$.

(iii) Obtain $\Pr[X \leqslant 0.5]$.

(iv) Obtain $E[Y|X = 0.5]$.

(v) Obtain $E[XY]$.

(b) Suppose we have a random sample $x_1, ..., x_n$ of size n from a distribution with probability mass function $f(x;\theta) = \theta^x(1+\theta)^{-(x+1)}$, $x = 0, 1, 2, 3, ...$, $\theta > 0$. X has mean θ and variance $\theta(1+\theta)$.

(i) Obtain the first-order conditions for the MLE of θ.

(ii) Is there an explicit solution for $\widehat{\theta}$? If so, give it.

(iii) Using standard results for the MLE, give the limit distibution of $\sqrt{n}(\widehat{\theta} - \theta)$.

(iv) Suppose $\widehat{\theta} = 1$ and $n = 100$. Do you reject $H_0 : \theta = 1.2$ against $H_a : \theta \neq 1.2$ at level 0.05?

(c) Each of these is a stand-alone question.

(i) Consider a random variable X with moment generating function $m_X(t) = (1-2t)^{-r}$ for $t < 1/2$. Obtain the mean and variance of X.

(ii) Suppose X has density $f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$. Obtain the density of $Y = \ln X$.

(iii) Voters are either Democrat or Republican and either favor or do not favor gun control. Suppose the probability of being a Democrat voter is 0.4, the probability of favoring gun control is 0.5, and the probability of being both a Democrat and favoring gun control is 0.1. What is the conditional probability of a Republican voter supporting gun control?

II. Linear Regression

Consider the model $y_i = x_i\beta + e_i$, where x_i is scalar, $E[e_i] = 0$, $E[x_i] > 0$, $E[e_i|x_i] = 0$, and $E[e_i^2|x_i] = \sigma^2$. You have a sample of size n. Consider the following estimators:

$$\tilde{\beta} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \qquad \text{and} \qquad \hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

Credit will be given for answers that avoid imposing unnecessarily strong assumptions.

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

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

(c) Following on from (b), find the asymptotic distribution of $\sqrt{n}(\tilde{\beta} - \beta)$ as $n \to \infty$. State any additional assumptions you need.

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

(e) Assuming sufficient conditions such that $\tilde{\beta} \xrightarrow{p} \beta$ and $\hat{\beta} \xrightarrow{p} \beta$, show that $\hat{\beta}$ has a smaller asymptotic variance than $\tilde{\beta}$.

(f) In part (e), you showed that $\hat{\beta}$ is asymptotically more efficient than $\tilde{\beta}$ in this setting. Describe in words the implications of this result for hypothesis testing.

(g) Given the model $y_i = x_i\beta + e_i$, where x_i is scalar, are there conditions under which $\tilde{\beta}$ is asymptotically more efficient than $\hat{\beta}$? If so, state sufficient conditions.

III. **Omitted Variables.**

Parts (a) and (b) are independent. State any assumptions you make.

(a) For $i = 1, \ldots, n$, let $y_i = \alpha_0 + \beta_0 x_i + \gamma_0 w_i + u_i$, where $E[u_i|x_i, w_i] = 0$. Suppose you observe a variable z_i with the properties $E[u_i|z_i] = 0$ and $cov[z_i, x_i] \neq 0$.

(i) Consider the OLS estimator $\hat{\beta} = \left(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})\right) / \left(\sum_{i=1}^n (x_i - \bar{x})^2\right)$. Derive the probability limit of $\hat{\beta}$. State any conditions under which this probability limit equals β_0.

(ii) Consider using z_i as an instrument in the IV estimator $\hat{\beta}_{IV} = \left(\sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})\right) / \left(\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})\right)$. Is $\hat{\beta}_{IV}$ consistent for β_0? If so, prove it. If not, state additional conditions you need for consistency and prove consistency under those conditions.

(iii) Consider using z_i as a control variable, i.e., regressing y_i on x_i and z_i. The OLS coefficient on x_i can be expressed as $\tilde{\beta} = (X'(I_n - Z(Z'Z)^{-1}Z')X)^{-1}X'(I_n - Z(Z'Z)^{-1}Z')y$, where X is a $n \times 1$ vector containing the elements of x and Z is a $n \times 2$ matrix containing a vector of ones in the first column and the elements of z_i in the second column. Is $\tilde{\beta}$ consistent for β_0? If so, prove it. If not, state additional conditions you need for consistency and prove consistency under those conditions.

(b) For $i = 1, \ldots, n$ and $t = 1, \ldots, T$, let $y_{it} = \beta_0 x_{it} + \gamma_0 w_{it} + a_i + u_{it}$, where $E[a_i|x_{i1}, \ldots, x_{iT}, w_{i1}, \ldots, w_{iT}] \neq 0$, $E[u_{it}|x_{i1}, \ldots, x_{iT}, w_{i1}, \ldots, w_{iT}, a_i] = 0$. All probability limits in this problem are taken as $n \to \infty$ with $T < \infty$.

(i) Derive the probability limit of $\hat{\beta}_{FE} = \left(\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)\right) / \left(\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)^2\right)$, which is the fixed effects estimator from a regression of y on x only. Is it equal to β_0? If not, explain the difference between its probability limit and β_0.

(ii) Now assume that $w_{it} = w_i$, i.e. it is constant across t but varies across i, what is the probability limit of $\hat{\beta}_{FE}$ in this case? Is it equal to β_0? If not, explain the difference between its probability limit and β_0.

(iii) Now assume that $w_{it} = w_t$, i.e. it is constant across i but varies across t, what is the probability limit of $\hat{\beta}_{FE}$ in this case? Is it equal to β_0? If not, explain the difference between its probability limit and β_0.

IV. **Estimation of Nonlinear Models with Endogeneity**.

Let $y_i = g(x_i; \delta_0) + \epsilon_i$, where $E[\epsilon_i|x_i] \neq 0$ and $g(x;\delta)$ is a known function given x and δ, where $dim(x) = d_x$ and $dim(\delta) = d_\delta$.

(a) Given instruments z_i, where $E[z_i\epsilon_i] = 0$ and $dim(z_i) = d_z$, propose a consistent estimator of δ_0 given data $\{y_i, x_i, z_i\}_{i=1}^n$. Be sure to specify the number of instruments required for identification.

(b) Now consider a special case of the above model, where $d_x = 1$, $d_z = 1$, and $g(x;\delta_0) = \delta_{1,0}x + \delta_{2,0}x^2$, where $\delta_0 = (\delta_{1,0}, \delta_{2,0})'$.

(i) Can you identify δ_0 given the instrument validity condition, $E[z_i\epsilon_i] = 0$? If yes, propose an estimator of δ_0?

(ii) Now suppose that a stronger condition $E[\epsilon_i|z_i] = 0$ holds, can you identify δ_0? If yes, propose an estimator of δ_0?

(c) Now, back to the general case. Give sufficient conditions for consistency of the estimator you proposed in (a) as $n \to \infty$. State primitive conditions wherever possible.

(d) Use the mean value expansion to derive an expression for $\sqrt{n}(\hat{\delta} - \delta_0)$, where $\hat{\delta}$ is the estimator you proposed in (a). Give sufficient conditions for its asymptotic normality as $n \to \infty$. State primitive conditions wherever possible and make sure to state its asymptotic distribution.

(e) Propose an estimator of the asymptotic variance of $\sqrt{n}(\hat{\delta} - \delta_0)$.

Notation. θ_0, Θ, y_i, x_i, $s(y_i, x_i; \theta)$ and $H(y_i, x_i; \theta)$ pertain to the objects defined in the 240B lecture notes.

Assumption ULLN 1 $\sup_{\theta\in\Theta} |\sum_{i=1}^n f(y_i, x_i; \theta)/n - E[f(y_i, x_i; \theta)]| \xrightarrow{p} 0$, if the following conditions hold,

(i) *(i.i.d.)* $\{y_i, x_i\}_{i=1}^n$ is an i.i.d. sequence of random variables;

(ii) *(Compactness)* Θ is compact;

(iii) *(Continuity)* $f(y_i, x_i; \theta)$ is continuous in θ for all $(y_i, x_i')'$;

(iv) *(Measurability)* $f(y_i, x_i; \theta)$ is measurable in $(y_i, x_i')'$ for all $\theta \in \Theta$;

(v) *(Dominance)* There exists a dominating function $d(y_i, x_i)$ such that $|f(y_i, x_i; \theta)| \leqslant d(y_i, x_i)$ for all $\theta \in \Theta$ and $E[d(y_i, x_i)] < \infty$.

Assumption ULLN 2 $\sup_{\theta\in\Theta} |\sum_{i=1}^n f(y_i, x_i; \theta)/n - E[f(y_i, x_i; \theta)]| \xrightarrow{p} 0$, if the following conditions hold,

(i) (*Law of Large Numbers*) $\{y_i, x_i\}$ is i.i.d., and $E[f(y_i, x_i; \theta)] < \infty$ for all $\theta \in \Theta$, which implies $\sum_{i=1}^n f(y_i, x_i; \theta)/n \xrightarrow{p} E[f(y_i, x_i; \theta)]$.

(ii) (*Compactness of* Θ) Θ is in a compact subset of $\mathbb{R}^k$.

(iii) *(Measurability in $(y_i, x_i')'$)* $f(y_i, x_i; \theta)$ is measurable in $(y_i, x_i')'$ for all $\theta \in \Theta$.

(iv) (*Lipschitz Continuity*) For all $\theta, \theta' \in \Theta$, there exists $g(y_i, x_i)$, such that $|f(y_i, x_i; \theta) - f(y_i, x_i; \theta')| \leqslant g(y_i, x_i)\|\theta - \theta'\|$, for some norm $\|.\|$, and $E[g(y_i, x_i)] < \infty$.

Formula for the score statistic

$$S \equiv \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n s_i(\hat\theta_R)\right)' A_{nR}^{-1}C_{nR}' \left\{\widehat{Avar}\left(C_{nR}A_{nR}^{-1}\sum_{i=1}^n s_i(\hat\theta_R)/\sqrt{n}\right)\right\}^{-1} C_{nR}A_{nR}^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n s_i(\hat\theta_R)\right)$$