Econometrics Preliminary Examination

Department of Agricultural and Resource Economics
University of California, Davis
July 2014

Answer each part of each question. The four questions are weighted equally.

Within each question, each part receives equal weight.

1. This question has three unrelated parts.

   (a) Let A denote the event corresponding to *"a person believes Bayes Theorem to be correct"* and $\overline{A}$ denotes the complement, the person does not believe the theorem to be correct. You may assume that these are the only two possibilities—everyone holds one or the other belief.

   You know that $P(A|B) = .5$, where B corresponds to *"the person is a California resident"*. $\overline{B}$ then corresponds to the other option, the person does not reside in California. Suppose $P(B) = 0.2$ for the population of interest.

      i. Indicate how you can place bounds on $P(A)$; justify your answer.
      ii. Suppose that $P(A|\overline{B}) = .4$. Find $P(\overline{B}|A)$.

   (b) Suppose you know that U is distributed as a uniform random variable with range 0 to 1. You also know that $W = e^U$.

      i. Find the cumulative distribution function and probability density function for W.
      ii. Find $Pr[W < 2]$.
      iii. Find $E(W)$.

   (c) $Y_1, Y_2, \ldots, Y_n$ denote n independent and identically distributed observations from $N(\mu, \sigma^2)$:

   $$f(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-(1/2\sigma^2)(y-\mu)^2\right\}$$

   μ is known but σ^2 is not.

      i. Find the maximum-likelihood estimator of σ^2.
      ii. Give a proof that indicates whether your estimator is unbiased.
      iii. Based on your results from (a) and (b), derive a 95% confidence interval for σ^2.
      iv. It turns out that you are really interested in the reciprocal of the variance, $1/\sigma^2$. Give a 95% confidence interval for this expression.

v. Derive a likelihood ratio test of the null hypothesis that $\sigma^2 = 1$, continuing to assume that μ is known, against the alternative hypothesis that $\sigma^2 = 4$.

2. You are interested in the following regression model:

$$y = \beta + X\delta + W\gamma + u.$$

X consists of K_1 columns, W consists of K_2 columns, and the parameter vectors δ and γ are conformable.

(a) Indicate a complete set of assumptions under which OLS is BLUE.

(b) Suppose that the variance-covariance matrix of u is a known, positive-definite matrix Σ. This matrix has non-zero values off of the main diagonal.

Derive an estimator that you would be prepared to prove to be BLUE.

(c) Explain precisely why the proof you would give, to establish the result that your estimator is BLUE, would no longer work if the elements in Σ were unknown to you.

(d) Returning to the part (b) setup, suppose you impose the restriction that $\gamma = 0$, but it is a false restriction. Derive the expected value and variance-covariance matrix of your vector of estimated coefficients. (Hint: you may retain any other assumptions you made beginning in part (a).)

(e) You have available a matrix Z that contains additional variables. Under the set of assumptions for the previous part of this question, including the assumption that you incorrectly imposed the restriction that $\gamma = 0$, describe how you could obtain a consistent estimator for the coefficients β and δ. (Hint: state any new assumptions you require.)

(f) Now you realize that it was a correct restriction to set $\gamma = 0$. Comment on the properties of the estimator you derived in the previous part of this question. Is it unbiased? It is consistent? Is it efficient?

3. Let $y_i = \Lambda(x_i'\beta) + u_i$ for $i = 1, 2, ..., n$, where y_i is a scalar binary regressor, $\Lambda(z) = \exp(z)/\{1 + \exp(z)\}$, x_i and β are $k \times 1$ vectors. We assume here that $E[u_i|x_i] = 0$.
<u>Notation</u>: In your answers below, you can denote the derivative of $\Lambda(z)$ by $\lambda(z) = \Lambda(z)(1 - \Lambda(z))$.

a. Give two extremum estimators for β, denoting them by $\hat{\beta}_1$ and $\hat{\beta}_2$, and give assumptions for their consistency and an intuitive explanation for how these assumptions imply consistency. Comment briefly on one key difference or similarity between the assumptions required for the consistency of $\hat{\beta}_1$ and $\hat{\beta}_2$.
<u>Note</u>: State clearly the types of extremum estimators you are proposing and give primitive conditions wherever possible.

b. For one of the estimators you proposed, perform the mean-value expansion for the sampling error and give conditions that imply its asymptotic normality. Make sure to write

down the asymptotic distribution.
Note: No need to give primitive conditions here.

c. Propose an estimator for the asymptotic variance of the estimator in (b).

d. Now suppose that x_i is endogenous, such that $E[u_i|x_i] \neq 0$. Fortunately, you observe w_i for which $E[w_i u_i] = 0$. What should the dimension of w_i have to be so that β_0 is identified? Give a brief explanation of your answer.

e. Propose an estimator for β_0 assuming $E[u_i|x_i] \neq 0$ and $E[w_i u_i] = 0$. Make sure you specify the dimension of w_i.

f. Suppose an empirical researcher suggests you perform two-stage least squares, where you regress x_i on w_i in the first stage and then plug in the fitted values of x_i into the model for y_i, which you can then estimate by any of the methods you proposed in (a). Is that procedure equivalent to the estimator you propose in (e)? Comment on the validity of the empirical researcher's suggestion.

4. Given the following panel data model, $y_{it} = x_{it}\beta + a_i + u_{it}$, where x_{it} and β are $k \times 1$ vectors and all other variables are scalar. Let $X_i = (x_{i1}, ..., x_{iT})$ and $u_i = (u_{i1}, ..., u_{iT})$. In the following, we will maintain that $E[u_{it}|X_i, a_i] = 0$ for all t.

   (a) Assume $E[a_i|X_i] = 0$, $E[a_i^2|X_i] = \sigma_a^2$, $E[u_{it}^2|X_i, a_i] = \sigma_u^2$ and $E[u_{it}u_{i,t-1}|X_i] = \rho$, and $E[u_{it}u_{is}|X_i] = 0$ for all $|s-t| > 1$.
   Under the above scenario:

      (i) Write down the conditional variance covariance matrix, $\Omega = E[v_i v_i|X_i]$, where $v_i = a_i + u_i$.

      (ii) Would the pooled OLS estimator yield a consistent estimator for β? Write down the estimator and justify your answer by showing the probability limit of the pooled OLS estimator. Make any additional assumptions you require to show the probability limit of the pooled OLS estimator

      (iii) Is the pooled OLS estimator asymptotically efficient? If not, propose an asymptotically efficient estimator assuming all elements in Ω are known.

      (iv) Describe briefly how you would estimate the elements in Ω in practice.

   (b) Now assume $E[a_i|X_i] \neq 0$, $E[u_{it}^2|X_i, a_i] = \sigma_u^2$ and $E[u_{it}u_{i,t-1}|X_i] = \rho$, and $E[u_{it}u_{is}|X_i] = 0$ for all $|s-t| > 1$.
   Under the above scenario:

      (i) Show whether the pooled OLS estimator is consistent for β or not. If not, propose an estimator that would be consistent.

      (ii) Is the consistent estimator (according to your answer in (i)) asymptotically efficient? If not, describe briefly how you could obtain an asymptotically efficient estimator. You can assume that σ_u^2 and ρ are known.

   (c) In practice, the empirical researcher does not know whether Scenario (a) or (b) hold in practice. How would you suggest to test which of the two scenarios hold? Clearly state the null hypothesis and discuss any issues with applying the test under the above scenarios.

Econometrics Preliminary Exam: Reference Sheet
Spring 2014

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

Assumption (Uniform Law of Large Numbers $\{\sup_{\theta\in\Theta} |\sum_{i=1}^{n} f(y_i, w_i; \theta)/n - E[f(y_i, w_i; \theta)]| \xrightarrow{p} 0\}$**)**

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

(ii) *(Compactness)* Θ is compact;

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

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

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

Assumption (Consistency of Sample Average of Hessian for M-Estimators)

(i) Each element of $H(y_i, x_i; \theta)$ is bounded in absolute value by a function $b(y_i, x_i)$, where $E[b(y_i, x_i)] < \infty$;

(ii) $A_0 = -E[H(y_i, x_i; \theta_0)]$ is positive definite.

Assumption (Asymptotic Normality of Sample Average of Score for M-Estimators)

(i) $E[s(y_i, x_i; \theta_0)] = 0$;

(ii) each element in $s(y_i, x_i; \theta_0)$ has finite second moment.

Formula for the score statistic

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

GMM Expression for the Sampling Error

$$\sqrt{n}(\hat{\theta} - \theta_0) \quad = \quad - \left(\mathcal{H}_0 W \frac{1}{n} \sum_{i=1}^{n} \left. \frac{\partial h(y_i, w_i; \theta)}{\partial \theta} \right|_{\theta = \theta^*} \right)^{-1} \mathcal{H}_0 W \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h(y_i, w_i; \theta_0) + o_p(1)$$

where

$$\mathcal{H}_0 \quad = \quad E \left[\left. \frac{\partial h(y_i, w_i; \theta)}{\partial \theta} \right|_{\theta = \theta_0} \right]$$

Econometrics Preliminary Examination

Department of Agricultural and Resource Economics
University of California, Davis
August 2014

Answer each part of each question. The three questions are weighted equally.

Within each question, each part receives equal weight.

1. This question has three unrelated parts.

   (a) Suppose the joint probability density function of X and Y is

   $$f(x,y) = \begin{cases} 2e^{-x-y} & 0 \le x \le y < \infty \\ 0 & \text{elsewhere} \end{cases}$$

   i. Find marginal density functions for X and Y.

   ii. Are X and Y independent? (Why or why not?)

   (b) The random variable U is distributed as Uniform$(0, 2\theta)$, with θ unknown and finite. You will have a random sample of size n from this probability distribution.

   i. Find the maximum-likelihood estimate of θ and demonstrate that it is consistent.

   ii. Your colleague suggests that $\tilde{\theta} = \overline{U}$ is also a consistent estimator. In fact, your colleague claims that $\tilde{\theta}$ is the best, linear, unbiased estimator of θ. Do you agree? Why or why not?

   (c) A random variable Y is known to be distributed on the interval between 0 and 2. Under the null hypothesis, Y is uniformly distributed, while under the alternative hypothesis,

   $$f(y) = \begin{cases} cy & 0 \le y \le 2 \\ 0 & \text{elsewhere} \end{cases}$$

   i. Assuming that you will have one observation available, find the most powerful test of your null hypothesis against your alternative hypothesis.

   ii. Find the power of your test and the probability of a type-II error.

2. A survey is used to construct a cross-sectional data set consisting of responses from both female and male employees of a local firm. OLS is used to estimate the parameters of a linear regression model

   $$y = X\beta + u,$$

   where X is an n by K matrix, y and u are n by 1 vectors, and all of the typical OLS assumptions are satisfied.

(a) Indicate how the typical OLS assumptions are used in establishing the result that the OLS estimator is BLUE.

(b) Explain in precise detail how to use OLS to test the null hypothesis that the vector of parameters, β, is the same for the two groups of employees, against the alternative hypothesis that the parameter vector differs in some unspecified way between the two groups.

(c) If your null hypothesis is false, what are the consequences for OLS estimates of a common parameter vector? (Hint: focus on the properties of your OLS estimator from part (a).)

(d) How would you estimate the parameters of your model if you believed that the parameter vectors differed between the two groups?

(e) Is your new estimator BLUE? (Indicate any assumptions required.)

(f) Suppose you used the approach you suggested in the previous part, but the null hypothesis is, in fact, true; what are the consequences for your answer in to the previous part of this question?

> *For the rest of this question, you may assume that the null hypothesis is correct; there is only one parameter vector, β, and it applies to all survey respondents. Any assumptions you required for part (a) above can be assumed to hold, with one exception, as follows:*
> *Suppose also that $V(u) = \Sigma$, a diagonal matrix.*
> *The elements on the main diagonal of Σ are either σ_F^2 (the variance of u_i for female observations) or σ_M^2 (the variance of u_i for male observations), where it is anticipated that $\sigma_F^2 \neq \sigma_M^2$.*

(g) What are the finite-sample consequences for OLS estimates of β?

(h) Under this same set of circumstances, what (if anything) would you do differently?

(i) How would you test the hypothesis that the variance of u_i is the same for every observation?

(j) Now you realize that the variance of each u_i does not depend on whether observation i is from a female or male employee, but instead that $V(u_i)$ is proportional to Z_i, a variable that varies continuously across the observations.

Repeat questions (e) through (g) under this new assumption.

3. Consider the following model,

$$y_i \quad = \quad g(x_i;\theta_0) + u_i$$

where y_i and u_i are scalar random variables, x_i and θ_0 are $k \times 1$ vectors. Assume $E[u_i|x_i] = 0$.

(a) Write down the objective function of the nonlinear least squares (NLS) estimator for θ_0. Denote the estimator by $\hat{\theta}_{NLS}$.

(b) Give primitive conditions that ensure that the identification condition for θ_0 holds.

(c) State the assumptions required for the consistency of the NLS estimator you proposed in (a). (No need to give primitive conditions.)

(d) Perform a mean-value expansion to derive an expression for $\sqrt{n}(\hat{\theta}_{NLS} - \theta_0)$.

(e) State the assumptions required for the asymptotic normality of $\sqrt{n}(\hat{\theta}_{NLS} - \theta_0)$ and write down its asymptotic variance. (No need to give primitive conditions.)

(f) Now let $k = 2$ and $g(x_i; \theta) = \theta_1 \exp(-\theta_2(x_1 + x_2))$.

   (i) You would like to test $H_0 : \theta_1 = \log(\theta_2)$. Write down the Wald statistic to test H_0.

   (ii) One of the issues with the Wald statistic is its lack of invariance, can you illustrate this for the hypothesis in (i)?

(g) Now back to the general nonlinear least squares problem. Suppose that $E[u_i|x_i] = 0$. Explain whether or not the identification condition in (b) would fail. If you conclude it would fail, describe briefly how you would identify θ_0 in this situation and what estimation methodology you could use in this setting.