Midterm Exam

ECON 837

Prof. Simon Woodcock, Spring 2009

You may use one page of hand-written notes, intution, and divine inspiration, but no electronic devices or collusion. There are 4 questions, each worth 25 points. You should answer all four questions. You have 110 minutes. Good luck, and remember: econometrics is fun!

1. [A Warmup] Consider the two regressions

$$\begin{aligned} E[y] &= x_1\beta_1 + x_2\beta_2 + x_3\beta_3 \\ E[y] &= z_1\alpha_1 + z_2\alpha_2 + z_3\alpha_3 \end{aligned}$$

   where $z_1 = x_1 - 2x_2$, $z_2 = x_1 + x_2$, and $z_3 = x_1 + 2x_2 + 2x_3$. What is the exact relationship between the least squares estimator of $\boldsymbol{\beta} = [\beta_1 \quad \beta_2 \quad \beta_3]'$ and $\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \quad \alpha_3]'$?

2. Suppose you draw three disjoint iid random samples from the same $N(\mu, \sigma^2)$ population: $(X_1, X_2, ..., X_{n_1})$, $(X_{n_1+1}, X_{n_1+2}, ..., X_{n_1+n_2})$, and $(X_{n_1+n_2+1}, X_{n_1+n_2+2}, ..., X_{n_1+n_2+n_3})$. Suppose we denote the sample mean and variance in the first sample as $\bar{x}_1$ and s_1^2, in the second sample as $\bar{x}_2$ and s_2^2, and in the third sample as $\bar{x}_3$ and s_3^2. Define a class of estimators:

$$s_c^2 = c_1 s_1^2 + c_2 s_2^2 + c_3 s_3^2$$

   where c_1, c_2 and c_3 are constants.

   (a) [5 points] Give a necessary and sufficient condition for s_c^2 to be an unbiased estimator of σ^2.

   (b) [10 points] Find the best unbiased estimator of σ^2 in the class of estimators s_c^2. What is its variance?

   (c) [10 points] Let s^2 denote the sample variance based on the pooled sample $X_1, X_2, ..., X_{n_1+n_2+n_3}$. Is s^2 more efficient than the estimator you found in part b? Prove your claim, and give the intuition for your result.

3. We say a random variable z has an exponential(γ) distribution if it has pdf:

$$f(z) = \frac{1}{\gamma} e^{-z/\gamma}, \quad \gamma > 0, \quad z \geq 0.$$

   (a) [5 points] Derive the moment generating function of z.

   (b) [10 points] Find the mean and variance of z.

   (c) [10 points] Suppose you specify the simple regression model $y_i = \alpha + \beta x_i + \varepsilon_i$, where the errors $\varepsilon_i \geq 0$ have an exponential(γ) distribution. Are the least squares estimators of α and β biased or unbiased? Prove your claim.

4. Suppose the data generating process is given by $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ where $\mathbf{X}_1$ is $n \times k_1$, $\mathbf{X}_2$ is $n \times k_2$, and the other quantities are vectors. Suppose you estimate this model (call it the "long" model) via OLS, and you also estimate the "short" model, which excludes $\mathbf{X}_2$.

   (a) [10 points] Derive the sum of squared residuals in both models, and sign their difference.

   (b) [10 points] Derive the **expected** sum of squared residuals in both models, and sign their difference.

   (c) [5 points] Suppose $\boldsymbol{\beta}_2 = \mathbf{0}$. Does this change your answers to parts a and b? Explain.

Some (possibly) useful results:

1. **Partitioned inverse formula:**

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1}\left(\mathbf{I} + \mathbf{A}_{12}\mathbf{F}_2\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\right) & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{F}_2 \\ -\mathbf{F}_2\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F}_2 \end{bmatrix}$$

   where $\mathbf{F}_2 = \left(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1}$. The upper left block can also be written $\mathbf{F}_1 = \left(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1}$.

2. **Integration by parts:** $\int_a^b f'(x)\,g(x)\,dx = [f(x)\,g(x)]_a^b - \int_a^b f(x)\,g'(x)\,dx$

3. **L'Hopital's Rule:** Consider the ratio $f(x)/g(x)$. If both f and g are continuous and differentiable, and if $\lim f(x) = \lim g(x) = 0$ or $\lim f(x) = \lim g(x) = \infty$, then:

$$\text{if } \lim \frac{f'(x)}{g'(x)} = L, \text{ then } \lim \frac{f(x)}{g(x)} = \lim \frac{f'(x)}{g'(x)} = L.$$