First Year Quantitative Comp Exam
Spring, 2012

Instruction: There are three parts. Answer every question in every part.

Part I - 203A

Question I-1

A random variable X is distributed with the marginal density:

$$f_X(x) = \begin{cases} 2(1-x) & if \quad 0 < x < 1 \\ \\ 0 & otherwise \end{cases}$$

The conditional cumulative distribution function of another random variable Y given X is

$$F_{Y|X=x}(y) = \begin{cases} 0 & if \quad y < x \\ \\ \frac{y-x}{1-x} & if \quad x \le y < 1 \\ \\ 1 & if \quad y \ge 1 \\ \end{cases}$$

1. Calculate the expectation of the random vector (Y, X).
2. Calculate the covariance of X and Y. Are X and Y independently distributed? Explain.
3. Calculate the probability that $X \in [.5, 1]$ conditional on $Y = .75$
4. Let $Z = Y - X$. Calculate the density of (Z, X).

Question I-2

An observable random variable Y is determined by an unobservable random variable α and an unobservable random variable ε, according to the model

$$Y = \begin{cases} 1 & if \quad \beta\,\alpha + \varepsilon > 0 \\ \\ 0 & otherwise \end{cases}$$

where β is a parameter of unknown value. The marginal density of the random variable ε is known to be $N\left(\mu_\varepsilon, \sigma_\varepsilon^2\right)$, for unknown values of μ_ε and σ_ε^2. The unobservable random variable α is known to be distributed $N\left(\mu_0, \sigma_0^2\right)$ when $Z = 0$ and $N\left(\mu_1, \sigma_1^2\right)$ when $Z = 1$, for unknown values of μ_0, μ_1, σ_0^2, and σ_1^2. The random variable Z is observable. Denote the probability that $Z = 0$ by the parameter p_0 and the probability that $Z = 1$ by the parameter p_1, where $p_0 + p_1 = 1$. Assume that (α, Z) and ε are independently distributed.

1. Obtain an expression for the probability that $Y = 1$ conditional on $Z = 0$ in terms of the unknown parameters.
2. Obtain an expression for the (marginal) probability that $Y = 1$.
3. Are p_0 and p_1 identified? Provide a proof for your answer.
4. Are μ_ε and σ_ε^2 identified? Provide a proof for your answer.

   Suppose next that the values of $\left(\mu_0, \sigma_0^2\right)$ and $\left(\mu_\varepsilon, \sigma_\varepsilon^2\right)$ are known, with $\mu_0 = 0$ and $\mu_\varepsilon \neq 0$.

5. Determine what parameters are identified. Provide proofs.
6. For the parameters that are not identified, can you provide bounds for their values? If your answer is YES, determine those bounds. If your answer is NO, explain.

Question I-3

Consider the following model:

$$Y = \begin{cases} 1 & if \quad \beta\, \alpha + X + \varepsilon > 0 \\ \\ 0 & \qquad otherwise \end{cases}$$

where the random variables Y, X, and Z are observable, the random variables α and ε are unobservable, and β is a parameter of unknown value. As in Question 2, the distribution of α depends on the value of Z. The random variable Z attains the value 0 with probability p_0 and the value 1 with probability p_1, where $p_0 + p_1 = 1$. Assume, further, that *(i)* the support of the continuous random variable X is R, *(ii)* the random variable ε is distributed $N(1, 4)$, and *(iii)* (α, Z), X, and ε are mutually independent.

1. Suppose first that the distribution of α when $Z = 0$ is $N(0, 16)$ and the distribution of α when $Z = 1$ is $N(\mu_1, \sigma_1^2)$, where the values of μ_1 and σ_1^2, as well as the values of p_0 and p_1 are unknown.
   (a) Determine the identified parameters. Provide proofs of your claims.
   (b) Given i.i.d. observations $\left\{Y^i, X^i, Z^i\right\}_{i=1}^N$, provide consistent estimators for the identified parameters.
   (c) Prove that your proposed estimators in 1.b are consistent. What can you say regarding the asymptotic distribution of your estimators?

2. Suppose now that the distribution $F_{\alpha|Z=1}$ of α when $Z = 1$ and the distribution $F_{\alpha|Z=0}$ of α when $Z = 0$ do not necessarily belong to a parametric family. They are only known to be strictly increasing and continuous functions. It is however still known that $Var(\alpha|Z = 0) = 16$. Answer the following questions and provide proofs.

   (a) Is β identified?

   (b) Is the distribution of $(\alpha + \varepsilon)$ identified?

   (c) Are $F_{\alpha|Z=1}$ and $F_{\alpha|Z=0}$ identified?

   (d) Is the (marginal) distribution of α identified?

Part II - 203B

Question II-1

Suppose that $X = (X_1, \ldots, X_n)'$ is such that X_i are iid $N(\theta_1, \theta_2)$. Compute the information for $\theta = (\theta_1, \theta_2)$ based on X. Let

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

Prove that s^2 is unbiased for θ_2. Prove that the finite sample variance of s^2 is strictly larger than the Cramer-Rao bound. (Let $n = 2$ when you make this comparison.) Derive the asymptotic distribution of $\sqrt{n}\left(s^2 - \theta_2\right)$ as $n \to \infty$, and show that the asymptotic variance of $\sqrt{n}\left(s^2 - \theta_2\right)$ is identical to the inverse of the Fisher information.

Question II-2

Suppose that

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \varepsilon_i$$

such that ε_i, x_{i1}, and x_{i2} are independent of each other with the common distribution $N(0,1)$. We assume that every variable is a scalar. Consider two estimators of β_1. The first estimator $\widetilde{\beta}_1$ is obtained by regressing y_i on x_{i1}:

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^{n} x_{i1} y_i}{\sum_{i=1}^{n} x_{i1}^2}$$

The second estimator $\widehat{\beta}_1$ is the first component when y_i is regressed on x_{i1} and x_{i2}:

$$\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1} x_{i2} \\ \sum_{i=1}^{n} x_{i1} x_{i2} & \sum_{i=1}^{n} x_{i2}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} x_{i1} y_i \\ \sum_{i=1}^{n} x_{i2} y_i \end{bmatrix}$$

Compute the asymptotic variances of $\sqrt{n}\left(\widetilde{\beta}_1 - \beta_1\right)$ and $\sqrt{n}\left(\widehat{\beta}_1 - \beta_1\right)$. Is it possible to determine which one is larger? If so, which one is more efficient?

Question II-3

Suppose that

$$y_i = x_i \beta_i + \varepsilon_i$$

where ε_i, x_i, and β_i are independent of each other. Note that β_i is a random variable. We assume that $(y_i, x_i)'$ $i = 1, 2, \ldots, n$ are observed, and they are iid. Let $\beta = E[\beta_i]$, and propose a consistent estimator of β. Prove why your estimator is consistent. Derive the asymptotic variance of $\sqrt{n}\left(\widehat{\beta} - \beta\right)$, where $\widehat{\beta}$ denotes your proposed estimator.

Part III - 203C

Question III-1

Suppose that $X_1, ..., X_n$ is an *iid* sample from a distribution having density function of the form

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & \text{if } x \in (0,1) \\ 0 & \text{otherwise} \end{cases}.$$

Show that a best critical region for testing $H_0 : \theta = 1$ against $H_1 : \theta = 2$ is

$$C = \left\{(X_1, ..., X_n) : c \leq \prod_{i=1}^{n} X_i\right\}.$$

Question III-2

Suppose that

$$y_t = \alpha y_{t-1} + u_t$$
$$u_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

where $\varepsilon_t \sim iid\ \left(0, \sigma_\varepsilon^2\right)$.

1. Suppose that $|\alpha| < 1$. Is the process $\{y_t\}$ covariance stationary? Derive the autocovariance function of $\{y_t\}$.

2. Under the assumption $|\alpha| < 1$, is the OLS estimate $\widehat{\alpha}_n$ defined as

$$\widehat{\alpha}_n = \frac{\sum_{t=1}^{n} y_t y_{t-1}}{\sum_{t=1}^{n} y_{t-1}^2} \tag{1}$$

   a consistent estimator of α?

3. Suppose $|\alpha| < 1$ still hold. Now we estimate α by using the Instrumental Variable (IV) method, with y_{t-2} being the instrument for y_{t-1}. The IV estimator $\widehat{\alpha}_n^{IV}$ is defined as

$$\widehat{\alpha}_n^{IV} = \frac{\sum_{t=2}^{n} y_t y_{t-2}}{\sum_{t=2}^{n} y_{t-1} y_{t-2}}.$$

   Is the IV estimator $\widehat{\alpha}_n^{IV}$ a consistent estimator of α? Derive its limiting distribution.

4. Now suppose that $\alpha = 1$, is the OLS estimate $\widehat{\alpha}_n$ defined above a consistent estimator of α?

5. Derive the limiting distribution of the OLS estimate $\widehat{\alpha}_n$ under $\alpha = 1$.

Question III-3

Suppose that

$$Y_t = X_t\beta + u_t + c\sum_{s=0}^{t-1} u_s$$

where $u_t \sim iid\ (0,1)$, c and β are some unknown constants and

$$X_t = X_{t-1} + \varepsilon_t$$

where ε_t is an $iid\ (0,1)$ sequence independent of u_s for all t and s, and $X_0 = 0$. You run a regression of Y_t on X_t and get the following OLS estimate

$$\widehat{\beta}_n = \frac{\sum_{t=1}^n X_t Y_t}{\sum_{t=1}^n X_t^2}.$$

1. Suppose that $c \neq 0$. Is $\widehat{\beta}_n$ a consistent estimate of β?
2. Suppose that $c = 0$. Is $\widehat{\beta}_n$ a consistent estimate of β? Derive its limiting distribution.
3. Using the results in (1) and (2), construct a statistic for testing $H_0 : c = 0$ against $H_1 : c \neq 0$. Why is your test consistent?
4. Let $\Delta X_t = X_t - X_{t-1}$ and $\Delta Y_t = Y_t - Y_{t-1}$. What's the limiting distribution of the following estimate?
$$\widehat{\beta}_n^* = \frac{\sum_{t=2}^n \Delta X_t \Delta Y_t}{\sum_{t=2}^n \Delta X_t^2}.$$
5. Which estimate, i.e. $\widehat{\beta}_n$ or $\widehat{\beta}_n^*$, do you prefer? Justify your choice (Hint: your answer may depend on the value of c).