Econometrics prelim, Spring 2021

1. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables. Suppose that $\mathbb{P}(X_1 \leq 0) = 0$ and $\lim_{h \downarrow 0} \mathbb{P}(X_1 \leq x + h \mid X_1 > x)/h = 1$ for all $x > 0$. Further, let $X_{(j)}$ be the j^{th} smallest variable among $X_1, X_2, \ldots, X_n$.

   (a) Please compute $\mathbb{E}(X_1)$ and $\mathbb{E}(X_{(1)})$.

   (b) Let F be the cumulative distribution function of X_1. Please derive the cumulative distribution function and the probability density function of $Y_{(j)} := F(X_{(j)})$ for $j = 1, 2, \ldots, n$.

2. Suppose that X and Y are random variables with $\mathbb{E}\{\exp(tX + sY)\} = \exp\{(t^2 + s^2)/2\}$ for all $t, s \in \mathbb{R}$. Please compute $\mathbb{E}(2X + Y \mid X + 2Y > 0)$.

3. Suppose that (X, Y) has a joint probability density function $f_{XY}(x, y) = 8xy\mathbb{I}(0 \leq x \leq y \leq 1)$.

   (a) Obtain the function $g : (0, 1) \to \mathbb{R}$ such that for any measurable function $h : (0, 1) \to \mathbb{R}$, we have $\mathbb{E}\big[\{Y - g(X)\}^2\big] \leq \mathbb{E}\big[\{Y - h(X)\}^2\big]$.

   (b) Let $Z = X^2$. Obtain the probability density function of Z.

   (c) Compute the correlation coefficient between $Y - g(X)$ and Z, where g and Z are defined in parts (a) and (b), respectively.

4. Suppose that (Y_1, Y_0) and T are independent, where T is Bernoulli. Define $Y := Y_1 T + Y_0(1 - T)$.

   (a) Can we identify the (marginal) distribution functions of Y_0 and Y_1 from the joint distribution of (Y, T)? Prove or disprove.

   (b) Suppose that Y_1 and Y_0 are also Bernoulli and that they are dependent such that $Y_1 \geq Y_0$ with probability one. Can we identify $\mathbb{P}(Y_1 = 1 \mid Y_0 = 0)$ from the joint distribution of (Y, T)? Prove or disprove.

5. Suppose that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with $\mathbb{E}(X_1^4) < \infty$. Let $\hat{\mu} = n^{-1} \sum_{i=1}^n X_i$ and $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n X_i^2 - \hat{\mu}^2$. You may assume that $\sigma^2 > 0$ and $\mathbb{E}\{(X_1 - \mu)^4\} > \sigma^4$, where $\mu = \mathbb{E}(X_1)$ and $\sigma^2 = \mathbb{V}(X_1)$.

   (a) Please derive the (non-degenerate) limit distribution of $n^r(\hat{\sigma}^2 - \sigma^2)$ for some appropriate number $r > 0$.

   (b) Please derive the (non-degenerate) limit distribution of $n^r(\hat{\mu}^2 - \mu^2)$ for some appropriate number $r > 0$.

6. Prove or disprove each of the following statements. You may assume that all the random variables are defined on the same probability space.

   (a) If $\mathbb{E}(|X_n|^3) = O(n)$, then we must have $X_n = O_p(n^{1/3})$.

   (b) If $X_n \stackrel{as}{\to} 0$, then we must have $\mathbb{E}(X_n) \to 0$.

10. Assume i.i.d. data and consider the model

$$\begin{cases} y_i^* = x_i^\top \beta_0 + q_i \gamma_0 + u_i, \\ q_i = w_i^\top \delta_0 + v_i, \end{cases} \tag{2}$$

where y_i^* is latent, u_i, v_i are mean zero jointly normal and independent of $z_i = [x_i^\top, w_i^\top]^\top$ and u_i has variance one. Assume that everything is well–behaved in terms of collinearity and the existence of moments and that w_i, x_i are vectors with multiple elements that have no elements in common. Instead of y_i^* you observe $y_i = \mathbb{1}(y_i^* \geq 0)$. It can be shown that

$$\mathbb{E}(y_i \mid z_i = z) = \Phi\left(\frac{x^\top \beta_0 + w^\top \delta_0 \gamma_0}{\sqrt{1 + \gamma_0^2 \sigma_0^2 + 2\gamma_0 \sigma_0 \rho_0}}\right), \tag{3}$$

where σ_0^2, ρ_0 are respectively the variance of v_i and the correlation of u_i, v_i. Now suppose that you form moment conditions

$$\begin{cases} \mathbb{E}\left[z_i \left\{ y_i - \Phi\left(\frac{x_i^\top \beta_0 + w_i^\top \delta_0 \gamma_0}{\sqrt{1 + \gamma_0^2 \sigma_0^2 + 2\gamma_0 \sigma_0 \rho_0}}\right)\right\}\right] = 0, \\ \mathbb{E}\{z_i (q_i - w_i^\top \delta_0)\} = 0, \\ \mathbb{E}\{(q_i - w_i^\top \delta_0)^2 - \sigma_0^2\} = 0. \end{cases} \tag{4}$$

(a) Do the moment conditions provide underidentification, exact identification, or overidentification? Explain briefly.

(b) Name one advantage of using GMM with an overidentified system compared to using an asymptotically equally efficient GMM estimator based on an exactly identified system.

(c) Finish the sentence: "Exactly identified GMM using a given fixed vector of instruments z_i is to OLS what GMM with optimal instruments is to ...."

11. Suppose that $y_i = x_i^\top \beta_0 + u_i$, where

$$\begin{cases} x_i = \Pi_0^\top z_i + v_i, \\ u_i = \exp(z_i^\top \gamma_0) e_i, \end{cases} \tag{5}$$

where e_i is mean zero unit variance and e_i, v_i are independent of z_i and all variables have finite fourth moments.

(a) Write down some moment conditions that provide identification under reasonable additional conditions and state what those conditions are.

(b) Derive the optimal instruments.

12. True or false? Please explain.

(a) This exam is hard.

....