Problem Set 1

Due February 1, 2021

Please include your name on the solutions

Exercises

Exercise 1. Suppose $(Y, X')'$ is a random vector with

$$Y = X'\beta_0 \cdot U$$

where $\mathbb{E}[U \mid X] = 1$, $\mathbb{E}[XX']$ is invertible, and $\mathbb{E}[Y^2 + \|X\|^2] < \infty$. Furthermore, suppose that $\{(Y_i, X_i')'\}_{i=1}^n$ is a random sample from the distribution of $(Y, X')'$ where $\frac{1}{n}\sum_{i=1}^n X_i X_i'$ is invertible and let $\hat{\beta}$ be the OLS estimator, i.e.,

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^n X_i Y_i$$

(i) Interpret the entries of β_0 in this model?

(ii) Show that $Y = X'\beta_0 + \tilde{U}$ where $\mathbb{E}[\tilde{U} \mid X] = 0$.

(iii) Show that $\mathbb{E}[X(Y - X'\beta)] = 0$ if and only if $\beta = \beta_0$ and use this to derive OLS as a method of moments estimator.

(iv) Show that the OLS estimator is conditionally unbiased, i.e., that

$$\mathbb{E}\left[\hat{\beta} \mid X_1, \ldots, X_n\right] = \beta_0$$

(v) Show that the OLS estimator is consistent, i.e., that $\hat{\beta} \xrightarrow{p} \beta_0$ as $n \to \infty$.

Exercise 2. Let X be a random variable with $\mathbb{E}[X^4] < \infty$ and $\mathbb{E}[X^2] > 0$. Furthermore, let $\{X_i\}_{i=1}^n$ be a random sample from the distribution of X.

(i) For which of the following four statistics can you use the law of large numbers and continuous mapping theorem to show convergence in probability as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^n X_i^3 \qquad \max_{1\le i\le n} X_i \qquad \frac{\sum_{i=1}^n X_i^3}{\sum_{i=1}^n X_i^2} \qquad 1\left\{\frac{1}{n}\sum_{i=1}^n X_i > 0\right\}$$

(ii) For which of the following three statistics can you use the central limit theorem and continuous mapping to show convergence in distribution as $n \to \infty$,

$$W_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i^2 - \mathbb{E}[X_1^2]) \qquad W_n^2 \qquad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(X_i^2 - \overline{X_n^2} \right)$$

where $\overline{X_n^2} = \frac{1}{n} \sum_{i=1}^{n} X_i^2$.

(iii) Show that $\max_{1 \le i \le n} X_i \xrightarrow{p} 1$ if $X \sim uniform(0,1)$.

(iv) Show that $\Pr\left(\max_{1 \le i \le n} X_i > M\right) \to 1$ for any $M \ge 0$ if $X \sim exponential(1)$.

Exercise 3. Suppose that $\{X_i\}_{i=1}^n$ is an *i.i.d.* sequence of $N(0,1)$ random variables. Let W be independent of $\{X_i\}_{i=1}^n$ with $\Pr(W = 1) = \Pr(W = -1) = \frac{1}{2}$. Let $Y_i = X_i W$.

(i) Show that $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{d} N(0,1)$ as $n \to \infty$.

(ii) Show that $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i \xrightarrow{d} N(0,1)$ as $n \to \infty$.

(iii) Show that $Cov(X_i, Y_i) = 0$.

(iv) Does $V := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i, Y_i)' \xrightarrow{d} N(0, I_2)$ as $n \to \infty$?

(v) How does this exercise relate to the Cramér-Wold device introduced in lecture 2?

Optional exercises

Exercise 4. Consider the measurement error model from lecture 3 in the presence of two noisy measurements:

$$Y = \beta_0 + X^* \beta_1 + \varepsilon$$

where X^* is exogenous, $\mathbb{E}[\varepsilon \mid X^*] = 0$, but X^* is unobserved. Instead of X^* we have two noisy measurements

$$X_1 = X^* + V_1 \qquad X_2 = X^* + V_2$$

where V_1 and V_2 are independent of X^* and ε.

(i) Can we construct an IV estimator based on this model and would we have to add any further conditions to ensure that the assumptions of the simple linear IV model are satisfied.

Exercise 5. Consider a simple linear regression model

$$Y = \beta_0 + X\beta_1 + U$$

where the object of interest is β_1 and we preferably want to give some causal interpretation to β_1. Suppose that X is endogenous in the sense that $Cov(X, U) \neq 0$ and Z is "nearly" a valid instrument in the sense that $Cov(Z, X) \neq 0$ and $Cov(Z, U)$ is close to zero.

(i) Find β_1^{OLS}, the unique solution to

$$Cov(X, (Y - X\beta)) = 0$$

(ii) Find β_1^{IV}, the unique solution to

$$Cov(Z, (Y - X\beta)) = 0$$

(iii) Discuss when we expect β_1^{IV} to be closer to β_1 than β_1^{OLS}.