Econ 710 Problem Set 1

Sarah Bass *

January 29, 2021

Question 1

(i)

The entries of β_0 in this model tells us how the conditional expectation of Y given X changes with X.

$$\begin{aligned}\mathbb{E}[Y|X] &= \mathbb{E}[X'\beta_0 \cdot U|X] \\ &= X'\beta_0 \mathbb{E}[U|X] \\ &= X'\beta_0\end{aligned}$$

(ii)

Let $\tilde{U} := X'\beta_0(U-1)$.

$$\begin{aligned}Y &= X'\beta_0 \cdot U \\ &= X'\beta_0 \cdot U + X'\beta_0 - X'\beta_0 \\ &= X'\beta_0 + \tilde{U}\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\tilde{U}|X] &= \mathbb{E}[X'\beta_0(U-1)|X] \\ &= \mathbb{E}[X'\beta_0(U-1)|X] \\ &= X'\beta_0\mathbb{E}[(U-1)|X] \\ &= 0\end{aligned}$$

*I have discussed this problem set with Emily Case, Michael Nattinger, Alex Von Hafften, Katherine Kwok, and Danny Edgel.

(iii)

$$
\begin{aligned}
\mathbb{E}[X(Y - X'\beta)] &= \mathbb{E}[X(X'\beta_0 + \tilde{U} - X'\beta)] \\
&= \mathbb{E}[XX'\beta_0] + \mathbb{E}[X\tilde{U}] - \mathbb{E}[X(X'\beta)] \\
&= \mathbb{E}[XX'\beta_0] + \mathbb{E}[X\mathbb{E}[\tilde{U}|X]] - \mathbb{E}[XX'\beta] \\
&= \mathbb{E}[XX'](\beta_0 - \beta) \\
&= 0 \text{ if and only if } \beta_0 = \beta
\end{aligned}
$$

Then we can derive OLS as a methods of moments estimator as follows. Consider $\hat{\beta}_{MM}$

$$
\begin{aligned}
0 &= \frac{1}{n}\sum_{i=1}^{n} X_i Y - X_i X_i' \hat{\beta}_{MM} \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i Y - \frac{1}{n}\sum_{i=1}^{n} X_i X_i' \hat{\beta}_{MM} \\
\Rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i Y &= \frac{1}{n}\sum_{i=1}^{n} X_i X_i' \hat{\beta}_{MM} \\
\Rightarrow \hat{\beta}_{MM} &= \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i Y = \hat{\beta}_{OLS}
\end{aligned}
$$

(iv)

$$\begin{aligned}
\mathbb{E}[\hat{\beta}|X_1,...,X_n] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i Y_i \middle| X_1,...,X_n\right] \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i (X_i'\beta_0 + \tilde{U}) \middle| X_1,...,X_n\right] \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i X_i'\beta_0 + X_i\tilde{U} \middle| X_1,...,X_n\right] \\
&= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\beta_0 + \frac{1}{n}\sum_{i=1}^{n} X_i\tilde{U}\right) \middle| X_1,...,X_n\right] \\
&= \mathbb{E}\left[\beta_0 + \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i\tilde{U} \middle| X_1,...,X_n\right] \\
&= \beta_0 + \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i \mathbb{E}\left[\tilde{U} \middle| X_1,...,X_n\right] \\
&= \beta_0
\end{aligned}$$

(v)

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

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} X_i X_i' &\to_p \mathbb{E}[XX'] \text{ by WLLN} \\
\frac{1}{n}\sum_{i=1}^{n} X_i Y &= \frac{1}{n}\sum_{i=1}^{n} X_i (X_i'\beta_0 + \tilde{U}) \\
&\to_p \mathbb{E}[XX']\beta_0 \text{ by WLLN}
\end{aligned}$$

Then by the continuous mapping theorem,

$$\begin{aligned}
\hat{\beta}_{OLS} &\to_p \mathbb{E}[XX']^{-1}\mathbb{E}[XX']\beta_0 \\
&= \beta_0
\end{aligned}$$

Question 2

(i)

We can apply the law of large numbers and continuous mapping theorem to show convergence in probability to the first and third statistics. Because we know $\mathbb{E}[X^4] < \infty$, we know that the expectation is finite for X^3, and the first statistic $\frac{1}{n}\sum_{i=1}^n X_i^3 \to_p \mathbb{E}[X_3]$. Similarly, we know that the expectation is finite for X^2, and $\frac{1}{n}\sum_{i=1}^n X_i^2 \to_p \mathbb{E}[X_2]$. Since $\mathbb{E}[X^2] > 0$, we know that the denominator of the third statistic is well defined, so $\frac{\frac{1}{n}\sum_{i=1}^n X_i^3}{\frac{1}{n}\sum_{i=1}^n X_i^2} \to_p \frac{\mathbb{E}[X_3]}{\mathbb{E}[X_2]}$.

The second statistic does not contain an average, so we cannot use the law of large numbers to show convergence in probability. Since we do not know that $\mathbb{E}[X] \neq 0$, there may be a point of discontinuity at 0 for the fourth statistic, so the continuous mapping theorem cannot be applied.

(ii)

Using the central limit theorem, we know that $W_n \to_d N(0, var(X_i^2))$ since $\mathbb{E}[X_i^2 - \mathbb{E}[X_1^2]] = 0$ and $\mathbb{E}[X^4] < \infty$. Using the continuous mapping theorem, we know that $W_n^2 \to_d N(0, var(X_i^2))^2$ (a scaled Chi-squared).

We can rewrite the third statistic as follows:

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i^2 - \bar{X_n^2}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i^2 - \frac{1}{n}\sum_{j=1}^n X_j^2) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i^2 - \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{1}{n}\sum_{j=1}^n X_j^2 \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i^2 - \frac{1}{\sqrt{n}}\sum_{j=1}^n X_j^2 \\
&= 0
\end{aligned}$$

Thus, the central limit theorem does not apply.

(iii)

$$\begin{aligned}
P(\left| \max_{1\le i\le n} X_i - 1 \right| < \varepsilon) &= P(\max_{1\le i\le n} X_i \ge 1 - \varepsilon) \\
&= 1 - P(\max_{1\le i\le n} X_i < 1 - \varepsilon) \\
&= 1 - (1-\varepsilon)^n \\
&\to_p 1 \text{ as } n \to \infty
\end{aligned}$$

(iv)

$$\begin{aligned} P(\max_{1\le i\le n} X_i > M) &= 1 - P(\max_{1\le i\le n} X_i \le M) \\ &= 1 - (1 - e^{-M})^n \\ &\to 1 \text{ as } n \to \infty \end{aligned}$$

Question 3

(i)

By the CLT,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i - \mathbb{E}[X_i] \\ &\to_d N(0,1) \end{aligned}$$

(ii)

First note that

$$\begin{aligned} \mathbb{E}[Y_i] &= \mathbb{E}[X_i]\mathbb{E}[W] = 0 \\ \mathbb{E}[Y_i^2] &= \mathbb{E}[X_i^2]\mathbb{E}[W^2] = 1 \end{aligned}$$

Then by the CLT

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i W \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i W - \mathbb{E}[X_i W] \\ &\to_d N(0,1) \end{aligned}$$

(iii)

$$\begin{aligned} Cov(X_i, Y_i) &= \mathbb{E}[X_i Y_i] - \mathbb{E}[X_i]\mathbb{E}[Y_i] \\ &= \mathbb{E}[X_i^2 W] - 0 \\ &= \mathbb{E}[X_i^2]\mathbb{E}[W] \\ &= 0 \end{aligned}$$

(iv)

No. We can prove that V does not converge in distribution to $N(0, I_2)$ using the Cramer-Wold device. Consider $t = \langle \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \rangle$. When $W = -1$, $t'V = 0$, which means that $P(t'V = 0) \geq \frac{1}{2} > 0$ (e.g. there is a unit point mass at 0, so $t'V$ is not normal). Thus $t'V$ does not converge in distribution to $N(0, t'I_2t)$

(v)