Econometrics HW1

Michael B. Nattinger*

January 27, 2021

1 Question 1

1.1 Part i

β_0 are the true coefficients of the conditional expectation function of Y conditional on X:

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

1.2 Part ii

Define $\bar{U} := X'\beta_0(U-1)$. Then,

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

Moreover, $E[\bar{U}|X] = E[X'\beta_0(U-1)|X] = X'\beta_0(E[U|X]-1) = 0$.

1.3 Part iii

Assume $\beta = \beta_0$. Then, by the conditioning theorem,

$$\begin{aligned} E[X(Y - X'\beta)] &= E[X(Y - X'\beta_0)] \\ &= E[XE[Y - X'\beta_0|X]] \\ &= E[XE[X'\beta_0 + \bar{U} - X'\beta_0|X]] \\ &= E[XE[\bar{U}|X]] \\ &= 0. \end{aligned}$$

Now, assume instead that $E[X(Y - X'\beta)] = 0$. Then, again using the conditioning theorem,

$$\begin{aligned} 0 &= E[X(Y - X'\beta)] \\ &= E[XE[Y - X'\beta|X]] \\ &= E[XE[X'\beta_0 + \bar{U} - X'\beta|X]] \\ &= E[XE[X'(\beta_0 - \beta)|X]] \\ &= E[XX'](\beta_0 - \beta) \end{aligned}$$

*I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, Katherine Kwok, and Danny Edgel.

We know $E[XX']$ is invertible so it must be the case that $(\beta_0 - \beta) = 0 \Rightarrow \beta_0 = \beta$.

We have proven both directions of the iff, ergo $E[X(Y - X'\beta)] = 0$ iff $\beta = \beta_0$. We can now define our method of moments estimator $\hat{\beta}^{MM}$ to be the unique solution to the following equation:

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

Therefore, the OLS estimator is a method of moments estimator.

1.4 Part iv

By application of various given definitions,

$$
\begin{aligned}
E[\hat{\beta}|X_1, \ldots, X_n] &= E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1}\frac{1}{n}\sum_{i=1}^{n} X_i Y_i \Big| X_1, \ldots, X_n\right] \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1}\frac{1}{n}\sum_{i=1}^{n} X_i E[Y_i|X_1, \ldots, X_n] \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1}\frac{1}{n}\sum_{i=1}^{n} X_i E[X_i'\beta_0 \cdot U_i|X_1, \ldots, X_n] \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1}\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\beta_0 E[U_i|X_1, \ldots, X_n] \\
&= \beta_0.
\end{aligned}
$$

Therefore, the OLS estimator is conditionally unbiased.

1.5 Part v

By the LLN and CMT, the following are true as $n \to \infty$:

$$
\begin{aligned}
\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} &\to_p E[XX']^{-1}, \\
\frac{1}{n}\sum_{i=1}^{n} X_i Y_i &= \frac{1}{n}\sum_{i=1}^{n} X_i(X_i'\beta_0 + \bar{U}) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n} X_i X_i'\right)\beta_0 + \frac{1}{n}\left(\sum_{i=1}^{n} X_i \bar{U}\right) \\
&\to_p E[XX']\beta_0 + E[X\bar{U}] \\
&= E[XX']\beta_0 + E[XE[\bar{U}|X]] \\
&= E[XX']\beta_0,
\end{aligned}
$$

where we have applied the conditioning theorem in the second-to-last line. By further application of the CMT, as $n \to \infty$,

$$\begin{aligned}\left(\frac{1}{n}\sum_{i=1}^{n} X_i X_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} X_i Y_i &\to_p E[XX']^{-1}E[XX']\beta_0 \\ &= \beta_0.\end{aligned}$$

2 Question 2

2.1 Part i

We can immediately apply LLN and CMT to show convergence in probability as $n \to \infty$ to the following statistics:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^3,$$
$$\frac{\sum_{i=1}^{n} X_i^3}{\sum_{i=1}^{n} X_i^2}$$

Note that this works for the second expression because we are given that $E[X_i^2] > 0$.

$\max_{1\leq i\leq n}\{X_i\}$ does not always converge in probability. For example, in part (iv) of this question we prove that, for $X_i \sim exponential(1)$, this function has no probability limit.[1] The reason why we cannot apply LLN and CMT is because this function does not involve averages.

In most cases we can apply LLN and CMT to $1\{\frac{1}{n}\sum_{i=1}^{n} X_i > 0\}$ so long as $E[X_i] \neq 0$. If instead $E[X_i] = 0$, the function is not continuous at the relevant moment and therefore CMT does not apply.

2.2 Part ii

We can apply the central limit theorem and continuous mapping theorem to the first two as all of the transformations to the random variables are continuous. The first distribution is converges in distribution to $N(0, Var(X_i^2))$ by application of the central limit theorem and we can further apply the continuous mapping theorem on that distribution to find the asymptotic distribution of the second statistic - that is, a scaled chi-squared. However, the third expression is trivially 0.[2] We can therefore not use the central limit theorem in this case.

2.3 Part iii

Define $M_n := \max_{1\leq i\leq n} X_i$ and let $X \sim uniform(0,1)$. Let $\epsilon > 0$ be arbitrary. If $\epsilon \geq 1$ then $P(|M_n - 1| \leq \epsilon) = 1 \geq 1 - \epsilon$ so the definition of convergence in probability is trivially satisfied. Assume instead that $\epsilon \in (0,1)$.

$$\begin{aligned}P(|M_n - 1| \leq \epsilon) &= P(1 - M_n \leq \epsilon) \\ &= 1 - \prod_{i=1}^{n} P(X_i < 1-\epsilon) \\ &= 1 - (1-\epsilon)^n\end{aligned}$$

[1]The only part of this statement not explicitly shown in part (iv) is that the proven statement shows that this function has no probability limit. However, it is trivial to see that for any candidate maximum M and any $N > M$, we can draw some X_i larger than N with probability approaching 1 as we let $n \to \infty$, so it clearly is impossible for any such M to be the probability limit of the maximum function.

[2]Here is a quick proof that the third expression is 0: $\sum_{i=1}^{n}(X_i^2 - (1/n)\sum_{j=1}^{n} X_j^2) = \sum_{i=1}^{n}(X_i^2) - \sum_{j=1}^{n} X_j^2 = 0$.

$1-(1-\epsilon)^n \to 1$ so $\exists N$ such that for all $n > N, 1-(1-\epsilon)^n > 1-\epsilon$. Therefore, for all $n > N$, $P(|M_n - 1| \leq \epsilon) = 1 - (1-\epsilon)^n \geq 1 - \epsilon$. Therefore, $M_n \to_p 1$.

2.4 Part iv

As before, define $M_n := \max_{1 \leq i \leq n} X_i$ but now let $X \sim exponential(1)$. Let $M \geq 0$.

$$\begin{aligned}
P(M_n > M) &= 1 - P(M_n < M) \\
&= 1 - \prod_{i=1}^{n} P(X_i < M) \\
&= 1 - (1 - exp(-M))^n \\
&\to_{n\to\infty} 1.
\end{aligned}$$

3 Question 3

3.1 Part i

By the central limit theorem,

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

where $V = Var(X_i) = 1$.

3.2 Part ii

$E[Y_i] = E[X_i|W = 1]P(W = 1) + E[-X_i|W = -1]P(W = -1) = 0(0.5) + 0(0.5) = 0$.
$E[Y_i^2] = E[X_i^2|W = 1]P(W = 1) + E[X_i^2|W = -1]P(W = -1) = 1(0.5) + 1(0.5) = 1$. By the central limit theorem,

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

where $W = Var(Y) = 1$.

3.3 Part iii

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

3.4 Part iv

No. Conditional on the draw of W, V converges in distribution to different distributions. We can show that it does not unconditionally converge to $N(0,1)$ by applying the C-W device. Let $t = (1/\sqrt{2}, 1/\sqrt{2})'$. Then, if $W = -1, t'V = 0$ so $P(t'V = 0) \geq 1/2 > 0$ so $t'V$ does not have a continuous distribution.

3.5 Part v

The naive answer to part (iv) is yes, but by simple application of Cramer-Wold we see instead that the joint asymptotic distribution does not actually exist. The 'problem' with this random vector's asymptotic properties is that there are actually two possible asymptotic distributions for V conditional on W, so there is no unconditional asymptotic distribution.