Economics 241B – 2015 Final

Question 1

(a) Find $\mathbb{E}\left(\mathbb{E}\left(\mathbb{E}\left(u|x_1, x_2, x_3\right)|x_1, x_2\right)|x_1\right)$.

(b) Prove: For any function $h(x)$ such that $\mathbb{E}\left|h(x)u\right| < \infty$, if $\mathbb{E}(u|x) = 0$ then $\mathbb{E}(h(x)u) = 0$. Can this result be used to establish that if $\mathbb{E}(u|x) = 0$ then $\mathbb{E}(xu) = 0$?

(c) If $y = x\beta + u$, $x \in \mathbb{R}$, then for each of the following statements either establish that they are true or provide a counterexample:

i) $\mathbb{E}(u|x) = 0$ implies $\mathbb{E}(x^2u) = 0$,

ii) $\mathbb{E}(xu) = 0$ implies $\mathbb{E}(x^2u) = 0$,

iii) $\mathbb{E}(u|x) = 0$ implies $\mathbb{E}(y|x) = x\beta$,

iv) $\mathbb{E}(xu) = 0$ implies $\mathbb{E}(y|x) = x\beta$.

(d) Recall that the conditional variance is $\sigma^2(x) = Var(y|x) = \mathbb{E}\left((y - \mathbb{E}(y|x))^2|x\right)$. Show that the conditional variance can be written as

$$\sigma^2(x) = \mathbb{E}\left(y^2|x\right) - \left(\mathbb{E}(y|x)\right)^2.$$

Question 2

You wish to measure the effect of school spending on student achievement. You collect data at the school level and propose the following model,

$$T_i = \beta_0 + \beta_1 Exp_i + \beta_2 Size_i + u_i,$$

where T_i is the average standardized test score for students at school i, Exp_i is the expediture-per-pupil, and $Size_i$ is the average class size. The coefficient of interest is β_1. Denote X_i as the 3x1 vector of covariates and β as the 3x1 vector of coefficients.

(a) Show that the vector β is identified and derive $\hat{\beta}^{OLS}$. Be clear about the assumptions needed to ensure identification.

Answer:

$$T_i = X_i'\beta + u_i$$
$$X_iT_i = X_iX_i'\beta + X_iu_i$$
$$\mathbb{E}[X_iT_i] = \mathbb{E}[X_iX_i']\beta + \mathbb{E}[X_iu_i]$$
$$\beta = \mathbb{E}[X_iX_i']^{-1}\mathbb{E}[X_iT_i] + \mathbb{E}[X_iX_i']^{-1}\mathbb{E}[X_iu_i]$$
$$\text{(Assuming that } \mathbb{E}[X_iX_i'] < \infty \text{ and is invertible)}$$
$$\beta = \mathbb{E}[X_iX_i']^{-1}\mathbb{E}[X_iT_i]$$
$$\text{(Assuming that } \mathbb{E}[X_iu_i] = 0\text{)}$$

Given that we can approximate both of the above population moments with sample moments, we have shown that β is identified. We can now derive the OLS estimator using those sample analogs because the OLS and Method of Moments estimators are equivalent in linear models.

$$\hat{\beta}^{OLS} = \left(\frac{1}{n}\sum_i X_iX_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_iT_i\right)$$

(b) Suppose that the data (T_i, X_i) are i.i.d. (note that this implies homoskedasticity, $\mathbb{E}[u_i^2|X] = \sigma^2$). Show that $\hat{\beta}^{OLS}$ is consistent and asymptotically normal.

Answer: First note that

$$\begin{aligned}
\hat{\beta}^{OLS} &= \left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_i T_i\right) \\
&= \left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_i \left(X_i'\beta + u_i\right)\right) \\
&= \beta + \left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_i u_i\right)
\end{aligned}$$

The Law of Large Numbers gives us that

$$\begin{aligned}
\frac{1}{n}\sum_i X_i X_i' &\to_p \mathbb{E}[X_i X_i'] \\
\frac{1}{n}\sum_i X_i u_i &\to_p \mathbb{E}[X_i u_i]
\end{aligned}$$

Continuous Mapping Theorem gives us

$$\left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1} \to_p \mathbb{E}[X_i X_i']^{-1}$$

Because $\mathbb{E}[X_i u_i] = 0$ by assumption

$$\begin{gathered}
\hat{\beta}^{OLS} \to_p \beta + \mathbb{E}[X_i X_i']^{-1} * 0 \text{ (Using Slutsky)} \\
\hat{\beta}^{OLS} \to_p \beta
\end{gathered}$$

Thus, $\hat{\beta}^{OLS}$ is consistent. Then we can use the Central Limit Theorem to show that

$$\begin{gathered}
\sqrt{n}\left(\left(\frac{1}{n}\sum_i X_i u_i\right) - \mathbb{E}[X_i u_i]\right) \to_d N(0, Var\left(X_i u_i\right)) \\
\sqrt{n}\left(\left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_i u_i\right) - 0\right) \\
\to_d \mathbb{E}[X_i X_i']^{-1} * N(0, Var(X_i u_i))
\end{gathered}$$

Because $\hat{\beta}^{OLS} = \beta + \left(\frac{1}{n}\sum_i X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_i X_i u_i\right)$ and $Var(X_i u_i) = \mathbb{E}[\mathbb{E}[X_i X_i' u_i^2|X]] = \mathbb{E}[X_i X_i' \mathbb{E}[u_i^2|X]] = \sigma^2 \mathbb{E}[X_i X_i']$, we can simplify to

$$\sqrt{n}\left(\hat{\beta}^{OLS} - \beta\right) \to_d N\left(0, \mathbb{E}[X_i X_i']^{-1}\sigma^2\right)$$

(c) You estimate the OLS regression and get an estimate, $b_1 = 2.7$. Suppose you also have a consistent estimator of the error variance, $\hat{S}^2 \to_p \sigma^2$, and estimate a value of $s = 1.2$. You want to know whether b_1 is "statistically significant." Derive a 95% confidence interval for β_1 and explain how you would determine whether the estimate is significant.

First, derive the test statistic and its distribution. By Slutsky and Continuous Mapping Theorem,

$$\left(\left(\frac{1}{n}\sum_i X_i X_i'\right)\hat{S}^2\right)^{-1/2} \to_p \left(\mathbb{E}[X_i X_i']\sigma^2\right)^{-1/2}$$

$$\frac{\sqrt{n}\left(\hat{\beta}^{OLS} - \beta\right)}{\sqrt{\frac{1}{n}\left(\sum_i X_i X_i'\right)\hat{S}^2}} \to_d \left(\mathbb{E}[X_i X_i']\sigma^2\right)^{-1/2} * N\left(0, \mathbb{E}[X_i X_i']^{-1}\sigma^2\right)$$

$$\frac{\sqrt{n}\left(\hat{\beta}^{OLS} - \beta\right)}{\sqrt{\frac{1}{n}\left(\sum_i X_i X_i'\right)\hat{S}^2}} \to_d N(0,1)$$

Now we can derive the 95% confidence interval. Denote X_{22} to be the 2, 2 element of the matrix $(\sum_i X_i X_i')$.

$$\frac{\left(\hat{\beta}_1^{OLS} - \beta_1\right)}{\sqrt{X_{22}}\hat{S}} \to_d N(0,1)$$

$$\mathbb{P}\left(-1.96 < \frac{\left(\hat{\beta}_1^{OLS} - \beta_1\right)}{\sqrt{X_{22}}\hat{S}} < 1.96\right) = 0.95$$

Once we plug in the estimates for $\hat{\beta}_1^{OLS}$ and $\hat{S}$ we convert the probability interval to a confidence interval.

$$\left[-1.96 < \frac{(b_1 - \beta_1)}{\sqrt{X_{22}}s} < 1.96\right]$$
$$\left[2.7 - 1.96 * (1.2\sqrt{X_{22}}) < \beta_1 < 2.7 + 1.96 * (1.2\sqrt{X_{22}})\right]$$

If this interval does not contain the null value of zero then we conclude that the null hypothesis is false and say that our estimate of β_1 is statistically significant at the 5% level. Otherwise, we fail to reject the null and cannot say that our estimate is significant.

You present these results to a helpful colleague who immediately points out that the error term for each school is likely inversely proportional to the number of students in that school (n_i) so that $\mathbb{E}[u_i^2|X] = \sigma_i^2 = \sigma^2/n_i$. She suggests that you estimate the model using Weighted Least Squares instead.

(d) Explain why the WLS estimator might be prefered to the OLS estimator in this case.

If the errors are not conditionally homoskedastic (as in this case) then OLS is no longer the Best Linear Unbiased Estimator. WLS will put more weight on the observations that are measured most precisely and will provide a more efficient estimator for β.

Following through with this suggestion, you transform the data and estimate

$$\hat{\beta}^{WLS} = (\sum_i X_i^* X_i^{*\prime})^{-1}(\sum_i X_i^* T_i^*)$$

(e) Define X_i^* and T_i^*.

$X_i^* = X_i/\sqrt{\sigma_i^2} = X_i\sqrt{n_i}/\sigma$ and $T_i^* = T_i\sqrt{n_i}/\sigma$. Note that when we calculate $\hat{\beta}^{WLS}$, the σ's in the numerator and denominator cancel out so that we could equivalently write $X_i^* = X_i\sqrt{n_i}$ and $T_i^* = T_i\sqrt{n_i}$.

(f) Show that $\hat{\beta}^{WLS}$ is consistent and asymptotically normal. (Hint: Is X_i^* covariance stationary?)

Note that the data is no longer i.i.d. because each X_i^* has a different mean that depends on the number of observations in that school, $\mathbb{E}[X_i\sqrt{n_i}]$. This obviously also implies that the data is not stationary. However, we can still use a Law of Large Numbers, for example, the Markov LLN only requires that a sequence of random variables be *uncorrelated* with finite means, $\mathbb{E}[X_i^*] = \mu_i$, and uniformly bounded variances, $Var(X_i^*) = \sigma_i^2 \leq M < \infty \ \forall i$. Then

$$\frac{1}{n}\sum_i X_i^* \to_p \lim_{n\to\infty} \frac{1}{n}\sum_i \mu_i.$$

This gives us

$$\frac{1}{n}\sum_i X_i^* X_i^{*\prime} \to_p \lim_{n\to\infty} \frac{1}{n}\sum_i \mathbb{E}[X_i^* X_i^{*\prime}]$$

$$\left(\frac{1}{n}\sum_i X_i^* X_i^{*\prime}\right)^{-1} \to_p \left(\lim_{n\to\infty} \frac{1}{n}\sum_i \mathbb{E}[X_i^* X_i^{*\prime}]\right)^{-1}$$

(By Continuous Mapping Theorem)

$$\frac{1}{n}\sum_i X_i^* u_i^* \to_p \lim_{n\to\infty} \frac{1}{n}\sum_i \mathbb{E}[X_i^* u_i^*]$$

$$\hat{\beta}^{WLS} = \beta + \left(\frac{1}{n}\sum_i X_i^* X_i^{*\prime}\right)^{-1} \left(\frac{1}{n}\sum_i X_i^* u_i^*\right)$$

$$\hat{\beta}^{WLS} \to_p \beta + \left(\lim_{n\to\infty} \frac{1}{n}\sum_i \mathbb{E}[X_i^* X_i^{*\prime}]\right)^{-1} \left(\lim_{n\to\infty} \frac{1}{n}\sum_i \mathbb{E}[X_i^* u_i^*]\right)$$

(By Slutsky)

$$\mathbb{E}[X_i^* u_i^*] = \mathbb{E}[X_i u_i n_i] = \mathbb{E}[X_i u_i] n_i = 0$$

(By orthogonality assumption)

$$\hat{\beta}^{WLS} \to_p \beta$$

Similarly, we can use the Lindeberg-Feller CLT, which applies to independent, but not identical, random variables and gives us that

$$\sqrt{n}\left(\frac{1}{n}\sum_i X_i^* u_i^* - \frac{1}{n}\sum_i \mathbb{E}[X_i^* u_i^*]\right) \to_d N\left(0, \lim_{n\to\infty}(\frac{1}{n}\sum_i Var(X_i^* u_i^*)\right)\right)$$

$$\sqrt{n}\left(\left(\frac{1}{n}\sum_i X_i^* X_i^{*\prime}\right)^{-1} \frac{1}{n}\sum_i X_i^* u_i^* - 0\right)$$

$$\to_d \left(\lim_{n\to\infty}\frac{1}{n}\sum_i \mathbb{E}[X_i^* X_i^{*\prime}]\right)^{-1} * N\left(0, \left(\lim_{n\to\infty}\frac{1}{n}\sum_i Var(X_i^* u_i^*)\right)\right)$$

Note: $\lim_{n\to\infty}\frac{1}{n}\sum_i \mathbb{E}[X_i^* X_i^{*\prime}] = \lim_{n\to\infty}\frac{1}{n}\sum_i n_i \mathbb{E}[X_i X_i'] = \mathbb{E}[X_i X_i']\lim_{n\to\infty}\frac{1}{n}\sum_i n_i = \mathbb{E}[X_i X_i']\lim_{n\to\infty}\frac{1}{n}n = \mathbb{E}[X_i X_i']$ and $Var(X_i^* u_i^*) = Var(X_i u_i n_i) = \mathbb{E}[X_i X_i' n_i^2 \mathbb{E}[u_i^2|X]] = n_i \sigma^2 \mathbb{E}[X_i X_i']$. So that $\lim_{n\to\infty}\frac{1}{n}\sum_i Var(X_i^* u_i^*) = \sigma^2 \mathbb{E}[X_i X_i']$ and

$$\sqrt{n}(\hat{\beta}^{WLS} - \beta) \to_d N(0, \sigma^2 \mathbb{E}[X_i X_i']^{-1})$$