Econometrics 710 - Problem Set # 5

March 1, 2015

TA: Cristian Hernandez
cahernandez3@wisc.edu

Exercise 4.7

$$
\begin{aligned}
\hat{V}_{\hat{\beta}}^{W} &= (X'X)^{-1}\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
\bar{V}_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-1}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
\tilde{V}_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-2}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1}
\end{aligned}
$$

Since $(1-h_{ii})^{-2} > (1-h_{ii})^{-1} > 1$,

$$
\begin{aligned}
\tilde{V}_{\hat{\beta}} - \bar{V}_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}((1-h_{ii})^{-2}-(1-h_{ii})^{-1})\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
&= (X'X)^{-1}(X'CX)(X'X)^{-1} \\
&= G'G
\end{aligned}
$$

where $C = \text{diag}(a_1,\cdots,a_n), a_i = ((1-h_{ii})^{-2}-(1-h_{ii})^{-1})\hat{e}^2$, and $G = C^{1/2}X(X'X)^{-1}$. Note that $C^{1/2} = \text{diag}(a_1^{1/2},\cdots,a_n^{1/2})$ is well defined since $a_i > 0$. Since G has full column rank k,$\tilde{V}_{\hat{\beta}} - \bar{V}_{\hat{\beta}}$ is positive definite. Similarly, we can show $\bar{V}_{\hat{\beta}} - \hat{V}_{\hat{\beta}}^{W}$ is positive definite.

Exercise 4.8

$$
\begin{aligned}
\mathbb{E}(\tilde{V}_{\hat{\beta}}|X) &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-2}\mathbf{x}_i\mathbf{x}_i'\mathbb{E}(\hat{e}_i^2|X)\right)(X'X)^{-1} \\
&= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-1}\mathbf{x}_i\mathbf{x}_i'\sigma^2\right)(X'X)^{-1} \\
&= (X'X)^{-1}\left(\sum_{i=1}^{n}x_ix_i'\sigma^2 + \sum_{i=1}^{n}((1-h_{ii})^{-1}-1)\mathbf{x}_i\mathbf{x}_i'\sigma^2\right)(X'X)^{-1}) \\
&= (X'X)^{-1}\sigma^2 + (X'X)^{-1}\left(\sum_{i=1}^{n}\frac{h_{ii}}{1-h_{ii}}\mathbf{x}_i\mathbf{x}_i'\sigma^2\right)(X'X)^{-1} \\
&= (X'X)^{-1}\sigma^2 + \sigma^2(X'X)^{-1}X'DX(X'X)^{-1}
\end{aligned}
$$

where $D = \text{diag}(\frac{h_{11}}{1-h_{11}}, \cdots, \frac{h_{nn}}{1-h_{nn}})$. Since $\frac{h_{ii}}{1-h_{ii}} \geq 0$, the second term in the last equality is positive semi-definite. Thus we have (4.33). Similarly, for (4.34);

$$\begin{aligned}
\mathbb{E}(\overline{V}_{\hat{\beta}}|X) &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-1}\mathbf{x}_i\mathbf{x}_i'\mathbb{E}(\hat{e}_i^2|X)\right)(X'X)^{-1} \\
&= (X'X)^{-1}\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\sigma^2\right)(X'X)^{-1} \\
&= (X'X)^{-1}\sigma^2
\end{aligned}$$

Exercise 4.9

Note that $(\sum_{i=1}^{n}(y_i-\mu))^3 = \sum_{i=1}^{n}(y_i-\mu)^3 + \sum_{i\neq j}(y_i-\mu)^2(y_j-\mu) + \sum_{i\neq j\neq k}(y_i-\mu)(y_j-\mu)(y_k-\mu)$. Because of independent observations, $\mathbb{E}(y_i-\mu)^2(y_j-\mu) = \mathbb{E}(y_i-\mu)^2\mathbb{E}(y_j-\mu) = 0, \mathbb{E}(y_i-\mu)(y_j-\mu)(y_k-\mu) = 0$

$$\begin{aligned}
\mathbb{E}(\bar{y}-\mu)^3 &= \mathbb{E}(\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu))^3 \\
&= \frac{1}{n^3}\mathbb{E}(\sum_{i=1}^{n}(y_i-\mu))^3 \\
&= \frac{1}{n^3}\sum_{i=1}^{n}\mathbb{E}(y_i-\mu)^3 + \sum_{i\neq j}\mathbb{E}(y_i-\mu)^2(y_j-\mu) + \sum_{i\neq j\neq k}\mathbb{E}(y_i-\mu)(y_j-\mu)(y_k-\mu)) \\
&= \frac{\mu_3}{n^2}
\end{aligned}$$

Exercise 4.10

Note that $\mathbb{E}(x_i^3e_i^3|X) = x_i^3\mathbb{E}(e_i^3|x_i) = x_i^3\mu_{3i}$, $\mathbb{E}((x_ie_i)^2(x_je_j)|X) = \mathbb{E}(x_i^2e_i^2|X)\mathbb{E}(x_je_j|X) = 0$, and $\mathbb{E}((x_ie_i)(x_je_j)(x_ke_k)|X) = 0$ for different i,j,k.

$$\begin{aligned}
\mathbb{E}\left((\hat{\beta}-\beta)^3|X\right) &= \mathbb{E}\left((\frac{\sum_{i=1}^{n}x_ie_i}{\sum_{i=1}^{n}x_i^2})^3|X\right) \\
&= \frac{1}{(\sum_{i=1}^{n}x_i^2)^3}\mathbb{E}\left((\sum_{i=1}^{n}x_ie_i)^3|X\right) \\
&= \frac{1}{(\sum_{i=1}^{n}x_i^2)^3}\mathbb{E}\left(\sum_{i=1}^{n}(x_ie_i)^3 + \sum_{i\neq j}(x_ie_i)^2(x_je_j)\sum_{i\neq j\neq k}(x_ie_i)(x_je_j)(x_ke_k)|X\right) \\
&= \frac{1}{(\sum_{i=1}^{n}x_i^2)^3}\left(\sum_{i=1}^{n}\mathbb{E}((x_ie_i)^3|X) + \sum_{i\neq j}\mathbb{E}((x_ie_i)^2(x_je_j)|X)\sum_{i\neq j\neq k}\mathbb{E}((x_ie_i)(x_je_j)(x_ke_k)|X)\right) \\
&= \frac{\sum_{i=1}^{n}x_i^3\mu_{3i}}{(\sum_{i=1}^{n}x_i^2)^3}
\end{aligned}$$

(Consider the special case where $x_i = 1$, $\hat{\beta} = \bar{y}, \mathbb{E}y_i = \beta, \mu_{3i} = \mathbb{E}(y_i-\beta)^3$, then result reduces to exercise 4.9)

Exercise 4.11

We calculate standard errors using five different covariance matrix estimators

$$
\begin{aligned}
(\text{Homoskedastic}) : \hat{V}^0_{\hat{\beta}} &= (X'X)^{-1}s^2 = \frac{1}{n-k}(X'X)^{-1}\sum_{i=1}^{n}\hat{e}_i^2 \\
(\text{White}) : \hat{V}^W_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
(\text{Scaled White}) : \hat{V}_{\hat{\beta}} &= \frac{n}{n-k}(X'X)^{-1}\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
(\text{Andrews}) : \tilde{V}_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-2}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1} \\
(\text{Horn-Horn-Duncan}) : \bar{V}_{\hat{\beta}} &= (X'X)^{-1}\left(\sum_{i=1}^{n}(1-h_{ii})^{-1}\mathbf{x}_i\mathbf{x}_i'\hat{e}_i^2\right)(X'X)^{-1}
\end{aligned}
$$

Variables	Coefficient Estimates	Homoske-dastic	White	Scaled White	Andrews	Horn-Horn-Duncan
Education	0.1443	0.0116	0.0117	0.0118	0.0121	0.0119
Experience	0.0426	0.0122	0.0124	0.0125	0.0128	0.0126
Experience squared/100	-0.0951	0.0349	0.0338	0.0341	0.0354	0.0346
constant	0.5309	0.1898	0.2001	0.2016	0.2054	0.2027
R^2	0.3893					
observations	267					

Table 1: OLS Estimates of Linear Equation for $log(wage)$ using the sub-sample of single Asian males with less than 45 years of experience ($n = 267$).

Exercise 4.12

Variables	Coefficient Estimates	Standard Errors
Education	0.0883	0.0029
Experience	0.0279	0.0028
Experience squared/100	-0.0365	0.0055
Regional dummy		
Northeast	0.0616	0.0361
South	-0.0675	0.0297
West	0.0201	0.0283
Marital Status dummy		
Married	0.1780	0.0250
Widowed	0.2430	0.1866
Divorced	0.0787	0.0450
Separated	0.0169	0.0528
constant	1.1918	0.0501
R^2	0.2492	
observations	4230	

Table 2: OLS Estimates of Linear Equation for $log(wage)$ using the sub-sample of white male Hispanics ($n = 4230$). Standard errors are heteroskedastic-robust (Horn-Horn-Duncan formula.)

Exercise 5.1

1. $\{a_n\} = \{1, 1/2, 1/3, \cdots\}$. $\liminf a_n = \limsup a_n = \lim a_n = 0$
2. $\{a_n\} = \{1, 0, -1, 0, 1, 0, -1, 0, \cdots\}$.$\liminf a_n = -1, \limsup a_n = 1$, $\lim a_n$ does not exists.
3. $\{a_n\} = \{1, 0, -1/3, 0, 1/5, 0, -1/7, 0, \cdots\}$. $\liminf a_n = \limsup a_n = \lim a_n = 0$

Exercise 5.2

1. $\mathbb{E}\bar{y}^* = \mathbb{E}(\frac{1}{n}\sum_{i=1}^n w_i y_i) = \frac{1}{n}\sum_{i=1}^n (w_i \mathbb{E} y_i) = (\frac{1}{n}\sum_{i=1}^n w_i)\mu = \mu$
2.

$$\begin{aligned}
\text{var}(\bar{y}^*) &= \text{var}\left(\frac{1}{n}\sum_{i=1}^n w_i y_i\right) = \frac{1}{n^2}\text{var}\left(\sum_{i=1}^n w_i y_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n w_i^2 \text{var}(y_i) + \sum_{i=1}^n \sum_{j\neq i} w_i w_j \text{cov}(y_i, y_j)\right) \\
&= \left(\frac{1}{n^2}\sum_{i=1}^n w_i^2\right)\text{var}(y_i)
\end{aligned}$$

3. By Chebyshev's Inequality, for any $\delta > 0$

$$Pr(|\bar{y}^* - \mu| > \delta) \leq \frac{\mathrm{var}(\bar{y}^*)}{\delta^2}$$

Thus, a sufficient condition for $\bar{y}^* \xrightarrow{p} \mu$ is $\mathrm{var}(\bar{y}^*) \longrightarrow 0$ as $n \to \infty$. By the calculations above, if (and only if) $(\frac{1}{n^2}\sum_{i=1}^n w_i^2) \longrightarrow 0$, then $\mathrm{var}(\bar{y}^*) \longrightarrow 0$ assuming that the second moment is finite. (or $\mathrm{var}(y_i) < \infty$)

4.
$$\frac{1}{n^2}\sum_{i=1}^n w_i^2 \leq \left(\frac{1}{n}\sum_{i=1}^n w_i\right)\frac{1}{n}\max_{1\leq i\leq n} w_i = n^{-1}\max_{1\leq i\leq n} w_i$$

Thus sufficient condition for the condition in part 3 is $n^{-1}\max_{1\leq i\leq n} w_i = o(1)$ or $\max_{1\leq i\leq n} w_i = o(n)$

Exercise 5.3

By Chebyshev's inequality, for any $\delta > 0$, $Pr(|Z| > \delta) \leq \frac{1}{\delta^2}$. Thus $\delta = \sqrt{20} = 4.472$ makes $Pr(|Z| > \delta) \leq \frac{1}{\delta^2} = 0.05$ for any random variable Z with $\mathbb{E}Z = 0, \mathbb{E}Z^2 = 1$. For standard normal random variable $Z \sim N(0,1)$, $\delta = z_{0.025} = 1.96$ makes $Pr(|Z| > \delta) = 0.05$. Chebyshev's inequality holds for *any* random variable, however it is in some sense not a tight bound. As we could see, for normal random variable, $Pr(|Z| > 4.472) < Pr(|Z| > 1.96) = 0.05$.

Exercise 5.4

Assume y_i are i.i.d observations. The moment estimator of third moment $\mathbb{E}y_i^3$ is sample moments $\hat{\mu}_3 = \frac{1}{n}\sum_{i=1}^n y_i^3$. By CLT, if $\mathbb{E}y_i^6 < \infty$ $\sqrt{n}(\hat{\mu}_3 - \mu_3) \xrightarrow{d} N(0, v^2)$ where $v^2 = \mathbb{E}(y_i^3 - \mathbb{E}y_i^3)^2 = \mathbb{E}y_i^6 - (\mathbb{E}y_i^3)^2$

Exercise 5.5

Let's solve a more general case and then apply the result to the function $g(z) = z^2$. Take a continuous function g. Continuity of this function implies that $\forall\epsilon > 0$, $\exists\delta > 0$ such that $|z_n - c| < \delta$ implies that $|g(z_n) - g(c)| < \epsilon$. Therefore, $|g(z_n) - g(c)| > \epsilon$ implies that $|z_n - c| > \delta$. This means that the event $A = |g(z_n) - g(c)| > \epsilon$ is a subset of the event $B = |z_n - c| > \delta$—every time that A occurs B also occurs, but not the other way around—, and hence

$$Pr\left(|g(z_n) - g(c)| > \epsilon\right) \leq Pr\left(|z_n - c| > \delta\right)$$

Since $z_n \xrightarrow{p} c$, for any $\delta > 0$ it holds that

$$\lim_{n\to\infty} Pr\left(|z_n - c| > \delta\right) = 0$$

, which implies that for any $\epsilon > 0$

$$\lim_{n\to\infty} Pr\left(|g(z_n) - g(c)| > \epsilon\right) = 0$$

Therefore, $g(z_n) \xrightarrow{p} g(c)$. Now we apply this proof to $g(z_n) = z_n^2$ and it holds that $z_n^2 \xrightarrow{p} c^2$.

Exercise 5.6

1. By the Delta method $\sqrt{n}(\hat{\beta} - \beta) = \sqrt{n}(\hat{\mu}^2 - \mu^2) \xrightarrow{d} 2\mu N(0, v^2) = N(0, 4\mu^2 v^2)$

2. If $\mu = 0$, then $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} 0$, i.e., $\sqrt{n}(\hat{\beta} - \beta)$ converges in distribution to a constant since the derivative of the function $\beta(\mu) = \mu^2$ is zero.

3. If $\mu = 0$, then we know that $\sqrt{n}\hat{\mu} \xrightarrow{d} N(0, v^2)$. Therefore, by the CMT $\frac{(\sqrt{n}\hat{\mu})^2}{v^2} \xrightarrow{d} \chi^2(1)$.

4. From the Delta Method we can decompose the asymptotic distribution of $\sqrt{n}(\hat{\beta} - \beta)$ as the derivative of $\beta(\mu)$—a deterministic component—, and $\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{d} Z \sim N(0, v^2)$—a stochastic component. The problem arises at $\mu = 0$ because the deterministic component happens to be zero at that point, so the random variable $\sqrt{n}(\hat{\beta} - \beta)$ converges in distribution to zero. The solution in part 3 takes care of this issue at $\mu = 0$ and allows to make inference.