ECON550: Problem Set 10

Nicholas Wu

Fall 2020

Problem 1

In order to use the mean-value expansion, we need $\tilde{\theta}_n$ in a neighborhood $B(\theta_0, \epsilon)$ of θ_0. Let

$$\bar{m} = n^{-1} \sum_{i=1}^{n} m(W_i, \tilde{\theta}_n)$$

$$Em = Em(W_i, \theta_0)$$

Then $P(|\bar{m} - Em| > k) \leq P((|\bar{m} - Em > k) \cap (\tilde{\theta_n} \in B(\theta_0, \epsilon))) + P(\tilde{\theta_n} \in B(\theta_0, \epsilon))$ We know the second probability term goes to 0 since $\tilde{\theta}_n \to \theta_0$. So we just need to show the first term also goes to 0. Specifically, it suffices to show that for $\tilde{\theta}_n \in B(\theta_0, \epsilon)$, $\bar{m} \to_p Em$.

Since $\tilde{\theta}_n \in B(\theta_0, \epsilon)$, take the mean-value expansion:

$$n^{-1} \sum_{i=1}^{n} m(W_i, \tilde{\theta}_n) = n^{-1} \sum_{i=1}^{n} m(W_i, \theta_0) + n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0)$$

where θ_n' is between θ_0 and $\tilde{\theta}_n$. By the WLLN, the first term converges to Em, and so we just need to show the second term converges to 0 in probability. By Cauchy-Schwarz,

$$0 \leq \left| n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0) \right| \leq n^{-1} \sum_{i=1}^{n} \left|\left| \frac{\partial m(W_i, \theta_n')}{\partial \theta'} \right|\right| ||(\tilde{\theta}_n - \theta_0)||$$

$$\leq n^{-1} \sum_{i=1}^{n} \left(\sup_{\theta \in B(\theta_0, \epsilon)} \left|\left| \frac{\partial m(W_i, \theta)}{\partial \theta} \right|\right| \right) ||(\tilde{\theta}_n - \theta_0)||$$

Now, by the WLLN,

$$n^{-1} \sum_{i=1}^{n} \left(\sup_{\theta \in B(\theta_0, \epsilon)} \left|\left| \frac{\partial m(W_i, \theta)}{\partial \theta} \right|\right| \right) \to_p E \left(\sup_{\theta \in B(\theta_0, \epsilon)} \left|\left| \frac{\partial m(W_i, \theta)}{\partial \theta} \right|\right| \right) < \infty$$

Since $||(\tilde{\theta}_n - \theta_0)|| \to_p 0$ as $\tilde{\theta}_n \to_p \theta_0$, we have that

$$n^{-1} \sum_{i=1}^{n} \left(\sup_{\theta \in B(\theta_0, \epsilon)} \left|\left| \frac{\partial m(W_i, \theta)}{\partial \theta} \right|\right| \right) ||(\tilde{\theta}_n - \theta_0)|| \to_p 0$$

But since

$$0 \leq \left| n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0) \right| \leq n^{-1} \sum_{i=1}^{n} \left|\left| \frac{\partial m(W_i, \theta_n')}{\partial \theta'} \right|\right| ||(\tilde{\theta}_n - \theta_0)||$$

we also get that

$$\left| n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0) \right| \to_p 0$$

and hence

$$n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0) \to_p 0$$

So all together

$$n^{-1} \sum_{i=1}^{n} m(W_i, \tilde{\theta}_n) = n^{-1} \sum_{i=1}^{n} m(W_i, \theta_0) + n^{-1} \sum_{i=1}^{n} \frac{\partial m(W_i, \theta_n')}{\partial \theta'} (\tilde{\theta}_n - \theta_0) \to_p Em(W_i, \theta_0) + 0 = Em(W_i, \theta_0)$$

and we are done.

Problem 2

The FOCs for maximization definition of $\hat{\theta}_n$, we get

$$0 = \left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum g(W_i, \hat{\theta}_n) \right)$$

Using the mean-value expansion on the rightmost parenthesized expression, we get

$$0 = \left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum g(W_i, \theta_0) + n^{-1} \sum \frac{\partial g(W_i, \theta_n')}{\partial \theta} (\hat{\theta}_n - \theta_0) \right)$$

where θ_n' is between θ_0 and $\hat{\theta}_n$

$$\left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum \frac{\partial g(W_i, \theta_n')}{\partial \theta} (\hat{\theta}_n - \theta_0) \right) = - \left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum g(W_i, \theta_0) \right)$$

$$\left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum \frac{\partial g(W_i, \theta_n')}{\partial \theta} \right) \sqrt{n} (\hat{\theta}_n - \theta_0) = - \left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \sqrt{n} \left(n^{-1} \sum g(W_i, \theta_0) \right)$$

$$\sqrt{n} (\hat{\theta}_n - \theta_0) = - \left(\left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1} \sum \frac{\partial g(W_i, \theta_n')}{\partial \theta} \right) \right)^{-1} \left(n^{-1} \sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta} \right)' B_n \left(n^{-1/2} \sum g(W_i, \theta_0) \right)$$

By the multivariate CLT:

$$\left(n^{-1/2} \sum g(W_i, \theta_0) \right) \to_p N(0, E[g(W_i, \theta_0) g(W_i, \theta_0)'])$$

Since $\hat{\theta}_n \to \theta_0$, $\theta'_n \to \theta_0$, applying Problem 1 element-wise, we get,

$$\left(n^{-1}\sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta}\right)' \to_p \Gamma'_0$$

$$\left(n^{-1}\sum \frac{\partial g(W_i, \theta'_n)}{\partial \theta}\right) \to_p \Gamma_0$$

Since $B_n \to B$, by Slutsky's,

$$\left(\left(n^{-1}\sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta}\right)' B_n \left(n^{-1}\sum \frac{\partial g(W_i, \theta'_n)}{\partial \theta}\right)\right)^{-1} \to_p (\Gamma'_0 B \Gamma_0)^{-1}$$

Then we get all together

$$\sqrt{n}(\hat{\theta}_n - \theta_0)$$

$$= -\left(\left(n^{-1}\sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta}\right)' B_n \left(n^{-1}\sum \frac{\partial g(W_i, \theta'_n)}{\partial \theta}\right)\right)^{-1} \left(n^{-1}\sum \frac{\partial g(W_i, \hat{\theta}_n)}{\partial \theta}\right)' B_n \left(n^{-1/2}\sum g(W_i, \theta_0)\right)$$

$$\to_p (\Gamma'_0 B \Gamma_0)^{-1}\Gamma_0 B N(0, E[g(W_i, \theta_0) g(W_i, \theta_0)'])$$

$$= N(0, (\Gamma'_0 B \Gamma_0)^{-1}\Gamma_0 B E[g(W_i, \theta_0) g(W_i, \theta_0)'] B' \Gamma'_0 (\Gamma'_0 B \Gamma_0)^{-1})$$

Problem 3

We know from the previous problem that since $\hat{\theta}_n \to_p \theta_0$,

$$\hat{\Gamma}_n = \left(n^{-1}\sum \frac{\partial g(W_i, \theta'_n)}{\partial \theta}\right) \to_p \Gamma_0$$

So we just need to find an estimator for $E[g(W_i, \theta_0) g(W_i, \theta_0)']$. Consider

$$\hat{E}_n = n^{-1}\sum g(W_i, \hat{\theta}_n) g(W_i, \hat{\theta}_n)'$$

Then as long as $E \sup ||\partial g(W_i, \hat{\theta}_n) g(W_i, \hat{\theta}_n)/\partial \theta|| < \infty$, we can apply problem 1, and we get

$$\hat{E}_n \to_p E[g(W_i, \theta_0) g(W_i, \theta_0)']$$

So applying Slutsky's and the rules of convergence in probability, we have

$$(\hat{\Gamma}'_n B_n \hat{\Gamma}_n)^{-1}\hat{\Gamma}_n B_n \hat{E}_n B'_n \hat{\Gamma}'_n (\hat{\Gamma}'_n B_n \hat{\Gamma}_n)^{-1} \to_p (\Gamma'_0 B \Gamma_0)^{-1}\Gamma_0 B E[g(W_i, \theta_0) g(W_i, \theta_0)'] B' \Gamma'_0 (\Gamma'_0 B \Gamma_0)^{-1}$$

Problem 4

We can apply the delta method.

$$g'(\rho) = \frac{1}{2(1+\rho)} + \frac{1}{2(1-\rho)}$$

$$= \frac{1-\rho+1+\rho}{2(1-\rho^2)}$$

$$= \frac{1}{1-\rho^2}$$

Since $\sqrt{n}(\hat{\rho}_n - \rho) \to N(0, (1-\rho^2)^2)$, by the delta method, we have

$$\sqrt{n}(g(\hat{\rho}_n) - g(\rho)) \to N(0, g'(\rho)^2(1-\rho^2)^2) = N(0,1)$$

Problem 5

Suppose, for sake of contradiction, $\exists \epsilon > 0$ such that

$$\inf_{\theta \notin B(\theta_0, \epsilon)} Q(\theta) \le Q(\theta_0)$$

This implies that $\exists$ a sequence of θ_n's such that $Q(\theta_n) \to Q^* \le Q(\theta_0)$. Since Θ is compact, by Heine-Borel it is bounded, and hence by Bolzano-Weierstrass we can pick a convergent subsequence, $\theta'_n \to \theta^* \neq \theta_0$ (since the sequence is not contained in $B(\theta_0, \epsilon)$). By continuity of Q, $Q(\theta'_n)$ also converges, and since θ'_n is a subsequence of θ_n and $Q(\theta_n) \to Q^*$, $Q(\theta'_n) \to Q^*$. Now, since Θ is compact, by Heine-Borel it is also closed, so $\theta^* \in \Theta$, and $Q(\theta^*) = Q^* \le Q(\theta_0)$. But this contradicts our assumption that θ_0 uniquely minimizes Q on Θ, and hence we are done.

Problem 6

(a) The log-likelihood is (dropping constant terms without θ)

$$-\sum \frac{(X_i - \theta)^2}{2\sigma^2}$$

Taking the FOC on θ:

$$0 = \frac{1}{\sigma^2}\left(\sum (X_i - \theta)\right)$$

Now, if $\bar{X}_n \ge 0$, we can just take $\hat{\theta}_n = \bar{X}_n$, and this will satisfy the FOC and maximize log-likelihood. If $\bar{X}_n < 0$, we note that the log-likelihood, while maximized at $\hat{\theta}_n = \bar{X}_n$, is decreasing in $\hat{\theta}_n$ on the range $[0, \infty)$. Hence, if $\bar{X}_n < 0$, the value of $\hat{\theta}_n$ in the allowable range that maximizes the log-likelihood is 0. Hence, the MLE is $\hat{\theta}_n = \max(0, \bar{X}_n)$.

(b) We have that due to normality,

$$P(X \le c) = P\left(\frac{X-\mu}{\sigma} \le \frac{c-\mu}{\sigma}\right) = \Phi\left(\frac{c-\mu}{\sigma}\right)$$

Since functions of $\hat{\theta}$ being an MLE for θ implies $g(\hat{\theta})$ is an MLE for $g(\theta)$, we get that if we take the MLEs for μ, σ as $\hat{\mu}, \hat{\sigma}$, then

$$\Phi\left(\frac{c-\hat{\mu}}{\hat{\sigma}}\right)$$

is an MLE for

$$\Phi\left(\frac{c-\mu}{\sigma}\right) = P(X \leq c)$$