BIOS 7731 Homework 9 Tim Vigers

7.1.2) 7.1.6. is essentially a high dimensional extension of 7.1.4, where $\vec{\theta}_0$ and t are both vectors. First, we define

$$Z_n(\vec{t}) \equiv l_n(\vec{\theta}_0 + \vec{t}/\sqrt{n}) - l_n(\vec{\theta}_0)$$

where $l_n(\theta) = \sum_{i=1}^{n} \log(p(x_i, \theta))$. Next, we can use a Taylor series (second-order) to approximate $Z_n(\vec{t})$.

$$Z_n(\vec{t}) = \sum_{i=1}^{n} \log\left(\frac{p(x_i, \vec{\theta}_0 + \vec{t}/\sqrt{n})}{p(x_i, \vec{\theta}_0)}\right) = \frac{\vec{t}}{\sqrt{n}} \sum_{i=1}^{n} \partial/\partial\vec{\theta}_0 \log(p(x_i, \theta_0)) +$$

this is a bit of notation abuse, but I found it helpful

$$\frac{\vec{t}^{\,2}}{2n} \sum_{i=1}^{n} \partial^2/\partial\theta_0^2 \log(p(x_i, \theta_0)) + \text{Remainder.}$$

The remainder will go to 0, so we can write it as $o_p(1)$. By definition, the first term in the expansion is $\vec{t} Z_n^0 = \vec{t} \dfrac{\sum_{i=1}^{n} l'(x_i, \vec{\theta}_0)}{\sqrt{n}}$. By the CLT and Slutsky's, $\vec{t} Z_n^0 \xrightarrow{\mathcal{L}} \vec{t} Z^0$. Next we look at the second term, which is a function of the observed information matrix:

$$\frac{\vec{t}}{2n} \sum_{i=1}^{n} \partial^2/\partial\theta_0^2 \log(p(x_i, \vec{\theta}_0)) \vec{t}$$

By the weak law of large numbers,

$$\frac{\sum^{n} l''(x_i, \vec{\theta}_0)}{n} \xrightarrow{\mathcal{L}} -I(\vec{\theta}_0)$$

So, again by Slutsky's and WLLN:

$$\frac{\vec{t}}{2n}\sum_{i=1}^{n} \ell''(x_i,\vec{\theta}_0)\vec{t} \xrightarrow{a.s.} -\frac{\vec{t}}{2} I(\vec{\theta}_0)\vec{t}$$

and $$Z_n(t) \longrightarrow \vec{t}\, Z^\circ - \frac{\vec{t}}{2} I(\vec{\theta}_0)\vec{t} \qquad \square.$$

7.1.3) Let A_{mj} be the event where
$\sup\{|Z_n(s) - Z_n(t)| : s,t \in T_{mj}\} > \varepsilon$, and let

$$p_m(\varepsilon) = \limsup_{n\to\infty} \max\{P[\sup\{|Z_n(s) - Z_n(t)| : s,t \in T_{mj}\} \geq \varepsilon] : 1 \leq j \leq k_m\}$$

$$= \limsup_{n\to\infty} \max\{P[A_{mj}] : 1 \leq j \leq k_m\}$$

Then, from the hint we know that:

$$P(\bigcup^{j} A_{mj}) \leq \sum^{j} P(A_{mj}) \leq k_m * \max\{P[A_{mj}] : 1 \leq j \leq k_m\}$$

Therefore, if $k_m p_m(\varepsilon) \to 0$, then $\lim_{n\to} P(\bigcup^{j} A_{mj})$ also goes to 0. So, by the definition of weak convergence,

$$Z_n \Rightarrow Z \qquad \square.$$

7.1.6) By the continuous mapping theorem for random variables, we know that

$$X_n \xrightarrow{\mathcal{L}} X \Rightarrow g(X_n) \xrightarrow{\mathcal{L}} g(X)$$

Thus, if $Z_n \xrightarrow{\mathcal{L}} Z$ then:

$$\mathcal{L}(g(Z_n(t_1)), \dots, g(Z_n(t_k))) \to \mathcal{L}(g(Z(t_1)), \dots, g(Z(t_k)))$$

for g a continuous map $\ell_\infty(T) \to \ell_\infty(S)$

And it follows that $g(Z_n) \xrightarrow{FIDI} g(Z)$ in $\ell_\infty(S)$

iii) For each m there exists a partition of T into a finite number m of sets and $Z^{(m)}(\cdot) \in \ell_\infty(T)$. Thus, by the continuous mapping theorem there also exists $g(Z^{(m)}(\cdot)) \in \ell_\infty(S)$. By the definition of continuity given in BD Definition 7.1.3, if $|h_n - h|_\infty \to 0$ then $g(h_n) \to g(h)$. Thus:

$$|g(Z^{(m)}) - g(Z)|_\infty \xrightarrow{P} 0 \quad \text{as } m \to \infty \quad \in \ell_\infty(S).$$

7.2.2) By BD remark 3.5.1, $SC(x, \hat{\theta}) = IF_{1/n}(x, \theta, \hat{F}_{n-1})$.

Where $IF_{\frac{1}{n}}(x, \theta, F) = n\left(\theta\left((1-1/n)F + \frac{1}{n}\delta_x\right) - \theta(F)\right)$ (*)

and δ_x is a point mass at x. Under certain regularity conditions, which we are assuming here, this can be calculated as

$$\psi_0(\delta_x, F) = \int \psi(x, F)\, \delta_x \, dx \qquad \text{where } \psi(x, F) \qquad (**)$$

is the influence function. Thus, setting $t = 1/n$ we can simply plug and chug using (*) and (* *) to easily see that:

$$SC(x) = n \int_0^{1/n} \left[\psi(x, (1-t)\hat{P}_{n-1} + t\delta_x) + \psi(\hat{P}_{n-1}, (1-t)\hat{P}_{n-1} + t\delta_x)\right] dt$$

b. If $P_m \xrightarrow{\mathcal{L}} P$ implies that $\psi(x, P_m) \to \psi(x, P)$ as $n \to \infty$, then

$$SC(x, \hat{v}) = n \int_0^{1/n} \left[\psi(x, (1-t)\hat{P}_{n-1} + t\delta_x) + \psi(\hat{P}_{n-1}, (1-t)\hat{P}_{n-1} + t\delta_x)\right] dt$$

$$\xrightarrow{\mathcal{L}} \psi(x, P) + \psi(P, P) \qquad \text{by Slutsky's (probably).}$$

Because $\psi(P,P) = 0$, we therefore have

$$SC_n(x, \hat{v}) \longrightarrow \psi(x, P)$$

7.2.18) I'm really sorry, I had to give up on this problem. I spent hours just on part a, and wasn't even able to finish that. Looking forward to seeing the solutions though!

7.2.19) a. Because we have a version of $\frac{\sum^n X_i}{n}$ in this problem, it suggests to me that we should use the CLT. By definition, $E_p[\Psi(X,P)] = 0$. and $E_p[\Psi^2(X,P)] < \infty$. So, let $Var[\Psi(x,P)] = \sigma^2$ and by the CLT we know:

$$\sqrt{n}\,\frac{\left(\frac{1}{n}\sum^n \Psi(X_i,P) - 0\right)}{\sigma} \xrightarrow{\mathcal{L}} N(0,1)$$

So $\frac{1}{n}\sum^n \Psi(X_i,P) \stackrel{\sim}{\sim} N(0, \sigma^2/n)$

By the definition of O-notation on page 516 of BD volume 1, if $E|Z|<\infty$ and $E[Z] = \mu$, then $\bar{Z}_n = \mu + o_p(1)$. Also, if $E|Z|^2 < \infty$, then $\bar{Z}_n = \mu + O_p(n^{-1/2})$ by the CLT. Since the CLT applies here, we know that

$$\frac{\sum^n \Psi(X_i,P)}{n} = 0 + O_p(n^{-1/2}) \qquad \square.$$

b. If $v(P) + \frac{1}{n} \sum^n \psi(x_i, P) + o_p(n^{-1/2}) = v(P) + \frac{1}{n} \sum^n \psi_2(x_i, P) + o_p(n^{-1/2})$

then from the hint we also know that

$$\frac{\sqrt{n}}{n} \left(\sum_{i=1}^n \psi_1(x_i, P) - \psi_2(x_i, P) \right) = o_p(1)$$

Or, to make the notation easier:

$$\frac{\sqrt{n}}{n} \left(\sum_{i=1}^n \Delta(x_i) \right) = o_p(1).$$

Then, by the CLT we have:

$$\frac{\sqrt{n}}{n} \left(\sum_{i=1}^n \Delta(x_i) \right) \xrightarrow{\mathcal{L}} N(0, \operatorname{Var}[\Delta(x_i)]).$$

In order for this limiting distribution to be $o_p(1)$, $\psi_1(x_i, P)$ must be equal to $\psi_2(x_i, P)$. $\square$.