Solution to HW2.3

(Hard; extremely optional) In class on September 11 you saw lots of technical details that are usually hidden in a course at this level. In particular, you saw the role played by the process $Z_n(t) = n^{-1/2}\sum_{i\le n}\dot{g}(x_i,t)$ for t near θ. Define $H_\theta(t) = \mathbb{E}_\theta\dot{g}(z,t)$ and

$$W_n(t) = Z_n(t) - \mathbb{E}_\theta Z_n(t) = n^{-1/2}\sum_{i\le n}\left(\dot{g}(x_i,t) - H_\theta(t)\right).$$

For each constant C show (under the assumptions described in class) that

$$\mathbb{E}_\theta \sup_{|h|\le C}|W_n(\theta + h/\sqrt{n}) - W_n(\theta)| \to 0 \qquad \text{as } n\to\infty.$$

What does this tell you about the behavior of $Z_n(\theta_n^*)$ (under $\mathbb{P}_\theta$) if θ_n^* is a root-n consistent estimator for θ? Hint: Use more Taylor, starting from

$$\sqrt{n}\left(Z_n(\theta+\delta) - Z_n(\theta)\right) = \sum_{i\le n}\delta\int_0^1 \ddot{g}(x_i,\theta+s\delta)\,ds.$$

I'll use primes rather than dots to indicate derivatives with respect to the parameter, because $\dddot{g}(z,t)$ looks weird. Maybe it would be better use another letter, say $q = \dot{g}$, because $\ddot{q}$ looks better than $\dddot{g}$.

0.1 Assumptions

The assumptions are

(i) $\mathbb{E}_\theta|g'(z,\theta)|^2 < \infty$ and $\mathbb{E}_\theta\dot{g}(z,\theta) = 0$

(ii) $\mathbb{E}_\theta|g''(z,\theta)| < \infty$

(iii) There exists a neighborhood $U = (\theta-\epsilon_0, \theta+\epsilon_0)$ of θ and a function M for which

$$|g'''(z,t)| \le M(z) \qquad \text{for all } t\in U \text{ and all } z.$$

Also, $\mathbb{E}_\theta M < \infty$.

0.2 Pointwise control

Taylor's theorem with remainders written in integral form make it easier to control the error terms in the approximations for W_n. Start with the pointwise approximation to g'. For any fixed δ with $|\delta| \le \epsilon_0$,

<1> $$g'(z,\theta+\delta) - g'(z,\theta) = \int_0^1 \frac{\partial}{\partial s} g''(z,\theta+s\delta)\,ds = \delta \int_0^1 g''(z,\theta+s\delta)\,ds.$$

Similarly,

$$g''(z,\theta+s\delta) - g''(z,\theta) = \delta \int_0^s g'''(z,\theta+t\delta)\,dt.$$

Combine the last two equalities to deduce that

$$\begin{aligned} r_\theta(z,\delta) &:= g'(z,\theta+r) - g'(z,\theta) - \delta g''(z,\theta) \\ &= \delta^2 \iint \{0<t<s<1\} g'''(z,\theta+t\delta)\,dt\,ds \end{aligned}$$

and

<2> $$|r_\theta(z,\delta)| \le \iint \{0<t<s<1\} M(z)\,dt\,ds = \tfrac{1}{2}\delta^2 M(z).$$

0.3 Approximations

Temporarily write a_n for $n^{-1/2}$. By definition,

$$\begin{aligned} H_\theta(\theta+\delta) &= \mathbb{E}_\theta g'(z,\theta+\delta) \\ &= \mathbb{E}_\theta \left(g'(z,\theta) + \delta g''(z,\theta) + r_\theta(z,\delta) \right) \\ &= 0 - \delta \mathbb{I}(\theta) + R(\delta) \end{aligned}$$

where

$$|R(\delta)| \le \mathbb{E}_\theta |r_\theta(z,\delta)| \le \tfrac{1}{2}\delta^2 \mathbb{E}_\theta M \qquad \text{if } |\delta| \le \epsilon_0.$$

In particular, if $|h| \le C$ and $Cn^{-1/2} \le \epsilon_0$, then

<3> $$\mathbb{E}_\theta Z_n(\theta+ha_n) = a_n^{-1} H_\theta(\theta+ha_n) = -h\mathbb{I}(\theta) + R(ha_n)$$

where $|R(ha_n)| \le Ka_n$ for $K := C^2 \mathbb{E}_\theta M/2$.

Similarly,

$$Z_n(\theta + ha_n) = a_n \sum\nolimits_{i\le n} \big(g'(x_i,\theta) + ha_n g''(x_i,\theta) + r_\theta(x_i, ha_n)\big)$$

<4> $$= Z_n(\theta) + hJ_n + R_n(ha_n)$$

where $J_n = n^{-1}\sum_{i\le n} g''(x_i,\theta)$ and

$$|R_n(ha_n)| \le \frac{a_n C^2}{2n} \sum\nolimits_{i\le n} M(x_i).$$

From <**3**> and <**4**>, for $|h| \le C$

<5> $$\Delta_n(C) := \sup_{|h|\le C} |W_n(\theta + ha_n) - W_n(\theta)| \le C|J_n + \mathbb{I}(\theta)| + R_n(ha_n) + Ka_n.$$

In HW2.3 I asked you to show that $\mathbb{E}_\theta \Delta_n(C) \to 0$ for each fixed C, in the mistaken belief that it would make the problem easier. The contributions from R_n is easily disposed of:

$$|\mathbb{E}_\theta R_n(ha_n)| \le a_n C^2 \mathbb{E}_\theta M.$$

The SLLN tells us that $J_n + \mathbb{I}(\theta) \to 0$ with $\mathbb{P}_\theta$-probability one. You would need to make some sort of uniform integrabilty argument to deduce that $\mathbb{E}_\theta |J_n + \mathbb{I}(\theta)| \to 0$. (Sorry about that.)

In fact we only need the weaker convergence in probability,

<6> $$\mathbb{P}_\theta\{\Delta_n(C) > \epsilon\} \to 0 \qquad \text{as } n\to\infty, \text{ for each fixed } \epsilon > 0 \text{ and } C < \infty,$$

for which a WLLN applied to $J_n + \mathbb{I}(\theta)$ suffices.

Let me show you why <**6**> suffices to control $Z_n(\theta_n^*)$ for a root-n consistent estimator θ_n^*. Define $h_n^* = \sqrt{n}(\theta^* - \theta)$. For each $\epsilon > 0$ there exists a constant C_ϵ for which

$$\mathbb{P}_\theta\{|h_n^*| > C_\epsilon\} < \epsilon \qquad \text{for all } n.$$

The equality

$$Z_n(\theta_n^*) - Z_n(\theta) = W_n(\theta_n^*) - W_n(\theta) h_n^* \mathbb{I}(\theta)$$

shows that

$$|Z_n(\theta_n^*) - Z_n(\theta) + h_n^*\mathbb{I}(\theta)| \le \Delta_n(C_\epsilon) + Ka_n \qquad \text{whenever } |h_n^*| \le C_\epsilon.$$

It follows that $\mathbb{P}_\theta\{|Z_n(\theta_n^*) - Z_n(\theta) + h_n^*\mathbb{I}(\theta)| > 2\epsilon\} \to 0$ as $n \to \infty$, for each $\epsilon > 0$.