Statistics 581, Problem Set 9 Solutions

Wellner; 12/2/2009

1. Lehmann and Casella, TPE, problem 6.3.22, page 503, reworded as follows. (In other words, prove (vi) of theorem 1.2, pages 5-6, chapter 4 notes). Suppose that $X_1, \ldots, X_n$ are i.i.d. with density p_θ, $\theta \in \Theta \subset R^k$, satisfying the hypotheses of theorem 4.1, page 463 (the Cramér conditions given in (A) - (D) on pages 462-463). Show that the following Local Asymptotic Normality (LAN) result holds for the (local) log- likelihood ratios: with

$$L_n(\theta) \equiv \log(\prod_{i=1}^n p_\theta(X_i)) = \sum_{i=1}^n \log p_\theta(X_i),$$

for a fixed $\theta_0 \in \Theta$,

$$\begin{aligned} L_n(\theta_0 + n^{-1/2}\underline{t}) - L_n(\theta_0) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \underline{t}^T \dot{\underline{l}}_\theta(X_i) - \frac{1}{2}\underline{t}^T I(\theta_0)\underline{t} + o_p(1) \\ &\to_d N(0, t^T I(\theta_0)t) - \frac{1}{2}t^T I(\theta_0)t \\ &\stackrel{d}{=} N(-\sigma^2/2, \sigma^2) \end{aligned}$$

under P_{θ_0} where $\sigma^2 \equiv \underline{t}^T I(\theta_0)\underline{t}$. (The convergence in the last display actually holds under the considerably weaker hypothesis of Hellinger differentiability of p_θ at θ_0, as stated in Corollary 3 of section 3.3, page 28, of the Chapter 3 notes.)

Solution: I will use the notation of the Chapter 4 notes: in particular, $l_n(\theta) = \log\left(\prod_{i=1}^n p_\theta(X_i)\right) = \sum_{i=1}^n \log p_\theta(X_i)$. As in the proof of Theorem 1.3 of the Chapter 4 notes, we can write, with $\theta_n = \theta_0 + n^{-1/2}t$,

$$\begin{aligned} l_n(\theta_n) - l_n(\theta_0) &= (\theta_n - \theta_0)^T \dot{\mathbf{l}}_n(\theta_0) - \frac{1}{2}(\theta_n - \theta_0)^T \left(-\ddot{\mathbf{l}}_n(\theta_0)\right)(\theta_n - \theta_0) \\ &\quad + \frac{1}{6}\sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d (\theta_{nj} - \theta_{0j})(\theta_{nk} - \theta_{0k})(\theta_{nl} - \theta_{0l})(\dddot{\mathbf{l}}_{n,jkl}(\theta_n^*) \end{aligned}$$

where $|\theta_n^* - \theta_0| \le |\theta_n - \theta_0| = n^{-1/2}|t|$. Note that

$$\dot{\mathbf{l}}_n(\theta_0) = \sum_{i=1}^n \dot{\mathbf{l}}_\theta(X_i; \theta_0), \qquad \text{and} \tag{0.1}$$

$$\ddot{\mathbf{l}}_n(\theta_0) = \sum_{i=1}^n \ddot{\mathbf{l}}_{\theta,\theta}(X_i; \theta_0). \tag{0.2}$$

Furthemore, by the hypothesis A3 (ii),

$$\dddot{\mathbf{l}}_{n,jkl}(\theta_n^*) = \sum_{i=1}^n \dddot{\mathbf{l}}_{j,k,l}(\theta_n^*) = \sum_{i=1}^n \gamma_{j,k,l}(X_i) M_{j,k,l}(X_i)$$

where $0 \le |\gamma_{j,k,l}(x)| \le 1$ and

$$\frac{1}{n}\sum_{i=1}^{n} M_{j,k,l}(X_i) \to_{p,a.s.} E_{\theta_0} M_{j,k,l}(X_1) \qquad (0.3)$$

for all $1 \le j,k,l \le d$. Thus we can write

$$\begin{aligned}
l_n(\theta_n) - l_n(\theta_0) &= t^T n^{-1/2}\sum_{i=1}^{n} \dot{\mathbf{l}}_\theta(X_i;\theta_0) - \frac{1}{2}t^T I(\theta_0)t \\
&\quad + \frac{1}{2}t^T I(\theta_0)t - \frac{1}{2}t^T\left(-n^{-1}\ddot{\mathbf{l}}_n(\theta_0)\right)t \\
&\quad + \frac{1}{6}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d} t_j t_k t_l n^{-3/2}\sum_{i=1}^{n}\gamma_{j,k,l}(X_i)M_{j,k,l}(X_i) \\
&= t^T n^{-1/2}\sum_{i=1}^{n} \dot{\mathbf{l}}_\theta(X_i;\theta_0) - \frac{1}{2}t^T I(\theta_0)t \\
&\quad + \frac{1}{2}t^T I(\theta_0)t - \frac{1}{2}t^T\left(-n^{-1}\ddot{\mathbf{l}}_n(\theta_0)\right)t \\
&\quad + + \frac{1}{6}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d} t_j t_k t_l n^{-3/2}\sum_{i=1}^{n}\gamma_{j,k,l}(X_i)M_{j,k,l}(X_i) \\
&\equiv t^T n^{-1/2}\sum_{i=1}^{n} \dot{\mathbf{l}}_\theta(X_i;\theta_0) - \frac{1}{2}t^T I(\theta_0)t \\
&\quad + R_{n,1}(\theta_0,t) + R_{n,2}(\theta_0,t)
\end{aligned}$$

where, in view of (0.2) and (0.3)

$$R_{n,1}(\theta_0,t) \to_p 0 \qquad \text{since} \qquad -n^{-1}\sum_{i=1}^{n}\ddot{\mathbf{l}}_{\theta,\theta}(X_i;\theta_0) \to_p I(\theta_0),$$

and

$$|R_{n,2}(\theta_0,t)| \le \frac{1}{6}n^{-1/2}\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{l=1}^{d}|t_j||t_k||t_l| n^{-1}\sum_{i=1}^{n} M_{j,k,l}(X_i) \to_p 0.$$

2. Lehmann and Casella, problem 6.2.14, page 501. [**Hint:** We did this in class on Friday 11/20 using the first display on page 26 of the Chapter 3 notes. What remains to prove is uniform integrability or some other justification of the interchange of limit and integration.]

First Solution: From the first display on page 26 of Chapter 3, for $\theta_n = cn^{-1/2}$,

$$\begin{aligned}
\sqrt{n}(T_n - \theta_n) &\overset{d}{=} Z1_{[|Z+c|>n^{1/4}]} + (aZ + c(a-1))1_{[|Z+c|\le n^{1/4}]} \\
&\equiv Z_n \equiv A_n + B_n \\
&\to_{a.s.} aZ + c(a-1) \sim N(c(a-1), a^2).
\end{aligned}$$

Moreover, by the c_r-inequality (Proposition 2.1.3, page 5, chapter 2 notes) applied twice,

$$\begin{aligned} E|Z_n|^4 &\le 2^3\left(E|A_n|^4 + E|B_n|^4\right) \\ &\le 2^3\left(E|Z|^4 + E|aZ + c(a-1)|^4\right) \\ &\le 2^3\left(3 + 2^3(E|aZ|^4 + c^4(1-a)^4)\right) \\ &= 2^3\left(3 + 2^3(3a^4 + c^4(1-a)^4)\right). \end{aligned}$$

This implies that the sequence $\{|Z_n|^2\}_{n\ge 1}$ is uniformly integrable, and hence by Vitali's theorem (Theorem 2.1.2, page 4, chapter 2 notes),

$$R_n(\theta_n) = E_{\theta_n}\{n[T_n - \theta_n]^2\} = EZ_n^2 \to E(aZ + c(a-1))^2 = a^2 + c^2(1-a)^2.$$

As we showed in class, this limiting risk exceeds 1, the risk of $\overline{X}_n$, if $|c| > \sqrt{(1+a)/(1-a)}$.

Second Solution: Now $Z_n \to_{a.s.} aZ + c(a-1)$ where $Z \sim N(0,1)$ and by the c_r-inequality with $r = 2$

$$|Z_n|^2 \le 2\{Z^2 + (aZ + c(a-1))^2\} \equiv Y$$

where $EY < \infty$. Thus it follows from the dominated convergence theorem that $E(Z_n^2) \to E((aZ + c(a-1))^2)$.

3. (a) Exercise 2.1.6, page 10, chapter 2 notes.
(b) Exercise 2.1.7, page 10, chapter 2 notes.

Solution: (a) From the proof of proposition 1.13, chapter 2 notes, page 9, we see that

$$\begin{aligned} d_{TV}(P,Q) &= \frac{1}{2}\int |p-q|d\mu = \int_{[p\ge q]} (p-q)\, d\mu = \int_{[p\ge q]} pd\mu - \int_{[p\ge q]} p\wedge q\; d\mu \\ &= \int_{[p\ge q]} pd\mu + \int_{[p<q]} pd\mu - \int_{[p\ge q]} p\wedge q\; d\mu - \int_{[p<q]} p\; d\mu \\ &= \int pd\mu - \int_{[p\ge q]} p\wedge q\; d\mu - \int_{[p<q]} p\wedge q\; d\mu \\ &= 1 - \int p\wedge q\; d\mu \equiv 1 - \eta(P,Q). \end{aligned}$$

Alternatively, use the identity $|a-b| = a + b - 2(a\wedge b)$ for all $a, b \in \mathbb{R}$ to deduce that

$$|p(x) - q(x)| = p(x) + q(x) - 2p(x)\wedge q(x)$$

for each fixed x, and hence

$$\begin{aligned} d_{TV}(P,Q) &= \frac{1}{2}\int |p-q|d\mu = \frac{1}{2}\left(\int pd\mu + \int qd\mu - 2\int p\wedge qd\mu\right) \\ &= 1 - \int p\wedge qd\mu \equiv 1 - \eta(P,Q). \end{aligned}$$

(b) To see the first inequality, note that $H^2(P,Q) = 1 - \rho(P,Q)$ where

$$\rho(P,Q) = \int \sqrt{pq}\; d\mu \geq \int p \wedge q\; d\mu \equiv \eta(P,Q)$$

since $\sqrt{p(x)q(x)} \geq p(x) \wedge q(x)$ for all x. Thus we have

$$H^2(P,Q) \;=\; 1 - \rho(P,Q) \leq 1 - \eta(P,Q) = d_T V(P,Q).$$

For the second inequality, write $|p-q| = |(\sqrt{p} - \sqrt{q})(\sqrt{p} + \sqrt{q})|$ and then apply the Cauchy-Schwarz inequality: thus

$$\begin{aligned}
2d_{TV}(P,Q) &= \int |p-q| d\mu = \int |(\sqrt{p} - \sqrt{q})(\sqrt{p} + \sqrt{q})| d\mu \\
&\leq \left(\int |\sqrt{p} - \sqrt{q}|^2 d\mu \right)^{1/2} \left(\int |\sqrt{p} + \sqrt{q}|^2 d\mu \right)^{1/2} \\
&= \sqrt{2} H(P,Q) \left\{ \int (p + 2\sqrt{pq} + q) d\mu \right\}^{1/2} \\
&= \sqrt{2} H(P,Q) \left\{ 2 + 2\rho(P,Q) \right\}^{1/2} \\
&= 2H(P,Q) \left\{ 1 + \rho(P,Q) \right\}^{1/2},
\end{aligned}$$

and this yields the claimed inequality. The third inequality is easy since $\rho(P,Q) \leq 1$ by Cauchy-Schwarz again.