Empirical Likelihood Inference of Variance Components in Linear Mixed-Effects Models

Jingru Zhang[1], Wei Guo[2], Joanne S. Carpenter[3], Andrew Leroux[2,4], Kathleen R. Merikangas[2], Nicholas G. Martin[5], Ian B. Hickie[3], Haochang Shou[1], Hongzhe Li[1,*]

[1] *University of Pennsylvania,* [2] *National Institutes of Health,*
[3] *University of Sydney,* [4] *University of Colorado,*
[5] *QIMR Berghofer Medical Research Institute*

Supplementary Material

S1 Proofs and complements

S1.1 Proof of Theorem ??

To prove Theorem **??**, we first consider the setting with known β^*. We define $Z_i(\theta_1)$, $L_1(\theta_1)$, $\text{ELR}_1(\theta_1^0)$, and $\tilde{\nu}_{1n}^2(\theta_1^0)$ in the same way as $\hat{Z}_i(\theta_1)$, $L(\theta_1)$, $\text{ELR}(\theta_1^0)$, and $\hat{\nu}_{1n}^2(\theta_1^0)$, respectively, with $\hat{R}_i$ replaced by R_i. Under Conditions **??** and **??**, we derive the asymptotic distribution of $\text{ELR}_1(\theta_1^0)$ in the following theorem.

Theorem 1. *Let* $\tilde{c}_n(\theta_1^0) = \tilde{\nu}_{2n}^2(\theta_1^0)/\tilde{\nu}_{1n}^2(\theta_1^0)$, *where* $\tilde{\nu}_{1n}^2(\theta_1^0)$ *is a consistent estimator of the asymptotic variance of* $n^{-1/2}\sum_{i=1}^n Z_i(\theta_1^0)$ *and* $\tilde{\nu}_{2n}^2(\theta_1^0) = n^{-1}\sum_{i=1}^n Z_i^2(\theta_1^0)$. *If* $\theta_{(1)}^* \in R_+^{d-1}$, *then under Conditions* **??** *and* **??**, *as* $n \to \infty$, $\tilde{c}_n(\theta_1^0)\big(-2\log \mathrm{ELR}_1(\theta_1^0)\big) \to \chi_1^2$ *in distribution when* $\theta_1^0 > 0$, *and* $\tilde{c}_n(0)(-2\log \mathrm{ELR}_1(0)) \to U_+^2$ *in distribution, where* $U \sim N(0,1)$ *and* $U_+ = \max(U, 0)$.

Proof. For simplicity, we sometimes use Z_i to denote $Z_i(\theta_1)$ when there is no confusion. Using the method of Lagrange multipliers, let

$$\mathcal{L} = -\sum_{i=1}^n \log p_i + \kappa(\sum_{i=1}^n p_i - 1) + \lambda_0 \sum_{i=1}^n p_i Z_i.$$

Since

$$\frac{\partial \mathcal{L}}{\partial p_i} = -\frac{1}{p_i} + \kappa + \lambda_0 Z_i = 0,$$

we have

$$p_i = \frac{1}{\kappa + \lambda_0 Z_i} \quad \text{and} \quad \kappa = n. \tag{S1.1}$$

Plugging $\lambda_0 = n\lambda$ into (S1.1), we obtain

$$p_i = \frac{1}{n(1+\lambda Z_i)}. \tag{S1.2}$$

Since

$$0 = \sum_{i=1}^n p_i Z_i = \sum_{i=1}^n \frac{Z_i}{n(1+\lambda Z_i)}, \tag{S1.3}$$

under Condition **??**, one can show that

$$\lambda = \Big(\sum_{i=1}^{n} Z_i^2\Big)^{-1}\sum_{i=1}^{n} Z_i + o_p(n^{-1/2}) \text{ by Taylor expansion.}$$

Let $W_1(\theta_1) = n^n L_1(\theta_1)$. We have

$$\begin{aligned}
-2\log(W_1(\theta_1)) &= -2\sum_{i=1}^{n}(\log p_i + \log n) = 2\sum_{i=1}^{n}\log(1+\lambda Z_i)\\
&= 2\sum_{i=1}^{n}\Big(\lambda Z_i - \frac{1}{2}(\lambda Z_i)^2\Big) + o_p(1) \text{ (by Taylor expansion)}\\
&= \Big(\sum_{i=1}^{n} Z_i\Big)^2\Big(\sum_{i=1}^{n} Z_i^2\Big)^{-1} + o_p(1). \qquad \text{(S1.4)}
\end{aligned}$$

(1) If $\theta_1^0 > 0$, then Lemma 1 implies

$$\begin{aligned}
&\tilde{c}_n(\theta_1^0)\Big(-2\log\frac{L_1(\theta_1^0)}{\max_{\theta_1\geq 0} L_1(\theta_1)}\Big) = \tilde{c}_n(\theta_1^0)\Big(-2\log W_1(\theta_1^0)\Big)\\
=&\Big(n^{-1/2}\sum_{i=1}^{n} Z_i(\theta_1^0)\Big)^2/\tilde{\nu}_{1n}^2(\theta_1^0) + o_p(1)\\
=&\frac{\Big(n^{-1/2}\alpha^{-1}\sum_{i=1}^{n}\big\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, R_i - \theta_1^0\Phi_{i1}\big\rangle\Big)^2}{n^{-1}\alpha^{-2}\sum_{i=1}^{n}\big\langle R_i - H_i((\theta_1^0, \tilde{\theta}_{(1)}^T)^T), \Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}\big\rangle^2} + o_p(1).
\end{aligned}$$

Since

$$\begin{aligned}
&n^{-1}\alpha^{-2}\sum_{i=1}^{n}\Big\langle R_i - H_i((\theta_1^0, \tilde{\theta}_{(1)}^T)^T), \Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}\Big\rangle^2\\
=&\mathrm{var}\Big(n^{-1/2}\alpha^{-1}\sum_{i=1}^{n}\Big\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, R_i - \theta_1^0\Phi_{i1}\Big\rangle\Big) + o_p(1)
\end{aligned}$$

under Condition **??**, it implies $\tilde{c}_n(\theta_1^0)\big(-2\log \mathrm{ELR}_1(\theta_1^0)\big) \to \chi_1^2$ in distribution when $\theta_1^0 > 0$.

(2) When $\theta_1^0 = 0$, Lemma 1 implies

$$\tilde{c}_n(0)\big(-2\log\frac{L_1(0)}{\max_{\theta_1\geq 0} L_1(\theta_1)}\big) = \tilde{c}_n(0)\big(-2\log W_1(0)\big) I\big(\sum_{i=1}^n Z_i(0) \geq 0\big)$$

as n is large enough. Therefore, $\tilde{c}_n(0)\big(-2\log \mathrm{ELR}_1(0)\big) \to U_+^2$ in distribution, where $U \sim N(0,1)$ and $U_+ = \max(U, 0)$.

□

Lemma 1. *Let* $W_1(\theta_1) = n^n L_1(\theta_1)$ *and* $\tilde{\theta}_1 = \arg\max_{\theta_1\geq 0} W_1(\theta_1)$. *If the true value* $\theta_1^* > 0$, *then* $W_1(\tilde{\theta}_1) = 1$ *as* n *is large enough. If the true value* $\theta_1^* = 0$, *then* $W_1(\tilde{\theta}_1) = I(\sum_{i=1}^n Z_i(0) \geq 0) + W_1(0) I(\sum_{i=1}^n Z_i(0) < 0)$ *as* n *is large enough.*

Proof. Let $\check{\theta}_1 = \arg\max_{\theta_1} W_1(\theta_1)$. We use a similar method in Qin and Lawless (1994). Since

$$\check{\theta}_1 = \arg\min_{\theta_1} -2\log W_1(\theta_1),$$

$$\begin{aligned} 0 = \frac{\partial(-2\log W_1(\theta_1))}{\partial\theta_1}\Big|_{\theta_1=\check{\theta}_1} &= 2\sum_{i=1}^n \frac{\frac{\partial\lambda}{\partial\theta_1} Z_i + \lambda\frac{\partial Z_i}{\partial\theta_1}}{1+\lambda Z_i}\Big|_{\theta_1=\check{\theta}_1} \\ &= 2\lambda\sum_{i=1}^n \frac{1}{1+\lambda Z_i}\frac{\partial Z_i}{\partial\theta_1}\Big|_{\theta_1=\check{\theta}_1} \text{ (by (S1.3)).} \quad \text{(S1.5)} \end{aligned}$$

Let $\check{\lambda} = \lambda(\check{\theta}_1)$. We note $\check{\theta}_1$ and $\check{\lambda}$ satisfy

$$Q_{1n}(\check{\theta}_1, \check{\lambda}) = 0, \quad Q_{2n}(\check{\theta}_1, \check{\lambda}) = 0,$$

where

$$Q_{1n}(\theta_1, \lambda) = \frac{1}{n}\sum_{i=1}^{n} \frac{Z_i(\theta_1)}{1+\lambda Z_i(\theta_1)} \text{ (by (S1.3))},$$
$$Q_{2n}(\theta_1, \lambda) = \frac{\lambda}{n}\sum_{i=1}^{n} \frac{1}{1+\lambda Z_i(\theta_1)} \frac{\partial Z_i(\theta_1)}{\partial \theta_1} \text{ (by (S1.5))}.$$

Taking derivatives about θ_1 and λ, we have

$$\frac{\partial Q_{1n}(\theta_1, 0)}{\partial \theta_1} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial Z_i(\theta_1)}{\partial \theta_1}, \qquad \frac{\partial Q_{1n}(\theta_1, 0)}{\partial \lambda} = -\frac{1}{n}\sum_{i=1}^{n} Z_i(\theta_1)^2,$$
$$\frac{\partial Q_{2n}(\theta_1, 0)}{\partial \theta_1} = 0, \qquad \frac{\partial Q_{2n}(\theta_1, 0)}{\partial \lambda} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial Z_i(\theta_1)}{\partial \theta_1}.$$

Expanding Q_{1n} and Q_{2n} at $(\theta_1 = \theta_1^*, \lambda = 0)$, we have

$$\begin{aligned} 0 =& Q_{1n}(\check{\theta}_1, \check{\lambda}) \\ =& Q_{1n}(\theta_1^*, 0) + \frac{\partial Q_{1n}(\theta_1^*, 0)}{\partial \theta_1}(\check{\theta}_1 - \theta_1^*) + \frac{\partial Q_{1n}(\theta_1^*, 0)}{\partial \lambda}\check{\lambda} + o_p(n^{-1/2}), \quad \text{(S1.6)} \\ 0 =& Q_{2n}(\check{\theta}_1, \check{\lambda}) \\ =& Q_{2n}(\theta_1^*, 0) + \frac{\partial Q_{2n}(\theta_1^*, 0)}{\partial \theta_1}(\check{\theta}_1 - \theta_1^*) + \frac{\partial Q_{2n}(\theta_1^*, 0)}{\partial \lambda}\check{\lambda} + o_p(n^{-1/2}). \quad \text{(S1.7)} \end{aligned}$$

(S1.6) and (S1.7) give

$$\begin{aligned} &(\check{\theta}_1 - \theta_1^*,\ \check{\lambda})^T \\ =& \begin{pmatrix} \frac{1}{n}\sum_{i=1}^n \frac{\partial Z_i(\theta_1^*)}{\partial \theta_1} & -\frac{1}{n}\sum_{i=1}^n Z_i(\theta_1^*)^2 \\ 0 & \frac{1}{n}\sum_{i=1}^n \frac{\partial Z_i(\theta_1^*)}{\partial \theta_1} \end{pmatrix}^{-1} \begin{pmatrix} -\frac{1}{n}\sum_{i=1}^n Z_i(\theta_1^*) + o_p(n^{-1/2}) \\ o_p(n^{-1/2}) \end{pmatrix} \end{aligned}$$

Hence,

$$\check{\theta}_1 - \theta_1^* = -\frac{n^{-1}\sum_{i=1}^n Z_i(\theta_1^*)}{n^{-1}\sum_{i=1}^n \frac{\partial Z_i(\theta_1^*)}{\partial \theta_1}} + o_p(n^{-1/2}), \quad \text{(S1.8)}$$

where $(\sum_{i=1}^n Z_i(\theta_1^*))/(\sum_{i=1}^n \partial Z_i(\theta_1^*)/\partial\theta_1) = O_p(n^{-1/2})$.

When $\theta_1^* > 0$, (S1.8) implies $\check{\theta}_1 > 0$ as n is large enough. Thus, $\tilde{\theta}_1 = \check{\theta}_1$ as n is large enough. Then $\tilde{\theta}_1$ satisfies (S1.5), i.e.,

$$2\lambda \sum_{i=1}^n \frac{1}{1+\lambda Z_i}(-\|\Phi_{i1}\|_F^2) = 0. \tag{S1.9}$$

Plugging (S1.2) into (S1.9), we have $\lambda = 0$. Then $p_i = n^{-1}$ and $W_1(\tilde{\theta}_1) = 1$.

When $\theta_1^* = 0$, $\tilde{\theta}_1 = \check{\theta}_1 I(\check{\theta}_1 \geq 0)$ as n is large enough. Since

$$\sum_{i=1}^n \partial Z_i(0)/\partial\theta_1 = -\sum_{i=1}^n \|\Phi_{i1}\|_F^2 < 0,$$

we have $\tilde{\theta}_1 = \check{\theta}_1 I(\sum_{i=1}^n Z_i(0) \geq 0)$ as n is large enough. So $W_1(\tilde{\theta}_1) = I(\sum_{i=1}^n Z_i(0) \geq 0) + W_1(0) I(\sum_{i=1}^n Z_i(0) < 0)$. □

Proof of Theorem **??**. Let Δ be a d-dimensional vector with the kth element $\Delta_k = \sum_{i=1}^n \mathrm{tr}(\Phi_{ik} E(\hat{\epsilon}_i))$. Let $\varsigma_i = (\mathrm{tr}(\Phi_{i1}\Phi_{i2}), \cdots, \mathrm{tr}(\Phi_{i1}\Phi_{id}))^T$. For $i = 1, \cdots, n$,

$$E(\hat{R}_i) = H_i(\theta^*) + E(\hat{\epsilon}_i),$$

$$E(\hat{\theta}_{(1)}) = \theta_{(1)}^* + \left(\Xi^{-1}\right)_{-1}^T \Delta,$$

so we have

$$E(\hat{Z}_i(\theta_1^0)) = \mathrm{tr}(\Phi_{i1} E(\hat{\epsilon}_i)) - \varsigma_i^T (\Xi^{-1})_{-1}^T \Delta = O(n^{-1}), \tag{S1.10}$$

$$E(n^{-1/2} \sum_{i=1}^n \hat{Z}_i(\theta_1^0)) = o(1) \tag{S1.11}$$

under Proposition **??**. Then with similar techniques in the proof of Theorem 1, Theorem **??** can be proved. □

S1.2 Proof of of Proposition ??

Proof. Since

$$n^{1/2}(\hat{\beta} - \beta^*) = n^{1/2}(X^T X)^{-1} X^T r = (n^{-1} X^T X)^{-1} n^{-1/2} X^T r,$$

we have $n^{1/2}(\hat{\beta} - \beta^*) \xrightarrow{d} \Sigma^{-1}\eta$.

Under Condition **??**, $E\|r_i\|_2^2$ and $E\|r_i\|_2^4$ are bounded uniformly. Since

$$\begin{aligned} nE(r_i(\beta^* - \hat{\beta})^T X_i^T) =& - E(r_i \sum_{k=1}^{n} r_k^T X_k (n^{-1} X^T X)^{-1} X_i^T) \\ =& - E(r_i r_i^T X_i (n^{-1} X^T X)^{-1} X_i^T) \to O(1), \end{aligned}$$

$$nE(X_i(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_i^T) \to O(1),$$

we have $E(\hat{\epsilon}_i) = O(n^{-1})$.

Note that for $i \neq j$,

$$\begin{aligned} & \mathrm{cov}(r_i r_i^T, \hat{\epsilon}_j) \\ =& \mathrm{cov}(r_i r_i^T, r_j(\beta^* - \hat{\beta})^T X_j^T) + \mathrm{cov}(r_i r_i^T, X_j(\beta^* - \hat{\beta}) r_j^T) \\ & + \mathrm{cov}(r_i r_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T) \\ =& \mathrm{cov}(r_i r_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T). \end{aligned}$$

Since

$$
\begin{aligned}
& n^2 \mathrm{cov}(r_i r_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T) \\
= \ & \mathrm{cov}(r_i r_i^T, X_j(n^{-1}X^TX)^{-1} \sum_{l=1}^n X_l^T r_l \sum_{k=1}^n r_k^T X_k (n^{-1}X^TX)^{-1} X_j^T) \\
\to \ & \mathrm{cov}(r_i r_i^T, X_j \Sigma^{-1} X_i^T r_i r_i^T X_i \Sigma^{-1} X_j^T) = O(1),
\end{aligned}
$$

$\mathrm{cov}(r_i r_i^T, \hat{\epsilon}_j) = O(n^{-2})$.

For $\mathrm{cov}(\hat{\epsilon}_i, \hat{\epsilon}_j),\ i \neq j$, we only analyze $\mathrm{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T, r_j(\beta^* - \hat{\beta})^T X_j^T)$, $\mathrm{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T)$ and $\mathrm{cov}(X_i(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_i^T,$ $X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T)$. Since

$$
\begin{aligned}
& n^2 \mathrm{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T, r_j(\beta^* - \hat{\beta})^T X_j^T) \\
= \ & \mathrm{cov}(r_i \sum_{l=1}^n r_l^T X_l (n^{-1}X^TX)^{-1} X_i^T, r_j \sum_{k=1}^n r_k^T X_k (n^{-1}X^TX)^{-1} X_j^T) \\
\to \ & \mathrm{cov}(r_i r_j^T X_j \Sigma^{-1} X_i^T, r_j r_i^T X_i \Sigma^{-1} X_j^T) = O(1),
\end{aligned}
$$

$$
n^2 \mathrm{cov}(X_i(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T) \to O(1),
$$

$$
\begin{aligned}
&\text{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T, X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T) \\
=& - n^{-3}\text{cov}(r_i \sum_{k=1}^{n} r_k^T X_k (n^{-1}X^T X)^{-1} X_i^T, X_j (n^{-1}X^T X)^{-1} \\
&\sum_{s=1}^{n}\sum_{t=1}^{n} X_s^T r_s r_t^T X_t (n^{-1}X^T X)^{-1} X_j^T) \\
=& - n^{-3}\text{cov}(r_i r_i^T X_i (n^{-1}X^T X)^{-1} X_i^T, X_j (n^{-1}X^T X)^{-1} X_i^T r_i r_i^T X_i (n^{-1}X^T X)^{-1} X_j^T) \\
&- n^{-3}\sum_{k\neq i}\text{cov}(r_i r_k^T X_k (n^{-1}X^T X)^{-1} X_i^T, X_j (n^{-1}X^T X)^{-1}(X_i^T r_i r_k^T X_k \\
&+ X_k^T r_k r_i^T X_i)(n^{-1}X^T X)^{-1} X_j^T),
\end{aligned}
$$

we have $\text{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T, r_j(\beta^* - \hat{\beta})^T X_j^T)$, $\text{cov}(r_i(\beta^* - \hat{\beta})^T X_i^T,$ $X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T)$, and $\text{cov}(X_i(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_i^T,$ $X_j(\beta^* - \hat{\beta})(\beta^* - \hat{\beta})^T X_j^T) = O(n^{-2})$. Hence, $\text{cov}(\hat{\epsilon}_i, \hat{\epsilon}_j) = O(n^{-2})$. □

S1.3 Proof of equation (??)

Proof. Rewrite Ξ as $\Xi = \begin{pmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix}$ with E_{11} being a scalar. Rewrite $\hat{\Upsilon}$ as $\hat{\Upsilon} = (\hat{\Upsilon}_1, \hat{\Upsilon}_{(1)})^T$. So

$$
\hat{\theta}_{(1)} = (\Xi^{-1})_{-1}^T \hat{\Upsilon} = -E_{22}^{-1}E_{21}q^{-1}\hat{\Upsilon}_1 + E_{22}^{-1}\hat{\Upsilon}_{(1)} + E_{22}^{-1}E_{21}q^{-1}E_{12}E_{22}^{-1}\hat{\Upsilon}_{(1)},
$$

where $q = E_{11} - E_{12}E_{22}^{-1}E_{21}$. Let $F = E_{22}^{-1}E_{21}$. We obtain

$$\begin{aligned}\sum_{i=1}^{n}\hat{Z}_i(\theta_1^0) =&\hat{\Upsilon}_1 - F^T\big(-q^{-1}E_{21}\hat{\Upsilon}_1 + \hat{\Upsilon}_{(1)} + E_{21}q^{-1}E_{12}E_{22}^{-1}\hat{\Upsilon}_{(1)}\big) - E_{11}\theta_1^0\\ =&(1+q^{-1}F^TE_{21})\hat{\Upsilon}_1 - (1+F^TE_{21}q^{-1})F^T\hat{\Upsilon}_{(1)} - E_{11}\theta_1^0\\ =&(1+q^{-1}F^TE_{21})\sum_{i=1}^{n}\Big\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, \hat{R}_i\Big\rangle - E_{11}\theta_1^0\\ =&\sum_{i=1}^{n}\hat{D}_i(\theta_1^0),\end{aligned}$$

where $\hat{D}_i(\theta_1^0) = \alpha^{-1}\big\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, \hat{R}_i - \theta_1^0\Phi_{i1}\big\rangle$. Note that for any $b = (b_1, \cdots, b_{d-1})^T$,

$$\sum_{i=1}^{n}\Big\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, \sum_{j=1}^{d-1}b_j\Phi_{ij+1}\Big\rangle = \sum_{j=1}^{d-1}b_j\Xi_{1j+1} - \sum_{j=1}^{d-1}b_j\sum_{q=1}^{d-1}F_q\Xi_{q+1j+1} = 0,$$

so we have

$$\sum_{i=1}^{n}\hat{D}_i(\theta_1^0) = \sum_{i=1}^{n}\hat{M}_i(\theta_1^0),$$

where $\hat{M}_i(\theta_1^0) = \alpha^{-1}\langle\Phi_{i1} - \sum_{q=1}^{d-1}F_q\Phi_{iq+1}, \hat{R}_i - H_i((\theta_1^0, \hat{\theta}_{(1)}^T)^T)\rangle$. □

S1.4 Proof of Proposition ??

Proof. Since $\xi_i^{(g)} \sim N(0,1)$ and (S1.10), property (i) holds.

To prove property (ii), we have

$$\mathrm{var}\Big(n^{-1/2}\sum_{i=1}^{n}\hat{M}_i(\theta_1^0,t)\xi_i^{(g)}\Big) = n^{-1}\sum_{i=1}^{n}\mathrm{var}(\hat{M}_i(\theta_1^0,t)\xi_i^{(g)}) = n^{-1}\sum_{i=1}^{n}E\hat{M}_i(\theta_1^0,t)^2,$$

and

$$\begin{aligned}\operatorname{var}\big(n^{-1/2}\sum_{i=1}^{n}\hat{Z}_i(\theta_1^0,t)\big) &= \operatorname{var}\big(n^{-1/2}\sum_{i=1}^{n}\hat{D}_i(\theta_1^0,t)\big)\\ &= n^{-1}\sum_{i=1}^{n}\operatorname{var}(\hat{D}_i(\theta_1^0,t)) + o(1)\\ &= n^{-1}\sum_{i=1}^{n}E\hat{M}_i(\theta_1^0,t)^2 + o(1).\end{aligned}$$

Thus, we have property (ii).

Since

$$\begin{aligned}&\operatorname{cov}\big(n^{-1/2}\sum_{i=1}^{n}\hat{M}_i(\theta_1^0,s)\xi_i^{(g)}, n^{-1/2}\sum_{j=1}^{n}\hat{M}_j(\theta_1^0,t)\xi_j^{(g)}\big)\\ &= n^{-1}\sum_{i=1}^{n}\operatorname{cov}(\hat{M}_i(\theta_1^0,s)\xi_i^{(g)}, \hat{M}_i(\theta_1^0,t)\xi_i^{(g)})\\ &= n^{-1}\sum_{i=1}^{n}E(\hat{M}_i(\theta_1^0,s)\hat{M}_i(\theta_1^0,t)),\end{aligned}$$

and

$$\begin{aligned}&\operatorname{cov}\big(n^{-1/2}\sum_{i=1}^{n}\hat{Z}_i(\theta_1^0,s), n^{-1/2}\sum_{j=1}^{n}\hat{Z}_j(\theta_1^0,t)\big)\\ &= \operatorname{cov}\big(n^{-1/2}\sum_{i=1}^{n}\hat{D}_i(\theta_1^0,s), n^{-1/2}\sum_{j=1}^{n}\hat{D}_j(\theta_1^0,t)\big)\\ &= n^{-1}\sum_{i=1}^{n}\operatorname{cov}(\hat{D}_i(\theta_1^0,s), \hat{D}_i(\theta_1^0,t)) + o(1)\\ &= n^{-1}\sum_{i=1}^{n}E(\hat{M}_i(\theta_1^0,s)\hat{M}_i(\theta_1^0,t)) + o(1),\end{aligned}$$

we see property (iii) holds. $\square$

References

Qin, J. and J. Lawless (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics 22*(1), 300–325.