Frequentist Inference on Random Effects Based on Summarizability

Myunghee Cho Paik[1], Youngjo Lee[1] and Il Do Ha[2]

[1] *Seoul National University and* [2] *Pukyong National University*

Supplement Material

This supplement material gives proofs of the theorems in the paper: Frequentist Inference on Random Effects Based on Summarizability.

1. Proof of Theorem 1

Conditioning on u_{0i}, we can apply the standard maximum likelihood theory. We see that $E\Big[h_i^{(1)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}|u_i=u_{0i}\Big]=n_i^{-1}l_{2i}^{(1)}\{\boldsymbol{\theta},\nu(u_0);\mathbf{Y}_i\}$ and $\lim_{n_i\to\infty}E\Big[h_i^{(1)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}|u_i=u_{0i}\Big]=0$. Coupled with positivity of $I(\boldsymbol{\theta},u_{0i})$, $\hat{u}_i$ is consistent estimator of u_{0i}. Conditioning on u_{0i}, $\sqrt{n_i}(\hat{u}_i-u_{0i})$ is asymptotically equivalent to $\sqrt{n_i}I(\boldsymbol{\theta},u_{0i})^{-1}h_i^{(1)}\{\boldsymbol{\theta},\nu(u_{0i});\mathbf{Y}_i\}$. Denoting the conditional mean and variance of $\sqrt{n_i}(\hat{u}_i-u_i)$ by $A_{n_i}(\boldsymbol{\theta},u_{0i})$ and $B_{n_i}(\boldsymbol{\theta},u_{0i})$, we can see that

$$\begin{aligned}A_{n_i}(\boldsymbol{\theta},u_{0i})&=E\Big[\sqrt{n_i}(\hat{u}_i-u_i)|u_i=u_{0i}\Big]=\sqrt{n_i}I(\boldsymbol{\theta},u_{0i})^{-1}l_{2i}^{(1)}\{\boldsymbol{\theta},\nu(u_{0i})\}\\&=n_i^{-\frac{1}{2}}I(\boldsymbol{\theta},u_{0i})^{-1}\Big[a_1(\boldsymbol{\alpha})-a_2(\boldsymbol{\alpha})c^{(1)}(u_i)+\frac{d}{du_i}\log(du_i/d\nu_i)\Big]\Big|_{u_i=u_{0i}}=O(n_i^{-\frac{1}{2}}),\end{aligned}$$

and

$$\begin{aligned}B_{n_i}(\boldsymbol{\theta},u_{0i})&=Var\Big[\sqrt{n_i}(\hat{u}_i-u_i)|u_i=u_{0i}\Big]\\&=I(\boldsymbol{\theta},u_{0i})^{-1}Var\Big[\sqrt{n_i}h_i^{(1)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}|u_i=u_{0i}\Big]I(\boldsymbol{\theta},u_{0i})^{-1}\\&=I(\boldsymbol{\theta},u_{0i})^{-1}Var\Big[\sqrt{n_i}l_{1i}^{(1)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}|u_i=u_{0i}\Big]I(\boldsymbol{\theta},u_{0i})^{-1}\\&=I(\boldsymbol{\theta},u_{0i})^{-1}E\Big[-l_{1i}^{(2)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}|u_i=u_{0i}\Big]I(\boldsymbol{\theta},u_{0i})^{-1}\\&=I(\boldsymbol{\theta},u_{0i})^{-1}E\Big[-l_{1i}^{(2)}\{\boldsymbol{\theta},\nu(u_i);\mathbf{Y}_i\}-l_{2i}^{(2)}\{\boldsymbol{\theta},\nu(u_i)\}|u_i=u_{0i}\Big]I(\boldsymbol{\theta},u_{0i})^{-1}\\&+I(\boldsymbol{\theta},u_{0i})^{-1}E\Big[l_{2i}^{(2)}\{\boldsymbol{\theta},\nu(u_i)\}|u_i=u_{0i}\Big]I\{\boldsymbol{\theta},u_{0i})^{-1}\\&=I(\boldsymbol{\theta},u_{0i})^{-1}+I(\boldsymbol{\theta},u_{0i})^{-1}E\Big[l_{2i}^{(2)}\{\boldsymbol{\theta},\nu(u_i)\}|u_i=u_{0i}\Big]I(\boldsymbol{\theta},u_{0i})^{-1}\\&=I(\boldsymbol{\theta},u_{0i})^{-1}+O(n_i^{-1})\end{aligned}$$

Due to conditional independence, conditioning on $u_i = u_{0i}$, $\sqrt{n_i}(\hat{u}_i - u_{0i})$ is asymptotically distributed as $N(0, I(\boldsymbol{\theta}, u_{0i})^{-1})$. □

2. Proof of Theorem 2

Marginally, as $n_i \to \infty$, $E\{A_{n_i}(\boldsymbol{\theta}, u_i)\} = 0$, and $Var\{\sqrt{n_i}(\hat{u}_i - u_i)\} = E\{B_{n_i}(\boldsymbol{\theta}, u_i)\} + Var\{A_{n_i}(\boldsymbol{\theta}, u_i)\} = E\{I(\boldsymbol{\theta}, u_i)^{-1}\}$. Furthermore, marginally the moment generating function of $\sqrt{n_i}(\hat{u} - u_i)$ is

$$\begin{aligned} E_{u_i} E_{\mathbf{Y}_i|u_i}\Big[\exp\{t\sqrt{n_i}(\hat{u}_i - u_i)\}|u_i\Big] &= E_{u_i}\Big[\exp\{tA_{n_i}(\boldsymbol{\theta}, u_i) + \frac{1}{2}t^2 B_{n_i}(\boldsymbol{\theta}, u_i)\}\Big] + o(1) \\ &= E_{u_i}\Big[\{1 + tA_{n_i}(\boldsymbol{\theta}, u_i)\}\exp\{\frac{1}{2}t^2 B_{n_i}(\boldsymbol{\theta}, u_i)\}\Big] + o(1) \\ &= E_{u_i}\Big[\exp\{\frac{1}{2}t^2 I(\boldsymbol{\theta}, u_i)^{-1}\}\Big] + o(1). \Box \end{aligned}$$

3. Proof of Theorem 3

Under (A1) in Section 3, we can write

$$U_i(\boldsymbol{\theta}, \hat{u}_i; \mathbf{Y}_i) = 0 = U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + (\hat{u}_i - u_i)U_i^{(1)}(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + \frac{1}{2}(\hat{u}_i - u_i)^2 U_i^{(2)}(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)$$

and

$$(\hat{u}_i - u_i) = \{-U_i^{(1)}(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + O_p(n_i^{-\frac{1}{2}})\}^{-1} U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i).$$

Since

$$\begin{aligned} U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) =& n_i^{-1} g(\boldsymbol{\theta}, u_i)^{-1} \sum_{j=1}^{n} \frac{\partial \psi_{ij}}{\partial u_i}\{Y_{ij} - \frac{\partial}{\partial \psi_{ij}} b(\psi_{ij})\}/\boldsymbol{\phi} + n_i^{-1} g(\boldsymbol{\theta}, u_i)^{-1}\Big[a_1(\boldsymbol{\alpha}) \\ &- a_2(\boldsymbol{\alpha})\frac{d}{du_i}c(u_i) + \frac{d}{du_i}\log\{du_i/d\nu(u_i)\}\Big], \end{aligned}$$

we have

$$EU_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) = E\Big[U_{2i}(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\Big] = O(n_i^{-1}),$$

where $U_{2i}(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) = n_i^{-1} g(\boldsymbol{\theta}, u_i)^{-1}\Big[a_1(\gamma) - a_2(\gamma)\frac{d}{du_i}c(u_i) + \frac{d}{du_i}\log\{du_i/d\nu(u_i)\}\Big]$.

Under the conditions (A1) and (A2) in Section 3, we can summarize $\sqrt{n_i}(\hat{u}_i - u_i) = \sqrt{n_i}\kappa_i^{-1} U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + o_p(1)$. Using the conditional independence, as $\sqrt{n_i} \to \infty$, the moment generating function of $\sqrt{n_i}\kappa_i^{-1} U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)$ conditioning on u_i is $\exp\Big[\frac{1}{2}t^2 n_i \kappa_i^{-2} Var\{U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\}\Big]$. Then marginally, the moment generating function of $\sqrt{n_i}(\hat{u}_i - u_i)$ equals

$$\begin{aligned} &E_{u_i} E_{\mathbf{Y}_i|u_i}\Big[\exp\{t\sqrt{n_i}\kappa_i^{-1} U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\}|u_i\Big] \\ &= E_{u_i}\left[\exp\Big[\frac{1}{2}t^2 n_i \kappa_i^{-2} Var\{U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\}\Big]\right] + o(1). \Box \end{aligned} \tag{1}$$

4. Proof of Theorem 4

Assume that $n_i/K \to O(1)$. We can write

$$h_i^{(1)}\{\hat{\boldsymbol{\theta}}, \nu(\hat{u}_i); \mathbf{Y}_i\} = 0 \approx h_i^{(1)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}_i\} + (\hat{u}_i - u_{0i})h_i^{(2)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}_i\} + \frac{1}{2}(\hat{u}_i - u_{0i})^2 h_i^{(3)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}_i\}$$

and

$$\sqrt{n_i}(\hat{u}_i - u_{0i}) = \{-h_i^{(2)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}\} + O_p(n_i^{-\frac{1}{2}})\}^{-1}\sqrt{n_i}h_i^{(1)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}_i\}.$$

Since

$$h_i^{(1)}\{\hat{\boldsymbol{\theta}}, \nu(u_{0i}); \mathbf{Y}_i\} = h_i^{(1)}\{\boldsymbol{\theta}, \nu(u_{0i}); \mathbf{Y}_i\} + B_{21i}A_{11}^{-1}\{\frac{\partial}{\partial\boldsymbol{\theta}}l(\boldsymbol{\theta}; \mathbf{Y})\},$$

where $A_{11} = E\left\{-\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}l(\boldsymbol{\theta}; \mathbf{Y})\right\}$ and $B_{21i} = E\left\{[\frac{\partial}{\partial\boldsymbol{\theta}^T}h_i^{(1)}\{\boldsymbol{\theta}, \nu(u_i); \mathbf{Y}_i\}]|u_i = u_{0i}\right\}$, we can summarize

$$\sqrt{n_i}(\hat{u}_i - u_{0i}) = I(\boldsymbol{\theta}, u_{0i})^{-1}\sqrt{n_i}\Big[h_i^{(1)}\{\boldsymbol{\theta}, \nu(u_{0i}); \mathbf{Y}_i\} + B_{21i}A_{11}^{-1}\{\frac{\partial}{\partial\boldsymbol{\theta}}l(\boldsymbol{\theta}; \mathbf{Y})\}\Big] + o_p(1).$$

We can see that

$$\begin{aligned} Var\{\sqrt{n_i}(\hat{u}_i - u_{0i})\} &= I(\boldsymbol{\theta}, u_{0i})^{-1} \\ &+ n_i I(\boldsymbol{\theta}, u_{0i})^{-1}B_{21i}A_{11}^{-1}Var\Big[\frac{\partial}{\partial\boldsymbol{\theta}}l(\boldsymbol{\theta}; \mathbf{Y})|u_i = u_{0i}\Big]A_{11}^{-1}B_{21i}^T I(\boldsymbol{\theta}, u_{0i})^{-1} \\ &- 2n_i I(\boldsymbol{\theta}, u_{0i})^{-1}B_{21i}^T A_{11}^{-1}Cov\Big[\frac{\partial}{\partial\boldsymbol{\theta}}l(\boldsymbol{\theta}; \mathbf{Y}), h_i^{(1)}\{\boldsymbol{\theta}, \nu(u_i)\mathbf{Y}_i\}|u_i = u_{0i}\Big]. \quad (2) \end{aligned}$$

Using the conditional independence, as $\sqrt{n_i} \to \infty$, $\sqrt{n_i}(\hat{u}_i - u_{0i})$ is distributed as normal with mean 0 and variance given by (2). □

5. Proof of Theorem 5

Under (A1) in Section 3, we can write

$$U_i(\hat{\boldsymbol{\theta}}, \hat{u}_i; \mathbf{Y}_i) = 0 = U_i(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i) + (\hat{u}_i - u_i)U_i^{(1)}(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i) + \frac{1}{2}(\hat{u}_i - u_i)^2 U_i^{(2)}(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i)$$

and

$$\sqrt{n_i}(\hat{u}_i - u_i) = \{-U_i^{(1)}(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i) + O_p(n_i^{-\frac{1}{2}})\}^{-1}\sqrt{n_i}U_i(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i).$$

Replacing $U_i(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i)$ with $U_i^*(\hat{\boldsymbol{\theta}}, u_i; \mathbf{Y}_i)$, where

$$U_i^*(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) = U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + \Big[\frac{\partial}{\partial\boldsymbol{\theta}^T}U_i(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\Big]A_{11}^{-1}\Big[\frac{\partial}{\partial\boldsymbol{\theta}}l(\boldsymbol{\theta}; \mathbf{Y})\Big],$$

we can summarize $\sqrt{n_i}(\hat{u}_i - u_i) = \sqrt{n_i}\kappa^{-1}U_i^*(\boldsymbol{\theta}, u_i; \mathbf{Y}_i) + o_p(1)$. Using the conditional independence, as $\sqrt{n_i} \to \infty$, the moment generating function of $\sqrt{n_i}\kappa_i^{-1}U_i^*(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)$ conditioning on u_i is $\exp\left[\frac{1}{2}t^2 n_i\kappa_i^{-2}Var\{U_i^*(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)|u_i\}\right]$. Then marginally, the moment generating function of $\sqrt{n_i}(\hat{u}_i - u_i)$ equals

$$E_{u_i}\left[\exp\Big[\frac{1}{2}t^2 n_i\kappa_i^{-2}Var\{U_i^*(\boldsymbol{\theta}, u_i; \mathbf{Y}_i)\}\Big]\right] + o(1). \quad (3)$$

Table S1. Average coverage probabilities of the nominal 95% intervals over K independent units and average computing time (sec.) of each replication in marginal inference under Poisson-gamma models with unknown $\boldsymbol{\theta}$ using 500 replications.

		Coverage probability		Computing time		
		HL	Bayesian	HL	Bayesian	Comparison
N	(K, n)	Mean (SD)	Mean (SD)	sec.	sec.	rate of times
100	(50,2)	0.954 (0.010)	0.951 (0.009)	0.339	23.95	70.6
200	(100,2)	0.954 (0.011)	0.957 (0.009)	0.733	46.819	63.9
250	(50,5)	0.953 (0.009)	0.950 (0.009)	0.404	41.088	101.7
500	(100,5)	0.951 (0.010)	0.952 (0.010)	3.025	83.090	27.5
500	(50,10)	0.952 (0.009)	0.950 (0.011)	1.140	71.076	62.3
1000	(50,20)	0.949 (0.008)	0.957 (0.010)	4.684	132.531	28.3
Average						59.1

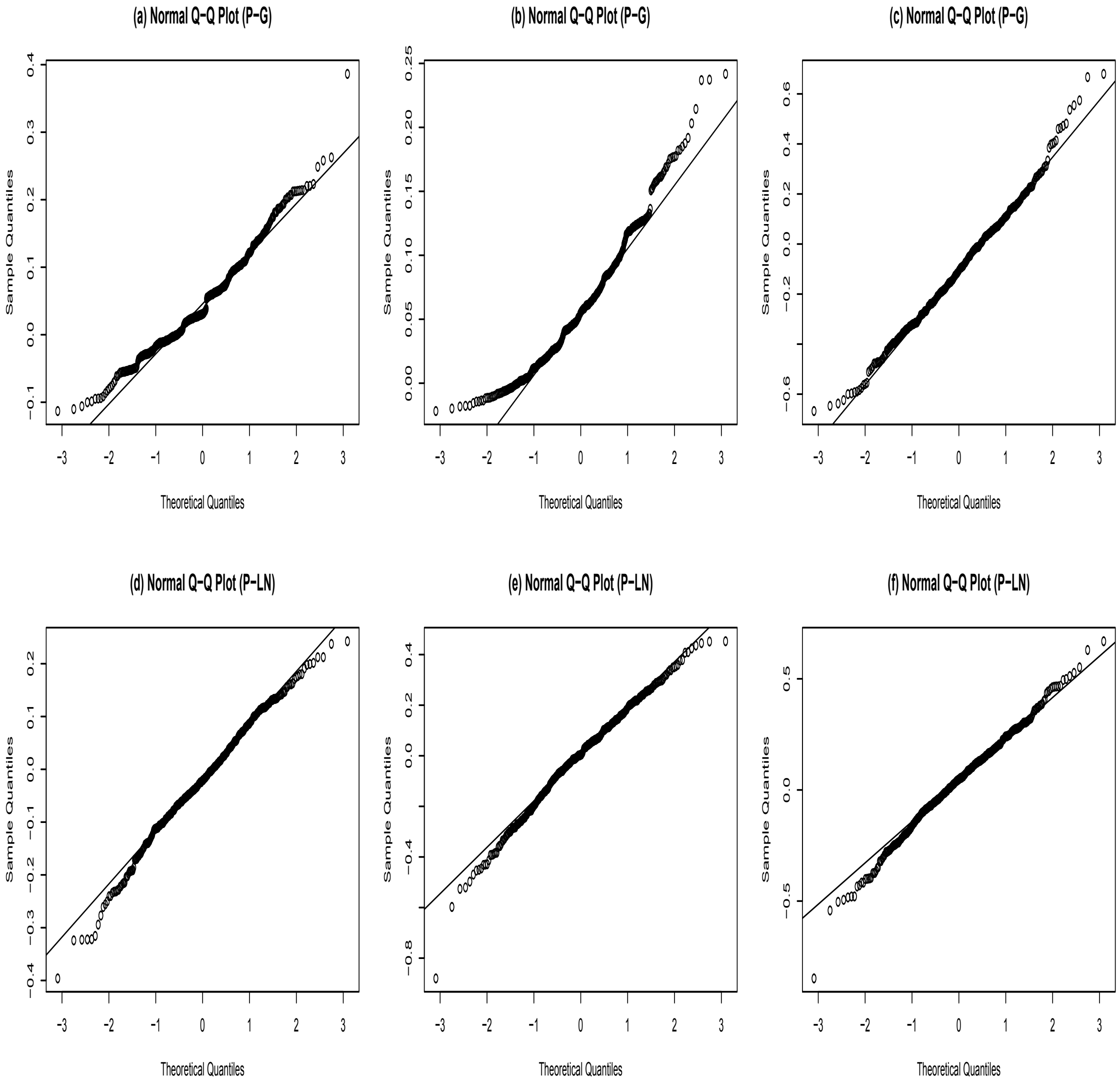

Figure S1. (a), (b) and (c), respectively, indicate the normal q-q plots of $\hat{u}_1 - u_{10}$, $\hat{u}_2 - u_{20}$ and $\hat{u}_3 - u_{30}$ using 500 replications under the Poisson-gamma model with sample size $(K, n) = (100, 5)$ and unknown fixed parameters; (d), (e) and (f) are also for the Poisson-lognormal model.