Economic Review (Otaru University of Commerce),
Vol.66, No.4, 1-46. March, 2016.

Expository supplement II to the paper "Asymptotic expansions for the estimators of Lagrange multipliers and associated parameters by the maximum likelihood and weighted score methods"

Haruhiko Ogasawara

This article gives the second half of an expository supplement to Ogasawara (2016).

4.5 Asymptotic cumulants of $t_{\mathrm{W}\theta}$ and $t_{\mathrm{W}\eta}$ by the weighted score method

$$t_{\mathrm{W}\theta} = \frac{n^{1/2}(\hat{\theta}_{\mathrm{W}} - \theta_0)}{\{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})\}^{1/2}} \text{ (recall that } \hat{\theta}_{\mathrm{W}} = \hat{\theta} + n^{-1}2k(1-2\hat{\theta}) + O_p(n^{-2})\text{)},$$

where

$$\{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})\}^{-1/2} = \{\hat{\theta}(1-\hat{\theta})\}^{-1/2} - \frac{1-2\hat{\theta}}{2\{\hat{\theta}(1-\hat{\theta})\}^{3/2}}(\hat{\theta}_{\mathrm{W}} - \hat{\theta}) + O_p(n^{-2})$$

$$= \{\hat{\theta}(1-\hat{\theta})\}^{-1/2} - n^{-1}\frac{k(1-2\hat{\theta})^2}{\{\hat{\theta}(1-\hat{\theta})\}^{3/2}} + O_p(n^{-2}),$$

consequently,

$$t_{\mathrm{W}\theta} = n^{1/2}\left[\{\hat{\theta}(1-\hat{\theta})\}^{-1/2} - n^{-1}\frac{k(1-2\hat{\theta})^2}{\{\hat{\theta}(1-\hat{\theta})\}^{3/2}}\right]\{\hat{\theta} - \theta_0 + n^{-1}2k(1-2\hat{\theta})\}$$

$$+ O_p(n^{-3/2})$$

$$= t_\theta + n^{-1/2}k\left[-(\hat{\theta} - \theta_0)\frac{(1-2\hat{\theta})^2}{\{\hat{\theta}(1-\hat{\theta})\}^{3/2}} + \frac{2(1-2\hat{\theta})}{\{\hat{\theta}(1-\hat{\theta})\}^{1/2}}\right] + O_p(n^{-3/2}).$$

Then, the asymptotic cumulants different from those for t_θ given earlier are

$$\kappa_1(t_{\mathrm{W}\theta}) = n^{-1/2}\left\{\alpha_{\theta 1}^{(t)} + \frac{2k(1-2\theta_0)}{\{\theta_0(1-\theta_0)\}^{1/2}}\right\} + O(n^{-3/2})$$

$$= n^{-1/2}\frac{\{2k-(1/2)\}(1-2\theta_0)}{\{\theta_0(1-\theta_0)\}^{1/2}} + O(n^{-3/2}) \equiv n^{-1}\alpha_{\mathrm{W}\theta 1}^{(t)} + O(n^{-3/2}),$$

where we find that when $k = 1/4$, $\alpha_{\mathrm{W}\theta 1}^{(t)} = 0$,

$$\kappa_2(t_{\mathrm{W}\theta}) = 1 + n^{-1}\underset{\text{(A)}}{\Bigg[}\ \alpha_{\theta\Delta 2}^{(t)} + 2k\Bigg[-\frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)}$$

$$-2\left\{\frac{2}{\{\theta_0(1-\theta_0)\}^{1/2}} + \frac{(1-2\theta_0)^2}{2\{\theta_0(1-\theta_0)\}^{3/2}}\right\}\theta_0(1-\theta_0)\frac{1}{\{\theta_0(1-\theta_0)\}^{1/2}}\Bigg]\underset{\text{(A)}}{\Bigg]}$$

$$+O(n^{-2})$$

$$= 1 + n^{-1}\left[\alpha_{\theta\Delta 2}^{(t)} - 2k\left\{2\frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)} + 4\right\}\right] + O(n^{-2})$$

$$= 1 + n^{-1}\left[\left(\frac{7}{4} - 4k\right)\frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)} + 3 - 8k\right] + O(n^{-2})$$

$$= 1 + n^{-1}\alpha_{\mathrm{W}\theta\Delta 2}^{(t)} + O(n^{-2})\ \ (\alpha_{\mathrm{W}\theta\Delta 2}^{(t)} \le \alpha_{\theta\Delta 2}^{(t)}).$$

$$t_{\mathrm{W}\eta} = n^{1/2}(-\hat{i}_{\mathrm{W}}^{\eta\eta})^{-1/2}(n^{-1}\hat{\eta}_{\mathrm{W}}) = n^{1/2}(-\hat{i}_{\mathrm{W}}^{\eta\eta})^{-1/2}\frac{n_{\mathrm{W}}}{n}\frac{c_{\mathrm{W}1}c_{\mathrm{W}2}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})}$$

$$= n^{1/2}\left(\frac{c_{\mathrm{W}1}c_{\mathrm{W}2}}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})}\right)^{-1/2}\frac{n_{\mathrm{W}}}{n}\frac{c_{\mathrm{W}1}c_{\mathrm{W}2}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})}$$

$$= n^{1/2}\frac{n_{\mathrm{W}}}{n}\frac{(c_{\mathrm{W}1}c_{\mathrm{W}2})^{1/2}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{\{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})\}^{1/2}},$$

where

$$c_{\mathrm{W}1}^{1/2} = c_1^{1/2} + \frac{n^{-1}}{2c_1^{1/2}} 2k(1-2c_1) + O(n^{-2})$$
$$= c_1^{1/2} + n^{-1}kc_1^{-1/2}(1-2c_1) + O(n^{-2}),$$
$$c_{\mathrm{W}2}^{1/2} = c_2^{1/2} + n^{-1}kc_2^{-1/2}(1-2c_2) + O(n^{-2}),$$
$$\hat{p}_{\mathrm{W}1} = \hat{p}_1 + n^{-1}kc_1^{-1}(1-2\hat{p}_1) + O_p(n^{-2}),$$
$$\hat{p}_{\mathrm{W}2} = \hat{p}_2 + n^{-1}kc_2^{-1}(1-2\hat{p}_2) + O_p(n^{-2}).$$

Then,

$$t_{\mathrm{W}\eta} = n^{1/2}(1+n^{-1}4k)\{c_1^{1/2} + n^{-1}kc_1^{-1/2}(1-2c_1)\}\{c_2^{1/2} + n^{-1}kc_2^{-1/2}(1-2c_2)\}$$
$$\times[\hat{p}_2 - \hat{p}_1 + n^{-1}k\{c_2^{-1} - c_1^{-1} - 2(c_2^{-1}\hat{p}_2 - c_1^{-1}\hat{p}_1)\}]$$
$$\times\left[\frac{1}{\{\hat{\theta}(1-\hat{\theta})\}^{1/2}} - n^{-1}\frac{k(1-2\hat{\theta})^2}{\{\hat{\theta}(1-\hat{\theta})\}^{3/2}}\right] + O_p(n^{-3/2})$$
$$= t_\eta + n^{-1}\left[\left\{4k + kc_1^{-1}(1-2c_1) + kc_2^{-1}(1-2c_2) - \frac{k(1-2\hat{\theta})^2}{\hat{\theta}(1-\hat{\theta})}\right\}t_\eta\right.$$
$$\left. + kn^{1/2}\{c_2^{-1} - c_1^{-1} - 2(c_2^{-1}\hat{p}_2 - c_1^{-1}\hat{p}_1)\}\frac{(c_1c_2)^{1/2}}{\{\hat{\theta}(1-\hat{\theta})\}^{1/2}}\right] + O_p(n^{-3/2})$$
$$= t_\eta + n^{-1}\left[k\left\{\frac{1}{c_1c_2} - \frac{(1-2\hat{\theta})^2}{\hat{\theta}(1-\hat{\theta})}\right\}t_\eta + \frac{n^{1/2}k\{c_1 - c_2 - 2(c_1\hat{p}_2 - c_2\hat{p}_1)\}}{\{c_1c_2\hat{\theta}(1-\hat{\theta})\}^{1/2}}\right]$$
$$+ O_p(n^{-3/2}).$$

Then, the asymptotic cumulants different from those for t_η given earlier are

$$k_1(t_{\mathrm{W}\eta}) = n^{-1/2}\left[\alpha_{\eta 1}^{(t)} + \frac{k(c_1-c_2)(1-2\theta_0)}{\{c_1c_2\theta_0(1-\theta_0)\}^{1/2}}\right] + O(n^{-3/2})$$

$$= n^{-1/2}\frac{k(c_1-c_2)(1-2\theta_0)}{\{c_1c_2\theta_0(1-\theta_0)\}^{1/2}} + O(n^{-3/2}) \equiv n^{-1/2}\alpha_{\mathrm{W}\eta 1}^{(t)} + O(n^{-3/2}) \quad (\alpha_{\eta 1}^{(t)} = 0),$$

$$k_2(t_{\mathrm{W}\eta}) = 1 + n^{-1}\Bigg[\ \alpha_{\eta\Delta 2}^{(t)} + 2k\left\{\frac{1}{c_1c_2} - \frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)}\right.$$

$$-\frac{(c_1-c_2)(1-2\theta_0)^2}{2(c_1c_2)^{1/2}\{\theta_0(1-\theta_0)\}^{3/2}}\theta_0(1-\theta_0)(c_1,c_2)\begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1}\end{pmatrix}\begin{pmatrix}-1 \\ 1\end{pmatrix}\left(\frac{c_1c_2}{\theta_0(1-\theta_0)}\right)^{1/2}$$

$$\left.\left.-\frac{2\theta_0(1-\theta_0)}{\{c_1c_2\theta_0(1-\theta_0)\}^{1/2}}\frac{\partial(c_1p_2-c_2p_1)}{\partial\mathbf{p}'}\begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1}\end{pmatrix}\begin{pmatrix}-1 \\ 1\end{pmatrix}\left(\frac{c_1c_2}{\theta_0(1-\theta_0)}\right)^{1/2}\right\}\right]$$

$$+O(n^{-2})$$

$$= 1 + n^{-1}\left[\alpha_{\eta\Delta 2}^{(t)} + 2k\left\{\frac{1}{c_1c_2} - \frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)} - \frac{2}{c_1c_2} + 4\right\}\right] + O(n^{-2})$$

$$= 1 + n^{-1}\left[\alpha_{\eta\Delta 2}^{(t)} + 2k\left\{4 - \frac{1}{c_1c_2} - \frac{(1-2\theta_0)^2}{\theta_0(1-\theta_0)}\right\}\right] + O(n^{-2})$$

$$= 1 + n^{-1}\alpha_{\mathrm{W}\eta\Delta 2}^{(t)} + O(n^{-2}),$$

where $(c_1,c_2)\begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1}\end{pmatrix}\begin{pmatrix}-1 \\ 1\end{pmatrix} = 0$ and

$$-\frac{2\theta_0(1-\theta_0)}{\{c_1c_2\theta_0(1-\theta_0)\}^{1/2}}\frac{\partial(c_1p_2-c_2p_1)}{\partial\mathbf{p}'}\begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1}\end{pmatrix}\begin{pmatrix}-1 \\ 1\end{pmatrix}\left(\frac{c_1c_2}{\theta_0(1-\theta_0)}\right)^{1/2}$$

$$= -2(-c_2,c_1)\begin{pmatrix}-c_1^{-1} \\ c_2^{-1}\end{pmatrix} = -2\left(\frac{c_2}{c_1} + \frac{c_1}{c_2}\right) = -2\frac{1-2c_1c_2}{c_1c_2} = -\frac{2}{c_1c_2} + 4$$

in the second and third terms, respectively in braces on the right-hand side of the first equation for $k_2(t_{\mathrm{W}\eta})$ are used (note that $\theta_0 = p = p_1 = p_2$).

Recall that $\hat{\theta}_{\mathrm{W}} = \hat{\theta} + n^{-1}2k(1-2\hat{\theta}) + O_p(n^{-2})$ and

$$\hat{\alpha}_{\mathrm{W}\theta 1}^{(t)} = \hat{\alpha}_{\theta 1}^{(t)} + \frac{2k(1-2\hat{\theta}_{\mathrm{W}})}{\{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})\}^{1/2}}\text{, then}$$

$$n\,\mathrm{acov}(\hat{\theta}_{\mathrm{W}},\hat{\alpha}_{\mathrm{W}\theta 1}^{(t)}) = n\,\mathrm{acov}(\hat{\theta},\hat{\alpha}_{\theta 1}^{(t)}) + n\,\mathrm{acov}\left[\hat{\theta},\frac{2k(1-2\hat{\theta})}{\{\hat{\theta}(1-\hat{\theta})\}^{1/2}}\right]$$

$$= n\,\mathrm{acov}(\hat{\theta},\hat{\alpha}_{\theta 1}^{(t)}) - k\alpha_{\theta 2}\left[\frac{4}{\{\theta_0(1-\theta_0)\}^{1/2}} + \frac{(1-2\theta_0)^2}{\{\theta_0(1-\theta_0)\}^{3/2}}\right]$$

$$= n\,\mathrm{acov}(\hat{\theta},\hat{\alpha}_{\theta 1}^{(t)}) - k\left[4\{\theta_0(1-\theta_0)\}^{1/2} + \frac{(1-2\theta_0)^2}{\{\theta_0(1-\theta_0)\}^{1/2}}\right].$$

From $\alpha_{\mathrm{W}\theta 3}^{(t)} = \alpha_{\theta 3}^{(t)}$, $n\,\mathrm{acov}(\hat{\theta}_{\mathrm{W}},\hat{\alpha}_{\mathrm{W}\theta 3}^{(t)}) = n\,\mathrm{acov}(\hat{\theta},\hat{\alpha}_{\theta 3}^{(t)})$.

Recall that $n^{-1}\hat{\eta}_{\mathrm{W}} = n^{-1}\hat{\eta} + O_p(n^{-1})$ and

$$\hat{\alpha}_{\mathrm{W}\eta 1}^{(t)} = \hat{\alpha}_{\eta 1}^{(t)} + \frac{k(c_1-c_2)(1-2\hat{\theta}_{\mathrm{W}})}{\{c_1c_2\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})\}^{1/2}}\text{, where } \hat{\alpha}_{\eta 1}^{(t)} = \alpha_{\eta 1}^{(t)} = 0\text{, then}$$

$$n\,\mathrm{acov}(n^{-1}\hat{\eta}_{\mathrm{W}},\hat{\alpha}_{\mathrm{W}\eta 1}^{(t)}) = \frac{k(c_1-c_2)}{(c_1c_2)^{1/2}}\left[\frac{-2}{\{\theta_0(1-\theta_0)\}^{1/2}} - \frac{(1-2\theta_0)^2}{2\{\theta_0(1-\theta_0)\}^{3/2}}\right]$$
$$\times n\,\mathrm{acov}(\hat{\theta},\hat{\mathbf{p}}')\frac{\partial n^{-1}\eta_0}{\partial \mathbf{p}}$$
$$= 0$$

where

$$n\,\mathrm{acov}(\hat{\theta},\hat{\mathbf{p}}')\frac{\partial n^{-1}\eta_0}{\partial \mathbf{p}} = \theta_0(1-\theta_0)(c_1,c_2)\begin{pmatrix} c_1^{-1} & 0 \\ 0 & c_2^{-1}\end{pmatrix}\begin{pmatrix}-1\\ 1\end{pmatrix}\frac{c_1c_2}{\theta_0(1-\theta_0)} = 0$$ is

used.

From $\alpha_{\mathrm{W}\eta 3}^{(t)} = \alpha_{\eta 3}^{(t)}$, $n\,\mathrm{acov}(\hat{\theta}_{\mathrm{W}},\hat{\alpha}_{\mathrm{W}\eta 3}^{(t)}) = n\,\mathrm{acov}(\hat{\theta},\hat{\alpha}_{\eta 3}^{(t)})$.

5. Results using special properties of Example 2.2

5.1 Asymptotic cumulants of $\hat{\theta}$ and $\hat{\eta}$ by maximum likelihood

Let $\theta_{12} = \theta_1 + \theta_2 = 2\theta_0$ and $\hat{p}_{12} = \hat{p}_1 + \hat{p}_2$. Then, $\hat{\theta} \equiv \hat{\theta}_{\mathrm{ML}} = (\hat{p}_1 + \hat{p}_2)/2 = \hat{p}_{12}/2$. Since $\hat{p}_{12}$ is the usual sample proportion for the combined category of the first two original ones,

$$\kappa_1(\hat{\theta}-\theta_0)=0,\ \kappa_2(\hat{\theta})=n^{-1}2^{-2}\theta_{12}(1-\theta_{12}),$$

$$\kappa_3(\hat{\theta})=n^{-2}2^{-3}\theta_{12}(1-\theta_{12})(1-2\theta_{12}),$$

$$\kappa_4(\hat{\theta})=n^{-3}2^{-4}\theta_{12}(1-\theta_{12})\{1-6\theta_{12}(1-\theta_{12})\}.$$

$$n^{-1}\hat{\eta}\equiv n^{-1}\hat{\eta}_{\mathrm{ML}}=\frac{\hat{p}_2-\hat{p}_1}{\hat{p}_1+\hat{p}_1}=\frac{\hat{p}_2-\hat{p}_1}{\hat{p}_{12}},\ \eta_0=0,$$

$$\hat{\mathbf{p}}=(\hat{p}_1,\hat{p}_2)',\mathbf{p}=(p_1,p_2)'=(\theta_1,\theta_2)'=(\theta_1,\theta_1)'=(\theta_2,\theta_2)'=\theta_0(1,1)',$$

$$n^{-1}\hat{\eta}=\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}(\hat{\mathbf{p}}-\mathbf{p})+\frac{1}{2}\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}(\hat{\mathbf{p}}-\mathbf{p})^{<2>}+\frac{1}{6}\frac{n^{-1}\partial^3\eta_0}{(\partial\mathbf{p}')^{<3>}}(\hat{\mathbf{p}}-\mathbf{p})^{<3>}$$

$$+O_p(n^{-2}),$$

where

$$\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}}=\frac{n^{-1}\partial\hat{\eta}}{\partial\hat{\mathbf{p}}}|_{\hat{\mathbf{p}}=\mathbf{p}}=\left\{\frac{1}{\hat{p}_{12}}\begin{pmatrix}-1\\1\end{pmatrix}-\frac{\hat{p}_2-\hat{p}_1}{\hat{p}_{12}^2}\begin{pmatrix}1\\1\end{pmatrix}\right\}|_{\hat{\mathbf{p}}=\mathbf{p}}=\frac{1}{p_{12}}\begin{pmatrix}-1\\1\end{pmatrix},$$

$$\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p})^{<2>}}=\left\{-\frac{1}{\hat{p}_{12}^2}\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}-\frac{1}{\hat{p}_{12}^2}\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}+\frac{2(\hat{p}_2-\hat{p}_1)}{\hat{p}_{12}^3}\begin{pmatrix}1\\1\end{pmatrix}^{<2>}\right\}|_{\hat{\mathbf{p}}=\mathbf{p}}$$

$$=-\frac{1}{p_{12}^2}\left\{\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\right\}=\frac{2}{p_{12}^2}(1,0,0,-1)',$$

$$\frac{n^{-1}\partial^3\eta_0}{(\partial\mathbf{p})^{<3>}}=\frac{2}{p_{12}^3}\left\{\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}^{<2>}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}^{<2>}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\right\}$$

$$=\frac{2}{p_{12}^3}(-3,-1,-1,1,-1,1,1,3)',$$

where

$$\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}^{<2>}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}^{<2>}\otimes\begin{pmatrix}-1\\1\end{pmatrix}$$
$$=\left\{\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\right\}\otimes\begin{pmatrix}1\\1\end{pmatrix}+(-1,1,-1,1,-1,1,-1,1)'$$
$$=(-2,0,0,2)'\otimes(1,1)'+(-1,1,-1,1,-1,1,-1,1)'$$
$$=(-2,-2,0,0,0,0,2,2)'+(-1,1,-1,1,-1,1,-1,1)'$$
$$=(-3,-1,-1,1,-1,1,1,3)'.$$

In elementwise expressions,

$$\kappa_1(\hat{p}_a)=p_a,$$
$$n\kappa_2(\hat{p}_a,\hat{p}_b)=\delta_{ab}p_a-p_ap_b,$$
$$n^2\kappa_3(\hat{p}_a,\hat{p}_b,\hat{p}_c)=\delta_{abc}p_a-\sum^{(3)}\delta_{ab}p_ap_c+2p_ap_bp_c,$$
$$n^3\kappa_4(\hat{p}_a,\hat{p}_b,\hat{p}_c,\hat{p}_d)=\delta_{abcd}p_a-\sum^{(4)}\delta_{abc}p_ap_d-\sum^{(3)}\delta_{ab}\delta_{cd}p_ap_c$$
$$+2\sum^{(6)}\delta_{ab}p_ap_cp_d-6p_ap_bp_cp_d\quad(a,b,c=1,...,A)$$

(Stuart & Ort, 1994, Equation (7.18); Ogasawara, 2010), where $\delta_{ab\cdots d}=\delta_{ab}\delta_{b\bullet}\cdots\delta_{\bullet d}$.

Alternatively,

$$\kappa_1(\hat{\mathbf{p}})=\mathbf{p},\ n\kappa_2(\hat{\mathbf{p}})=\mathrm{vec}\,\mathrm{diag}(\mathbf{p})-\mathbf{p}^{<2>},$$
$$n^2\kappa_3(\hat{\mathbf{p}})=\mathbf{e}^{(3)}(\mathbf{p})-\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}\otimes\mathbf{p}-\mathbf{p}\otimes\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}$$
$$-(\mathbf{I}_{(2)}\otimes\mathbf{K}_2)[\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}\otimes\mathbf{p}]+2\mathbf{p}^{<3>},$$

where $\mathbf{K}_A(\mathbf{a}\otimes\mathbf{b})=\mathbf{b}\otimes\mathbf{a}$ (**a** and **b** are $A\times 1$ vectors) and $\mathbf{e}^{(j)}(\mathbf{a})=\sum_{i=1}^{A}\mathbf{e}_{(i)}^{<j>}a_i, \mathbf{e}_{(i)}=(\mathbf{0}',1,\mathbf{0}')'$, $\mathbf{e}_{(i)}$ is the $A\times 1$ vector whose *i*-th element is 1 and the remaining ones are 0. Note that $\mathbf{e}^{(2)}(\mathbf{a})=\mathrm{vec}\,\mathrm{diag}(\mathbf{a})$.
Let

$$\sum^{(3)}\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}\otimes\mathbf{p}=\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}\otimes\mathbf{p}$$

$$+\mathbf{p}\otimes\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\}+(\mathbf{I}_{(2)}\otimes\mathbf{K}_2)\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})\otimes\mathbf{p}\}.$$

Then,

$$n^2\kappa_3(\hat{\mathbf{p}})=\mathbf{e}^{(3)}(\mathbf{p})-\sum^{(3)}\mathbf{e}^{(2)}(\mathbf{p})\otimes\mathbf{p}+2\mathbf{p}^{<3>},$$

$$n^3\kappa_4(\hat{\mathbf{p}})=\mathbf{e}^{(4)}(\mathbf{p})-\sum^{(4)}\mathbf{e}^{(3)}(\mathbf{p})\otimes\mathbf{p}-\sum^{(3)}\mathbf{e}^{(2)}(\mathbf{p})\otimes\mathbf{e}^{(2)}(\mathbf{p})$$

$$+2\sum^{(6)}\mathbf{e}^{(2)}(\mathbf{p})\otimes\mathbf{p}^{<2>}-6\mathbf{p}^{<4>}.$$

In the above expressions, noting that $\mathbf{p}=(p_1,p_2)'$ and using the lexicographical order for eight quantities,
(111, 112, 121, 122, 211, 212, 221, 222)

$$n^2\kappa_3(\hat{\mathbf{p}})=\mathbf{e}^{(3)}(\mathbf{p})-\sum^{(3)}\mathbf{e}^{(2)}(\mathbf{p})\otimes\mathbf{p}+2\mathbf{p}^{<3>}$$

$$=(p_1,\mathbf{0}_{(6)}',p_2)'-(3p_1^2,p_1p_2,p_1p_2,p_1p_2,p_1p_2,p_1p_2,p_1p_2,3p_2^2)'$$

$$+2(p_1^3,p_1^2p_2,p_1^2p_2,p_1p_2^2,p_1^2p_2,p_1p_2^2,p_1p_2^2,p_2^3).$$

Using the lexicographical order for 16 quantities,
(1111, 1112, 1121, 1122, 1211, 1212, 1221, 1222,
2111, 2112, 2121, 2122, 2211, 2212, 2221, 2222),

$$n^3\kappa_4(\hat{\mathbf{p}}) = (p_1, \mathbf{0}_{(14)}{}', p_2)' - (4p_1^2, p_1p_2, p_1p_2, 0, p_1p_2, 0, 0, p_1p_2,$$
$$p_1p_2, 0, 0, p_1p_2, 0, p_1p_2, p_1p_2, 4p_2^2)'$$
$$-(3p_1^2, 0, 0, p_1p_2, 0, p_1p_2, p_1p_2, 0, 0, p_1p_2, p_1p_2, 0, p_1p_2, 0, 0, 3p_2^2)'$$
$$+2(6p_1^3, 3p_1^2p_2, 3p_1^2p_2, p_1p_2^2 + p_1^2p_2, 3p_1^2p_2, p_1p_2^2 + p_1^2p_2,$$
$$p_1p_2^2 + p_1^2p_2, 3p_1p_2^2,$$
$$3p_1^2p_2, p_1p_2^2 + p_1^2p_2, p_1p_2^2 + p_1^2p_2, 3p_1p_2^2, p_1p_2^2 + p_1^2p_2, 3p_1p_2^2,$$
$$3p_1p_2^2, 6p_2^3)'$$
$$-6(p_1^4, p_1^3p_2, p_1^3p_2, p_1^2p_2^2, p_1^3p_2, p_1^2p_2^2, p_1^2p_2^2, p_1p_2^3$$
$$p_1^3p_2, p_1^2p_2^2, p_1^2p_2^2, p_1p_2^3, p_1^2p_2^2, p_1p_2^3, p_1p_2^3, p_2^4)'.$$

For $n^{-1}\hat{\eta} = \dfrac{\hat{p}_2 - \hat{p}_1}{\hat{p}_1 + \hat{p}_2}$ $(p \equiv p_1 = p_2 = p_{12}/2)$,

$$\kappa_1(n^{-1}\hat{\eta}) = \frac{1}{2}\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\mathrm{E}_\theta(\hat{\mathbf{p}} - \mathbf{p})^{<2>} + O(n^{-2})$$
$$= \frac{n^{-1}}{p_{12}^2}(1, 0, 0, -1)\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p}) - \mathbf{p}^{<2>}\} + O(n^{-2})$$
$$= \frac{1}{p_{12}^2}(p_1 - p_2 - p_1^2 + p_2^2) + O(n^{-2})$$
$$= O(n^{-2})\ (\alpha_{\eta 1} = 0),$$

$$\kappa_2(n^{-1}\hat{\eta}) = \frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}n^{-1}\{\mathrm{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'\}\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}}$$
$$+\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\kappa_3(\hat{\mathbf{p}}, \hat{\mathbf{p}}'^{<2>})\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p})^{<2>}} + \frac{1}{4}\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\kappa_2(\hat{\mathbf{p}}^{<2>}, \hat{\mathbf{p}}'^{<2>})\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p})^{<2>}}$$
$$-n^{-2}\alpha_{\eta 1}^2 + \frac{1}{3}\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\kappa_2(\hat{\mathbf{p}}, \hat{\mathbf{p}}'^{<3>})\frac{n^{-1}\partial^3\eta_0}{(\partial\mathbf{p})^{<3>}} + O(n^{-3})$$

$$
\begin{aligned}
&= \frac{n^{-1}}{p_{12}^2}(-1,1)\{\text{diag}(\mathbf{p})-\mathbf{pp}'\}\begin{pmatrix}-1\\1\end{pmatrix} + \frac{2n^{-2}}{p_{12}^3}\{(-1,1)\otimes(1,0,0,-1)\}n^2\kappa_3(\hat{\mathbf{p}}) \\
&\quad + \frac{2n^{-2}}{p_{12}^4}\text{tr}\left[\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\{\text{diag}(\mathbf{p})-\mathbf{pp}'\}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\{\text{diag}(\mathbf{p})-\mathbf{pp}'\}\right] \\
&\quad + \frac{2n^{-2}}{p_{12}^4}[(-1,1)\{\text{diag}(\mathbf{p})-\mathbf{pp}'\}]\otimes\{\text{vec}'\,\text{diag}(\mathbf{p})-\mathbf{p}'^{<2>}\} \\
&\qquad \times(-3,-1,-1,1,-1,1,1,3)' + O(n^{-3}) \\
&= \frac{n^{-1}}{p_{12}^2}(p_1+p_2) + \frac{2n^{-2}}{p_{12}^3}\{(-p_1-p_2)-(-3p_1^2+p_1p_2+p_1p_2-3p_2^2) \\
&\qquad\qquad\qquad\qquad + 2(-p_1^3+p_1p_2^2+p_1^2p_2-p_2^3)\} \\
&\quad + \frac{2n^{-2}}{p_{12}^4}\text{tr}\left\{\begin{pmatrix}p_1-p_1^2 & -p_1p_2\\ p_1p_2 & -(p_2-p_2^2)\end{pmatrix}\begin{pmatrix}p_1-p_1^2 & -p_1p_2\\ p_1p_2 & -(p_2-p_2^2)\end{pmatrix}\right\} \\
&\quad + \frac{2n^{-2}}{p_{12}^4}[\{-(p_1-p_1^2)-p_1p_2,\ p_1p_2+(p_2-p_2^2)\} \\
&\qquad \otimes(p_1-p_1^2,-p_1p_2,-p_1p_2,p_2-p_2^2)](-3,-1,-1,1,-1,1,1,3)' + O(n^{-3}) \\
&= \frac{n^{-1}}{p_{12}} + \frac{2n^{-2}}{p_{12}^3}(-p_{12}+4p^2) + \frac{2n^{-2}}{p_{12}^4}\{(p-p^2)^2-p^4-p^4+(p-p^2)^2\} \\
&\quad + \frac{2n^{-2}}{p_{12}^4}2(p-p^2+p^2)\{3(p-p^2)-p^2-p^2-(p-p^2)\} + O(n^{-3}) \\
&= \frac{n^{-1}}{2p} + \frac{2n^{-2}}{8p^3}(-2p+4p^2) + \frac{2n^{-2}}{16p^4}(2p^2-4p^3) + \frac{4n^{-2}}{16p^4}p(2p-4p^2) \\
&\quad + O(n^{-3}) \\
&= \frac{n^{-1}}{2p} + n^{-2}\frac{1-2p}{4p^2} + O(n^{-3}) = n^{-1}\alpha_{\eta 2} + n^{-2}\alpha_{\eta\Delta 2} + O(n^{-3}),
\end{aligned}
$$

$$\kappa_3(n^{-1}\hat{\eta}) = n^{-2}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<3>} n^2\kappa_3(\hat{\mathbf{p}})$$

$$+n^{-2}\frac{3}{2}\left\{\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\otimes\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<2>}\right\}\sum^{(3)}\{\mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<2>}$$

$$-n^{-2}\frac{3}{2}\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}})\alpha_{\eta 2} + O(n^{-3})$$

$$= \frac{n^{-2}}{p_{12}^3}(-1,1)^{<3>} n^2\kappa_3(\hat{\mathbf{p}}) + 3n^{-2}\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'} n\,\mathrm{cov}(\hat{\mathbf{p}})\frac{n^{-1}\partial^2\eta_0}{\partial\mathbf{p}\,\partial\mathbf{p}'} n\,\mathrm{cov}(\hat{\mathbf{p}})\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}}$$

$$+O(n^{-3})$$

$$= \frac{n^{-2}}{p_{12}^3}(-1,1,1,-1,1,-1,-1,1)\{(p,\mathbf{0}_{(6)}',p) - p^2(3,\mathbf{1}_{(6)}',3) + 2p^3\mathbf{1}_{(8)}'\}'$$

$$+\frac{3n^{-2}}{p_{12}^4}2(-1,1)\{\mathrm{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'\}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\{\mathrm{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}'\}\begin{pmatrix}-1\\ 1\end{pmatrix} + O(n^{-3})$$

$$= \frac{6n^{-2}}{p_{12}^4}\{-(p-p^2)-p^2, p^2+(p-p^2)\}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}-(p-p^2)-p^2\\ p^2+(p-p^2)\end{pmatrix}$$

$$+O(n^{-3})$$

$$= \frac{6n^{-2}}{p_{12}^4}(-p,p)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}-p\\ p\end{pmatrix} + O(n^{-3}) = O(n^{-3})\ (\alpha_{\eta 3} = 0),$$

where $(-1,1)^{<3>} = (1,-1,-1,1)\otimes(-1,1) = (-1,1,1,-1,1,-1,-1,1)$.

$$\kappa_4(n^{-1}\hat{\eta}) = n^{-3}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<4>} n^3\kappa_4(\hat{\mathbf{p}})$$

$$+2n^{-3}\left\{\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<3>} \otimes \frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\right\}\sum^{(10)} \mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}}) \otimes n^2\kappa_3(\hat{\mathbf{p}})$$

$$+n^{-3}\left\{\frac{3}{2}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<2>} \otimes \left(\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\right)^{<2>} + \frac{2}{3}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<3>} \frac{n^{-1}\partial^3\eta_0}{(\partial\mathbf{p}')^{<3>}}\right\}$$

$$\times\sum^{(15)}\{\mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>} - 6\alpha_{\eta 2}\alpha_{\eta\Delta 2} + O(n^{-4}),$$

where the first term is

$$n^{-3}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<4>} n^3\kappa_4(\hat{\mathbf{p}})$$

$$= \frac{n^{-3}}{p_{12}^4}(1,-1,-1,1,-1,1,1,-1,-1,1,1,-1,1,-1,-1,1)$$

$$\times\{(p,\mathbf{0}_{(14)}',p) - p^2(4,1,1,0,1,0,0,1,1,0,0,1,0,1,1,4)$$

$$-p^2(3,0,0,1,0,1,1,0,0,1,1,0,1,0,0,3)$$

$$+2p^3(6,3,3,2,3,2,2,3,3,2,2,3,2,3,3,6)\}'$$

$$= \frac{n^{-3}}{p_{12}^4}\{2p - 2p^2(4-1-1-1-1) - 2p^2(3+1+1+1)$$

$$+4p^3(6-3-3+2-3+2+2-3)\}$$

$$= \frac{n^{-3}}{p_{12}^4}(2p - 12p^2) = \frac{n^{-3}}{8p^3}(1-6p),$$

with

$$(-1,1)^{<4>} = (-1,1)^{<3>} \otimes (-1,1)$$

$$= (1,-1,-1,1,-1,1,1,-1,-1,1,1,-1,1,-1,-1,1),$$

the second term is

$$2n^{-3}\left\{\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<3>}\otimes\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\right\}\sum^{(10)}\mathrm{vec}\,n\,\mathrm{cov}(\hat{\mathbf{p}})\otimes n^2\kappa_3(\hat{\mathbf{p}})$$

$$=n^{-3}\frac{4}{p_{12}^5}\{(-1,1,1,-1,1,-1,-1,1)\otimes(1,0,0,-1)\}$$

$$\times\sum^{(10)}\mathrm{vec}\,n\,\mathrm{cov}(\hat{\mathbf{p}})\otimes n^2\kappa_3(\hat{\mathbf{p}}),$$

and the third term is

$$n^{-3}\left\{\frac{3}{2}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<2>}\otimes\left(\frac{n^{-1}\partial^2\eta_0}{(\partial\mathbf{p}')^{<2>}}\right)^{<2>}+\frac{2}{3}\left(\frac{n^{-1}\partial\eta_0}{\partial\mathbf{p}'}\right)^{<3>}\frac{n^{-1}\partial^3\eta_0}{(\partial\mathbf{p}')^{<3>}}\right\}$$

$$\times\sum^{(15)}\{\mathrm{vec}\,n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}$$

$$=n^{-3}\left\{\frac{3}{2}\frac{4}{p_{12}^6}(-1,1)^{<2>}\otimes(1,0,0,-1)^{<2>}\right.$$

$$\left.+\frac{2}{3}\frac{2}{p_{12}^6}(-1,1)^{<3>}\otimes(-3,-1,-1,1,-1,1,1,3)\right\}\sum^{(15)}\{\mathrm{vec}\,n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}$$

$$=n^{-3}\frac{1}{p_{12}^6}\left\{\;6(-1,1)^{<2>}\otimes(1,0,0,-1)^{<2>}\right.$$

$$\left.+\frac{4}{3}(-1,1)^{<3>}\otimes(-3,-1,-1,1,-1,1,1,3)\right\}\sum^{(15)}\{\mathrm{vec}\,n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}.$$

Then,

$$\kappa_4(n^{-1}\hat{\eta}) = n^{-3}\Bigg[\ \frac{1-6p}{8p^3} + \frac{1}{8p^5}\{(-1,1,1,-1,1,-1,-1,1)\otimes(1,0,0,-1)\}$$

$$\times \sum^{(10)} \mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}}) \otimes n^2\kappa_3(\hat{\mathbf{p}}) + \frac{1}{p^6}\left\{\frac{3}{32}(-1,1)^{<2>} \otimes (1,0,0,-1)^{<2>}\right.$$

$$\left. + \frac{1}{48}(-1,1)^{<3>} \otimes (-3,-1,-1,1,-1,1,1,3)\right\} \sum^{(15)} \{\mathrm{vec}\, n\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}$$

$$-6\alpha_{\eta 2}\alpha_{\eta\Delta 2} \ \Bigg] + O(n^{-4})$$

$$= n^{-3}\alpha_{\eta 4} + O(n^{-4}).$$

5.2 The information matrix

We use $\theta_0 \equiv \theta_1 = \theta_2 = p_1 = p_2 = p$ after differentiation.

$$\mathbf{I}_0^* = \begin{pmatrix} \mathbf{I}_0 & -\mathbf{H}_0 \\ -\mathbf{H}_0{}' & 0 \end{pmatrix}, \ \mathbf{H}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \ \text{where}$$

$$\mathbf{I}_0 = -\mathrm{E}_\theta\left(\frac{\partial^2 \overline{l}}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}'}\right) = -\mathrm{E}_\theta\begin{pmatrix} -\frac{m_1/n}{\theta_1^2} - \frac{m_3/n}{\theta_3^2} & -\frac{m_3/n}{\theta_3^2} \\ -\frac{m_3/n}{\theta_3^2} & -\frac{m_2/n}{\theta_2^2} - \frac{m_3/n}{\theta_3^2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\theta_1} + \frac{1}{\theta_3} & \frac{1}{\theta_3} \\ \frac{1}{\theta_3} & \frac{1}{\theta_2} + \frac{1}{\theta_3} \end{pmatrix} = \begin{pmatrix} \frac{1}{\theta_1} & 0 \\ 0 & \frac{1}{\theta_2} \end{pmatrix} + \frac{1}{\theta_3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1,1),$$

$$\mathbf{I}_0^{*-1} = \begin{Bmatrix} \mathbf{I}_0^{-1} - \mathbf{I}_0^{-1}\mathbf{H}_0(\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\mathbf{I}_0^{-1} & -\mathbf{I}_0^{-1}\mathbf{H}_0(\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1} \\ -(\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\mathbf{I}_0^{-1} & -(\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1} \end{Bmatrix}$$

where

$$\mathbf{I}_0^{-1} = \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix} - \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \left\{ (1,1) \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \theta_3 \right\}^{-1} (1,1) \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix}$$

$$= \begin{pmatrix} \theta_1 - \theta_1^2 & -\theta_1\theta_2 \\ -\theta_1\theta_2 & \theta_2 - \theta_2^2 \end{pmatrix} = \begin{pmatrix} \theta_0 - \theta_0^2 & -\theta_0^2 \\ -\theta_0^2 & \theta_0 - \theta_0^2 \end{pmatrix},$$

$$\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0 = (1,-1) \begin{pmatrix} \theta_1 - \theta_1^2 & -\theta_1\theta_2 \\ -\theta_1\theta_2 & \theta_2 - \theta_2^2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= \theta_1 - \theta_1^2 + 2\theta_1\theta_2 + \theta_2 - \theta_2^2 = 2\theta_0,$$

$$\mathbf{I}_0^{-1}\mathbf{H}_0 = \begin{pmatrix} \theta_0 - \theta_0^2 & -\theta_0^2 \\ -\theta_0^2 & \theta_0 - \theta_0^2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \theta_0 \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Consequently,

$$\mathbf{I}_0^{*-1} = \begin{Bmatrix} \begin{pmatrix} \theta_0 - \theta_0^2 & -\theta_0^2 \\ -\theta_0^2 & \theta_0 - \theta_0^2 \end{pmatrix} - \dfrac{\theta_0^2}{2\theta_0} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1,-1) & -\dfrac{\theta_0}{2\theta_0} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ -\dfrac{\theta_0}{2\theta_0} (1,-1) & -\dfrac{1}{2\theta_0} \end{Bmatrix}$$

$$= \begin{pmatrix} \dfrac{\theta_0}{2} - \theta_0^2 & \dfrac{\theta_0}{2} - \theta_0^2 & -\dfrac{1}{2} \\ \dfrac{\theta_0}{2} - \theta_0^2 & \dfrac{\theta_0}{2} - \theta_0^2 & \dfrac{1}{2} \\ -\dfrac{1}{2} & \dfrac{1}{2} & -\dfrac{1}{2\theta_0} \end{pmatrix}.$$

Since $\theta_1 + \theta_2 = 2\theta_0 = \theta_{12}$,

$n\,\mathrm{avar}(\hat{\theta}_1) = n\,\mathrm{avar}(\hat{\theta}_2) = \dfrac{\theta_0}{2} - \theta_0^2 = \dfrac{\theta_{12}}{4} - \dfrac{\theta_{12}^2}{4} = \dfrac{1}{4}\theta_{12}(1-\theta_{12})$ and

$n\,\mathrm{avar}(n^{-1}\hat{\eta}) = -i^{\eta\eta} = \dfrac{1}{2\theta_0} = \dfrac{1}{2p}$, which are equal to the results in Subsection 5.1.

5.3 Asymptotic cumulants of t_θ and t_η by ML

$$t_\theta \equiv \frac{n^{1/2}(\hat{\theta}-\theta_0)}{\{\hat{\theta}_{12}(1-\hat{\theta}_{12})/4\}^{1/2}} = \frac{n^{1/2}\{(\hat{\theta}_{12}/2)-(\theta_{12}/2)\}}{\{\hat{\theta}_{12}(1-\hat{\theta}_{12})/4\}^{1/2}} = \frac{n^{1/2}(\hat{\theta}_{12}-\theta_{12})}{\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{1/2}}.$$

$$\kappa_1(t_\theta) = -n^{-1/2}\frac{\{\theta_{12}(1-\theta_{12})\}^{-1/2}}{2}(1-2\theta_{12}) + O(n^{-3/2})$$

$$\equiv n^{-1/2}\alpha_{\theta 1}^{(t)} + O(n^{-3/2}),$$

$$\kappa_2(t_\theta) = 1 + n^{-1}\left\{\frac{7}{4}\frac{(1-2\theta_{12})^2}{\theta_{12}(1-\theta_{12})} + 3\right\} + O(n^{-2}) = 1 + n^{-1}\alpha_{\theta\Delta 2}^{(t)} + O(n^{-2})$$

$$(\alpha_{\theta 2}^{(t)} = 1),$$

$$\kappa_3(t_\theta) = -n^{-1/2}2\{\theta_{12}(1-\theta_{12})\}^{-1/2}(1-2\theta_{12}) + O(n^{-3/2})$$

$$\equiv n^{-1/2}\alpha_{\theta 3}^{(t)} + O(n^{-3/2}),$$

$$\kappa_4(t_\theta) = n^{-1}[\{\theta_{12}(1-\theta_{12})\}^{-1} + 9(1-2\theta_{12})^2\{\theta_{12}(1-\theta_{12})\}^{-1} + 6] + O(n^{-2})$$

$$\equiv n^{-1}\alpha_{\theta 4}^{(t)} + O(n^{-2}),$$

$$n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 1}^{(t)}) = \frac{1}{4}(1-2\theta_{12})^2\{\theta_{12}(1-\theta_{12})\}^{-1/2} + \{\theta_{12}(1-\theta_{12})\}^{1/2},$$

$$n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 3}^{(t)}) = (1-2\theta_{12})^2\{\theta_{12}(1-\theta_{12})\}^{-1/2} + 4\{\theta_{12}(1-\theta_{12})\}^{1/2}.$$

$$t_\eta = \frac{n^{1/2}(n^{-1}\hat{\eta})}{\{1/(2\hat{p})\}^{1/2}} = \frac{n^{1/2}(\hat{p}_2-\hat{p}_1)}{(\hat{p}_1+\hat{p}_2)(\hat{p}_1+\hat{p}_2)^{-1/2}} = \frac{n^{1/2}(\hat{p}_2-\hat{p}_1)}{(\hat{p}_1+\hat{p}_2)^{1/2}},$$

where $n^{-1}\hat{\eta} = \dfrac{\hat{p}_2-\hat{p}_1}{\hat{p}_1+\hat{p}_2}$, $\hat{p} = \dfrac{\hat{p}_1+\hat{p}_2}{2}, 1/(2\hat{p}) = -\hat{i}^{\eta\eta}$.

Since $n\,\mathrm{var}(\hat{p}_2-\hat{p}_1) = p_2 - p_2^2 + p_1 - p_1^2 + 2p_1p_2 = p_1 + p_2$, t_η is the studentized $\hat{p}_2 - \hat{p}_1$.

$$\frac{\partial t_\eta}{\partial \hat{\mathbf{p}}} = n^{1/2}\left\{\frac{(-1,1)'}{(\hat{p}_1+\hat{p}_2)^{1/2}} - \frac{(\hat{p}_2-\hat{p}_1)(1,1)'}{2(\hat{p}_1+\hat{p}_2)^{3/2}}\right\},$$

$$\frac{\partial^2 t_\eta}{(\partial \hat{\mathbf{p}})^{<2>}} = n^{1/2}\left[-\frac{1}{2(\hat{p}_1+\hat{p}_2)^{3/2}}\left\{\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\right\} + \frac{3}{4}\frac{\hat{p}_2-\hat{p}_1}{(\hat{p}_1+\hat{p}_2)^{5/2}}\begin{pmatrix}1\\1\end{pmatrix}^{<2>}\right],$$

$$\frac{\partial^3 t_\eta}{(\partial \mathbf{p})^{<3>}} = n^{1/2}\frac{3}{4(p_1+p_2)^{5/2}}\left\{\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}^{<2>}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix}^{<2>}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\right\}$$

$$= n^{1/2}\frac{3}{4p_{12}^{5/2}}(-3,-1,-1,1,-1,1,1,3)',$$

where

$$\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix} = (-2,0,0,2)',$$

$$\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}^{<2>}+\begin{pmatrix}1\\1\end{pmatrix}\otimes\begin{pmatrix}-1\\1\end{pmatrix}\otimes\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}^{<2>}\otimes\begin{pmatrix}-1\\1\end{pmatrix}$$

$$= (-2,0,0,2)'\otimes(1,1)' + (-1,1,-1,1,-1,1,-1,1)'$$

$$= (-2,-2,0,0,0,0,2,2)' + (-1,1,-1,1,-1,1,-1,1)'$$

$$= (-3,-1,-1,1,-1,1,1,3)',$$

$$\frac{\partial t_\eta}{\partial \mathbf{p}} = \frac{n^{1/2}}{p_{12}^{1/2}}(-1,1)', \quad \frac{\partial^2 t_\eta}{(\partial \mathbf{p})^{<2>}} = \frac{n^{1/2}}{p_{12}^{3/2}}(1,0,0,-1)'.$$

$$\kappa_1(t_\eta)=\frac{1}{2}\frac{\partial^2 t_\eta}{(\partial \mathbf{p}')^{<2>}}\mathrm{E}_\theta\{(\hat{\mathbf{p}}-\mathbf{p})^{<2>}\}+O(n^{-3/2})$$

$$=\frac{n^{-1/2}}{2p_{12}^{3/2}}(1,0,0,-1)'\{\mathrm{vec}\,\mathrm{diag}(\mathbf{p})-\mathbf{p}^{<2>}\}+O(n^{-3/2})$$

$$=\frac{n^{-1/2}}{2p_{12}^{3/2}}(p_1-p_2+p_1^2-p_2^2)+O(n^{-3/2})=O(n^{-3/2})\ \ (\alpha_{\eta 1}^{(t)}=0),$$

$$\kappa_2(t_\eta)=\frac{\partial t_\eta}{\partial \mathbf{p}'}\mathrm{cov}(\hat{\mathbf{p}})\frac{\partial t_\eta}{\partial \mathbf{p}}+\frac{\partial t_\eta}{\partial \mathbf{p}'}\kappa_3(\hat{\mathbf{p}},\hat{\mathbf{p}}'^{<2>})\frac{\partial^2 t_\eta}{(\partial \mathbf{p})^{<2>}}$$

$$+\frac{1}{4}\frac{\partial^2 t_\eta}{(\partial \mathbf{p}')^{<2>}}\kappa_2(\hat{\mathbf{p}}^{<2>},\hat{\mathbf{p}}'^{<2>})\frac{\partial^2 t_\eta}{(\partial \mathbf{p})^{<2>}}+\frac{1}{3}\frac{\partial t_\eta}{\partial \mathbf{p}'}\kappa_2(\hat{\mathbf{p}},\hat{\mathbf{p}}'^{<3>})\frac{\partial^3 t_\eta}{(\partial \mathbf{p})^{<3>}}$$

$$-n^{-1}(\alpha_{\eta 1}^{(t)})^2+O(n^{-2}),$$

where the first term is

$$\frac{1}{p_{12}}(-1,1)n\,\mathrm{cov}(\hat{\mathbf{p}})\begin{pmatrix}-1\\1\end{pmatrix}=\frac{1}{p_{12}}(-1,1)\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}=\frac{2p}{p_{12}}=1$$

(recall that $p=p_1=p_2=p_{12}/2$),

the second term is

$$\frac{n}{p_{12}^2}\{(-1,1)\otimes(1,0,0,-1)\}\kappa_3(\hat{\mathbf{p}})$$

$$=\frac{n^{-1}}{p_{12}^2}(-1,0,0,1,1,0,0,-1)\{(p,\mathbf{0}_{(6)}',p)-p^2(3,\mathbf{1}_{(6)}',3)+2p^3\mathbf{1}_{(8)}'\}'$$

$$=\frac{n^{-1}}{p_{12}^2}(-2p+6p^2-2p^2)=\frac{n^{-1}}{p_{12}^2}(-2p+4p^2)=-\frac{n^{-1}}{2p}(1-2p),$$

the third term is

$$\frac{n^{-1}}{2}\mathrm{tr}\left\{\frac{\partial^2 t_\eta}{\partial\mathbf{p}\ \partial\mathbf{p}'}n\,\mathrm{cov}(\hat{\mathbf{p}})\frac{\partial^2 t_\eta}{\partial\mathbf{p}\ \partial\mathbf{p}'}n\,\mathrm{cov}(\hat{\mathbf{p}})\right\}$$

$$=\frac{n^{-1}}{2}\frac{1}{p_{12}^3}\mathrm{tr}\left\{\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\right\}$$

$$=\frac{n^{-1}}{2p_{12}^3}\mathrm{tr}\left\{\begin{pmatrix}p-p^2 & -p^2\\ p^2 & -(p-p^2)\end{pmatrix}\begin{pmatrix}p-p^2 & -p^2\\ p^2 & -(p-p^2)\end{pmatrix}\right\}$$

$$=\frac{n^{-1}}{2p_{12}^3}\{(p-p^2)^2-p^4-p^4+(p-p^2)^2\}$$

$$=\frac{n^{-1}}{16p^3}(2p^2-4p^3)=\frac{n^{-1}}{8p}(1-2p),$$

and the fourth term is

$$\frac{n^{-1}}{p_{12}^3}\frac{3}{4}(-1,1)[n\,\mathrm{cov}(\hat{\mathbf{p}})\otimes\mathrm{vec}'\{n\,\mathrm{cov}(\hat{\mathbf{p}})\}](-3,-1,-1,1,-1,1,1,3)'$$

$$=\frac{n^{-1}}{p_{12}^3}\frac{3}{4}(-1,1)\left\{\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\otimes(p-p^2,-p^2,-p^2,p-p^2)\right\}$$
$$\times(-3,-1,-1,1,-1,1,1,3)'$$

$$=\frac{n^{-1}}{p_{12}^3}\frac{3}{4}\{(-p,p)\otimes(p-p^2,-p^2,-p^2,p-p^2)\}(-3,-1,-1,1,-1,1,1,3)'$$

$$=\frac{n^{-1}}{p_{12}^3}\frac{3}{4}p^2\{-(1-p),p,p,-(1-p),(1-p),-p,-p,(1-p)\}$$
$$\times(-3,-1,-1,1,-1,1,1,3)'$$

$$=\frac{n^{-1}}{p_{12}^3}\frac{3}{4}p^2 2\{3(1-p)-p-p-(1-p)\}$$

$$=\frac{3n^{-1}}{16p}(2-4p)=\frac{3n^{-1}}{8p}(1-2p).$$

Then,

$$\kappa_2(t_\eta)=1+\frac{n^{-1}}{p}(1-2p)\left(-\frac{1}{2}+\frac{1}{8}+\frac{3}{8}\right)+O(n^{-2})$$

$$=1+O(n^{-2})\quad(\alpha_{\eta 2}^{(t)}=1,\ \alpha_{\eta\Delta 2}^{(t)}=0).$$

$$\kappa_3(t_\eta)=\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<3>}\kappa_3(\hat{\mathbf{p}})$$

$$+\frac{3}{2}\left\{\frac{\partial^2 t_\eta}{(\partial \mathbf{p}')^{<2>}}\otimes\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<2>}\right\}\sum^{(3)}\{\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}-n^{-1/2}3\alpha_{\eta 1}^{(t)}+O(n^{-3/2})$$

$$=\frac{n^{-1/2}}{p_{12}^{3/2}}(-1,1)^{<3>}n^2\kappa_3(\hat{\mathbf{p}})$$

$$+n^{-1/2}3\frac{n^{-1/2}\partial t_\eta}{\partial \mathbf{p}'}n\,\mathrm{cov}(\hat{\mathbf{p}})\frac{n^{-1/2}\partial^2 t_\eta}{\partial \mathbf{p}\,\partial \mathbf{p}'}n\,\mathrm{cov}(\hat{\mathbf{p}})\frac{n^{-1/2}\partial t_\eta}{\partial \mathbf{p}}+O(n^{-3/2})$$

$$=\frac{n^{-1/2}}{p_{12}^{3/2}}(-1,1,1,-1,1,-1,-1,1)\{(p,\mathbf{0}_{(6)}',p)-p^2(3,\mathbf{1}_{(6)}',3)+2p^3\mathbf{1}_{(8)}'\}'$$

$$+\frac{n^{-1/2}}{p_{12}^{5/2}}3(-1,1)\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}p-p^2 & -p^2\\ -p^2 & p-p^2\end{pmatrix}\begin{pmatrix}-1\\ 1\end{pmatrix}+O(n^{-3/2})$$

$$=\frac{n^{-1/2}}{p_{12}^{5/2}}3(-p,p)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}-p\\ p\end{pmatrix}+O(n^{-3/2})=O(n^{-3/2})\ \ (\alpha_{\eta 3}^{(t)}=0),$$

$$\kappa_4(t_\eta)=\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<4>}\kappa_4(\hat{\mathbf{p}})$$

$$+2\left\{\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<3>}\otimes\frac{\partial^2 t_\eta}{(\partial \mathbf{p}')^{<2>}}\right\}\sum^{(10)}\{\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\otimes\kappa_3(\hat{\mathbf{p}})\}$$

$$+\left\{\frac{3}{2}\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<2>}\otimes\left(\frac{\partial^2 t_\eta}{(\partial \mathbf{p}')^{<2>}}\right)^{<2>}+\frac{2}{3}\left(\frac{\partial t_\eta}{\partial \mathbf{p}'}\right)^{<3>}\otimes\frac{\partial^3 t_\eta}{(\partial \mathbf{p}')^{<3>}}\right\}$$

$$\times\sum^{(15)}\{\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}-n^{-1}6\alpha_{\eta\Delta 2}^{(t)}+O(n^{-2}),$$

where the first term is

$$\frac{n^{-1}}{p_{12}^{2}}(1,-1,-1,1,-1,1,1,-1,-1,1,1,-1,1,-1,-1,1)$$

$$\times\{(p,\mathbf{0}_{(14)}{}',p)-p^{2}(4,1,1,0,1,0,0,1,1,0,0,1,0,1,1,4)$$

$$-p^{2}(3,0,0,1,0,1,1,0,0,1,1,0,1,0,0,3)$$

$$+2p^{3}(6,3,3,2,3,2,2,3,3,2,2,3,2,3,3,6)\}'$$

$$=\frac{n^{-1}}{p_{12}^{2}}2\{p-p^{2}(4-1-1-1-1)-p^{2}(3+1+1+1)$$

$$+2p^{3}(6-3-3+2-3+2+2-3)\}$$

$$=\frac{n^{-1}}{p_{12}^{2}}2(p-6p^{2})=\frac{n^{-1}}{2p}(1-6p),$$

the second term is

$$2\frac{n^{-1}}{p_{12}^{3}}\{(-1,1,1,-1,1,-1,-1,1)\otimes(1,0,0,-1)\}n^{3}\sum^{(10)}\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\otimes\kappa_{3}(\hat{\mathbf{p}})\,,$$

the third term is

$$n^{-1}\left\{\frac{3}{2}\frac{1}{p_{12}^{4}}(1,-1,-1,1)\otimes(1,0,0,-1)^{<2>}\right.$$

$$\left.+\frac{1}{p_{12}^{4}}\frac{1}{2}(-1,1,1,-1,1,-1,-1,1)\otimes(-3,-1,-1,1,-1,1,1,3)\right\}$$

$$\times n^{3}\sum^{(15)}\{\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\}^{3>}.$$

Consequently,

$$\kappa_4(t_\eta) = n^{-1}\left[\frac{1-6p}{2p} + \frac{1}{4p^3}\{(-1,1,1,-1,1,-1,-1,1)\otimes(1,0,0,-1)\}\right.$$

$$\times n^3 \sum^{(10)} \mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\otimes\kappa_3(\hat{\mathbf{p}})$$

$$+\frac{1}{32p^4}\{3(1,-1,-1,1)\otimes(1,0,0,-1)^{<2>}$$

$$+(-1,1,1,-1,1,-1,-1,1)\otimes(-3,-1,-1,1,-1,1,1,3)\}$$

$$\left.\times n^3 \sum^{(15)} \{\mathrm{vec}\,\mathrm{cov}(\hat{\mathbf{p}})\}^{<3>}\right] + O(n^{-2})$$

$$= n^{-1}\alpha_{\eta 4}^{(t)} + O(n^{-2}).$$

$n\,\mathrm{acov}(n^{-1}\hat{\eta}, \hat{\alpha}_{\eta 1}^{(t)}) = 0$ (note that $\alpha_{\eta 1}^{(t)} = 0$),

$n\,\mathrm{acov}(n^{-1}\hat{\eta}, \hat{\alpha}_{\eta 3}^{(t)}) = 0$ (note that $\alpha_{\eta 3}^{(t)} = 0$),

5.4 Asymptotic cumulants $\hat{\theta}_{\mathrm{W}}$ and $\hat{\eta}_{\mathrm{W}}$ by the weighted score method

$$\hat{\theta}_{\mathrm{W}} = \frac{\hat{\theta} + kn^{-1}}{1+3kn^{-1}} = \hat{\theta}\left(1 - \frac{3kn^{-1}}{1+3kn^{-1}}\right) + \frac{kn^{-1}}{1+3kn^{-1}}$$

$$= \hat{\theta} - n^{-1}3k\hat{\theta} + n^{-1}k + O_p(n^{-2})$$

$$= \hat{\theta} + n^{-1}k(1-3\hat{\theta}) + O_p(n^{-2}),$$

$$\kappa_1(\hat{\theta}_{\mathrm{W}} - \theta_0) = n^{-1}\alpha_{\theta 1} + n^{-1}k(1-3\theta_0) + O(n^{-2})$$

$$= n^{-1}k(1-3\theta_0) + O(n^{-2}) = n^{-1}k\left(1 - \frac{3}{2}\theta_{12}\right) + O(n^{-2})$$

$$\equiv n^{-1}\alpha_{\mathrm{W}\theta 1} + O(n^{-2})\ (\alpha_{\theta 1} = 0),$$

$$\kappa_2(\hat{\theta}_{\mathrm{W}}) = n^{-1}\alpha_{\theta 2} + n^{-2}(\alpha_{\theta\Delta 2} - 6k\alpha_{\theta 2}) + O(n^{-3})$$

$$\equiv n^{-1}\alpha_{\theta 2} + n^{-2}\alpha_{\mathrm{W}\theta\Delta 2} + O(n^{-3})\quad (\alpha_{\mathrm{W}\theta 2} = \alpha_{\theta 2}, \alpha_{\mathrm{W}\theta\Delta 2} \le \alpha_{\theta\Delta 2}),$$

$$n^{-1}\hat{\eta}_{\mathrm{W}} = \frac{(1+3kn^{-1})(\hat{p}_2 - \hat{p}_1)}{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}} = \frac{(1+3kn^{-1})\dfrac{\hat{p}_2 - \hat{p}_1}{\hat{p}_1 + \hat{p}_2}}{1 + \dfrac{2kn^{-1}}{\hat{p}_1 + \hat{p}_2}} = \frac{(1+3kn^{-1})n^{-1}\hat{\eta}}{1 + \dfrac{2kn^{-1}}{\hat{p}_1 + \hat{p}_2}}$$

$$= n^{-1}\hat{\eta}\left\{1 + n^{-1}k\left(3 - \frac{2}{\hat{p}_1 + \hat{p}_2}\right)\right\} + O_p(n^{-3/2}),$$

$$\kappa_1(n^{-1}\hat{\eta}_{\mathrm{W}} - n^{-1}\eta_0) = O(n^{-2}) \ \ (\alpha_{\mathrm{W}\eta 1} = 0),$$

where $\eta_0 = 0$ and $\alpha_{\eta 1} = 0$ are used,

$$\kappa_2(n^{-1}\hat{\eta}_{\mathrm{W}}) = n^{-1}\alpha_{\eta 2} + n^{-2}\left\{\alpha_{\eta\Delta 2} + 2k\left(3 - \frac{2}{\theta_{12}}\right)\alpha_{\eta 2}\right\} + O(n^{-3})$$

$$= n^{-1}\alpha_{\eta 2} + n^{-2}\left\{\alpha_{\eta\Delta 2} + 2k\left(3 - \frac{2}{\theta_{12}}\right)\frac{1}{\theta_{12}}\right\} + O(n^{-3}),$$

where $\alpha_{\eta 2} = 1/\theta_{12}$.

5.5 Asymptotic cumulants $t_{\mathrm{W}\theta}$ and $t_{\mathrm{W}\eta}$ by the weighted score method

$$t_{\mathrm{W}\theta} \equiv \frac{n^{1/2}(\hat{\theta}_{\mathrm{W}} - \theta_0)}{\{\hat{\theta}_{\mathrm{W}12}(1-\hat{\theta}_{\mathrm{W}12})/4\}^{1/2}}.$$

Since $\hat{\theta}_{\mathrm{W}} = \dfrac{\hat{\theta} + kn^{-1}}{1+3kn^{-1}} = \dfrac{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}{2(1+3kn^{-1})}$, we have

$\hat{\theta}_{\mathrm{W}12} = 2\hat{\theta}_{\mathrm{W}} = \dfrac{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}{1+3kn^{-1}}$ and $t_{\mathrm{W}\theta} = \dfrac{n^{1/2}(\hat{\theta}_{\mathrm{W}12} - \theta_{12})}{\{\hat{\theta}_{\mathrm{W}12}(1-\hat{\theta}_{\mathrm{W}12})\}^{1/2}}$, where

$$\hat{\theta}_{\mathrm{W}12} = \hat{\theta}_{12}\left(1 - \frac{3kn^{-1}}{1+3kn^{-1}}\right) + \frac{2kn^{-1}}{1+3kn^{-1}} = \hat{\theta}_{12} + n^{-1}k(2-3\hat{\theta}_{12}) + O_p(n^{-2}),$$

$$\{\hat{\theta}_{\mathrm{W}12}(1-\hat{\theta}_{\mathrm{W}12})\}^{-1/2} = \{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{-1/2} - \frac{1-2\hat{\theta}_{12}}{2\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{3/2}}(\hat{\theta}_{\mathrm{W}12} - \hat{\theta}_{12})$$

$$+ O_p(n^{-2})$$

$$= \{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{-1/2} - n^{-1}\frac{k(1-2\hat{\theta}_{12})(2-3\hat{\theta}_{12})}{2\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{3/2}} + O_p(n^{-2}).$$

Consequently,

$$t_{\mathrm{W}\theta} = n^{1/2}\left[\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{-1/2} - n^{-1}\frac{k(1-2\hat{\theta}_{12})(2-3\hat{\theta}_{12})}{2\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{3/2}}\right]$$

$$\times\{\hat{\theta}_{12} - \theta_{12} + n^{-1}k(2-3\hat{\theta}_{12})\} + O_p(n^{-3/2})$$

$$= t_\theta + n^{-1/2}k\left[\frac{2-3\hat{\theta}_{12}}{\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{1/2}} - \frac{(1-2\hat{\theta}_{12})(2-3\hat{\theta}_{12})(\hat{\theta}_{12}-\theta_{12})}{2\{\hat{\theta}_{12}(1-\hat{\theta}_{12})\}^{3/2}}\right] + O_p(n^{-3/2}).$$

The asymptotic cumulants different from those by ML given earlier are

$$\kappa_1(t_{\mathrm{W}\theta}) = n^{-1/2}\left[\alpha_{\theta 1}^{(t)} + \frac{k(2-3\theta_{12})}{\{\theta_{12}(1-\theta_{12})\}^{1/2}}\right] + O(n^{-3/2})$$

$$\equiv n^{-1/2}\alpha_{\mathrm{W}\theta 1}^{(t)} + O(n^{-3/2}),$$

$$\kappa_2(t_{\mathrm{W}\theta}) = 1$$

$$+n^{-1}\left[\alpha_{\theta\Delta 2}^{(t)} + 2kn\,\mathrm{var}(\hat{\theta}_{12})\left\{\frac{-3}{\theta_{12}(1-\theta_{12})} - \frac{(1-2\theta_{12})(2-3\theta_{12})}{\{\theta_{12}(1-\theta_{12})\}^2}\right\}\right] + O(n^{-2})$$

$$= 1 + n^{-1}\left[\alpha_{\theta\Delta 2}^{(t)} - 2k\left\{3 + \frac{(1-2\theta_{12})(2-3\theta_{12})}{\theta_{12}(1-\theta_{12})}\right\}\right] + O(n^{-2})$$

$$\equiv 1 + n^{-1}\alpha_{\mathrm{W}\theta\Delta 2}^{(t)} + O(n^{-2}).$$

Note that $\hat{\theta}_{\mathrm{W}12} = 2\hat{p}_{\mathrm{W}} = \hat{p}_{\mathrm{W}1} + \hat{p}_{\mathrm{W}2} = 2\hat{\theta}_{\mathrm{W}}$,

$$t_{W\eta} \equiv n^{1/2}(-\hat{i}_{W}^{\eta\eta})^{-1/2}(n^{-1}\hat{\eta}_{W}),$$

$$-\hat{i}_{W}^{\eta\eta} = \frac{1}{2\hat{p}_{W}} = \frac{1+3kn^{-1}}{\hat{p}_1+\hat{p}_2+2kn^{-1}}$$

$$= \frac{1}{\hat{p}_1+\hat{p}_2}(1+3kn^{-1}) - \frac{2kn^{-1}}{(\hat{p}_1+\hat{p}_2)^2} + O_p(n^{-2})$$

$$= \frac{1}{\hat{p}_1+\hat{p}_2} + n^{-1}k\left\{\frac{3}{\hat{p}_1+\hat{p}_2} - \frac{2}{(\hat{p}_1+\hat{p}_2)^2}\right\} + O_p(n^{-2})$$

$$= -\hat{i}^{\eta\eta} + n^{-1}k\left(\frac{3}{\hat{\theta}_{12}} - \frac{2}{\hat{\theta}_{12}^2}\right) + O_p(n^{-2}) \qquad \left(-\hat{i}^{\eta\eta} = \frac{1}{\hat{p}_1+\hat{p}_2} = \frac{1}{\hat{\theta}_{12}}\right),$$

$$(-\hat{i}_{W}^{\eta\eta})^{-1/2} = (-\hat{i}^{\eta\eta})^{-1/2} - \frac{1}{2}(-\hat{i}^{\eta\eta})^{-3/2}(-\hat{i}_{W}^{\eta\eta} + \hat{i}^{\eta\eta}) + O_p(n^{-2})$$

$$= \hat{\theta}_{12}^{1/2} - \frac{n^{-1}k}{2}\hat{\theta}_{12}^{3/2}\left(\frac{3}{\hat{\theta}_{12}} - \frac{2}{\hat{\theta}_{12}^2}\right) + O_p(n^{-2})$$

$$= \hat{\theta}_{12}^{1/2} - \frac{n^{-1}k}{2}\left(3\hat{\theta}_{12}^{1/2} - \frac{2}{\hat{\theta}_{12}^{1/2}}\right) + O_p(n^{-2}),$$

$$n^{-1}\hat{\eta}_{W} = \frac{(1+3kn^{-1})(\hat{p}_2-\hat{p}_1)}{\hat{p}_1+\hat{p}_2+2kn^{-1}}$$

$$= \frac{\hat{p}_2-\hat{p}_1}{\hat{p}_1+\hat{p}_2}(1+3kn^{-1}) - 2kn^{-1}\frac{\hat{p}_2-\hat{p}_1}{(\hat{p}_1+\hat{p}_2)^2} + O_p(n^{-2})$$

$$= \frac{\hat{p}_2-\hat{p}_1}{\hat{p}_1+\hat{p}_2} + n^{-1}k(\hat{p}_2-\hat{p}_1)\left\{\frac{3}{\hat{p}_1+\hat{p}_2} - \frac{2}{(\hat{p}_1+\hat{p}_2)^2}\right\} + O_p(n^{-2}),$$

$$= \frac{\hat{p}_2-\hat{p}_1}{\hat{\theta}_{12}} + n^{-1}k(\hat{p}_2-\hat{p}_1)\left(\frac{3}{\hat{\theta}_{12}} - \frac{2}{\hat{\theta}_{12}^2}\right) + O_p(n^{-2}).$$

Then,

$$t_{\mathrm{W}\eta} = n^{1/2}(-\hat{i}_{\mathrm{W}}^{\eta\eta})^{-1/2}(n^{-1}\hat{\eta}_{\mathrm{W}})$$

$$= n^{1/2}\left\{\hat{\theta}_{12}^{1/2} - \frac{n^{-1}k}{2}\left(3\hat{\theta}_{12}^{1/2} - \frac{2}{\hat{\theta}_{12}^{1/2}}\right)\right\}$$

$$\times\left\{\frac{\hat{p}_2 - \hat{p}_1}{\hat{\theta}_{12}} + n^{-1}k(\hat{p}_2 - \hat{p}_1)\left(\frac{3}{\hat{\theta}_{12}} - \frac{2}{\hat{\theta}_{12}^2}\right)\right\} + O_p(n^{-3/2})$$

$$= t_\eta + n^{-1/2}k(\hat{p}_2 - \hat{p}_1)\left\{-\frac{1}{2}\left(\frac{3}{\hat{\theta}_{12}^{1/2}} - \frac{2}{\hat{\theta}_{12}^{3/2}}\right) + \frac{3}{\hat{\theta}_{12}^{1/2}} - \frac{2}{\hat{\theta}_{12}^{3/2}}\right\} + O_p(n^{-3/2})$$

$$= t_\eta + n^{-1/2}k\,\frac{\hat{p}_2 - \hat{p}_1}{2}\left(\frac{3}{\hat{\theta}_{12}^{1/2}} - \frac{2}{\hat{\theta}_{12}^{3/2}}\right) + O_p(n^{-3/2}).$$

The asymptotic cumulants different from those by ML given earlier are

$\kappa_1(t_{\mathrm{W}\eta}) = O_p(n^{-3/2})$ $(\alpha_{\mathrm{W}\eta 1}^{(t)} = 0)$, where $\alpha_{\eta 1}^{(t)} = 0$ is used,

$$\kappa_2(t_{\mathrm{W}\eta}) = 1 + n^{-1}\left\{\alpha_{\eta\Delta 2}^{(t)} + 2kn\,\mathrm{var}(\hat{p}_2 - \hat{p}_1)\frac{1}{2}\left(\frac{3}{\theta_{12}} - \frac{2}{\theta_{12}^2}\right)\right\} + O(n^{-2})$$

$$= 1 + n^{-1}k\left(3 - \frac{3}{\theta_{12}}\right) + O(n^{-2}) \equiv 1 + n^{-1}\alpha_{\mathrm{W}\eta\Delta 2}^{(t)} + O(n^{-2}),$$

where $\alpha_{\eta\Delta 2}^{(t)} = 0$ and $n\,\mathrm{var}(\hat{p}_2 - \hat{p}_1) = p_1 + p_2 = \theta_{12}$ are used.

$$n\,\mathrm{acov}(\hat{\theta}_{\mathrm{W}12}, \hat{\alpha}_{\mathrm{W}\theta 1}^{(t)})$$

$$= n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 1}^{(t)}) + kn\,\mathrm{var}(\hat{\theta}_{12})\left[\frac{-3}{\{\theta_{12}(1-\theta_{12})\}^{1/2}} - \frac{(1-2\theta_{12})(2-3\theta_{12})}{2\{\theta_{12}(1-\theta_{12})\}^{3/2}}\right]$$

$$+ O(n^{-1})$$

$$= n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 1}^{(t)}) - k\left[3\{\theta_{12}(1-\theta_{12})\}^{1/2} + \frac{(1-2\theta_{12})(2-3\theta_{12})}{2\{\theta_{12}(1-\theta_{12})\}^{1/2}}\right] + O(n^{-1}),$$

$n\,\mathrm{acov}(\hat{\theta}_{\mathrm{W}12}, \hat{\alpha}_{\mathrm{W}\theta 3}^{(t)}) = n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 3}^{(t)})$ $(\alpha_{\mathrm{W}\theta 3}^{(t)} = \alpha_{\theta 3}^{(t)})$,

where for $n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 1}^{(t)})$ and $n\,\mathrm{acov}(\hat{\theta}_{12}, \hat{\alpha}_{\theta 3}^{(t)})$, see Subsection 5.3,

$$n\,\mathrm{acov}(n^{-1}\hat{\eta}_{\mathrm{W}},\hat{\alpha}^{(t)}_{\mathrm{W}\eta1}) = n\,\mathrm{acov}(n^{-1}\hat{\eta},\hat{\alpha}^{(t)}_{\eta1}) = 0,$$

$$n\,\mathrm{acov}(n^{-1}\hat{\eta}_{\mathrm{W}},\hat{\alpha}^{(t)}_{\mathrm{W}\eta3}) = n\,\mathrm{acov}(n^{-1}\hat{\eta},\hat{\alpha}^{(t)}_{\eta3}) = 0,$$

where $\alpha^{(t)}_{\mathrm{W}\eta1} = \alpha^{(t)}_{\eta1} = 0$ and $\alpha^{(t)}_{\mathrm{W}\eta3} = \alpha^{(t)}_{\eta3} = 0$ are used.

6. Asymptotic expansions for the estimators in logistic regression

The parameters in logistic regression are canonical parameters in the exponential family yielding $-\mathbf{\Lambda}_0^* = \mathbf{I}_0^*$ and $\partial^{(i+1)}\overline{l} / \partial\mathbf{\theta}_0(\partial\mathbf{\theta}_0')^{<i>} = \mathrm{E_T}\{\partial^{(i+1)}\overline{l} / \partial\mathbf{\theta}_0(\partial\mathbf{\theta}_0')^{<i>}\} \equiv -\mathbf{I}_0^{(i+1)}$ $(i = 2,3,\ldots)$.

The restriction used is linear with respect to parameters i.e., $h_0 = \beta_1 - \beta_2 = 0$, which gives $\mathbf{H}_0 = (1,\ -1)'$. Then, (A1.6) under correct model misspecification becomes

$$\begin{pmatrix}\hat{\mathbf{\theta}}_{\mathrm{W}} - \mathbf{\theta}_0 \\ n^{-1}\hat{\mathbf{\eta}}_{\mathrm{W}}\end{pmatrix} = -\left(\mathbf{\Lambda}_0^{(\bullet 1)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right)_{O_p(n^{-1/2})} - (n^{-1}\mathbf{\Lambda}_0^{(\bullet 1)}\mathbf{q}_0^*)_{O(n^{-1})}$$

$$-\left\{\frac{1}{2}\mathbf{\Lambda}_0^{*-1}\begin{pmatrix}\mathrm{E_T}(\mathbf{J}_0^{(3)}) \\ \mathbf{O}\end{pmatrix}\left(\mathbf{\Lambda}_0^{(11)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\right\}_{O_p(n^{-1})}$$

$$+\underset{(\mathrm{A})}{\Bigg[}\mathbf{\Lambda}_0^{*-1}\begin{pmatrix}\mathrm{E_T}(\mathbf{J}_0^{(3)}) \\ \mathbf{O}\end{pmatrix}\underset{(\mathrm{B})}{\Bigg[}\left(\mathbf{\Lambda}_0^{(11)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right)\otimes\Bigg\{-n^{-1}\mathbf{\Lambda}_0^{(11)}\mathbf{q}_0^* +$$

$$-\frac{1}{2}(\mathbf{\Lambda}_0^{(11)}\mathbf{\Lambda}_0^{(12)})\begin{pmatrix}\mathrm{E_T}(\mathbf{J}_0^{(3)}) \\ \mathbf{O}\end{pmatrix}\left(\mathbf{\Lambda}_0^{(11)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right)^{<2>}\Bigg\}\Bigg]_{(\mathrm{B})}\Bigg]_{(\mathrm{A})\,O_p(n^{-3/2})}$$

$$+\left\{\mathbf{\Lambda}_0^{(\bullet 1)}n^{-1}\frac{\partial\mathbf{q}_0^*}{\partial\mathbf{\theta}_0'}\mathbf{\Lambda}_0^{(11)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right\}_{O_p(n^{-3/2})}$$

$$+\left\{\frac{1}{6}\mathbf{\Lambda}_0^{*-1}\begin{pmatrix}\mathrm{E_T}(\mathbf{J}_0^{(4)}) \\ \mathbf{O}\end{pmatrix}\left(\mathbf{\Lambda}_0^{(11)}\frac{\partial\overline{l}}{\partial\mathbf{\theta}_0}\right)^{<3>}\right\}_{O_p(n^{-3/2})} + O_p(n^{-2})$$

$$
=\left(\mathbf{I}_0^{(\bullet 1)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)_{O_p(n^{-1/2})} + (n^{-1}\mathbf{I}_0^{(\bullet 1)}\mathbf{q}_0^*)_{O(n^{-1})} - \left\{\frac{1}{2}\mathbf{I}_0^{(\bullet 1)}\mathbf{I}_0^{(3)}\left(\mathbf{I}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<2>}\right\}_{O_p(n^{-1})}
$$

$$
-\Big[_{(\mathrm{A})}\ \mathbf{I}_0^{(\bullet 1)}\mathbf{I}_0^{(3)}
$$

$$
\times\left[\left(\mathbf{I}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\otimes\left\{n^{-1}\mathbf{I}_0^{(11)}\mathbf{q}_0^* + \frac{1}{2}\mathbf{I}_0^{(11)}\mathbf{I}_0^{(3)}\left(\mathbf{I}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<2>}\right\}\right]\Big]_{(\mathrm{A})O_p(n^{-3/2})}
$$

$$
+\left\{\mathbf{I}_0^{(\bullet 1)}n^{-1}\frac{\partial \mathbf{q}_0^*}{\partial \boldsymbol{\theta}_0{}'}\mathbf{I}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right\}_{O_p(n^{-3/2})}
$$

$$
-\left\{\frac{1}{6}\mathbf{I}_0^{(\bullet 1)}\mathbf{I}_0^{(4)}\left(\mathbf{I}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)^{<3>}\right\}_{O_p(n^{-3/2})} + O_p(n^{-2})
$$

$$
\equiv \sum_{i=1}^{3}\boldsymbol{\Lambda}_{\mathrm{W}}^{(i)}\mathbf{l}_0^{(i)} + n^{-1}\mathbf{I}_0^{(\bullet 1)}\mathbf{q}_0^* + O_p(n^{-2})
$$

$$
(\boldsymbol{\Lambda}_{\mathrm{W}}^{(i)} = O(1),\ \mathbf{l}_0^{(i)} = O_p(n^{-i/2}),\ \ i = 1,\ 2,\ 3),
$$

where

$$
\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} = \frac{\partial \bar{l}}{\partial \boldsymbol{\beta}_0} = n^{-1}\sum_{i=1}^{n}\left(\frac{U_i}{P_i} - \frac{1-U_i}{Q_i}\right)\frac{\partial P_i}{\partial \boldsymbol{\beta}_0}
$$

$$
= n^{-1}\sum_{i=1}^{n}\frac{U_i - P_i}{P_i Q_i}\frac{\partial P_i}{\partial \boldsymbol{\beta}_0} = n^{-1}\sum_{i=1}^{n}(U_i - P_i)\mathbf{x}_i,
$$

$$\mathbf{I}_0 = -\frac{\partial^2 \overline{l}}{\partial \boldsymbol{\beta}_0 \partial \boldsymbol{\beta}_0'} = n^{-1}\sum_{i=1}^{n} P_i Q_i \mathbf{x}_i \mathbf{x}_i',$$

$$(\mathbf{I}_0^{(3)})_{abc} = -\frac{\partial^3 \overline{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \beta_{0c}} n^{-1}\sum_{i=1}^{n}(1-2P_i)P_i Q_i x_{ia} x_{ib} x_{ic}$$

$$(\mathbf{I}_0^{(4)})_{abcd} = -\frac{\partial^4 \overline{l}}{\partial \beta_{0a} \partial \beta_{0b} \partial \beta_{0c} \partial \beta_{0d}} n^{-1}\sum_{i=1}^{n}(1-6P_i+6P_i^2)P_i Q_i x_{ia} x_{ib} x_{ic} x_{id},$$

$$x_{ia} \equiv (\mathbf{x}_i)_a \quad (i=1,...,n;\ a,\ b,\ c,\ d=1,...,q).$$

7. Some properties of the estimators of Lagrange multipliers

From the first-order condition of the estimators of parameters (see (A1.1)), we have

$$\frac{\partial \overline{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{H}}_{\mathrm{W}}\hat{\boldsymbol{\eta}}_{\mathrm{W}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*} = \mathbf{0}. \quad \text{(S7.1)}$$

Define $\mathbf{C}^*$ as a $q \times r$ fixed matrix of order $O(1)$ with $\mathbf{C}^*{}'\hat{\mathbf{H}}_{\mathrm{W}}$ being non-singular. From (S7.1)

$$n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}} = -(\mathbf{C}^*{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{C}^*{}'\left(\frac{\partial \overline{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*}\right) \quad \text{(S7.2)}$$

follows. We have

Lemma S6. *Assume that* $\mathbf{C}^*{}'\hat{\mathbf{H}}_{\mathrm{W}}$ *is non-singular, then* $n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}$ *in (S7.2) does not depend on* $\mathbf{C}^*$.

Proof. Let $\mathbf{E}_{ji}$ be a $r \times q$ matrix whose (*j*, *i*)th element is 1 with the remaining ones being 0. Then,

$$\frac{\partial n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}}{\partial c_{ij}^{*}} = (\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{E}_{ji}\hat{\mathbf{H}}_{\mathrm{W}}(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{C}^{*}{}'\left(\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*}\right)$$

$$- (\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{E}_{ji}\left(\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*}\right) \tag{S7.3}$$

$$= -(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{E}_{ji}\left(\hat{\mathbf{H}}_{\mathrm{W}}n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}} + \frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*}\right) = \mathbf{0},$$

which gives the required result. Q.E.D.

The asymptotic bias of $n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}$ up to order $O(n^{-1})$ is given by (5.14) or (5.17) as

$$n^{-1}\boldsymbol{\alpha}_{\eta \mathrm{W1}} = n^{-1}(\boldsymbol{\alpha}_{\eta \mathrm{ML1}} - \boldsymbol{\Lambda}_{0}^{(21)}\mathbf{q}_{0}^{*}), \tag{S7.4}$$

which is also given from (S7.3) as follows. Noting that

$$n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}} = -(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{ML}})^{-1}\mathbf{C}^{*}{}'\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}},$$

$$\hat{\mathbf{H}}_{\mathrm{W}} = \hat{\mathbf{H}}_{\mathrm{ML}} + O_p(n^{-1}) \ \text{(use (4.2))},$$

$$\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} = \frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}} + \frac{\partial^{2} \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{W}} - \hat{\boldsymbol{\theta}}_{\mathrm{ML}}) + O_p(n^{-2})$$

$$= \frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}} + O_p(n^{-1}), \tag{S7.5}$$

$$\hat{\mathbf{q}}_{\mathrm{W}}^{*} = \mathbf{q}_{0}^{*} + O_p(n^{-1/2}),$$

we have

$$\mathrm{E_T}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}) = -\mathrm{E_T}\left\{(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{W}})^{-1}\mathbf{C}^{*}{}'\left(\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{W}}} + n^{-1}\hat{\mathbf{q}}_{\mathrm{W}}^{*}\right)\right\}$$

$$= -\mathrm{E_T}\left\{(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{ML}})^{-1}\mathbf{C}^{*}{}'\left(\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}} + n^{-1}\mathbf{q}_{0}^{*}\right)\right\} + O(n^{-2}) \qquad \text{(S7.6)}$$

$$= n^{-1}\{\boldsymbol{\alpha}_{\eta \mathrm{ML1}} - (\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}'\mathbf{q}_0^{*}\} + O(n^{-2}).$$

Let $\mathbf{C}^{*} = \boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0 = O(1)$, then (S7.6) becomes

$$n^{-1}\{\boldsymbol{\alpha}_{\eta \mathrm{ML1}} - (\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{q}_0^{*}\} + O(n^{-2})$$

$$= n^{-1}(\boldsymbol{\alpha}_{\eta \mathrm{ML1}} - \boldsymbol{\Lambda}_0^{(21)}\mathbf{q}_0^{*}) + O(n^{-2}) \qquad \text{(S7.7)}$$

$$= n^{-1}\boldsymbol{\alpha}_{\eta \mathrm{W1}} + O(n^{-2}),$$

which shows (S7.4). In (S7.7), since

$$n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}} = -(\mathbf{C}^{*}{}'\hat{\mathbf{H}}_{\mathrm{ML}})^{-1}\mathbf{C}^{*}{}'\frac{\partial \bar{l}}{\partial \hat{\boldsymbol{\theta}}_{\mathrm{ML}}}$$

$$= \left\{-(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1} + (\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}'\sum_{i=1}^{q}\frac{\partial \mathbf{H}_0}{\partial \theta_{0i}}(\hat{\boldsymbol{\theta}}_{\mathrm{ML}} - \boldsymbol{\theta}_0)_i(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\right\}$$

$$\times \mathbf{C}^{*}{}'\left\{\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0} + \frac{\partial^2 \bar{l}}{\partial \boldsymbol{\theta}_0 \partial \boldsymbol{\theta}_0{}'}(\hat{\boldsymbol{\theta}}_{\mathrm{ML}} - \boldsymbol{\theta}_0) + \frac{1}{2}\frac{\partial^3 \bar{l}}{\partial \boldsymbol{\theta}_0 (\partial \boldsymbol{\theta}_0{}')^{<2>}}(\hat{\boldsymbol{\theta}}_{\mathrm{ML}} - \boldsymbol{\theta}_0)^{<2>}\right\}$$

$$+ O_p(n^{-3/2}), \qquad \text{(S7.8)}$$

we have

$$
\begin{aligned}
&\mathrm{E_T}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}}) = -(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}' \\
&\times\left\{\boldsymbol{\Lambda}_0\boldsymbol{\alpha}_{\theta\mathrm{ML1}} + \frac{1}{2}\mathrm{E_T}(\mathbf{J}_0^{(3)})\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}) - n\mathrm{E_T}\left(\mathbf{M}\boldsymbol{\Lambda}_0^{(11)}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\} \\
&+(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}'\sum_{i=1}^{q}\frac{\partial\mathbf{H}_0}{\partial\theta_{0i}}(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}'(-\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)} + \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)})_{\bullet i} \\
&+O_p(n^{-2}) \\
&= n^{-1}\boldsymbol{\alpha}_{\eta\mathrm{ML1}} + O_p(n^{-2}),
\end{aligned}
\tag{S7.9}
$$

which is an alternative expression of $\boldsymbol{\alpha}_{\eta\mathrm{ML1}}$ given by (5.17). The algebraic equivalence is shown as follows. From (5.17),

$$
\begin{aligned}
\boldsymbol{\alpha}_{\eta\mathrm{ML1}} &= \boldsymbol{\Lambda}_0^{(21)} n\mathrm{E_T}\left(\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0{}'}\otimes\mathbf{M}\right)\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}) \\
&\quad - \boldsymbol{\Lambda}_0^{(21)}\sum_{a=1}^{r}\frac{\partial(\mathbf{H}_0)_{\bullet a}}{\partial\boldsymbol{\theta}_0{}'}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(12)})_{\bullet a} \\
&\quad - \frac{n^{-1}}{2}(\boldsymbol{\Lambda}_0^{(21)}\boldsymbol{\Lambda}_0^{(22)})\begin{pmatrix}\mathrm{E_T}(\mathbf{J}_0^{(3)}) \\ \dfrac{\partial^2\mathbf{h}_0}{(\partial\boldsymbol{\theta}_0{}')^{<2>}}\end{pmatrix}\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}).
\end{aligned}
\tag{S7.10}
$$

On the other hand, let $\mathbf{C}^{*} = \boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0$. Then, using $\mathbf{H}_0\boldsymbol{\Lambda}_0^{(11)} = \mathbf{O}$, we have from (S7.9),

$$
\begin{aligned}
\boldsymbol{\alpha}_{\eta\mathrm{ML}1} &= -(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1} \\
&\quad\times\left\{\boldsymbol{\Lambda}_0\boldsymbol{\alpha}_{\theta\mathrm{ML}1} + \frac{1}{2}\mathrm{E_T}(\mathbf{J}_0^{(3)})\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}) - n\mathrm{E_T}\left(\mathbf{M}\boldsymbol{\Lambda}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\} \\
&\quad+(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\sum_{i=1}^{q}\frac{\partial \mathbf{H}_0}{\partial \theta_{0i}}(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1} \\
&\qquad\qquad\times(-\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)} + \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)})_{\bullet i} \\
&= -(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\left\{-\frac{1}{2}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{(12)}\frac{\partial^2 \mathbf{h}_0}{(\partial \boldsymbol{\theta}_0{}')^{<2>}}\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)})\right. \\
&\quad\left.+\frac{1}{2}\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathrm{E_T}(\mathbf{J}_0^{(3)})\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}) - \mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}n\mathrm{E_T}\left(\mathbf{M}\boldsymbol{\Lambda}_0^{(11)}\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0}\right)\right\} \\
&\quad-\boldsymbol{\Lambda}_0^{(21)}\sum_{i=1}^{q}\frac{\partial \mathbf{H}_0}{\partial \theta_{0i}}(\boldsymbol{\Lambda}_0^{(21)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)})_{\bullet i} \\
&= \left\{\frac{1}{2}(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}\frac{\partial^2 \mathbf{h}_0}{(\partial \boldsymbol{\theta}_0{}')^{<2>}} - \frac{1}{2}\boldsymbol{\Lambda}_0^{(21)}\mathrm{E_T}(\mathbf{J}_0^{(3)})\right\}\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}) \\
&\quad+\boldsymbol{\Lambda}_0^{(21)}n\mathrm{E_T}\left(\frac{\partial \bar{l}}{\partial \boldsymbol{\theta}_0{}'}\otimes\mathbf{M}\right)\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}) \\
&\quad-\boldsymbol{\Lambda}_0^{(21)}\sum_{a=1}^{r}\frac{\partial (\mathbf{H}_0)_{\bullet a}}{\partial \boldsymbol{\theta}_0{}'}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(12)})_{\bullet i},
\end{aligned}
\tag{S7.11}
$$

where $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}'\otimes\mathbf{A})\mathrm{vec}(\mathbf{B})$ is used. Noting $\boldsymbol{\Lambda}_0^{(22)} = -(\mathbf{H}_0{}'\boldsymbol{\Lambda}_0^{-1}\mathbf{H}_0)^{-1}$, we find that (S7.11) is algebraically equal to (S7.10).

When $\mathbf{h}_0$ is linear with respect to $\boldsymbol{\theta}_0$, $\hat{\mathbf{H}}_{\mathrm{ML}}$ becomes a fixed $\mathbf{H}_0$ in (S7.8), and (S7.9) is simplified as

$$
\begin{aligned}
&\mathrm{E}_{\mathrm{T}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}}) = -(\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}' \\
&\times\left\{\boldsymbol{\Lambda}_0\boldsymbol{\alpha}_{\theta\mathrm{ML1}} + \frac{1}{2}\mathrm{E}_{\mathrm{T}}(\mathbf{J}_0^{(3)})\mathrm{vec}(\boldsymbol{\Lambda}_0^{(11)}\boldsymbol{\Gamma}\boldsymbol{\Lambda}_0^{(11)}) - n\mathrm{E}_{\mathrm{T}}\left(\mathbf{M}\boldsymbol{\Lambda}_0^{(11)}\frac{\partial\bar{l}}{\partial\boldsymbol{\theta}_0}\right)\right\} \\
&+O_p(n^{-2}) \\
&= n^{-1}\boldsymbol{\alpha}_{\eta\mathrm{ML1}} + O_p(n^{-2}),
\end{aligned}
\tag{S7.12}
$$

Let $\mathbf{C}^{*} = \mathbf{I}_0^{-1}\mathbf{H}_0$. The asymptotic covariance matrix of $n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}})$ is given from (S7.8) with (S7.2) as

Theorem S4. *In the general case with* $\mathbf{h}_0$ *being possibly nonlinear with respect to* $\boldsymbol{\theta}_0$, *define* $\mathrm{acov}_{\mathrm{T}}(\cdot)$ *as the asymptotic covariance matrix of order* $O(n^{-1})$ *for the vector of the argument under possible model misspecification, then*

$$
\begin{aligned}
&n\,\mathrm{acov}_{\mathrm{T}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}) = n\,\mathrm{acov}_{\mathrm{T}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}}) \\
&= (\mathbf{C}^{*}{}'\mathbf{H}_0)^{-1}\mathbf{C}^{*}{}'\boldsymbol{\Gamma}\mathbf{C}^{*}(\mathbf{H}_0{}'\mathbf{C}^{*})^{-1} \\
&= (\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1}\mathbf{H}_0{}'\mathbf{I}_0^{-1}\boldsymbol{\Gamma}\mathbf{I}_0^{-1}\mathbf{H}_0(\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1}.
\end{aligned}
\tag{S7.13}
$$

Under correct model specification with $\boldsymbol{\Gamma} = \mathbf{I}_0$, (S7.13) becomes

$$
n\,\mathrm{acov}_{\mathrm{T}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{W}}) = n\,\mathrm{acov}_{\mathrm{T}}(n^{-1}\hat{\boldsymbol{\eta}}_{\mathrm{ML}}) = (\mathbf{H}_0{}'\mathbf{I}_0^{-1}\mathbf{H}_0)^{-1}, \tag{S7.14}
$$

which is known (see e.g., Ogasawara, 2016, Corollary 2).

8. Types of restrictions with examples in maximum likelihood estimation

Denote parameters in a statistical model by $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_q)'$ and restrictions by $\mathbf{h} = \mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}$, where **h** is an $r \times 1$ vector with its elements $h_i = h_i(\boldsymbol{\theta})(i = 1,...,r)$ being functions of $\boldsymbol{\theta}$. The $q \times r$ matrix $\mathbf{H} \equiv \partial\mathbf{h}'/\partial\boldsymbol{\theta}$ is also used. Let $L(\boldsymbol{\theta}\,|\,\mathbf{X})$ be the likelihood of $\boldsymbol{\theta}$ in a statistical model when n independent observations denoted generically by **X** are given. Let $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ with $\boldsymbol{\theta}^{(1)} \neq \boldsymbol{\theta}^{(2)}$ be in the neighborhood of $\boldsymbol{\theta}$. If $L(\boldsymbol{\theta}^{(1)}\,|\,\mathbf{X}) = L(\boldsymbol{\theta}^{(2)}\,|\,\mathbf{X})$, the statistical model is said to be unidentified. A typical model without model identification is that of exploratory factor analysis. Two rotated solutions e.g., varimax and promax, give the same value of the Wishart likelihood, where different rotation criteria to be optimized are used for identification (see e.g., Ogasawara, 2004). These cases without model identification typically give singular information matrices (Silvey, 1959, Section 6; Silvey, 1975, Subsection 4.7.5; Lee, 1979, Property 2B).

Assume that $L(\boldsymbol{\theta}^{(1)} \mid \mathbf{X}) \neq L(\boldsymbol{\theta}^{(2)} \mid \mathbf{X})$ for arbitrary $\boldsymbol{\theta}^{(1)} \neq \boldsymbol{\theta}^{(2)}$ in the neighborhood with the restriction(s) $\mathbf{h}(\boldsymbol{\theta}) = \mathbf{0}$ on $\boldsymbol{\theta}$. Let $\overline{r}$ be the minimum number of restrictions selected from $h_i = 0 \; (i = 1, ..., r)$ to satisfy the above inequality. Define the $\overline{r}$ restrictions as $\mathbf{h}^{(\mathrm{id})} = \mathbf{h}^{(\mathrm{id})}(\boldsymbol{\theta}) = (h_1^{(\mathrm{id})}(\boldsymbol{\theta}), ..., h_{\overline{r}}^{(\mathrm{id})}(\boldsymbol{\theta}))' = \mathbf{0}$.

Definition S1. *[1] The restriction(s) for model identification are defined as* $\mathbf{h}^{(\mathrm{id})} = \mathbf{0}$ *to remove the model unidentification in statistical models. [2] The restriction(s) for genuine constraint(s) are defined as those that impose added constraints on parameters, when the model can be specified as an identified statistical model without the genuine restriction(s). For the model with the restrictions* $\mathbf{h} = \mathbf{0}$ *including* $\mathbf{h}^{(\mathrm{id})} = \mathbf{0}$, *the* $r - \overline{r}$ *genuine restrictions are denoted by* $\mathbf{h}^{(\mathrm{gr})} = \mathbf{h}^{(\mathrm{gr})}(\boldsymbol{\theta}) = (h_{\overline{r}+1}^{(\mathrm{gr})}(\boldsymbol{\theta}), ..., h_r^{(\mathrm{gr})}(\boldsymbol{\theta}))' = \mathbf{0}$ with $\mathbf{h} = (\mathbf{h}^{(\mathrm{id})}{}', \mathbf{h}^{(\mathrm{gr})}{}')'$, *where the elements of* $\mathbf{h}^{(\mathrm{gr})}$ *are functionally independent of the elements of* $\mathbf{h}^{(\mathrm{id})}$. *When a model is identified without* $\mathbf{h}^{(\mathrm{id})} = \mathbf{0}$, $\mathbf{h} = \mathbf{h}^{(\mathrm{gr})}$.

Consider the example of testing the equality of the binomial proportions in two independent groups, which was used by Silvey (1959, pp. 405-407). Assume that the sample sizes for the two groups, denoted by n_1 and n_2 with $n \equiv n_1 + n_2$, are non-stochastic. The situation is summarized in the following table.

		1	0	Total
Group 1:	X_1	m_{11}	m_{12}	n_1
Group 2:	X_2	m_{21}	m_{22}	n_2
Total		m_1	m_2	n

In the table, X_1 and X_2 are dichotomous variables taking values of 1 and 0 for the two groups, respectively, while stochastic m_{11} and m_{12} are observed frequencies in Group 1 for the cases of 1 and 0 respectively, with

$m_{11} + m_{12} = n_1$. Stochastic m_{21} and m_{22} in Group 2 are similarly defined, which gives stochastic $m_1 \equiv m_{11} + m_{21}$ and $m_2 \equiv m_{12} + m_{22}$.

Next, parametrizations with equal proportions in this example are given in two ways to illustrate the definitions of the different types of restrictions on parameters. For illustration, we use the ML estimators (MLEs) rather than the WS estimators (WSEs) (the WSEs will be illustrated in the next section).

Example 1.1 (Silvey, 1959) Define $\theta_1 / (\theta_1 + \theta_2) \equiv \Pr(X_1 = 1)$, $\theta_2 / (\theta_1 + \theta_2) \equiv \Pr(X_1 = 0)$, $\theta_3 / (\theta_3 + \theta_4) \equiv \Pr(X_2 = 1)$ and $\theta_4 / (\theta_3 + \theta_4) \equiv \Pr(X_2 = 0)$. Three restrictions $\theta_1 + \theta_2 = 1$, $\theta_3 + \theta_4 = 1$ and $\theta_1 = \theta_3$ are summarized as $\mathbf{h} = (\theta_1 + \theta_2 - 1,\ \theta_3 + \theta_4 - 1, \theta_1 - \theta_3)' = \mathbf{0}$. Note that the unit value in $\theta_1 + \theta_2 = 1$ and $\theta_3 + \theta_4 = 1$ can be replaced by other nonzero values with equal signs for θ_1 and θ_2; and similarly for θ_3 and θ_4. It is also to be noted that $\mathbf{h} = \mathbf{0}$ can be replaced by $\mathbf{Ch} = \mathbf{0}$, where **C** is an $r \times r$ fixed nonsingular matrix whose order can be other than $O(1)$, if necessary, e.g., $O(n)$. For simplicity, we use $\mathbf{C} = \mathbf{I}_{(r)}$ as usual, where $\mathbf{I}_{(r)}$ is the $r \times r$ identity matrix.

We find that the first two restrictions $\theta_1 + \theta_2 = 1$ and $\theta_3 + \theta_4 = 1$ are for model identification, and the third restriction $\theta_1 = \theta_3$ is a genuine constraint as defined earlier. Let l_η be the log likelihood with the vector $\boldsymbol{\eta}$ of Lagrange multipliers. Then,

$$\begin{aligned} l_\eta &= m_{11} \log \theta_1 + m_{12} \log \theta_2 - n_1 \log(\theta_1 + \theta_2) \\ &\quad + m_{21} \log \theta_3 + m_{22} \log \theta_4 - n_2 \log(\theta_3 + \theta_4) + \mathbf{h}'\boldsymbol{\eta}, \end{aligned} \tag{S8.1}$$

which gives

$$\begin{aligned} \frac{\partial l_\eta}{\partial \boldsymbol{\theta}} &= \left(\frac{m_{11}}{\theta_1} - \frac{n_1}{\theta_1 + \theta_2}, \frac{m_{12}}{\theta_2} - \frac{n_1}{\theta_1 + \theta_2}, \frac{m_{21}}{\theta_3} - \frac{n_2}{\theta_3 + \theta_4}, \frac{m_{22}}{\theta_4} - \frac{n_2}{\theta_3 + \theta_4} \right)' + \mathbf{H}\boldsymbol{\eta} \\ &= \mathbf{0}, \end{aligned} \tag{S8.2}$$

where $\mathbf{H} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$. Subtracting the second element on the right-hand side of (S8.2) from the first element and similarly for the third and fourth elements,

we have

$$\frac{m_{11}}{\theta_1} - \frac{m_{12}}{1-\theta_1} = -\eta_3 \text{ and } \frac{m_{21}}{\theta_3} - \frac{m_{22}}{1-\theta_3} = \eta_3, \qquad \text{(S8.3)}$$

where $\theta_2 = 1-\theta_1$ and $\theta_4 = 1-\theta_3$ are used. Summing the two equations of (S8.3) and using $\theta_1 = \theta_3$, the MLEs are given as

$$\begin{aligned}
&\hat{\theta}_1 = \hat{\theta}_3 = m_1/(m_1+m_2) = m_1/n, \\
&\hat{\theta}_2 = \hat{\theta}_4 = m_2/n, \\
&\hat{\eta}_3 = -\frac{m_{11}}{\hat{\theta}_1} + \frac{m_{12}}{\hat{\theta}_2} = -n\left(\frac{m_{11}}{m_1} - \frac{m_{12}}{m_2}\right) = n\left(\frac{m_{21}}{m_1} - \frac{m_{22}}{m_2}\right), \\
&\hat{\eta}_1 = n_1 - \frac{m_{12}}{\hat{\theta}_2} = n\left(\frac{n_1}{n} - \frac{m_{12}}{m_2}\right), \\
&\hat{\eta}_2 = n_2 - \frac{m_{22}}{\hat{\theta}_4} = n\left(\frac{n_2}{n} - \frac{m_{22}}{m_2}\right).
\end{aligned} \qquad \text{(S8.4)}$$

From (S8.4), it is seen that when $n_1 \to +\infty$ and $n_2 \to +\infty$, the limiting values of $n^{-1}\hat{\eta}_i$ (i = 1, 2, 3) are zero. Since $\boldsymbol{\eta}$ can be redefined as $n^{-1}\boldsymbol{\eta}$ as addressed earlier, the probability limit of $\hat{\boldsymbol{\eta}}$, denoted by $\boldsymbol{\eta}_0$, which is defined when infinitely many observations are available both for Groups 1 and 2, is $\boldsymbol{\eta}_0 = \mathbf{0}$.

Example 1.2 When $\theta_2 = 1-\theta_1$ and $\theta_4 = 1-\theta_3$ are used in the likelihood,

$$l_\eta = m_{11}\log\theta_1 + m_{12}\log(1-\theta_1) + m_{21}\log\theta_3 + m_{22}\log(1-\theta_3) + h\eta, \qquad \text{(S8.5)}$$

where $h = \theta_1 - \theta_3 = 0$ and η are scalars with $\mathbf{H} = (1,-1)'$. From (S8.5),

$$\frac{\partial l_\eta}{\partial \boldsymbol{\theta}} = \left(\frac{m_{11}-n_1\theta_1}{\theta_1(1-\theta_1)}, \frac{m_{21}-n_2\theta_3}{\theta_3(1-\theta_3)}\right)' + \mathbf{H}\eta = \mathbf{0}, \qquad \text{(S8.6)}$$

giving

$$\frac{m_{11}-n_1\theta_1}{\theta_1(1-\theta_1)} + \eta = 0 \text{ and } \frac{m_{21}-n_2\theta_3}{\theta_3(1-\theta_3)} - \eta = 0. \qquad \text{(S8.7)}$$

From (S8.7), we find that $\hat{\theta}_1 = \hat{\theta}_3 = m_1/n$ is unchanged from that in Example

1.1 and

$$\hat{\eta} = -\frac{m_{11} - n_1\hat{\theta}_1}{\hat{\theta}_1(1-\hat{\theta}_1)} = -\frac{n(nm_{11} - n_1 m_1)}{m_1 m_2}$$
$$= \frac{m_{21} - n_2\hat{\theta}_1}{\hat{\theta}_1(1-\hat{\theta}_1)} = \frac{n(nm_{21} - n_2 m_1)}{m_1 m_2}, \quad \text{(S8.8)}$$

is different from $\hat{\eta}_3$ in Example 1.1. The only restriction in Example 1.2 is thus a genuine one as defined earlier. Assume that

$$c_1 \equiv n_1 / n = O(1) \text{ and } c_2 \equiv n_2 / n = O(1) \quad \text{(S8.9)}$$

are fixed even when $n_1 \to +\infty$ and $n_2 \to +\infty$. Note that the limiting value of $n^{-1}\hat{\eta}$ when $n_1 \to +\infty$ and $n_2 \to +\infty$ under (S8.9) is 0.

When c_1 and c_2 are used, $\hat{\theta} \equiv \hat{\theta}_1 = \hat{\theta}_3$ and $\hat{\eta}$ in Example 1.2 are rewritten as

$$\hat{\theta} = \frac{m_1}{n} = \frac{m_{11} + m_{21}}{n} = \frac{n_1}{n}\frac{m_{11}}{n_1} + \frac{n_2}{n}\frac{m_{21}}{n_2} \equiv c_1\hat{p}_1 + c_2\hat{p}_2, \quad \text{(S8.10)}$$

where $\hat{p}_1$ and $\hat{p}_2$ are the usual sample proportions in Groups 1 and 2, respectively, and

$$\hat{\eta} = -\frac{n\{(m_{11}/n) - (n_1/n)(m_1/n)\}}{(m_1/n)(m_2/n)} = -\frac{n(c_1\hat{p}_1 - c_1\hat{\theta})}{\hat{\theta}(1-\hat{\theta})}$$
$$= -\frac{nc_1(\hat{p}_1 - c_1\hat{p}_1 - c_2\hat{p}_2)}{\hat{\theta}(1-\hat{\theta})} = n\frac{c_1c_2(\hat{p}_2 - \hat{p}_1)}{\hat{\theta}(1-\hat{\theta})} \quad \text{(S8.11)}$$
$$= n\frac{c_1c_2(\hat{p}_2 - \hat{p}_1)}{(c_1\hat{p}_1 + c_2\hat{p}_2)(1 - c_1\hat{p}_1 - c_2\hat{p}_2)}.$$

The limiting value of $n^{-1}\hat{\eta}$ is clearly seen from (S8.11).

For another type of restrictions with an example, see Section 11 of this supplement.

9. Examples by the weighted score method

In this section, the estimators by the weighted score method are illustrated, where the weights are given by the derivatives of typical log priors. An example using the weight to remove the asymptotic biases of the estimators of restricted parameters and a Lagrange multiplier is given in Ogasawara (2016) using penalized logistic regression.

Example 2.1 This example has the same model specification as in Example 1.2. The two parameters θ_1 and θ_3 in Example 1.2 are redefined as θ_1 and θ_2. Using the independent beta priors for θ_1 and θ_2 with the same fixed parameter $k+1$, the posterior or the weighted likelihood of $\boldsymbol{\theta}$ using the single Lagrange multiplier η is written as

$$L_{\eta k} \equiv \text{constant} \times \{\prod_{i=1}^{n_1} \theta_1^{x_{1i}}(1-\theta_1)^{1-x_{1i}}\}\{\prod_{i=1}^{n_2} \theta_2^{x_{2i}}(1-\theta_2)^{1-x_{2i}}\} \times \{\theta_1(1-\theta_1)\}^k\{\theta_2(1-\theta_2)\}^k \exp(h\eta), \tag{S9.1}$$

where $x_{1i}(i=1,...,n_1)$ and $x_{2i}(i=1,...,n_2)$ are the observed values of X_1 and X_2 in Groups 1 and 2 shown in the previous section, respectively; and $h=\theta_1-\theta_2=0$ as before. Then, using $l_{\eta k} \equiv \log L_{\eta k}$,

$$\frac{\partial l_{\eta k}}{\partial \boldsymbol{\theta}} = \left(\frac{m_{11}+k-(n_1+2k)\theta_1}{\theta_1(1-\theta_1)}, \frac{m_{21}+k-(n_2+2k)\theta_2}{\theta_2(1-\theta_2)} \right)' + \mathbf{H}\eta = \mathbf{0}, \tag{S9.2}$$

where $\mathbf{H}=(1,-1)'$ is as before.

Define $\hat{\boldsymbol{\theta}}_{\mathrm{W}} = (\hat{\theta}_{\mathrm{W1}}, \hat{\theta}_{\mathrm{W2}})'$ as the vector of the WSEs for $\boldsymbol{\theta}_0$. Then, from (S9.2),

$$\frac{m_{11}+k-(n_1+2k)\hat{\theta}_{\mathrm{W1}}}{\hat{\theta}_{\mathrm{W1}}(1-\hat{\theta}_{\mathrm{W1}})} + \hat{\eta}_{\mathrm{W}} = 0 \quad \text{and} \quad \frac{m_{21}+k-(n_2+2k)\hat{\theta}_{\mathrm{W2}}}{\hat{\theta}_{\mathrm{W2}}(1-\hat{\theta}_{\mathrm{W2}})} - \hat{\eta}_{\mathrm{W}} = 0, \tag{S9.3}$$

where $\hat{\eta}_{\mathrm{W}}$ is the WSE of $\eta_0(=0)$. Summing (S9.3) and using $\hat{\theta}_{\mathrm{W1}} = \hat{\theta}_{\mathrm{W2}}$ give

$$\hat{\theta}_{\mathrm{W}} \equiv \hat{\theta}_{\mathrm{W1}} = \hat{\theta}_{\mathrm{W2}} = \frac{m_1+2k}{n+4k} \tag{S9.4}$$

and

$$\begin{aligned} \hat{\eta}_{\mathrm{W}} &= -\frac{m_{11}+k-(n_1+2k)\hat{\theta}_{\mathrm{W}}}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})} \\ &= -(n+4k)\frac{(n+4k)(m_{11}+k)-(n_1+2k)(m_1+2k)}{(m_1+2k)(m_2+2k)} \\ &= (n+4k)\frac{(n+4k)(m_{21}+k)-(n_2+2k)(m_1+2k)}{(m_1+2k)(m_2+2k)}. \end{aligned} \tag{S9.5}$$

From (S9.5), we find that *k* is an added pseudocount in each cell of the associated 2×2 contingency table in the previous section.

Define

$$\hat{p}_{\mathrm{W}1} \equiv \frac{m_{11}+k}{n_1+2k},\ \hat{p}_{\mathrm{W}2} \equiv \frac{m_{21}+k}{n_2+2k},\ n_{\mathrm{W}} \equiv n+4k, \tag{S9.6}$$

$c_{\mathrm{W}1} \equiv \frac{n_1+2k}{n+4k} = \frac{n_1+2k}{n_{\mathrm{W}}}$ and $c_{\mathrm{W}2} \equiv \frac{n_2+2k}{n+4k} = \frac{n_2+2k}{n_{\mathrm{W}}}$ with $c_{\mathrm{W}1}+c_{\mathrm{W}2}=1$.

Then,

$$\hat{\theta}_{\mathrm{W}} = \frac{m_1+2k}{n+4k} = \frac{n_1+2k}{n+4k}\frac{m_{11}+k}{n_1+2k} + \frac{n_2+2k}{n+4k}\frac{m_{21}+k}{n_2+2k} = c_{\mathrm{W}1}\hat{p}_{\mathrm{W}1} + c_{\mathrm{W}2}\hat{p}_{\mathrm{W}2} \tag{S9.7}$$

and

$$\begin{aligned}\hat{\eta}_{\mathrm{W}} &= n_{\mathrm{W}}\frac{n_{\mathrm{W}}(m_{21}+k)-(n_2+2k)(m_1+2k)}{(m_1+2k)(m_2+2k)} = \frac{n_{\mathrm{W}}c_{\mathrm{W}2}\hat{p}_{\mathrm{W}2} - n_{\mathrm{W}}c_{\mathrm{W}2}\hat{\theta}_{\mathrm{W}}}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})} \\ &= \frac{n_{\mathrm{W}}c_{\mathrm{W}2}(\hat{p}_{\mathrm{W}2} - c_{\mathrm{W}1}\hat{p}_{\mathrm{W}1} - c_{\mathrm{W}2}\hat{p}_{\mathrm{W}2})}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})} = \frac{n_{\mathrm{W}}c_{\mathrm{W}1}c_{\mathrm{W}2}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{\hat{\theta}_{\mathrm{W}}(1-\hat{\theta}_{\mathrm{W}})} \\ &= \frac{n_{\mathrm{W}}c_{\mathrm{W}1}c_{\mathrm{W}2}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{(c_{\mathrm{W}1}\hat{p}_{\mathrm{W}1} + c_{\mathrm{W}2}\hat{p}_{\mathrm{W}2})(1 - c_{\mathrm{W}1}\hat{p}_{\mathrm{W}1} + c_{\mathrm{W}2}\hat{p}_{\mathrm{W}2})}\end{aligned} \tag{S9.8}$$

follow.

Example 2.2 So far, the examples are for two-group cases. In this example, a single group is used, where the categorical distribution (a generalization of the Bernoulli distribution to more than two categories) with three categories is shown. A restriction of equal probabilities of occurrences of the first two categories is imposed. For the three original parameters θ_1, θ_2, and θ_3 corresponding to the three probabilities for the categories, the Dirichlet prior proportional to $(\theta_1\,\theta_2\,\theta_3)^k$ is used.

Define

$$\begin{aligned}L_{\eta k} &\equiv \text{constant}\times\{\textstyle\prod_{i=1}^{n}\theta_1^{y_{1i}}\theta_2^{y_{2i}}\theta_3^{y_{3i}}\}(\theta_1\theta_2\theta_3)^k\exp(h\eta) \\ &= \text{constant}\times\{\textstyle\prod_{i=1}^{n}\theta_1^{y_{1i}}\theta_2^{y_{2i}}(1-\theta_1-\theta_2)^{1-y_{3i}}\{\theta_1\theta_2(1-\theta_1-\theta_2)\}^k\exp(h\eta),\end{aligned} \tag{S9.9}$$

where y_{1i}, y_{2i}, and $y_{3i}(i=1,...,n)$ are observed values of the dichotomous variables Y_1, Y_2, and Y_3, respectively, which take values of 1 and 0 with $Y_1+Y_2+Y_3=1$, $h=\theta_1-\theta_2=0$, and $\mathbf{H}=(1,-1)'$. Define $l_{\eta k}\equiv\log L_{\eta k}$ and

$\boldsymbol{\theta} = (\theta_1, \theta_2)'$. Then,

$$\frac{\partial l_{\eta k}}{\partial \boldsymbol{\theta}} = \left(\frac{m_1 + k}{\theta_1} - \frac{m_3 + k}{1 - \theta_1 - \theta_2},\ \frac{m_2 + k}{\theta_2} - \frac{m_3 + k}{1 - \theta_1 - \theta_2} \right)' + \mathbf{H}\eta = \mathbf{0}, \qquad \text{(S9.10)}$$

where stochastic m_1, m_2, and m_3 are frequencies for the three categories with $m_1 + m_2 + m_3 = n$. From (S9.10),

$$\frac{m_1 + k}{\hat{\theta}_{\mathrm{W1}}} - \frac{m_3 + k}{1 - \hat{\theta}_{\mathrm{W1}} - \hat{\theta}_{\mathrm{W2}}} + \hat{\eta}_{\mathrm{W}} = 0 \quad \text{and} \quad \frac{m_2 + k}{\hat{\theta}_{\mathrm{W2}}} - \frac{m_3 + k}{1 - \hat{\theta}_{\mathrm{W1}} - \hat{\theta}_{\mathrm{W2}}} - \hat{\eta}_{\mathrm{W}} = 0 . \qquad \text{(S9.11)}$$

In a similar manner as before with $\hat{\theta}_{\mathrm{W}} \equiv \hat{\theta}_{\mathrm{W1}} = \hat{\theta}_{\mathrm{W2}}$,

$$\frac{m_1 + m_2 + 2k}{\hat{\theta}_{\mathrm{W}}} - 2\frac{m_3 + k}{1 - 2\hat{\theta}_{\mathrm{W}}} = 0 \quad \text{gives}$$

$$\hat{\theta}_{\mathrm{W}} = \frac{m_1 + m_2 + 2k}{2(n + 3k)} = \frac{(m_1 + m_2)n^{-1} + 2kn^{-1}}{2(1 + 3kn^{-1})}. \qquad \text{(S9.12)}$$

Define the MLE $\hat{\theta}$ as $\hat{\theta}_{\mathrm{W}}$ when $k = 0$. Then,

$$\hat{\theta}_{\mathrm{W}} = \frac{\hat{\theta} + kn^{-1}}{1 + 3kn^{-1}}. \qquad \text{(S9.13)}$$

Using the usual sample proportions $\hat{p}_i = m_i / n \ (i = 1,2,3)$,

$$\hat{\theta}_{\mathrm{W}} = \frac{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}{2(1 + 3kn^{-1})}. \qquad \text{(S9.14)}$$

On the other hand, (S9.11) gives $\hat{\eta}_{\mathrm{W}} = -\frac{m_1 + k}{\hat{\theta}_{\mathrm{W}}} + \frac{m_3 + k}{1 - 2\hat{\theta}_{\mathrm{W}}}$, and consequently,

$$\begin{aligned} n^{-1}\hat{\eta}_{\mathrm{W}} &= -\frac{2(1 + 3kn^{-1})(\hat{p}_1 + kn^{-1})}{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}} + 1 + 3kn^{-1} \\ &= (1 + 3kn^{-1})\left\{1 - \frac{2(\hat{p}_1 + kn^{-1})}{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}\right\} = \frac{(1 + 3kn^{-1})(\hat{p}_2 - \hat{p}_1)}{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}. \end{aligned} \qquad \text{(S9.15)}$$

Define $n_{\mathrm{W}} \equiv n(1 + 3kn^{-1}) = n + 3k$. Then, (S9.15) becomes

$$\hat{\eta}_{\mathrm{W}} = \frac{n_{\mathrm{W}}(\hat{p}_2 - \hat{p}_1)}{\hat{p}_1 + \hat{p}_2 + 2kn^{-1}}. \qquad \text{(S9.16)}$$

Define $\hat{p}_{\mathrm{W}i} \equiv \frac{m_i + k}{n_{\mathrm{W}}} = \frac{n}{n_{\mathrm{W}}} \frac{m_i + k}{n} = \frac{n}{n_{\mathrm{W}}} (\hat{p}_i + kn^{-1}).$ Then, (S9.16) is rewritten as

$$\hat{\eta}_{\mathrm{W}} = \frac{n_{\mathrm{W}}(\hat{p}_{\mathrm{W}2} - \hat{p}_{\mathrm{W}1})}{\hat{p}_{\mathrm{W}1} + \hat{p}_{\mathrm{W}2}} = \frac{n(\hat{p}_2 - \hat{p}_1)}{\hat{p}_{\mathrm{W}1} + \hat{p}_{\mathrm{W}2}}. \quad \text{(S9.17)}$$

10. The restrictions for model specification

The third type of restrictions is for model specification. The restrictions are defined as those to specify a part or whole of the associated model so that $f(\mathbf{X} | \boldsymbol{\theta})\ (= (L(\boldsymbol{\theta} | \mathbf{X}))$ is written as a probability density or mass of $\mathbf{X}$, which is seen as a set of random variables when $\boldsymbol{\theta}$ is given. This type is different from the two types of restrictions in Definition S1. Note that the model becomes meaningless unless the restrictions for model specification are imposed. Recall that the model without the restrictions for model identification is still of interest since the likelihood in unchanged. Recall also that the models without the genuine restrictions are regular ones by definition.

Example 1.3 When the restrictions $\theta_1 + \theta_2 = 1$ and $\theta_3 + \theta_4 = 1$ in Example 1.1 are used in the likelihood as $\theta_1 / (\theta_1 + \theta_2) = \theta_1,\ \theta_2 / (\theta_1 + \theta_2) = \theta_2,\ \theta_3 / (\theta_3 + \theta_4) = \theta_3$ and $\theta_4 / (\theta_3 + \theta_4) = \theta_4$ we have

$$l_\eta = m_{11} \log \theta_1 + m_{12} \log \theta_2 + m_{21} \log \theta_3 + m_{22} \log \theta_4 + \mathbf{h}'\boldsymbol{\eta}, \quad \text{(S10.1)}$$

where $\mathbf{h} = \begin{pmatrix} \theta_1 + \theta_2 - 1 \\ \theta_3 + \theta_4 - 1 \\ \theta_1 - \theta_3 \end{pmatrix}$ and $\mathbf{H} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ are as before. Equation (S10.1) gives

$$\frac{\partial l_\eta}{\partial \boldsymbol{\theta}} = \left(\frac{m_{11}}{\theta_1}, \frac{m_{12}}{\theta_2}, \frac{m_{21}}{\theta_3}, \frac{m_{22}}{\theta_4} \right)' + \mathbf{H}\boldsymbol{\eta} = \mathbf{0}. \quad \text{(S10.2)}$$

Consequently, (S8.3) also holds, which gives

$$\hat{\eta}_3 = -n \left(\frac{m_{11}}{m_1} - \frac{m_{12}}{m_2} \right) = n \left(\frac{m_{21}}{m_1} - \frac{m_{22}}{m_2} \right). \quad \text{(S10.3)}$$

On the other hand,

$$\hat{\eta}_1 = -\frac{m_{12}}{\hat{\theta}_2} = -n\frac{m_{12}}{m_2} \text{ and } \hat{\eta}_2 = -\frac{m_{22}}{\hat{\theta}_4} = -n\frac{m_{22}}{m_2}. \quad \text{(S10.4)}$$

Assume that (S8.9) holds. Then, while $(\boldsymbol{\eta}_0)_3$, where $(\cdot)_i$ is the *i*-th element of a vector, is the same as in Example 1.1, the limiting values of $n^{-1}\hat{\eta}_1$ and $n^{-1}\hat{\eta}_2$ when $n_1 \to +\infty$ and $n_2 \to +\infty$ under (S8.9), are

$$n^{-1}(\boldsymbol{\eta}_0)_1 = -c_1 \text{ and } n^{-1}(\boldsymbol{\eta}_0)_2 = -c_2, \quad \text{(S10.5)}$$

respectively. The results of (S10.5) are not equal to those in Example 1.1.

The difference stems from the differences of the types of restrictions $h_1 = 0$ and $h_2 = 0$ from those in Example 1.1. In Example 1.1, the restrictions are for model identification whereas in Example 1.3, they are for model specification in order that $\theta_i\,(i = 1,...,4)$ are probabilities. Note that the unit value in $\theta_1 + \theta_2 = 1$ and $\theta_3 + \theta_4 = 1$ cannot be replaced by other ones, which was possible in Example 1.1

Note also that in Example 1.3 unless the restrictions for model specification are imposed, the likelihood becomes infinitely large when $\theta_1,...,\theta_4$ go to infinity.

11. Some additional numerical results

Tables S1 to S3 give additional numerical results for Ogasawara (2016, Section 5).

12. Errata

The expression $(-|\mathbf{I}_0^*|)^{1/2}$ after (3.18) on page 24 of Ogasawara (2016) should be $(|\mathbf{I}_0^*|^2)^{1/4}$ as in (3.18).

References

Lee, S.-Y. (1979). Constrained estimation in covariance structure analysis. *Biometrika, 66*, 539-545.

Ogasawara, H. (2004). Asymptotic biases in exploratory factor analysis and structural equation modeling. *Psychometrika, 69,* 235-256.

Ogasawara, H. (2010). *Errata of the paper "Asymptotic cumulants of the parameter estimators in item response theory"*. Unpublished document. Available at http://www.res.otaru-uc.ac.jp/~hogasa/, https://barrel.repo.nii.ac.jp/.

Ogasawara, H. (2016). Asymptotic expansions for the estimators of Lagrange multipliers and associated parameters by the maximum likelihood and weighted score methods. *Journal of Multivariate Analysis, 147,* 20-37.

Stuart, A., & Ord, J. K. (1994). *Kendall's advanced theory of statistics: Distribution theory* (6th ed., Vol.1). London: Arnold.

Table S1. Accurate and higher-order asymptotic standard errors (SEs and HASEs) of the studentized estimators and their transformations (c_1 = .4 and c_2 = .6)

		(n) ML						(n) WS (k = .5)					
Ac.		(25)		(100)		(400)		(25)		(100)		(400)	
order		SE	HASE	SE	HASE	SE	HASE	SE	HASE	SE	HASE	SE	HASE
p = .1,	studentized sample proportion												
1	t_W	1.15	1.27	1.11	1.07	1.021	1.019	.84	.94	.99	.99	.997	.997
2	$t_{(1)}$	.74	.77	1.03	.95	.985	.987	.91	.83	.96	.96	.990	.990
2	t_{Ha}	.76	.81	1.09	.96	.988	.989	.93	.87	.97	.97	.992	.992
3	$t_{(2)}$	.84	1.00	731	1.00	1.27	1.000	1.19	1.00	5.9	1.00	1.009	1.000
p = .1,	studentized sample Lagrange multiplier												
1	t_W	.98	1.02	1.005	1.005	1.001	1.001	.86	.87	.967	.968	.992	.992
2	$t_{(1)}$	.97	1.01	1.003	1.003	1.001	1.001	.85	.85	.963	.964	.991	.991
2	t_{Ha}	.97	1.01	1.003	1.004	1.001	1.001	.85	.85	.964	.965	.991	.991
3	$t_{(2)}$	.92	1.00	.998	1.000	1.000	1.000	.97	1.00	1.002	1.000	1.000	1.000
p = .3,	studentized sample proportion												
1	t_W	1.13	1.08	1.023	1.021	1.006	1.005	.98	.98	.994	.994	.999	.999
2	$t_{(1)}$	1.03	.95	.984	.987	.997	.997	.95	.95	.988	.988	.997	.997
2	t_{Ha}	1.00	.95	.986	.988	.997	.997	.96	.96	.989	.989	.997	.997
3	$t_{(2)}$	202	1.00	3.42	1.000	1.000	1.000	1.67	1.00	1.006	1.000	1.001	1.000
p = .3,	studentized sample Lagrange multiplier												
1	t_W	1.021	1.020	1.005	1.005	1.001	1.001	.998	1.001	1.000	1.000	1.000	1.000
2	$t_{(1)}$	1.019	1.019	1.005	1.005	1.001	1.001	.995	1.000	1.000	1.000	1.000	1.000
2	t_{Ha}	1.019	1.019	1.005	1.005	1.001	1.001	.996	1.000	1.000	1.000	1.000	1.000
3	$t_{(2)}$	.998	1.000	1.000	1.000	1.000	1.000	1.001	1.000	1.000	1.000	1.000	1.000

Note. Ac. order = accuracy order, ML = maximum likelihood, WS = weighted score, SE = accurate standard error, HASE = $(1+n^{-1}\alpha^{(t)}_{\Delta 2})^{1/2}$ ($\alpha^{(t)}_{\Delta 2}$ is $\alpha^{(t)}_{\theta \mathrm{ML}\Delta 2}$, $\alpha^{(t)}_{\theta \mathrm{W}\Delta 2}$, $\alpha^{(t)}_{\eta \mathrm{ML}\Delta 2}$ or $\alpha^{(t)}_{\eta \mathrm{W}\Delta 2}$).

Table S2. Accurate and asymptotic third cumulants ($\alpha^{(t)}_{\theta\mathrm{ML3}}$, $\alpha^{(t)}_{\theta\mathrm{W3}}$, $\alpha^{(t)}_{\eta\mathrm{ML3}}$ and $\alpha^{(t)}_{\eta\mathrm{W3}}$) of the studentized estimators and their transformations (c_1 = .4 and c_2 = .6)

Ac.		(*n*) Acc.									
		ML			WS (k = .5)			WS (k = 1)			
order		(25)	(100)	(400)	(25)	(100)	(400)	(25)	(100)	(400)	Asy.
p = .1,	studentized sample proportion										
1	t_{W}	-4.2	-12.5	-6.2	-1.5	-5.8	-5.4	-.1	-3.2	-4.8	-5.3
2	$t_{(1)}$	1.0	117	.7	.5	2.5	.4	-.1	.1	.1	0
2	t_{Ha}	1.5	-322	.5	.2	.4	.2	-.1	-.1	.1	0
3	$t_{(2)}$	1.6	8e11	1e10	-4.3	4e5	8e4	-4.9	-2.5	-.4	0
p = .1,	studentized sample Lagrange multiplier										
1	t_{W}	-1.24	-1.18	-1.11	-.63	-1.02	-1.08	-.42	-.90	-1.05	-1.09
2	$t_{(1)}$	-.60	-.14	-.03	-.40	-.20	-.05	-.34	-.23	-.07	0
2	t_{Ha}	-.60	-.14	-.03	-.40	-.20	-.05	-.34	-.24	-.07	0
3	$t_{(2)}$	-.64	-.18	-.03	-.58	-.18	-.05	-.57	-.19	-.06	0
p = .3,	studentized sample proportion										
1	t_{W}	-5.5	-2.1	-1.8	-1.9	-1.8	-1.8	-.8	-1.5	-1.7	-1.7
2	$t_{(1)}$	35	.3	.1	.9	.2	.1	.1	.1	.0	0
2	t_{Ha}	-25	.3	.1	.3	.1	.0	-.0	.0	.0	0
3	$t_{(2)}$	4e9	1.9	.0	431	1e5	-.0	-1.3	-.1	-.1	0
p = .3,	studentized sample Lagrange multiplier										
1	t_{W}	-.44	-.37	-.36	-.38	-.37	-.36	-.34	-.36	-.36	-.36
2	$t_{(1)}$	-.07	-.01	-.00	-.10	-.02	-.01	-.12	-.03	-.01	0
2	t_{Ha}	-.07	-.01	-.00	-.10	-.02	-.01	-.12	-.03	-.01	0
3	$t_{(2)}$	-.10	-.01	-.00	-.10	-.02	-.01	-.10	-.03	-.01	0

Note. Ac. order = accuracy order, ML = maximum likelihood, WS = weighted score, Acc. = accurate values multiplied by $n^{1/2}$, Asy. = asymptotic values, xey = $x \times 10^y$.

Table S3. Accurate and asymptotic fourth cumulants ($\alpha^{(t)}_{\theta \mathrm{ML4}}$, $\alpha^{(t)}_{\theta \mathrm{W4}}$, $\alpha^{(t)}_{\eta \mathrm{ML4}}$ and $\alpha^{(t)}_{\eta \mathrm{W4}}$) of the studentized estimators and their transformations (c_1 = .4 and c_2 = .6)

Ac.		(*n*) Acc.									
		ML			WS (*k* = .5)			WS (*k* = 1)			
order		(25)	(100)	(400)	(25)	(100)	(400)	(25)	(100)	(400)	Asy.
p = .1,	studentized sample proportion										
1	t_W	11	457	106	2	116	87	-1	40	71	81
2	$t_{(1)}$	-5	7e4	-48	-10	230	-42	-6	-24	-37	-39
2	t_{Ha}	-2	3e5	-35	-9	1	-34	-6	-19	-30	-33
3	$t_{(2)}$	2e6	1e18	3e20	11	4e9	3e13	21	16	1e6	0
p = .1,	studentized sample Lagrange multiplier										
1	t_W	-15.8	-18.1	-15.9	-6.1	-14.6	-15.3	-3.1	-11.9	-14.7	-15.4
2	$t_{(1)}$	-17.0	-21.7	-19.2	-6.4	-17.2	-18.4	-3.2	-13.9	-17.6	-18.5
2	t_{Ha}	-17.0	-21.5	-19.0	-6.4	-17.1	-18.2	-3.2	-13.8	-17.4	-18.3
3	$t_{(2)}$	-7.9	-6.5	-1.1	-18.9	-35.6	-41.4	-19.6	-51.5	-74.8	0
p = .3,	studentized sample proportion										
1	t_W	111	24	19	24	19	18	7	15	17	18
2	$t_{(1)}$	6e3	-12	-10	5	-11	-10	-6	-9	-10	-10
2	t_{Ha}	4e3	4	-9	-7	-9	-9	-5	-8	-9	-9
3	$t_{(2)}$	2e15	3e19	-0	3e8	4e12	0	10	1e5	1	0
p = .3,	studentized sample Lagrange multiplier										
1	t_W	-5.0	-4.0	-3.8	-4.2	-3.9	-3.8	-3.5	-3.8	-3.7	-3.7
2	$t_{(1)}$	-5.7	-4.4	-4.1	-4.6	-4.3	-4.1	-3.9	-4.2	-4.1	-4.1
2	t_{Ha}	-5.6	-4.4	-4.1	-4.6	-4.2	-4.1	-3.9	-4.1	-4.1	-4.0
3	$t_{(2)}$	-1.7	-.2	-4.2	-6.1	-5.8	-5.6	-8.6	-10.6	-11.0	0

Note. Ac. order = accuracy order, ML = maximum likelihood, WS = weighted score, Acc. = accurate values multiplied by *n*, Asy. = asymptotic values, $x\mathrm{e}y = x \times 10^y$.