Annals of the Institute of Statistical Mathematics manuscript No.
(will be inserted by the editor)

A test for the presence of stochastic ordering under censoring: the k-sample case (Supplementary material)

Hammou El Barmi

Received: date / Revised: date

To prove Lemma 1, we make use of the following result from Aitchinson and Silvey (1958).
Result 1 : If $\boldsymbol{\psi}(\boldsymbol{\beta})$ is a continuous function mapping $\mathbf{R}^k$ into itself with the property that, for every $\boldsymbol{\beta}$ such that $||\boldsymbol{\beta}|| = \mathbf{1}, \boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{\psi}(\boldsymbol{\beta}) < \mathbf{0},$ then there exists a point $\hat{\boldsymbol{\beta}}^{(0)}$ such that $||\hat{\boldsymbol{\beta}}^{(0)}|| < 1$ and $\boldsymbol{\psi}(\hat{\boldsymbol{\beta}}^{(0)}) = \mathbf{0}$.
Proof of Lemma 1: Fix $\boldsymbol{\mu} \in \mathcal{U}_{a_n}$ and consider

$$\boldsymbol{\psi}(\boldsymbol{\beta}) = \begin{pmatrix} \sum_{j=1}^{n(t)} \log\left(1 - \dfrac{d_{1j}/n_1}{n_{1j}/n_1 + n^{-1/3-\delta/2}\beta_1(t)}\right) - \mu_1 \\ \sum_{j=1}^{n(t)} \log\left(1 - \dfrac{d_{2j}/n_2}{n_{2j}/n_2 + n^{-1/3-\delta/2}\beta_2(t)}\right) - \mu_2 \\ \vdots \\ \sum_{j=1}^{n(t)} \log\left(1 - \dfrac{d_{kj}/n_k}{n_{kj}/n_k + n^{-1/3-\delta/2}\beta_k(t)}\right) - \mu_k \end{pmatrix}.$$

Note that by Gijbels and Wang (1993), we have

$$\sup_{0\le s\le \tau_2} |\log(\bar{F}_i(s)) - \log(\hat{\bar{F}}_i(s))| = O(\sqrt{\log\log(n_i)/n_i}).$$

Using this, we get

$$\begin{aligned} \boldsymbol{\beta}^T\boldsymbol{\psi}(\boldsymbol{\beta}) &= \boldsymbol{\beta}^T(\log(\hat{\bar{\mathbf{F}}}(t)) - \boldsymbol{\mu}) - n^{-1/3-\delta/2}\boldsymbol{\beta}^T\hat{C}\boldsymbol{\beta} + O(n^{-2/3-\delta}) \\ &\le O(n^{-1/3-\delta}) - \min_{1\le i\le k}(c_i(t))n^{-1/3-\delta/2} + O(n^{-2/3-\delta}) \\ &< 0 \end{aligned}$$

Hammou El Barmi
Paul H. Chook Department of Information Systems and Statistics, Baruch College, The City University of New York. One Baruch Way New York, NY 10010, U.S.A. E-mail: hammou.elbarmi@baruch.cuny.edu

almost surely. Now Result 1 implies that there exists $\hat{\boldsymbol{\beta}} \equiv \hat{\boldsymbol{\beta}}(\boldsymbol{\mu})$ such that $\boldsymbol{\psi}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$, The rest of the lemma's claim follows using the implicit function theorem. $\square$

Proof of Lemma 2: Lemma 1 shows that for any $\boldsymbol{\mu}$ such that $||\boldsymbol{\mu} - \boldsymbol{\mu}^{(0)}|| \leq a_n$, there exists $\hat{\boldsymbol{\beta}}$, a differentiable function of $\boldsymbol{\mu}$, such that $\boldsymbol{\psi}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$. It easily follows from this that

$$\frac{\partial \boldsymbol{\psi}(\hat{\boldsymbol{\beta}})}{\partial \boldsymbol{\mu}^T} = \mathbf{0} \Rightarrow \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \boldsymbol{\mu}^T} = C^{-1} + O(a_n).$$

Also arguments similar to those in Aitchison and Silvey (1958) show that equations $Q_{i\mathbf{n}}(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \mathbf{0}, i = 1, 2,$ have solutions $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\lambda}}$ in $\{(\boldsymbol{\mu}, \boldsymbol{\lambda}) : ||\boldsymbol{\mu} - \boldsymbol{\mu}^{(0)}|| + ||\boldsymbol{\lambda}|| \leq a_n\}$.

Next we show that any solution to (8) maximizes $L(\boldsymbol{\mu})$ in $\mathcal{U}_{a_n}$ subject to $\mathbf{h}(\boldsymbol{\mu}) = \mathbf{0}$. By differentiating (6) with respect to μ_i we get

$$\frac{d}{d\mu_i}\beta_i = \frac{1}{\displaystyle\sum_{j=n(s)+1}^{n(t)} \frac{n_i d_{ij}}{(n_{ij} + n_i\beta_i - d_{ij})(n_{ij} + n_i\beta_i)}}, i = 1, 2, \ldots, k. \qquad (S.1)$$

Using this, we can easily show that

$$\nabla \log L(\boldsymbol{\mu}) = -N\boldsymbol{\beta} \qquad (S.2)$$

where $N = diag(n_1, n_2, \ldots, n_k)$. Therefore, we have for some $\boldsymbol{\mu}^* \in \mathcal{U}_{a_n}$,

$$\begin{aligned}\log(L(\boldsymbol{\mu})) &= \log(L(\hat{\boldsymbol{\mu}})) + \nabla \log L(\hat{\boldsymbol{\mu}})(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \\ &\quad + \frac{1}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^T \frac{\partial^2 \log(L(\boldsymbol{\mu}^*))}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \\ &= \log(L(\hat{\boldsymbol{\mu}})) - \hat{\boldsymbol{\beta}}^T N(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) + \frac{1}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^T \frac{\partial^2 \log(L(\boldsymbol{\mu}^*))}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}).\end{aligned}$$

Since

$$Q_{2\mathbf{n}}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\mu}}) = \mathbf{0} \Leftrightarrow N\hat{\boldsymbol{\beta}} = nH\hat{\boldsymbol{\lambda}}, \qquad (S.3)$$

we have

$$\hat{\boldsymbol{\beta}}^T(\hat{\boldsymbol{\mu}})N(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) = n\hat{\boldsymbol{\lambda}}^T H^T(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) = n(\mathbf{h}(\hat{\boldsymbol{\mu}}) - \mathbf{h}(\boldsymbol{\mu})) = \mathbf{0}.$$

Also, we have from (S.1) and (S.2)

$$\frac{\partial^2 \log(L(\boldsymbol{\mu}^*))}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} = -N\frac{\partial \hat{\boldsymbol{\beta}}(\boldsymbol{\mu}^*)}{\partial \boldsymbol{\mu}} = -N\hat{C}^{-1} + O(a_n) = n\hat{\tilde{C}}^{-1} + O(a_n).$$

As a result

$$\log(L(\boldsymbol{\mu})) - \log(L(\hat{\boldsymbol{\mu}})) = -n\frac{1}{2}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})^T \tilde{C}^{-1}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) + o(1)$$

and hence $L(\boldsymbol{\mu}) < L(\hat{\boldsymbol{\mu}})$ almost surely. $\square$

Proof of Theorem 1. First note that

$$Q_{2\mathbf{n}}(\hat{\boldsymbol{\beta}}^{(0)}, \hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)}) = \mathbf{0} \Leftrightarrow -\hat{\Gamma}\hat{\boldsymbol{\beta}}^{(0)} + H\hat{\boldsymbol{\lambda}}^{(0)} = \mathbf{0}$$

where $\hat{\Gamma} = diag(n_1/n, n_2/n, \ldots, n_k/n)$. Also

$$\mathbf{0} = Q_{3\mathbf{n}}(\hat{\boldsymbol{\beta}}^{(0)}, \hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)}) = H^T\boldsymbol{\mu}^{(0)} = H^T(\hat{\boldsymbol{\mu}}^{(0)} - \boldsymbol{\mu}^{(0)}).$$

Expanding $Q_{1\mathbf{n}}(\hat{\boldsymbol{\beta}}^{(0)}, \hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)})$ about $(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0})$, we get

$$\begin{aligned}
\mathbf{0} &= Q_{1\mathbf{n}}(\hat{\boldsymbol{\beta}}^{(0)}, \hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)}) \\
&= Q_{1\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0}) + \frac{\partial Q_{1\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0})}{\partial \boldsymbol{\beta}^T}\hat{\boldsymbol{\beta}}^{(0)} + \frac{\partial Q_{1\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0})}{\partial \boldsymbol{\mu}^T}(\hat{\boldsymbol{\mu}}^{(0)} - \boldsymbol{\mu}^{(0)}) \\
&+ \frac{\partial Q_{i\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0})}{\partial \boldsymbol{\lambda}^T}\hat{\boldsymbol{\lambda}}^{(0)} + o_p(\delta_n) \\
&= \hat{C}\hat{\boldsymbol{\beta}}^{(0)} - (\hat{\boldsymbol{\mu}}^{(0)} - \boldsymbol{\mu}^{(0)}) + o_p(\delta_n)
\end{aligned}$$

where $\delta_n = ||\hat{\boldsymbol{\mu}}^{(0)} - \boldsymbol{\mu}^{(0)}|| + ||\hat{\boldsymbol{\beta}}^{(0)}|| + ||\hat{\boldsymbol{\lambda}}^{(0)}||$. This gives

$$\begin{pmatrix} -Q_{1\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, 0) + o_p(\delta_n) \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \hat{C} & -I & \mathbf{0} \\ -\hat{\Gamma} & \mathbf{0} & H \\ \mathbf{0} & H^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}}^{(0)} \\ \hat{\boldsymbol{\mu}}^{(0)} - \boldsymbol{\mu}^{(0)} \\ \hat{\boldsymbol{\lambda}}^{(0)} \end{pmatrix}. \qquad (S.4)$$

Notice that

$$Q_{1n}(\mathbf{0}, \boldsymbol{\mu}_0, \mathbf{0}) = \begin{pmatrix} \log(\hat{\bar{F}}_1(t)) - \log(\bar{F}(t)) \\ \log(\hat{\bar{F}}_2(t)) - \log(\bar{F}(t)) \\ \vdots \\ \log(\hat{\bar{F}}_k(t)) - \log(\bar{F}(t)) \end{pmatrix}.$$

If we take $\delta_n = n^{-1/2}$, we get applying the Delta method

$$\sqrt{n}[-Q_{1\mathbf{n}}(\mathbf{0}, \boldsymbol{\mu}_0, \mathbf{0}) + o(\delta_n)] \xrightarrow{d} N(\mathbf{0}, \tilde{C}). \qquad (S.5)$$

Let $\mathrm{A_{11} = C, A_{12} = (I, 0), A_{21} = \Gamma A_{12}^T}$ and

$$A_{22} = \begin{pmatrix} 0 & \mathrm{H} \\ \mathrm{H^T} & 0 \end{pmatrix}$$

where $\Gamma = \lim_{n\to\infty} \hat{\Gamma} = diag(\gamma_1, \gamma_2, \ldots, \gamma_k)$. Then (S.4) implies that

$$\sqrt{n}\begin{pmatrix} \hat{\boldsymbol{\beta}}^{(0)} \\ \hat{\boldsymbol{\xi}}^{(0)} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} \sqrt{n}Q_{1n}(\mathbf{0}, \boldsymbol{\mu}^{(0)}, \mathbf{0}) \\ \mathbf{0} \end{pmatrix} + o_p(1).$$

Therefore

$$\sqrt{n}\hat{\boldsymbol{\xi}}^{(0)} = \sqrt{n}(0,I)\begin{pmatrix}\hat{\boldsymbol{\beta}}^{(0)}\\ \hat{\boldsymbol{\xi}}^{(0)}\end{pmatrix} = \sqrt{n}A_{22.1}^{-1}A_{21}A_{11}^{-1}Q_{1n}(\mathbf{0},\boldsymbol{\mu}^{(0)},\mathbf{0}) + o_p(1) \quad (S.6)$$

where

$$A_{22.1} = A_{22} - A_{21}A_{11}^{-1}A_{12} = \begin{pmatrix}\tilde{C}^{-1} & H\\ H^T & 0\end{pmatrix}.$$

Now, using (S.5), we get

$$\begin{aligned}\sqrt{n}A_{21}A_{11}^{-1}Q_{1\mathbf{n}}(\mathbf{0},\boldsymbol{\mu}^{(0)},\mathbf{0}) &= \begin{pmatrix}\tilde{C}^{-1}Q_{1n}(\mathbf{0},\boldsymbol{\mu}^{(0)},\mathbf{0})\\ \mathbf{0}\end{pmatrix} + o_P(1)\\ &\xrightarrow{d} N\left(\mathbf{0},\begin{pmatrix}\tilde{C}^{-1} & 0\\ 0 & 0\end{pmatrix}\right).\end{aligned}$$

Moreover, it is easy to check that

$$A_{22.1}^{-1} = \begin{pmatrix}P & Q\\ Q^T & -R\end{pmatrix}$$

where $Q = \tilde{C}HR$. Using (S.6), we have

$$\begin{aligned}\sqrt{n}\hat{\boldsymbol{\xi}}^{(0)} &\xrightarrow{d} N\left(\mathbf{0}, A_{22.1}^{-1}\begin{pmatrix}\tilde{C}^{-1} & 0\\ 0 & 0\end{pmatrix}A_{22.1}^{-1}\right)\\ &\xrightarrow{d} N\left(\mathbf{0},\begin{pmatrix}P & 0\\ 0 & R\end{pmatrix}\right)\end{aligned}$$

and we have the desired conclusion. □

Proof of Lemma 4: To prove (9), it suffices to show that for any $t \in [\tau_1,\tau_2]$ and any $1 \le \ell \le k$, with probability at least $1-\epsilon$,

$$|\hat{\beta}_\ell^{(0)}(t)| \le \frac{-2}{\log(\bar{F}(\tau_1))}(\max_{j\neq\ell}(-(\log(\hat{\bar{F}}_j(t)) + \hat{\Lambda}_\ell(t)), \max_{j\neq\ell}(-(\log(\hat{\bar{F}}_\ell(t)) + \hat{\Lambda}_j(t))))$$

where for any $r, \hat{\Lambda}_r(t) = \sum_{j=1}^{n(t)} d_{rj}/n_{rj}$ is the Nelson-Aalen estimator of $\Lambda_r(t)$. This will give the desired conclusion since the right hand side is $O_p(n^{-1/2})$ uniformly in t over $[\tau_1,\tau_2]$ under $\mathcal{H}_0$. First notice that, since $n_1\hat{\beta}_1^{(0)}(t) = n\hat{\lambda}_1^{(0)}(t), n_2\hat{\beta}_2^{(0)}(t) = n(\lambda_2^{(0)}(t) - \lambda_1^{(0)}(t)), \ldots, n_{k-1}\hat{\beta}_{k-1}^{(0)}(t) = n(\hat{\lambda}_{k-1}^{(0)}(t) - \hat{\lambda}_k^{(0)}(t))$ and $n_k\hat{\beta}_k^{(0)}(t) = -n\hat{\lambda}_{k-1}^{(0)}(t)$, $\sum_{i=1}^{k} n_i\hat{\beta}_i^{(0)}(t) = 0$. Let $B^-(t) = \{i, \hat{\beta}_i^{(0)}(t) < 0\}$ and $B^+(t) = \{i, \hat{\beta}_i^{(0)}(t) > 0\}$. Clearly, $B^-(t) \neq \emptyset$ iff $B^+(t) \neq \emptyset$. Suppose $B^-(t) \neq \emptyset$ and suppose $\ell \in B^-(t)$ and $\ell' \in B^+(t)$. It follows from Li (1995) that

$$-\log\prod_{j=1}^{n(t)}\left(1 - \frac{d_{\ell j}}{n_{\ell j} + n_\ell\hat{\beta}_\ell^{(0)}(t)}\right) \ge \hat{\Lambda}_\ell(t)\left(\frac{n_\ell}{n_\ell - n_\ell|\hat{\beta}_\ell^{(0)}(t)|}\right)$$

and

$$\log \prod_{j=1}^{n(t)} \left(1 - \frac{d_{\ell j}}{n_{\ell' j} + n_{\ell'} \hat{\beta}_{\ell'}^{(0)}(t)}\right) \geq -\hat{\Lambda}_{\ell'}(t) \left(\frac{n_{\ell'}}{n_{\ell'} + n_{\ell'} \hat{\beta}_{\ell'}^{(0)}(t)}\right) + \log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell'}(t).$$

Since the sum of the terms on the left hand side of the two inequalities above is 0, we have

$$0 \geq \hat{\Lambda}_{\ell}(t) \left(\frac{n_{\ell}}{n_{\ell} - n_{\ell}|\beta_{\ell}^{(0)}(t)|}\right) - \hat{\Lambda}_{\ell'}(t) \left(\frac{n_{\ell'}}{n_{\ell'} + n_{\ell'} \beta_{\ell'}^{(0)}(t)}\right) + \log(\hat{\bar{F}}_{\ell}(t)) + \hat{\Lambda}_{\ell'}(t)$$

Since $\log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell'}(t) \leq 0$, we have

$$-|\hat{\beta}_{\ell}^{(0)}(t)| \log(\hat{\bar{F}}_{\ell'}(t)) + \beta_{\ell'}^{(0)}(t)(\hat{\Lambda}_{\ell}(t) + \log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell'}(t)) \leq -(\log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell}(t)).$$

Clearly the uniform strong consistency of the Kaplan-Meier estimator implies that $\sup_{0 \leq u \leq \tau_2} | - \log(\hat{\bar{F}}_j(u)) + \log(\bar{F}(u)))| \to 0$ with probabability one. In addition, by the consistency of the Nelson-Aalen estimator, we have $\sup_{0 \leq u \leq \tau_2} |\hat{\Lambda}_j(u) + \log(\bar{F}(u))| \xrightarrow{P} 0$ where $\xrightarrow{P}$ is used to denote convergence in probability. Consequently, $\sup_{0 \leq u \leq \tau_2} |\hat{\Lambda}_j(u) + \log(\hat{\bar{F}}_j(u))| \xrightarrow{P} 0$. Therefore, for sufficiently large n , we have $-\log(\hat{\bar{F}}_{\ell'}(t)) \geq -\frac{1}{2} \log(\bar{F}(\tau_1)) > 0$ and $\hat{\Lambda}_{\ell}(t) + \log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell'}(t)) \geq -\frac{1}{2} \log(\bar{F}(\tau_1)) > 0$ for all $t \in [\tau_1, \tau_2]$ with a probability of at least $1 - \epsilon$. This implies that, for sufficiently large n and with probability of at least $1 - \epsilon$,

$$0 < |\hat{\beta}_{\ell}^{(0)}(t)| \leq \frac{2}{-\log(\bar{F}(\tau_1))} [-(\log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell}(t))]$$

and

$$0 < \hat{\beta}_{\ell'}^{(0)}(t) < \frac{2}{-\log(\bar{F}(\tau_1))} [-(\log(\hat{\bar{F}}_{\ell'}(t)) + \hat{\Lambda}_{\ell}(t))].$$

The two inequalities above clearly imply that for any $1 \leq \ell \leq k$, for sufficiently large n

$$|\hat{\beta}_{\ell}(t)| \leq \frac{2}{-\log(\bar{F}(\tau_1))} \left\{ \max_{j \neq \ell}(-(\log(\hat{\bar{F}}_j(t)) + \hat{\Lambda}_{\ell}(t)), \max_{j \neq \ell}(-(\log(\hat{\bar{F}}_{\ell}(t)) + \hat{\Lambda}_j(t))) \right\}$$

for any $t \in [\tau_1, \tau_2]$ with probability at least $1 - \epsilon$ and we have (9).

Next we prove (10). Using the Taylor expansion of $\log(1 + x)$ to the fifth order we get

$$\hat{\mu}_i^{(0)}(t) \equiv \sum_{j=1}^{n(t)} \log \left(1 - \frac{d_{ij}}{n_{ij} - n_i \hat{\beta}^{(0)}(t)}\right)$$

$$= \log(\hat{F}_i(t)) + \left[\sum_{j=1}^{n(t)} \log\left(1 + \frac{n_i\hat{\beta}_i^{(0)}(t)}{n_{ij} - d_{ij}}\right) - \sum_{j=1}^{n(t)} \log\left(1 + \frac{n_i\hat{\beta}_i^{(0)}(t)}{n_{ij}}\right)\right]$$

$$= \log(\hat{F}_i(t)) + \hat{\beta}_i^{(0)}(t)\hat{c}_i(t) - \frac{[n_i\hat{\beta}_i^{(0)}(t)]^2}{2}\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^2} - \frac{1}{n_{ij}^2}\right\}$$

$$+ \frac{[n_i\hat{\beta}_i^{(0)}(t)]^3}{3}\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^3} - \frac{1}{n_{ij}^3}\right\}$$

$$- \frac{[n_i\hat{\beta}_i^{(0)}(t)]^4}{4}\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^4} - \frac{1}{n_{ij}^4}\right\}$$

$$+ \frac{[n_i\hat{\beta}_i^{(0)}(t)]^5}{5}\sum_{j=1}^{n(t)}\left\{\frac{1}{(1 + \hat{\psi}_{1ij}(t))^5(n_{ij} - d_{ij})^5} - \frac{1}{(1 + \hat{\psi}_{2ij}(t))^5 n_{ij}^5}\right\}$$

where

$$|\hat{\psi}_{1ij}(t)| \leq \frac{n_i|\hat{\beta}_i^{(0)}(t)|}{n_{ij} - d_{ij}} \quad \text{and} \quad |\hat{\psi}_{2\mathrm{ij}}(\mathrm{t})| \leq \frac{\mathrm{n_i}|\hat{\beta}_{\mathrm{i}}^{(0)}(\mathrm{t})|}{\mathrm{n_{ij}}}$$

Now

$$[n_i\hat{\beta}_i^{(0)}(t)]^2\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^2} - \frac{1}{n_{ij}^2}\right\} = [\hat{\beta}_i^{(0)}(t)]^2\hat{v}_i(t)$$

where

$$\hat{v}_i(t) = \sum_{j=1}^{n(t)}\left\{\frac{n_i^2 d_{ij}}{n_{ij}^2(n_{ij} - d_{ij})} + \frac{n_i^2 d_{ij}}{n_{ij}(n_{ij} - d_{ij})^2}\right\}.$$

Since $\hat{v}_i(t)$ is a uniformly strongly consistent estimator of $2\int_0^t \frac{d\Lambda_i(s)}{\pi_i^2(s)}$ on $[0, \tau_2]$, we have using (9)

$$[n_i\hat{\beta}_i^{(0)}(t)]^2\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^2} - \frac{1}{n_{ij}^2}\right\} = O_p(n_i^{-1})$$

uniformly on $[\tau_1, \tau_2]$.

Let $Y_i(\tau_2) = \sum_{j=1}^{n_i} I[Z_{ij} \geq \tau_2]$ and notice that $(n_{ij} - d_{ij}) \geq Y_i(\tau_2)$ for $j \leq n(t)$ and because $\pi_i(\tau_2) > 0, Y_i(\tau_2) = O_p(n_i^{-1})$. Using this we get

$$\left|[n_i\hat{\beta}_i^{(0)}(t)]^3\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^3} - \frac{1}{n_{ij}^3}\right\}\right| =$$

$$n_i^2|\hat{\beta}_i^{(0)}(t)|^3\sum_{j=1}^{n(t)}\frac{n_i d_{ij}}{n_{ij}(n_{ij} - d_{ij})}\left(\frac{1}{(n_{ij} - d_{ij})^2} + \frac{1}{n_{ij}(n_{ij} - d_{ij})} + \frac{1}{(n_{ij} - d_{ij})^2}\right) \leq$$

$$\frac{3n_i^2|\hat{\beta}_i^{(0)}(t)|^3}{[Y_i(\tau_2)]^2}\hat{c}_i(t) = O_p(n_i^{-3/2})$$

uniformly on $[\tau_1, \tau_2]$ by (9) and the strong uniform consistency of $\hat{c}_i(t)$ on $[0, \tau_2]$. Also

$$n_i^4[\hat{\beta}_i^{(0)}(t)]^4 \sum_{j=1}^{n(t)} \left\{ \frac{1}{(n_{ij}-d_{ij})^4} - \frac{1}{n_{ij}^4} \right\} =$$
$$n_i^4[\hat{\beta}_i^{(0)}(t)]^4 \sum_{j=1}^{n(t)} \left(\frac{1}{(n_{ij}-d_{ij})^2} - \frac{1}{n_{ij}^2} \right) \left(\frac{1}{(n_{ij}-d_{ij})^2} + \frac{1}{n_{ij}^2} \right) \le$$
$$\frac{2n_i^2[\hat{\beta}_i^{(0)}(t)]^4}{Y_i^2(\tau_2)}\hat{v}_i(t) = O_p(n_i^{-2})$$

uniformly in t over $[\tau_1, \tau_2]$ by (9) and the uniform strong consistency of $\hat{v}_i(t)$ on $[0, \tau_2]$. Finally, we have

$$\left| n_i^5[\hat{\beta}_i^{(0)}(t)]^5 \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5(n_{ij}-d_{ij})^5} \right| \le$$
$$\frac{2n_i^5|\hat{\beta}_i^{(0)}(t)|^5}{5Y_i^5(\tau_2)} \left| \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5} \right| = O_p(n_i^{-3/2})$$

uniformly over t in $[\tau_1, \tau_2]$ since $\max_{1\le i\le k} \max_{1\le j\le n(t)} \sup_{\tau_1\le t\le\tau_2} \hat{\psi}_{1ij}(t) = o_p(1)$ by (9) and $\sup_{\tau_1\le t\le\tau_2} n(t) = O_p(n_i)$ (this holds beacuse $n_i/n \to \gamma_i > 0$). Similarly, we have

$$\left| \frac{2n_i^5[\hat{\beta}_i^{(0)}(t)]^5}{5} \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5 n_{ij}^4} \right| \le$$
$$\frac{2n_i^5|\hat{\beta}_i^{(0)}(t)|^5}{5Y_i^4(\tau_2)} \left| \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5} \right| = O_p(n^{-3/2})$$

uniformly over t in $[\tau_1, \tau_2]$. Therefore, under $\mathcal{H}_0$,

$$\hat{\mu}_i^{(0)}(t) = \log(\hat{\bar{F}}_i(t)) + \hat{c}_i(t)\hat{\beta}_i^{(0)}(t) - \frac{\hat{v}_i(t)}{2}[\hat{\beta}_i^{(0)}(t)]^2 + O_p(n_i^{-3/2}) \qquad (S.7)$$

uniformly in t over $[\tau_1, \tau_2]$. Since by (S.3) we have $n_1\hat{\beta}_1^{(0)}(t) = n\hat{\lambda}_1^{(0)}(t)$, $n_i\hat{\beta}_i^{(0)}(t) = n(\hat{\lambda}_i^{(0)}(t) - \hat{\lambda}_{i-1}^{(0)}(t)), i = 2, 3, \ldots, k$, and because $\hat{\mu}_i^{(0)}(t) - \hat{\mu}_{i+1}^{(0)}(t) =$

$0, i = 1, 2, \ldots, k-1$, and $n_i/n \to \gamma_i > 0$, we get using (S.7)

$$\left(\frac{n}{n_1}\hat{c}_1(t) + \frac{n}{n_2}\hat{c}_2(t)\right)\hat{\lambda}_1^{(0)}(t) - \frac{n}{n_2}\hat{c}_2(t)\hat{\lambda}_2^{(0)}(t) = \log\left(\frac{\hat{\bar{F}}_1(t)}{\hat{\bar{F}}_2(t)}\right) + \hat{v}_1(t)\frac{n^2[\hat{\lambda}_1^{(0)}(t)]^2}{n_1^2} - \frac{\hat{v}_2(t)}{2}\frac{n^2(\hat{\lambda}_2^{(0)}(t) - \hat{\lambda}_1^{(0)}(t))^2}{n_2^2} + O_p(n^{-3/2})$$

and

$$-\frac{n}{n_i}\hat{c}_i(t)\hat{\lambda}_{i-1}^{(0)}(t) + \left(\frac{n}{n_i}\hat{c}_i(t) + \frac{n}{n_{i+1}}\hat{c}_{i+1}(t)\right)\hat{\lambda}_i^{(0)}(t) - \frac{n}{n_{i+1}}\hat{c}_{i+1}(t)\hat{\lambda}_{i+1}^{(0)}(t) = \log\left(\frac{\hat{\bar{F}}_i(t)}{\hat{\bar{F}}_{i+1}(t)}\right) + \frac{\hat{v}_i(t)}{2}\frac{n^2(\hat{\lambda}_i^{(0)}(t) - \hat{\lambda}_{i-1}^{(0)}(t))^2}{n_i^2} - \frac{\hat{v}_{i+1}}{2}(t)\frac{n^2(\hat{\lambda}_{i+1}^{(0)}(t) - \hat{\lambda}_i^{(0)}(t))^2(t)}{n_{i+1}^2} + O_p(n^{-3/2})$$

$i = 2, 3, \ldots, k-1$. This can be expressed as

$$H^T\hat{\tilde{C}}(t)H\hat{\boldsymbol{\lambda}}^{(0)}(t) = \left(\log\left(\frac{\hat{\bar{F}}_1(t)}{\hat{\bar{F}}_2(t)}\right), \log\left(\frac{\hat{\bar{F}}_2(t)}{\hat{\bar{F}}_3(t)}\right), \ldots, \log\left(\frac{\hat{\bar{F}}_{k-1}(t)}{\hat{\bar{F}}_k(t)}\right)\right) + O_p(n^{-1}) = H^T\log(\hat{\bar{\mathbf{F}}}(t)) + O_p(n^{-1})$$

where $\hat{\tilde{C}}(t) = diag(\frac{n}{n_1}\hat{c}_1(t), \frac{n}{n_2}\hat{c}_2(t), \ldots, \frac{n}{n_k}\hat{c}_k(t))$. Since the elements of $H^T\hat{\tilde{C}}(t)H$ are uniformly strongly consistent estimator of the corresponding elements of $H^T\tilde{C}(t)H = R^{-1}(t)$, we have

$$\hat{\boldsymbol{\lambda}}^{(0)}(t) = R(t)H^T\log(\hat{\bar{\mathbf{F}}}(t)) + O_p(n^{-1})$$

uniformly on $[\tau_1, \tau_2]$ and we have (10).

To prove (11), notice that equations in (S.7) are equivalent to

$$\hat{\boldsymbol{\mu}}^{(0)}(t) = \log(\hat{\bar{\mathbf{F}}}(t)) + \hat{C}(t)\hat{\boldsymbol{\beta}}^{(0)}(t) + O_p(n^{-1}).$$

Since $\hat{\boldsymbol{\beta}}^{(0)}(t) = \hat{\Gamma}^{-1}H\hat{\boldsymbol{\lambda}}^{(0)}(t)$ by (S.3), we have using (9)

$$\hat{\boldsymbol{\mu}}^{(0)}(t) = (I - \hat{\tilde{C}}(t)HR(t)H^T)\log(\hat{\bar{\mathbf{F}}}(t)) + O_p(n^{-1}) = P(t)\tilde{C}^{-1}(t)\log(\hat{\mathbf{F}}(t)) + O_p(n^{-1})$$

where I is the identity matrix and we have (11). □

Proof of Lemma 4: Using the functional delta method, we can easily show that

$$\sqrt{n}(\log(\hat{\bar{\mathbf{F}}}) - \log(\bar{\mathbf{F}})) \overset{w}{\Rightarrow} \mathbf{V}$$

on $[0, \tau_2]$. Since $H^T \log(\mathbf{F}) = \mathbf{0}$ under $\mathcal{H}_0$, equation (10) implies

$$\begin{aligned}\mathbf{U}_{1n}(t) &= \sqrt{n}R(t)H^T \log(\hat{\bar{\mathbf{F}}}(t)) + O_p(n^{-1/2}) \\ &= \sqrt{n}R(t)H^T[\log(\hat{\bar{\mathbf{F}}}(t)) - \log(\mathbf{F}(t))] + O_p(n^{-1/2})\end{aligned}$$

uniformly on $[\tau_1, \tau_2]$. Therefore

$$\{\mathbf{U}_{1n}(t), \tau_1 \leq t \leq \tau_2\} \overset{w}{\Rightarrow} \{R(t)H^T\mathbf{V}(t), \tau_1 \leq t \leq \tau_2\}.$$

Similarly, we have using (11)

$$\mathbf{U}_{2n}(t) = \sqrt{n}(I - \hat{\tilde{C}}(t)HR(t)H^T)[\log(\hat{\bar{\mathbf{F}}}(t)) - \log(\bar{\mathbf{F}}(t))] + O_p(n^{-1/2})$$

uniformly on $[\tau_1, \tau_2]$. Therefore

$$\begin{aligned}\{\mathbf{U}_{2n}(t), , \tau_1 \leq t \leq \tau_2\} &\overset{w}{\Rightarrow} \{(I - \tilde{C}(t)HR(t)H^T)\mathbf{V}(t), \tau_1 \leq t \leq \tau_2\} \\ &\overset{w}{\Rightarrow} \{P(t)\tilde{C}^{-1}(t)\mathbf{V}(t), \tau_1 \leq t \leq \tau_2\}.\end{aligned}$$

Finally, to show independence, we have

$$\mathrm{Cov}(\mathrm{R(t)H^T}\mathbf{V}(\mathrm{t}), \mathrm{P(t)\tilde{C}^{-1}(t)}\mathbf{V}(\mathrm{t})) = \mathrm{R(t)H^T\tilde{C}\tilde{C}^{-1}(\tilde{C} - \tilde{C}HR(t)H^T\tilde{C})} = \mathbf{0}.$$

This completes the proof. □

Proof of Theorem 2: By Taylor series expansion

$$\begin{aligned}-2\log(\mathcal{R}_{02}(t)) &= -2\sum_{i=1}^{k}\sum_{j=1}^{n(t)}(n_{ij} - d_{ij})\log\left(1 + \frac{n_i\hat{\beta}_i^{(0)}(t)}{n_{ij} - d_{ij}}\right) \\ &\quad + 2\sum_{i=1}^{k}\sum_{j=1}^{n(t)} n_{ij}\log\left(1 + \frac{n_i\hat{\beta}_i^{(0)}(t)}{n_{ij}}\right) \\ &= \sum_{i=1}^{k} n_i[\hat{\beta}_i^{(0)}(t)]^2\hat{c}_i(t) + A_n + B_n + C_n\end{aligned}$$

where

$$A_n = -\sum_{i=1}^{k}\frac{2n_i^3[\hat{\beta}_i^{(0)}(t)]^3}{3}\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^2} - \frac{1}{n_{ij}^2}\right\}$$

$$B_n = \sum_{i=1}^{k}\frac{n_i^4[\hat{\beta}_i^{(0)}(t)]^4}{2}\sum_{j=1}^{n(t)}\left\{\frac{1}{(n_{ij} - d_{ij})^3} - \frac{1}{n_{ij}^3}\right\}$$

$$C_n = -\sum_{i=1}^{k}\frac{2n_i^5[\hat{\beta}_i^{(0)}(t)]^5}{5}\sum_{j=1}^{n(t)}\left\{\frac{1}{(1 + \hat{\psi}_{1ij}(t))^5(n_{ij} - d_{ij})^4} - \frac{1}{(1 + \hat{\psi}_{2ij}(t))^5 n_{ij}^4}\right\}$$

and

$$|\hat{\psi}_{1ij}(t)| \leq \frac{n_i\hat{\beta}_i^{(0)}(t)|}{n_{ij} - d_{ij}} \quad \text{and} \quad |\hat{\psi}_{\mathrm{2ij}}(\mathrm{t})| \leq \frac{\mathrm{n_i}|\hat{\beta}_{\mathrm{i}}^{(0)}(\mathrm{t})|}{\mathrm{n_{ij}}}.$$

Using the same notation and the same arguments as in the proof of Lemma 4, we get

$$|A_n| \leq \frac{2n_i^3|\hat{\beta}_i^{(0)}(t)|^3}{3Y_i(\tau_2)} \sum_{j=1}^{n(t)} \left(\frac{1}{n_{ij}-d_{ij}} - \frac{1}{n_{ij}} \right) \leq \frac{2n_i^2[\hat{\beta}_i^{(0)}(t)]^3}{3Y_i(\tau_2)} \hat{c}_i(t) = O_p(n_i^{-1/2})$$

$$\begin{aligned} 0 \leq B_n &= \frac{n_i^4[\hat{\beta}_i^{(0)}(t)]^4}{2} \sum_{j=1}^{n(t)} \frac{d_{ij}}{n_{ij}(n_{ij}-d_{ij})} \left(\frac{1}{(n_{ij}-d_{ij})^2} + \frac{1}{n_{ij}(n_{ij}-d_{ij})} + \frac{1}{n_{ij}^2} \right) \\ &\leq \frac{3n_i^3[\hat{\beta}_i^{(0)}(t)]^4}{2Y_i^2(\tau_2)} \hat{c}_i(t) = O_p(n_i^{-1}). \end{aligned}$$

We also have $|D_n| \leq D_{1n} + D_{2n}$ where

$$D_{1n} = \left| \frac{2n_i^5[\hat{\beta}_i^{(0)}(t)]^5}{5} \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5(n_{ij}-d_{ij})^4} \right|$$

and

$$D_{2n} = \left| \frac{2n_i^5[\hat{\beta}_i^{(0)}(t)]^5}{5} \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5 n_{ij}^4} \right|.$$

In addition,

$$D_{1n} \leq \frac{2n_i^5|\hat{\beta}_i^{(0)}(t)|^5}{5Y_i^4(\tau_2)} \left| \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5} \right| = O_p(n_i^{-1/2})$$

and

$$D_{2n} \leq \frac{2n_i^5|\hat{\beta}_i^{(0)}(t)|^5}{5Y_i^4(\tau_2)} \left| \sum_{j=1}^{n(t)} \frac{1}{(1+\hat{\psi}_{1ij}(t))^5} \right| = O_p(n_i^{-1/2})$$

uniformly over t in $[\tau_1, \tau_2]$. Since $n_i/n \to \gamma_i > 0$, under $\mathcal{H}_0$,

$$\begin{aligned} -2\log(\mathcal{R}_{02}(t)) &= \sum_{i=1}^{k} n_i \hat{c}_i(t)[\hat{\beta}_i^{(0)}(t)]^2 + O_p(n^{-1/2}) \\ &= n[\hat{\boldsymbol{\lambda}}^{(0)}(t)]^T H^T \hat{\hat{C}}(t) H \hat{\boldsymbol{\lambda}}^{(0)}(t) + O_p(n^{-1/2}) \end{aligned}$$

uniformly in t over interval $[\tau_1, \tau_2]$. Now Lemma 5 implies that

$$\mathcal{T}_{02} \xrightarrow{d} \sup_{\tau_1 \leq t \leq \tau_2} [\mathbf{V}(t)]^T H [R(t)] H^T \mathbf{V}(t).$$

It is easy to verify that

$$
\begin{aligned}
[\mathbf{V}(t)]^T HR(t)H^T\mathbf{V}(t) &= \min\{||\mathbf{V}(t)-\boldsymbol{\mu}||^2_{\tilde{\mathbf{w}}(t)},\ H^T\boldsymbol{\mu}=\mathbf{0}\} \\
&= \min\{||\mathbf{V}(t)-\boldsymbol{\mu}||^2_{\tilde{\mathbf{w}}(t)},\ \mu_1=\mu_2=\cdots=\mu_k\} \\
&= \sum_{i=1}^k \tilde{w}_i(t)(V_i(t)-\bar{V}(t))^2 = ||\mathbf{V}(t)-E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)}
\end{aligned}
$$

where

$$\bar{V}(t)=\frac{\sum_{i=1}^k \tilde{w}_i(t)V_i(t)}{\sum_{i=1}^k \tilde{w}_i(t)} = E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0].$$

Therefore

$$\mathcal{T}_{02} \xrightarrow{d} \sup_{\tau_1\le t\le\tau_2} ||\mathbf{V}(t)-E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)}.$$

When all the samples are censored by the same distribution, $c_1(t)=c_2(t)=\ldots=c_k(t)\equiv c(t)$ and $\tilde{w}_i(t)=\gamma_i/c(t), i=1,2,\ldots,k$. This implies that

$$
\begin{aligned}
&||\mathbf{V}(t)-E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)} = \\
&\sum_{i=1}^k \frac{\gamma_i}{c(t)}\left(\gamma_i^{-1/2}W_i(c(t))-\sum_{j=1}^k \gamma_j^{1/2}W_j((c(t))\right)^2 = \\
&\sum_{i=1}^k \frac{\gamma_i(1+c(t))^2}{c(t)}\left(\gamma_i^{-1/2}\frac{W_i(c(t))}{(1+c(t))}-\sum_{j=1}^k \gamma_j^{1/2}\frac{W_j(c(t))}{(1+c(t))}\right)^2 \stackrel{d}{=} \\
&\sum_{i=1}^k \frac{\gamma_i}{u(1-u)}\left(\gamma_i^{-1/2}B_i(u)-\sum_{j=1}^k \gamma_j^{1/2}B_j(u)\right)^2
\end{aligned}
$$

where $u\equiv u(t)=c(t)/(1+c(t))$ and $B_1,B_2,\ldots,B_k$ are independent standard Brownian bridges. The desired conclusion follows immediately from this. □

To prove Lemma 6, we make use of the following result. Let $\mathcal{R}_{01:\pi}(t)$ denote the local empirical likelihood ratio test statistic for testing $\mathcal{H}_0^t$ against $\mathcal{H}_{1:\pi}^t-\mathcal{H}_0^t$ where, for π, a proper subset of $\{1,2,\ldots,k-1\}$,

$$\mathcal{H}_{1:\pi}^t : (\bar{F}_1(t),\bar{F}_2(t),\ldots,\bar{F}_k(t))^T\in\mathcal{I}(\pi)$$

and $\mathcal{I}(\pi)=\{\mathbf{x}\in\mathbf{R}^k,\ x_i=x_j,(i,j)\in\pi^2\}$. That is,

$$\mathcal{R}_{01:\pi}(t)=\frac{\sup\left\{\prod_{i=1}^k L(\bar{F}_i):(\bar{F}_1(t),\bar{F}_2(t),\ldots,\bar{F}_k(t))^T\in\mathcal{I}_0\right\}}{\sup\left\{\prod_{i=1}^k L(\bar{F}_i):(\bar{F}_1(t),\bar{F}_2(t),\ldots,\bar{F}_k(t))^T\in\mathcal{I}(\pi)\right\}}.$$

The following result holds.
Result 2.: Under $\mathcal{H}_0$

$$-2\log(\mathcal{R}_{01:\pi}(t)) \xrightarrow{d} ||E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}(\pi)]-E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0||^2_{\tilde{\mathbf{w}}(t)}\sim\chi^2_{\mathrm{k-1-card}(\pi)}$$

where card(π) is used to denote the cardinality of π.

Proof: Let $\hat{\boldsymbol{\lambda}}^{(0)}(t)$ be as defined before and let $\hat{\boldsymbol{\lambda}}_\pi^{(0)}(t)$ be the corresponding value when H is replaced by $H(\pi)$ where $H(\pi)$ is the sub-matrix H corresponding to constraints in $\mathcal{H}_{1:\pi}^t$. Lemma 4 implies that

$$\mathbf{U}_{1n}^\pi(t) \equiv \sqrt{n}\hat{\boldsymbol{\lambda}}_\pi^{(0)}(t) = \sqrt{n}R_\pi(t)H^T(\pi)\log(\hat{\bar{F}}(t)) + O_p(n^{-1/2}) \qquad (S.8)$$

uniformly on $[\tau_1, \tau_2]$. under $\mathcal{H}_0$. Here $R_\pi(t)$ is the value of $R(t)$ when H is replaced by $H(\pi)$ in Theorem 1. Clearly

$$-2\log(\mathcal{R}_{01:\pi}(t)) = -2\log(\mathcal{R}_{02}(t)) + 2\log(\mathcal{R}_{12:\pi}(t))$$

where $\mathcal{R}_{12:\pi}(t)$ is the local EL statistic for testing $\mathcal{H}_{1:\pi}^t$ against $\mathcal{H}_2^\tau - H_{1:\pi}^t$. The proof of Theorem 2 implies that

$$\begin{aligned}
-2\log(\mathcal{R}_{01:\pi}(t)) &= n[\hat{\boldsymbol{\lambda}}^{(0)}(t)]^T H^T\tilde{C}H\hat{\boldsymbol{\lambda}}^{(0)}(t) - n[\hat{\boldsymbol{\lambda}}_\pi^{(0)}(t)]^T H^T(\pi)\tilde{C}H(\pi)\hat{\boldsymbol{\lambda}}_\pi^{(0)}(t)\\
&+\ O_p(n^{-1/2})\\
&\xrightarrow{d} [\mathbf{V}(t)]^T HR(t)H^T\mathbf{V}(t) - [\mathbf{V}(t)]^T H(\pi)R_\pi(t)H^T(\pi)\mathbf{V}(t).
\end{aligned}$$

To complete the proof we note that

$$\begin{aligned}
&[\mathbf{V}(t)]^T HR(t)H^T\mathbf{V}(t) - [\mathbf{V}(t)]^T H(\pi)R_\pi(t)H^T(\pi)\mathbf{V}(t) =\\
&\min\{||\mathbf{V}(t) - \boldsymbol{\mu}||^2_{\tilde{\mathbf{w}}(t)},\ H^T\boldsymbol{\mu} = \mathbf{0}\} - \min\{||\mathbf{V}(t) - \boldsymbol{\mu}||^2_{\tilde{\mathbf{w}}(t)},\ H^T(\pi)\boldsymbol{\mu} = \mathbf{0}\} =\\
&||\mathbf{V} - E[\mathbf{V}|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)} - ||\mathbf{V}(t) - E[\mathbf{V}(t)|\mathcal{I}(\pi)]||^2_{\tilde{\mathbf{w}}(t)} =\\
&||E_{\tilde{\mathbf{w}}}[\mathbf{V}(t)|\mathcal{I}(\pi)] - E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)} =\\
&\sum_{i=1}^k \tilde{w}_i(t)\left(E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}(\pi)]_i - \bar{V}(t)\right)^2 \sim \chi^2_{k-card(\pi)-1}
\end{aligned}$$

where the third equality follows from the properties of the least squared projection onto nested linear subspaces. □

Proof of Lemma 6: Let $\mathcal{F}$ denote the class of all the subsets of $\{1, 2, \ldots, k-1\}$ and let π be an element of $\mathcal{F}$ and π^c be its complement. Let $\hat{\boldsymbol{\mu}}^{(1)}(t)$ and $\hat{\boldsymbol{\mu}}_\pi^{(0)}(t)$ be the maximizers of (7) under $\mathcal{H}_1^t$ and $\mathcal{H}_{1:\pi}^t$, respectively. Then $\hat{\boldsymbol{\mu}}^{(1)}(t)$ equals $\hat{\boldsymbol{\mu}}_\pi^{(0)}(t)$ for exactly one π. Moreover $\hat{\boldsymbol{\mu}}^{(1)}(t) = \hat{\boldsymbol{\mu}}_\pi^{(0)}(t)$ if and only if (Wollan and Dykstra (1985) and El Barmi and Dyksta (1999))

$$h_s(\hat{\boldsymbol{\mu}}_\pi^{(0)}(t)) < 0, \quad \forall s \in \pi^c \quad \text{and} \quad \hat{\boldsymbol{\lambda}}_\pi^{(0)}(\mathrm{t}) > 0.$$

Using the same steps as in the proof of Lemma 4, under $\mathcal{H}_0$, (S8) holds and

$$\sqrt{n}(\hat{\boldsymbol{\mu}}_\pi^{(0)}(t) - \boldsymbol{\mu}^{(0)}(t)) = P_\pi(t)\tilde{C}^{-1}(t)[\log(\hat{\bar{\mathbf{F}}}(t)) - \log(\bar{\mathbf{F}}(t))] + O_p(n^{-1/2}) \quad (S.9)$$

uniformly on $[\tau_1, \tau_2]$ where $P_\pi(t)$ is value of P is Theorem 1 when H is replaced by $H(\pi)$. Moreover, if $\mathbf{U}_{2n}^\pi(t) = \sqrt{n}(\hat{\boldsymbol{\mu}}_\pi^{(0)}(t) - \boldsymbol{\mu}^{(0)}(t))$, then

$$(\mathbf{U}_{1n}^\pi, \mathbf{U}_{2n}^\pi)^T \overset{w}{\Rightarrow} (\mathbf{U}_1^\pi, \mathbf{U}_2^\pi)^T$$

on $[\tau_1, \tau_2]$ where

$$\mathbf{U}_1^\pi(t) = R_\pi(t)H^T(\pi)\mathbf{V}(t) \quad \text{and} \quad \mathbf{U}_2^\pi(\mathrm{t}) = \mathrm{P}_\pi(\mathrm{t})\tilde{\mathrm{C}}^{-1}(\mathrm{t})\mathbf{V}(\mathrm{t})$$

and $\mathbf{U}_1^\pi$ and $\mathbf{U}_2^\pi$ are independent processes. Consequently, using (S.8), (S.9), independence of $\mathbf{U}_1^\pi$ and $\mathbf{U}_2^\pi$ and Result 2, we get

$$\begin{aligned}
-2\log(\mathcal{R}_{01}(t)) &= -2\sum_{\pi\in\mathcal{F}} \log\left(\mathcal{R}_{01:\pi}(t)\right) I[h_s(\hat{\boldsymbol{\mu}}_\pi^{(0)}(t)) < 0, \forall s \in \pi^c, \hat{\boldsymbol{\lambda}}_\pi^{(0)}(t) > 0] \\
&= -2\sum_{\pi\in\mathcal{F}} \log\left(\mathcal{R}_{01:\pi}(t)\right) I[\sqrt{n}H^T(\pi^c)(\hat{\boldsymbol{\mu}}_\pi^{(0)}(t) - \boldsymbol{\mu}^{(0)}(t)) < \mathbf{0}, \\
&\quad \sqrt{n}\hat{\boldsymbol{\lambda}}_\pi^{(0)}(t) > \mathbf{0}] \\
&= -2\sum_{\pi\in\mathcal{F}} \log\left(\mathcal{R}_{01:\pi}(t)\right) I[\sqrt{n}H^T(\pi^c)P_\pi(t)\tilde{C}^{-1}(t)[\log(\hat{\bar{\mathbf{F}}}(t)) \\
&- \log(\bar{\mathbf{F}}(t))] + O_p(n^{-1/2}) < \mathbf{0}, \sqrt{n}R_\pi(t)H^T(\pi)[\log(\hat{\bar{\mathbf{F}}}(t)) \\
&- \log(\bar{\mathbf{F}}(t))] + O_p(n^{-1/2}) > \mathbf{0}] \\
&\xrightarrow{d} \sum_{\pi\in\mathcal{F}} \left[[\mathbf{V}(t)]^T \left[HR(t)H^T - H(\pi)R_\pi(t)H^T(\pi)\right] \mathbf{V}(t)\right] \\
&\times\ I[H^T(\pi^c)P_\pi(t)\tilde{C}^{-1}\mathbf{V}(t) < \mathbf{0}, R_\pi(t)H^T(\pi)\mathbf{V}(t) \geq \mathbf{0}] \\
&\xrightarrow{d} \sum_{\pi\in\mathcal{F}} \left[[\mathbf{V}(t)]^T HR(t)H^T\mathbf{V}(t) - [\mathbf{V}H(\pi)(t)]^T R_\pi(t)H^T(\pi)\mathbf{V}(t)\right] \\
&\times\ I[H^T(\pi^c)\mathbf{U}_2^\pi(t) < \mathbf{0}, \mathbf{U}_1^\pi(t) \geq \mathbf{0}] \\
&= \sum_{\pi\in\mathcal{F}} ||E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}(\pi)] - E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_0]||^2_{\tilde{\mathbf{w}}(t)} \\
&\times\ I[H^T(\pi^c)P(\pi)\tilde{C}^{-1}\mathbf{V}(t) < \mathbf{0}, R_\pi(t)H^T(\pi)\mathbf{V}(t) \geq \mathbf{0}] \\
&\stackrel{d}{=} \sum_{i=1}^k \tilde{w}_i(t)\left(E_{\tilde{\mathbf{w}}(t)}[\mathbf{V}(t)|\mathcal{I}_1]_i - \bar{V}(t)\right)^2
\end{aligned}$$

where the last equality follows from Lemma 3.3 in El Barmi (1996). □

Proof of Theorem 3: It follows from the proof of Theorem 2 that the result in Lemma 3 hold uniformly on any interval $[\tau_1, \tau_2]$ where τ_1 and τ_2 satisfy the conditions of Theorem 3. Therefore we have the desired conclusion. □

References

Aitchison, J. and Silvey, S. D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics*, 29, 813-828.

El Barmi, H. (1996). Emipirical likelihood ratio test for or against a set if inequality constraints, *Journal of Statistical Planning and Inference*, 55, 191-204.

El Barmi, H. and Dykstra, R. (1999). Likelihood ratio test against a set of inequality constraints. *Journal of Nonparametric Statistics*, 11, 233-250.

Giljbels, I. and Wang, J. L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *Journal of Multivariate Analysis*, 47, 210-229.

Wollan P. C. and Dykstra, R. L. (1986). Conditional tests with an order restriction as a null hypothesis. In R. L. Dykstra and T. Robertson and F. T. Wright (eds). *Advances in Order Restricted Statistical Inference*, Springer-Verlag, 279-295.

Li, G. (1995). On nonparamteric likelihood estimation of survival probabilities for censored data. *Statistics and Probability Letters*, 25, 95-104.