Supplementary File to the Manuscript "Empirical Likelihood Based on Synthetic Right Censored Data"

Wei Liang
School of Mathematical Sciences, Xiamen University, Xiamen, China
Hongsheng Dai
Department of Mathematical Sciences, University of Essex, Colchester, UK

1. Detailed Proof of Lemma A.1

Lemma 1.1. *Assume* $\boldsymbol{E}\left(\frac{\delta\, g(Z,\theta_0)}{1-G(Z)}\right)^4 < \infty$, $\boldsymbol{E}\left(\frac{\delta}{1-G(Z)}\right)^4 < \infty$, *we have*

$$(i) \qquad \max_{1\le k\le N} |W_{nk}| = o_p(n^{1/2}).$$

$$(ii) \qquad \sqrt{n}\left(\frac{1}{N}\sum_{k=1}^{N} W_{nk}\right) \to N(0,\, \sigma_1^2), \quad \textit{in dist.}$$

$$(iii) \qquad \frac{1}{N}\sum_{k=1}^{N} W_{nk}^2 = O_p(1).$$

$$(iv) \qquad \frac{1}{N}\sum_{k=1}^{N} W_{nk}^2 = \frac{1}{2}\hat{\sigma}_{2A}^2 + o_p(1), \qquad \frac{1}{N}\sum_{k=1}^{N} W_{nk}^2 = \frac{1}{2}\hat{\sigma}_{2B}^2 + o_p(1).$$

Proof. (i) Lemma 11.2 in Owen (2001) shows that under the condition $\mathbf{E}\,Y_i^4 < \infty$, $Z_n = \max_{1\le i\le n} |Y_i| = o(n^{1/4})$. Therefore,

$$\begin{aligned}
\max_{1\le k\le N} |W_k| &= \max_{i<j} \frac{1}{2}\left|\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-G_i)(1-G_j)}\right| \le \frac{1}{2}\max_{i,j}\left|\frac{\delta_i\delta_j g_i(\theta_0)}{(1-G_i)(1-G_j)}\right| \\
&\le \frac{1}{2}\max_i\left|\frac{\delta_i g_i}{1-G_i}\right| \max_j\left|\frac{\delta_j}{1-G_j}\right| = o(n^{1/2}), \\
\max_{1\le k\le N} |W_{nk}| &\le \max_{1\le k\le N} |W_{nk}-W_k| + \max_{1\le k\le N} |W_k|
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{2}\max_{i<j}\left|\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-\hat{G}_i)(1-\hat{G}_j)}-\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-\hat{G}_i)(1-G_j)}\right| \\
&\quad +\frac{1}{2}\max_{i<j}\left|\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-\hat{G}_i)(1-G_j)}-\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-G_i)(1-G_j)}\right|+o(n^{1/2}) \\
&= \frac{1}{2}\max_{i<j}\left|\frac{1-G_i}{1-\hat{G}_i}\right|\left|\frac{\hat{G}_j-G_j}{1-\hat{G}_j}\right|\left|\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-G_i)(1-G_j)}\right| \\
&\quad +\frac{1}{2}\max_{i<j}\left|\frac{\hat{G}_i-G_i}{1-\hat{G}_i}\right|\left|\frac{\delta_i\delta_j(g_i(\theta_0)+g_j(\theta_0))}{(1-G_i)(1-G_j)}\right|+o(n^{1/2}) \\
&\leq \frac{1}{2}\left(\max_i\left|\frac{1-G_i}{1-\hat{G}_i}\right|\max_j\left|\frac{\hat{G}_j-G_j}{1-\hat{G}_j}\right|+\max_i\left|\frac{\hat{G}_i-G_i}{1-\hat{G}_i}\right|\right)\max_k|W_k|+o(n^{1/2})
\end{aligned}
$$

Following Zhou (1992),

$$
\sup_{0\leq z\leq Z_{n:n}}\left|\frac{\hat{G}(z)-G(z)}{1-\hat{G}(z)}\right|=O_p(1),
$$

we get $\max_k|W_{nk}|=o_p(n^{1/2})$.

(ii) We rewrite $N^{-1}\sum_{k=1}^N W_{nk}$ as

$$
\begin{aligned}
\frac{1}{N}\sum_{k=1}^N W_{nk} &= \frac{1}{2N}\sum_{i\neq j}\frac{\delta_i\delta_j g_i(\theta_0)}{(1-\hat{G}_i)(1-\hat{G}_j)}=\frac{1}{2N}\sum_{i=1}^n\frac{\delta_i g_i(\theta_0)}{1-\hat{G}_i}\left(\sum_{j=1}^n\frac{\delta_j}{1-\hat{G}_j}-\frac{\delta_i}{1-\hat{G}_i}\right) \\
&= \left(\frac{1}{n}\sum_{i=1}^n\frac{\delta_i g_i(\theta_0)}{1-\hat{G}_i}\right)\left(\frac{1}{n-1}\sum_{i=1}^n\frac{\delta_i}{1-\hat{G}_i}\right)-\frac{1}{n(n-1)}\sum_{i=1}^n\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}.
\end{aligned}
$$

Noting that

$$
\left|\frac{1}{n(n-1)}\sum_{i=1}^n\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}\right|\leq\left|\frac{1}{n(n-1)}\sum_{i=1}^n\frac{\delta_i g_i(\theta_0)}{(1-G_i)^2}\right|\max_i\left|\frac{1-G_i}{1-\hat{G}_i}\right|^2=O_p(n^{-1}),
$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{1-\hat{G}_i} = \int_0^\infty g(t,\,\theta_0)\,\mathrm{d}\hat{F}(t) = \int_0^\infty g(t,\,\theta_0)\,\mathrm{d}(\hat{F}(t)-F(t)),$$

and $\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{1-\hat{G}_i} = \int_0^\infty \mathrm{d}\hat{F}(t) = 1$, we get

$$\sqrt{n}\left(\frac{1}{N}\sum_{k=1}^{N}W_{nk}\right) = \sqrt{n}\int_0^\infty g(t,\,\theta_0)\,\mathrm{d}(\hat{F}(t)-F(t)) + o_p(1).$$

Following Corollary 1.2 in Stute (1995) we have the conclusion (ii).

(iii) We rewrite $N^{-1}\sum_{k=1}^{N}W_{nk}^2$ as

$$\begin{aligned}
\frac{1}{N}\sum_{k=1}^{N}W_{nk}^2 &= \frac{1}{4N}\sum_{i<j}\frac{\delta_i\delta_j(g_i^2(\theta_0)+2g_i(\theta_0)g_j(\theta_0)+g_j^2(\theta_0))}{(1-\hat{G}_i)^2(1-\hat{G}_j)^2} \\
&= \frac{1}{4N}\sum_{i\neq j}\frac{\delta_i\delta_j g_i^2(\theta_0)}{(1-\hat{G}_i)^2(1-\hat{G}_j)^2} + \frac{1}{4N}\sum_{i\neq j}\frac{\delta_i\delta_j g_i(\theta_0)g_j(\theta_0)}{(1-\hat{G}_i)^2(1-\hat{G}_j)^2} \\
&= \frac{1}{4N}\sum_{i=1}^{n}\left(\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^2}\left(\sum_{j=1}^{n}\frac{\delta_j}{(1-\hat{G}_j)^2} - \frac{\delta_i}{(1-\hat{G}_i)^2}\right)\right) \\
&\quad + \frac{1}{4N}\left(\left(\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}\right)^2 - \sum_{i=1}^{n}\left(\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}\right)^2\right) \\
&= \frac{n}{2(n-1)}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^2}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{(1-\hat{G}_i)^2}\right) \\
&\quad + \frac{n}{2(n-1)}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}\right)^2 - \frac{1}{n(n-1)}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^4}. \quad (1.1)
\end{aligned}$$

The last term of (1.1) is bounded above by,

$$\left|\frac{1}{n(n-1)}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-G_i)^4}\right| \max_i \left|\frac{1-G_i}{1-\hat{G}_i}\right|^4 = O_p(n^{-1})$$

and can be omitted. In addition, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-G_i)^2}\right| = \left|\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{G}_i-G_i)(1-\hat{G}_i+1-G_i)}{(1-\hat{G}_i)^2}\frac{\delta_i g_i^2(\theta_0)}{(1-G_i)^2}\right|$$
$$\leq \left(\max_i\left|\frac{\hat{G}_i-G_i}{1-\hat{G}_i}\right|+\max_i\left|\frac{(\hat{G}_i-G_i)(1-G_i)}{(1-\hat{G}_i)^2}\right|\right)\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-G_i)^2}=O_p(1),$$

and

$$\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-G_i)^2}\right|=O_p(1),\quad \left|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{(1-\hat{G}_i)^2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{(1-G_i)^2}\right|=O_p(1).$$

Therefore, part (iii) is proved.

(iv) Since $\hat{\theta}$ is a consistent estimator of θ_0 and

$$\frac{1}{N}\sum_{k=1}^{N}W_{nk}^2-\frac{1}{2}\hat{\sigma}_{2A}^2=\frac{1}{4N}\sum_{i<j}\frac{\delta_i\,\delta_i}{(1-\hat{G}_i)^2(1-\hat{G}_j)^2}\left[(g_i(\theta_0)+g_j(\theta_0))^2-(g_i(\hat{\theta})+g_j(\hat{\theta}))^2\right],$$

we have $\frac{1}{N}\sum_{k=1}^{N}W_{nk}^2=\frac{1}{2}\hat{\sigma}_{2A}^2+o_p(1)$.

From equation (1.1) we have,

$$\frac{1}{N}\sum_{k=1}^{N}W_{nk}^2-\frac{1}{2}\hat{\sigma}_{2B}^2 = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^2}-\hat{\sigma}_{22}^2\right)\hat{\sigma}_{20}^2 \tag{1.2}$$
$$+\frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}-\hat{\sigma}_{21}^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i(\theta_0)}{(1-\hat{G}_i)^2}+\hat{\sigma}_{21}^2\right)+o_p(1).$$

Since $\hat{\theta}$ is a consistent estimator of θ_0, then the first term of (1.2) becomes

$$\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i g_i^2(\theta_0)}{(1-\hat{G}_i)^2}-\hat{\sigma}_{22}^2\right|=\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i\, g_i^2(\theta_0)}{(1-\hat{G}_i)^2}-\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i\, g_i^2(\hat{\theta})}{(1-\hat{G}_i)^2}\right|$$

$$\leq \max_i \left| g_i(\theta_0) - g_i(\hat{\theta}) \right| \max_i \left| \frac{1 - G_i}{1 - \hat{G}_i} \right|^2 \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i \left(g_i(\theta_0) + g_i(\hat{\theta}) \right)}{(1 - G_i)^2} \right| = o_p(1),$$

and the second term of (1.2) becomes

$$\left| \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i g_i(\theta_0)}{(1 - \hat{G}_i)^2} - \hat{\sigma}_{21}^2 \right| = \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i \, g_i(\theta_0)}{(1 - \hat{G}_i)^2} - \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i \, g_i(\hat{\theta})}{(1 - \hat{G}_i)^2} \right|$$
$$\leq \max_i \left| g_i(\theta_0) - g_i(\hat{\theta}) \right| \max_i \left| \frac{1 - G_i}{1 - \hat{G}_i} \right|^2 \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i}{(1 - G_i)^2} \right| = o_p(1).$$

Therefore we have

$$\frac{1}{N} \sum_{k=1}^{N} W_{nk}^2 = \frac{1}{2} \hat{\sigma}_{2B}^2 + o_p(1).$$

This implies that both of $\hat{\sigma}_{2A}^2$ and $\hat{\sigma}_{2B}^2$ are asymptotically equivalent to $\frac{2}{N} \sum_{k=1}^{N} W_{nk}^2$, i.e. they converge to the same value in probability as $n \to \infty$.

□

2. Comparison for, the original EL, pairwise EL and three-point-mean EL

2.1. Theoretical Comparisons

Here, we will illustrate the difference of original EL, pair-wise mean EL estimator and three-point-mean EL by calculating the derivatives of their empirical likelihood functions. We focus on i.i.d. complete data with estimating equation $\mathbf{E}\, T_i(\theta_0) = \mathbf{E}\,(T_i - \theta_0) = 0$.

Recall the definitions of these empirical likelihood functions first. The empirical log-likelihood ratio statistics based on original data is $\mathcal{L}_O(\theta) = -2 \log \mathcal{R}_O(\theta)$, where

$$\mathcal{R}_O(\theta) = \sup \left\{ \prod_{i=1}^{n} np_i \,\middle|\, \sum_{i=1}^{n} p_i T_i(\theta) = 0, \, \sum_{i=1}^{n} p_i = 1, \, p_i \geq 0, \, i = 1, \, 2, \, \cdots, \, n \right\}.$$

Denote the corresponding log-likelihood ratio statistics based on pairwise mean data and three-point-mean data as

$$\mathcal{L}(\theta) = -2\log\mathcal{R}(\theta)/n \quad \text{and} \quad \mathcal{L}'(\theta) = \frac{-2\log\mathcal{R}'(\theta)}{(n+1)(n+2)/2},$$

respectively, where

$$\begin{aligned}
\mathcal{R}(\theta) &= \sup\left\{\prod_{k=1}^{N} Np_k \,\middle|\, \sum_{k=1}^{N} p_k M_k(\theta) = 0, \sum_{k=1}^{N} p_k = 1, p_k \geq 0, k = 1, 2, \cdots, N\right\}, \\
\mathcal{R}'(\theta) &= \sup\left\{\prod_{k=1}^{N'} N'p_k \,\middle|\, \sum_{k=1}^{N'} p_k M'_k(\theta) = 0, \sum_{k=1}^{N'} p_k = 1, p_k \geq 0, k = 1, 2, \cdots, N'\right\},
\end{aligned}$$

$M_k(\theta)$ is the pairwise mean estimating equation, $N = n(n+1)/2$, and $M'_k(\theta)$ is the three-point-mean estimating equation, $N' = n(n+1)(n+2)/6$.

By using the Lagrange multiplier, we have

$$\begin{aligned}
\mathcal{L}_O(\theta) &= -2\log\mathcal{R}_O(\theta) = 2\sum_{i=1}^{n} \log\left(1 + \lambda_O(\theta)T_i(\theta)\right), \\
\mathcal{L}(\theta) &= -2n^{-1}\log\mathcal{R}(\theta) = 2n^{-1}\sum_{k=1}^{N} \log\left(1 + \lambda(\theta)M_k(\theta)\right), \\
\mathcal{L}'(\theta) &= \frac{-2\log\mathcal{R}(\theta)}{(n+1)(n+2)/2} = \frac{4}{(n+1)(n+2)}\sum_{k=1}^{N}{}' \log\left(1 + \lambda'(\theta)M'_k(\theta)\right),
\end{aligned}$$

where $\lambda_O(\theta)$ satisfies

$$\sum_{i=1}^{n} \frac{T_i(\theta)}{1 + \lambda_O(\theta)T_i(\theta)} = 0, \tag{2.1}$$

$\lambda(\theta)$ satisfies

$$\sum_{k=1}^{N} \frac{M_k(\theta)}{1 + \lambda(\theta)M_k(\theta)} = 0, \tag{2.2}$$

and $\lambda'(\theta)$ satisfies

$$\sum_{k=1}^{N'} \frac{M'_k(\theta)}{1+\lambda'(\theta)M'_k(\theta)} = 0. \tag{2.3}$$

Denote $\hat{\theta} = \bar{T}$ which is the minimum point of these likelihood functions, such that $\lambda_O(\hat{\theta}) = \lambda(\hat{\theta}) = \lambda'(\hat{\theta}) = 0$. Then we have the following proposition.

Proposition 2.1. *For any $\tilde{\theta}$ and a constant c such that $|\tilde{\theta} - \hat{\theta}| \leq cn^{-1/2}$, the second derivative of the above three empirical log-likelihood ratio function are given by*

$$\begin{aligned}
D_1 &= \frac{\mathrm{d}^2\mathcal{L}_O(\theta)}{\mathrm{d}\theta^2}\Big|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{1}{n}\sum_{i=1}^n T_i^2(\tilde{\theta})} + O_p(n^{-1/2}),\\
D_2 &= \frac{\mathrm{d}^2\mathcal{L}(\theta)}{\mathrm{d}\theta^2}\Big|_{\theta=\tilde{\theta}} = \frac{2(n+1)}{\frac{n+2}{n+1}\left(\frac{1}{n}\sum_{i=1}^n T_i^2(\tilde{\theta})\right) + \frac{n}{n+1}\left(\frac{1}{n}\sum_{i=1}^n T_i(\tilde{\theta})\right)^2} + O_p(n^{-1/2}),\\
D_3 &= \frac{\mathrm{d}^2\mathcal{L}'(\theta)}{\mathrm{d}\theta^2}\Big|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{n+3}{n+1}\left(\frac{1}{n}\sum_{i=1}^n T_i^2(\tilde{\theta})\right) + \frac{2n}{n+1}\left(\frac{1}{n}\sum_{i=1}^n T_i(\tilde{\theta})\right)^2} + O_p(n^{-1/2}).
\end{aligned}$$

□

Remark 2.1. *Notice that, for any $|\tilde{\theta} - \hat{\theta}| \leq cn^{-1/2}$, we have*

$$\left(\frac{1}{n}\sum_{i=1}^n (T_i - \tilde{\theta})\right)^2 = \left(\frac{1}{n}\sum_{i=1}^n (T_i - \hat{\theta}) + (\hat{\theta} - \tilde{\theta})\right)^2 = \left(\hat{\theta} - \tilde{\theta}\right)^2 = O_p(n^{-1}).$$

Therefore, from the above proposition, we know that the difference between D_2 and D_1 is $O_p(1)$, within a small neighborhood of $\hat{\theta}$.

This means, comparing to D_1, the value of D_2 is slightly smaller which will lead to a wider confidence interval and achieve more accurate coverage probability. Meanwhile, comparing to D_1 and D_2, the value of D_3 is too small and will lead to a much larger confidence interval.

In summary, the above proposition tells us pair-wise mean data will give approximate *coverage probabilities, not too large (as the three-value mean data) and not too small (as the original EL methods). This is also confirmed by the simulation studies, provided in the following section.*

□

We now provide the proof of this proposition.

Proof. The first and second derivatives of $\mathcal{L}_O(\theta)$ are

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{L}_O(\theta)}{\mathrm{d}\theta} &= 2\sum_{i=1}^{n}\frac{1}{1+\lambda_O(\theta)T_i(\theta)}\left(\frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta}T_i(\theta)+\lambda_O(\theta)\frac{\mathrm{d}T_i(\theta)}{\mathrm{d}\theta}\right)\\
&= 2\sum_{i=1}^{n}\frac{T_i(\theta)}{1+\lambda_O(\theta)T_i(\theta)}\frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta} - 2\sum_{i=1}^{n}\lambda_O(\theta)\frac{1}{1+\lambda_O(\theta)T_i(\theta)}\\
&= -2n\lambda_O(\theta),\\
\frac{\mathrm{d}^2\mathcal{L}_O(\theta)}{\mathrm{d}\theta^2} &= -2n\frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta}.
\end{aligned}$$

Similarly, we have

$$\frac{\mathrm{d}\mathcal{L}(\theta)}{\mathrm{d}\theta} = -(n+1)\lambda(\theta), \qquad \frac{\mathrm{d}^2\mathcal{L}(\theta)}{\mathrm{d}\theta^2} = -(n+1)\frac{\mathrm{d}\lambda(\theta)}{\mathrm{d}\theta},$$
$$\frac{\mathrm{d}\mathcal{L}'(\theta)}{\mathrm{d}\theta} = -\frac{2n}{3}\lambda'(\theta), \qquad \frac{\mathrm{d}^2\mathcal{L}'(\theta)}{\mathrm{d}\theta^2} = -\frac{2n}{3}\frac{\mathrm{d}\lambda'(\theta)}{\mathrm{d}\theta},$$

Since $\lambda_O(\theta)$ satisfies (2.1), we have

$$\frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta} = -\left(\frac{1}{n}\sum_{i=1}^{n}\frac{T_i^2(\theta)}{(1+\lambda_O(\theta)T_i(\theta))^2}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(1+\lambda_O(\theta)T_i(\theta))^2}\right).$$

Similarly,

$$\frac{\mathrm{d}\lambda(\theta)}{\mathrm{d}\theta} = -\left(\frac{1}{N}\sum_{k=1}^{N}\frac{M_k^2(\theta)}{(1+\lambda(\theta)M_k(\theta))^2}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\frac{1}{(1+\lambda(\theta)M_k(\theta))^2}\right),$$

$$\frac{\mathrm{d}\lambda'(\theta)}{\mathrm{d}\theta} \quad = \quad -\left(\frac{1}{N}\sum_{k=1}^{N}\frac{M_k'^2(\theta)}{(1+\lambda'(\theta)M_k'(\theta))^2}\right)^{-1}\left(\frac{1}{N}\sum_{k=1}^{N}\frac{1}{(1+\lambda'(\theta)M_k'(\theta))^2}\right).$$

Note that,

$$\left.\frac{\mathrm{d}\mathcal{L}_O(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}} = \left.\frac{\mathrm{d}\mathcal{L}(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}} = \left.\frac{\mathrm{d}\mathcal{L}'(\theta)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}} = 0.$$

Consider $\tilde{\theta}$ is in a neighborhood of $\hat{\theta}$, i.e. $|\tilde{\theta}-\hat{\theta}| \le cn^{-1/2}$. Since $\mathrm{d}\lambda_O(\theta)/\mathrm{d}\theta = O_p(1)$, then

$$|\lambda_O(\tilde{\theta}) - \lambda_O(\hat{\theta})| = |\lambda_O(\tilde{\theta}) - 0| = O_p(n^{-1/2}). \tag{2.4}$$

Denote

$$\Delta(\lambda,\, \theta) = \frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta} = -\left(\frac{1}{n}\sum_{i=1}^{n}\frac{T_i^2(\theta)}{(1+\lambda T_i(\theta))^2}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{(1+\lambda T_i(\theta))^2}\right), \tag{2.5}$$

then

$$\Delta(\lambda,\, \theta) = \Delta(0,\, \theta) + \frac{\partial\Delta(\lambda^*,\, \theta)}{\partial\lambda}(\lambda - 0) = -\left(\frac{1}{n}\sum_{i=1}^{n}T_i^2(\theta)\right)^{-1} + \frac{\partial\Delta(\lambda^*,\, \theta)}{\partial\lambda}(\lambda-0). \tag{2.6}$$

where λ^* is between λ and 0. Therefore, (2.4), (2.5) and (2.6) imply

$$\left.\frac{\mathrm{d}\lambda_O(\theta)}{\mathrm{d}\theta}\right|_{\theta=\tilde{\theta}} = -\left(\frac{1}{n}\sum_{i=1}^{n}T_i^2(\tilde{\theta})\right)^{-1} + O_p(n^{-1/2}).$$

Hence, we get

$$\left.\frac{\mathrm{d}^2\mathcal{L}_O(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{1}{n}\sum_{i=1}^{n}T_i^2(\tilde{\theta})} + O_p(n^{-1/2}).$$

Similarly, we have

$$\left.\frac{\mathrm{d}^2\mathcal{L}(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{n+1}{\frac{1}{N}\sum_{k=1}^{N}M_k^2(\tilde{\theta})} + O_p(n^{-1/2}), \quad \left.\frac{\mathrm{d}^2\mathcal{L}'(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{3}{N'}\sum_{k=1}^{N'}M_k'^2(\tilde{\theta})} + O_p(n^{-1/2}).$$

Because of

$$
\begin{aligned}
\frac{1}{N}\sum_{k=1}^{N} M_k^2(\theta) &= \frac{n}{2(n+1)}\left(\frac{1}{n}\sum_{i=1}^{n} T_i(\theta)\right)^2 + \frac{n+2}{2(n+1)}\left(\frac{1}{n}\sum_{i=1}^{n} T_i^2(\theta)\right), \\
\frac{1}{N'}\sum_{k=1}^{N'} M_k'^2(\theta) &= \frac{2n}{3(n+1)}\left(\frac{1}{n}\sum_{i=1}^{n} T_i(\theta)\right)^2 + \frac{n+3}{3(n+1)}\left(\frac{1}{n}\sum_{i=1}^{n} T_i^2(\theta)\right),
\end{aligned}
$$

the second derivatives of likelihood functions at their minimum are

$$
\begin{aligned}
D_1 &= \left.\frac{\mathrm{d}^2\mathcal{L}_O(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{1}{n}\sum_{i=1}^{n} T_i^2(\tilde{\theta})} + O_p(n^{-1/2}), \\
D_2 &= \left.\frac{\mathrm{d}^2\mathcal{L}(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{2(n+1)}{\frac{n+2}{n+1}\left(\frac{1}{n}\sum_{i=1}^{n} T_i^2(\tilde{\theta})\right) + \frac{n}{n+1}\left(\frac{1}{n}\sum_{i=1}^{n} T_i(\tilde{\theta})\right)^2} + O_p(n^{-1/2}), \\
D_3 &= \left.\frac{\mathrm{d}^2\mathcal{L}'(\theta)}{\mathrm{d}\theta^2}\right|_{\theta=\tilde{\theta}} = \frac{2n}{\frac{n+3}{n+1}\left(\frac{1}{n}\sum_{i=1}^{n} T_i^2(\tilde{\theta})\right) + \frac{2n}{n+1}\left(\frac{1}{n}\sum_{i=1}^{n} T_i(\tilde{\theta})\right)^2} + O_p(n^{-1/2}).
\end{aligned}
$$

□

2.2. Simulation Comparison

In this subsection, we use simulation studies to compare the performance of the original EL method, the pair-wise mean method and the three-value mean method. The parameter of interests, θ, is the mean of T, therefore the estimating equation is $g(T, \theta) = T - \theta$. Exponential, LogNorm, Normal, t and χ^2 distributions are considered as the underlying lifetime distribution F. Based on 20,000 sets of simulated data, we construct three different types of confidence intervals.

The simulation results in Table 1 of this supplementary file verify the explanation in Remark 2.1, i.e. pair-wise mean gives the best coverage probability estimates, but original EL provides much lower coverage probabilities. The coverage probabilities of three-value mean method are too large. It confirms that using the pair-wise data is the best choice.

Table 1: Coverage probabilities under different distributions

sample size	nomial level = 0.90			nomial level = 0.95			nomial level = 0.99		
	EL	pairwiseEL	threeEL	EL	pairwiseEL	threeEL	EL	pairwiseEL	threeEL
	Exp(1)								
20	0.8604	0.8860	0.9442	0.9138	0.9336	0.9758	0.9638	0.9788	0.9954
40	0.8858	0.8996	0.9368	0.9342	0.9496	0.9746	0.9820	0.9890	0.9974
60	0.8898	0.9008	0.9300	0.9370	0.9484	0.9690	0.9840	0.9912	0.9972
80	0.8942	0.9026	0.9286	0.9464	0.9554	0.9706	0.9870	0.9912	0.9968
100	0.8926	0.8988	0.9194	0.9428	0.9504	0.9676	0.9880	0.9914	0.9966
	LogNorm(0, 1)								
20	0.8052	0.8362	0.9122	0.8666	0.8964	0.9566	0.9354	0.9598	0.9900
40	0.8420	0.8592	0.9084	0.8982	0.9186	0.9606	0.9622	0.9802	0.9948
60	0.8624	0.8794	0.9134	0.9184	0.9322	0.9622	0.9718	0.9838	0.9934
80	0.8706	0.8816	0.9150	0.9264	0.9366	0.9618	0.9776	0.9868	0.9946
100	0.8732	0.8836	0.9138	0.9296	0.9418	0.9626	0.9808	0.9862	0.9932
	Norm(0, 1)								
20	0.8804	0.9028	0.9608	0.9298	0.9510	0.9908	0.9810	0.9922	0.9992
40	0.8918	0.9048	0.9416	0.9390	0.9532	0.9802	0.9856	0.9926	0.9986
60	0.9008	0.9084	0.9340	0.9506	0.9576	0.9810	0.9908	0.9926	0.9988
80	0.9042	0.9092	0.925	0.9506	0.9556	0.9720	0.9900	0.9920	0.9980
100	0.9018	0.9064	0.9200	0.9496	0.9534	0.9684	0.9916	0.9946	0.9994
	t(5)								
20	0.8700	0.8940	0.9556	0.9232	0.9468	0.9858	0.9782	0.9894	0.9980
40	0.8868	0.9000	0.9382	0.9388	0.9498	0.9784	0.9844	0.9914	0.9980
60	0.8880	0.8984	0.9274	0.9426	0.9504	0.9726	0.9854	0.9906	0.9968
80	0.8908	0.8978	0.9192	0.9434	0.9490	0.9698	0.9872	0.9908	0.9962
100	0.8918	0.8952	0.9146	0.9448	0.9498	0.9666	0.9886	0.9914	0.9970
	$\chi^2(3)$								
20	0.8578	0.8810	0.9464	0.9130	0.9358	0.9800	0.9678	0.9818	0.9972
40	0.8926	0.9046	0.9388	0.9404	0.9542	0.9790	0.9844	0.9908	0.9982
60	0.8888	0.8970	0.9264	0.939	0.9482	0.9694	0.9850	0.9898	0.9964
80	0.8934	0.9026	0.9274	0.9462	0.9556	0.9724	0.9886	0.9932	0.9976
100	0.9042	0.9092	0.9284	0.9504	0.9566	0.9726	0.9908	0.9932	0.9958

References

Owen, A. B. (2001). *Empirical Likelihood*. Chapman and Hall, London.

Stute, W. (1995). The central limit theorem under random censorship. *Annals of Statistics* **23(2)**, 422–439.

Zhou M. (1992). Asymptotic normality of the synthetic data regression estimator for censored survival data. *Annals of Statistics*. **20**, 1002-1021.