Heteroscedastic survival data analysis with accelerated failure time model

Lili Yu[1] and Zhezhen Jin[2]

Georgia Southern University[1] *and Columbia University*[2]

Supplementary Material

In the supplementary Material, we will show the proofs of the Lemmas and Theorem, as well as additional simulation studies.

S1 Proofs

In this section, we sketch the proofs of the Lemmas and Theorem 1.

Proof of Lemma 1: Let

$$\mathbf{U}(\boldsymbol{\beta}) = \sum_{i=1}^{n} (\tilde{y}_{i0} - \mu_i)\mathbf{X}_i / \sigma_n^2(\tilde{y}_0, \mu_i),$$

where $\tilde{y}_{i0} = (\delta_i + (1-\delta_i)\tilde{\lambda}_{i0})y_i + (1-\delta_i)(1-\tilde{\lambda}_{i0}))\mu_i, i = 1, \cdots, n$ in which $\tilde{\lambda}_{i0} = I(y_i > \mu_{i0})$ is a known indicator function, $\mu_{i0} = \boldsymbol{\beta}_0^T \mathbf{X}_i$ and $\sigma_n^2(\tilde{y}_0, \mu_i)$ is $\sigma_n^2(\mu_i)$ based on the observations $\tilde{y}_{i0}$. Based on the work of Chiou and Muller (1999), $\mathbf{U}(\hat{\boldsymbol{\beta}}) - \mathbf{U}(\boldsymbol{\beta}_0) = O_p(n^{1/2})$, where $\hat{\boldsymbol{\beta}}$ is the solution of $\mathbf{U}(\boldsymbol{\beta}) = 0$. Because $\mathbf{U}(\boldsymbol{\beta}_0) = \tilde{\mathbf{U}}(\boldsymbol{\beta}_0)$ and both

$\mathbf{U}(\hat{\boldsymbol{\beta}})$ and $\tilde{\mathbf{U}}(\tilde{\boldsymbol{\beta}})$ are equal to 0, we can obtain $\tilde{\mathbf{U}}(\tilde{\boldsymbol{\beta}}) - \tilde{\mathbf{U}}(\boldsymbol{\beta}_0) = O_p(n^{1/2})$ as well.

Now we write $\kappa\{\tilde{\mathbf{U}}(\mathbf{b})\} = \mathbf{b}$ for any $\mathbf{b} \in \mathcal{B}$, because $\tilde{\mathbf{U}}(\mathbf{b})$ is a monotone function. By Taylor expansions of $\kappa\{\tilde{\mathbf{U}}(\tilde{\boldsymbol{\beta}})\}$ around $\tilde{\mathbf{U}}(\boldsymbol{\beta}_0)$, we get

$$\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + B(\boldsymbol{\beta}_0)^{-1}(\tilde{\mathbf{U}}(\tilde{\boldsymbol{\beta}}) - \tilde{\mathbf{U}}(\boldsymbol{\beta}_0)) + o_p(n^{-1/2}),$$

where $B(\mathbf{b}) = -\partial\kappa\{\tilde{\mathbf{U}}(\mathbf{b})\}/\partial\tilde{\mathbf{U}}(\mathbf{b})$. Then the proof is complete.

Proof of Lemma 2: Observed that

$$\sup_{u \in I} |\tilde{\sigma}_n^2(u) - \sigma^2(u)| \quad = \quad \sup_{u \in I} |\tilde{\sigma}_n^2(u) - \sigma_n^2(u)| + \sup_{u \in I} |\sigma_n^2(u) - \sigma^2(u)|.$$

The second term is $O_p\left([\frac{\log n}{nb}]^{1/2} + b^2\right)$ based on Lemma 4.1 in Chiou and Muller (1999).

Consider the first term, set $\hat{y}_i = (\delta_i + (1-\delta_i)\tilde{\lambda}_i)y_i + (1-\delta_i)(1-\tilde{\lambda}_i)\mu_i$, where $\tilde{\lambda}_i = I(C_i > \tilde{\mu}_i)$, then

$$\begin{aligned} \tilde{\epsilon}_i &= \hat{y}_i - \tilde{\mu}_i \\ &= \tilde{y}_i - \tilde{\mu}_i + O_p(n^{-1/2}) \end{aligned} \qquad \text{(S1.1)}$$

Now write

$$\begin{aligned}
\sup_{u\in I}|\tilde{\sigma}_n^2(u)-\sigma_n^2(u)| &= \sup_{u\in I}|\sum_{i=1}^{n}(\tilde{a}_i(u)\tilde{\epsilon}_i^2 - a_i(u)\epsilon_i^2)| \\
&\leq \sup_{u\in I}|\tilde{a}_i(u)-a_i(u)|\epsilon_i^2 \\
&\quad +\sup_{u\in I}|a_i(u)|(\tilde{\epsilon}_i-\epsilon_i)^2 \\
&\quad +\sup_{u\in I}|\tilde{a}_i(u)-a_i(u)|(\tilde{\epsilon}_i-\epsilon_i)^2 \\
&\quad +2\sup_{u\in I}|\tilde{a}_i(u)\epsilon_i(\tilde{\epsilon}_i-\epsilon_i)| \\
&\quad +2\sup_{u\in I}|\tilde{a}_i(u)-a_i(u)\epsilon_i(\tilde{\epsilon}_i-\epsilon_i)| \\
&= I+II+III+IV+V
\end{aligned}$$

Based on (S1.1) and Theorem 4.1 of Chiou and Muller (1999), the proof is complete.

Proof of Theorem 1: Consider the estimating equation $\mathbf{U}(\boldsymbol{\beta})$ and the resulting estimator $\hat{\boldsymbol{\beta}}$ in the proof of Lemma 1. Based on Chiou and Muller (1999),

$$\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) \to_d N(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}).$$

Then it suffices to prove that the Laplace approximated estimator $\tilde{\boldsymbol{\beta}}$ has the same limiting distribution as $\hat{\boldsymbol{\beta}}$. Let $\bar{\boldsymbol{\beta}}$ lies between $\boldsymbol{\beta}_0$ and $\hat{\boldsymbol{\beta}}$, then the observed information matrix of $\mathbf{U}(\boldsymbol{\beta})$ evaluated at $\bar{\boldsymbol{\beta}}$ can be written as

$$\begin{aligned} \mathbf{H}(\bar{\boldsymbol{\beta}}) &= -\frac{\partial \mathbf{U}(\bar{\boldsymbol{\beta}})}{\partial \bar{\boldsymbol{\beta}}} \\ &= \sum_{i=1}^{n} \frac{1}{\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)} \mathbf{X}_i \mathbf{X}_i^T + \sum_{i=1}^{n} (\tilde{y}_{i0} - \bar{\mu}_i) \frac{[\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]'}{[\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]^2} \mathbf{X}_i \mathbf{X}_i^T, \end{aligned}$$

where $\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)$ is $\sigma_n^2(\tilde{y}_0, \bar{\mu}_i)$ evaluated at $\bar{\boldsymbol{\beta}}$.

Now consider $\tilde{\boldsymbol{\beta}}$, which is obtained by solving $\tilde{\mathbf{U}}(\boldsymbol{\beta}) = 0$. Similarly, if $\bar{\boldsymbol{\beta}}$ lies between $\boldsymbol{\beta}_0$ and $\tilde{\boldsymbol{\beta}}$, then the observed information matrix of $\tilde{\mathbf{U}}(\boldsymbol{\beta})$ evaluated at $\bar{\boldsymbol{\beta}}$ is

$$\begin{aligned} \tilde{\mathbf{H}}(\bar{\boldsymbol{\beta}}) &= \frac{\partial \tilde{\mathbf{U}}(\bar{\boldsymbol{\beta}})}{\partial \bar{\boldsymbol{\beta}}} \\ &= -\frac{\partial \sum_{i=1}^{n} [(\delta_i y_i + (1-\delta_i) \max\{y_i, \bar{\mu}_i\}) - \bar{\mu}_i] \mathbf{X}_i / \tilde{\sigma}_n^2(\bar{\mu}_i)}{\partial \bar{\beta}} \\ &= \sum_{i=1}^{n} \frac{1}{\tilde{\sigma}_n^2(\bar{\mu}_i)} \mathbf{X}_i \mathbf{X}_i^T + \sum_{i=1}^{n} (\bar{y}_i - \bar{\mu}_i) (\frac{[\tilde{\sigma}_n^2(\bar{\mu}_i)]'}{[\tilde{\sigma}_n^2(\bar{\mu}_i)]^2} \mathbf{X}_i \mathbf{X}_i^T, \end{aligned}$$

where $\bar{y}_i = (\delta_i + (1-\delta_i)\bar{\lambda}_i) y_i + (1-\delta_i)(1-\bar{\lambda}_i)\bar{\mu}_i$, where $\bar{\lambda}_i = I(C_i > \bar{\mu}_i)$

To prove

$$\sqrt{n}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \rightarrow_d N(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}),$$

it suffices to prove $\mathbf{U}(\boldsymbol{\beta}) - \tilde{\mathbf{U}}(\boldsymbol{\beta}) = o_p(n^{1/2})$ and $\mathbf{H}(\bar{\boldsymbol{\beta}}) - \tilde{\mathbf{H}}(\bar{\boldsymbol{\beta}}) = o_p(n)$. It is obvious that $\mathbf{U}(\boldsymbol{\beta})$ and $\tilde{\mathbf{U}}(\boldsymbol{\beta})$ is equivalent at $\mu_{i0}, i = 1, \cdots, n$.

Next, we prove $1/n(\mathbf{H}(\bar{\boldsymbol{\beta}}) - \tilde{\mathbf{H}}(\bar{\boldsymbol{\beta}})) = o_p(1)$. Because $||\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}|| \leq ||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0|| + ||\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}_0|| = O_p(1/\sqrt{n}), I(y_i > \bar{\mu}_i) - I(y_i > \mu_{i0}) = O_p(1/\sqrt{n})$. Let $G_i = 1/(\tilde{\sigma}_n^2(\bar{\mu}_i)) -$

$1/(\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i))$ and $E_i = [\tilde{\sigma}_n^2(\bar{\mu}_i)]'/[\tilde{\sigma}_n^2(\bar{\mu}_i)]^2 - [\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]'/[\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]^2$. Note $\sigma_n^2(.) = \sigma_n^2(\tilde{y}_0, .)$. Then because $G_i \leq |1/(\tilde{\sigma}_n^2(\bar{\mu}_i)) - 1/(\sigma_n^2(\bar{\mu}_i))| + |1/(\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)) - 1/(\sigma_n^2(\bar{\mu}_i))|$, $\max_{1\leq i\leq n} G_i = O_p(\alpha_n/\zeta_n)$, where $\alpha_n = [\log n/nb]^{1/2} + b^2 + (\sqrt{n}b)^{-1}$ based on Theorem 4.2 of Chiou and Muller (1999). Similarly, $\max_{1\leq i\leq n} E_i = O_p(\alpha_n/\zeta_n b)$. Therefore,

$$
\begin{aligned}
& 1/n|(\mathbf{H}(\bar{\boldsymbol{\beta}}) - \tilde{\mathbf{H}}(\bar{\boldsymbol{\beta}}))| \\
= \;& 1/n|\sum_{i=1}^{n} G_i \mathbf{X}_i \mathbf{X}_i^T| \\
& +1/n|\sum_{i=1}^{n}(\delta_i + (1-\delta_i))(I(y_i > \mu_{i0}) - I(y_i > \bar{\mu}_i))\frac{[\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]'}{[\hat{\sigma}_n^2(\tilde{y}_0, \bar{\mu}_i)]^2}\mathbf{X}_i \mathbf{X}_i^T| \\
& +1/n|\sum_{i=1}^{n}(\bar{y}_i - \bar{\mu}_i)E_i \mathbf{X}_i \mathbf{X}_i^T| \\
\leq \;& 1/n\sum_{i=1}^{n}\max_{1\leq i\leq n}|G_i| + O_p(n^{-1/2}\zeta_n^{-1}) + 1/n\sum_{i=1}^{n}|\bar{y}_i - \bar{\mu}_i|\max_{1\leq i\leq n}|E_i| \\
= \;& o_p(1)
\end{aligned}
$$

using the assumptions for ζ_n. Then the proof is complete.

S2 Additional simulations

In this section, we provide the simulation results for scenario 2 and 3 in the main text with extreme value distribution.

Table 1: Simulation results for scenario 2 with extreme value distribution

n	CP	variance	para	Laplace				WLS		LBJ		LS	
				bias	SE	SEE	Cov	bias	SE	bias	SE	bias	SE
200	20%	σ_1	β_1	0.007	0.029	0.046	0.966	0.003	0.029	-0.014	0.166	-0.056	0.178
			β_2	-0.011	0.046	0.077	0.976	-0.004	0.045	0.013	0.200	0.098	0.238
			β_3	-0.004	0.031	0.049	0.964	0.000	0.032	0.004	0.147	0.048	0.162
			β_4	0.006	0.032	0.048	0.974	0.003	0.032	-0.008	0.142	-0.050	0.156
		σ_2	β_1	0.002	0.009	0.010	0.926	0.000	0.011	0.000	0.009	0.000	0.009
			β_2	-0.001	0.008	0.009	0.914	0.001	0.010	-0.001	0.008	-0.001	0.008
			β_3	-0.001	0.008	0.009	0.932	0.015	0.023	0.000	0.008	0.000	0.008
			β_4	0.001	0.008	0.010	0.940	0.016	0.023	0.000	0.009	0.000	0.009
400	20%	σ_1	β_1	0.000	0.027	0.041	0.966	-0.002	0.027	-0.017	0.126	-0.059	0.141
			β_2	0.001	0.045	0.072	0.978	0.004	0.044	0.014	0.148	0.097	0.187
			β_3	0.001	0.030	0.042	0.962	0.004	0.030	0.013	0.103	0.055	0.119
			β_4	-0.001	0.029	0.042	0.962	-0.003	0.029	-0.012	0.101	-0.054	0.120
		σ_2	β_1	0.001	0.006	0.007	0.964	0.001	0.007	0.000	0.006	-0.001	0.006
			β_2	-0.001	0.005	0.006	0.920	-0.001	0.006	0.000	0.005	0.000	0.006
			β_3	0.000	0.005	0.007	0.934	0.015	0.020	0.001	0.006	0.001	0.006
			β_4	0.000	0.005	0.007	0.936	0.014	0.019	0.000	0.006	0.000	0.006
200	40%	σ_1	β_1	-0.007	0.058	0.064	0.930	0.009	0.040	-0.027	0.191	-0.130	0.231
			β_2	0.014	0.092	0.103	0.948	-0.014	0.059	0.035	0.219	0.243	0.340
			β_3	0.005	0.054	0.065	0.946	-0.006	0.041	0.013	0.168	0.118	0.215
			β_4	-0.005	0.058	0.065	0.960	0.008	0.043	-0.018	0.163	-0.118	0.208
		σ_2	β_1	0.001	0.010	0.011	0.916	0.000	0.013	0.000	0.011	-0.001	0.011
			β_2	-0.001	0.009	0.011	0.926	0.002	0.013	-0.001	0.009	-0.001	0.010
			β_3	-0.001	0.009	0.011	0.942	0.019	0.027	0.001	0.009	0.001	0.009
			β_4	0.001	0.010	0.011	0.956	0.019	0.027	0.000	0.010	0.000	0.010
400	40%	σ_1	β_1	-0.005	0.043	0.052	0.950	0.004	0.032	-0.028	0.143	-0.136	0.193
			β_2	0.008	0.067	0.087	0.976	-0.008	0.047	0.031	0.161	0.239	0.286
			β_3	0.005	0.043	0.053	0.958	-0.003	0.033	0.023	0.117	0.125	0.172
			β_4	-0.006	0.043	0.053	0.954	0.003	0.032	-0.024	0.115	-0.129	0.172
		σ_2	β_1	0.000	0.007	0.008	0.938	0.000	0.008	-0.001	0.007	-0.001	0.007
			β_2	-0.001	0.007	0.007	0.916	-0.001	0.008	-0.001	0.007	0.000	0.007
			β_3	0.001	0.006	0.008	0.944	0.011	0.017	0.001	0.007	0.001	0.007
			β_4	0.000	0.007	0.008	0.932	0.010	0.016	0.000	0.007	0.000	0.007

Table 2: Simulation results for scenario 3 with extreme value distribution

n	CP	variance	para	Laplace				WLS		LBJ		LS	
				bias	SE	SEE	Cov	bias	SE	bias	SE	bias	SE
200	20%	σ_1	β_1	0.012	0.036	0.048	0.950	0.009	0.031	-0.009	0.170	-0.056	0.183
			β_2	-0.019	0.059	0.080	0.954	-0.013	0.049	0.008	0.204	0.101	0.245
			β_3	-0.008	0.039	0.050	0.946	-0.005	0.033	0.004	0.148	0.069	0.171
			β_4	0.010	0.038	0.050	0.962	0.007	0.033	-0.004	0.146	-0.050	0.163
		σ_3	β_1	0.006	0.029	0.032	0.948	0.010	0.025	-0.002	0.090	-0.019	0.096
			β_2	-0.010	0.044	0.049	0.944	-0.017	0.039	-0.001	0.109	0.033	0.127
			β_3	-0.005	0.029	0.034	0.950	-0.007	0.027	-0.002	0.077	0.021	0.086
			β_4	-0.001	0.021	0.028	0.970	0.003	0.019	0.009	0.079	-0.009	0.080
400	20%	σ_1	β_1	-0.004	0.029	0.043	0.964	-0.002	0.025	-0.014	0.129	-0.060	0.144
			β_2	0.008	0.049	0.075	0.980	0.002	0.040	0.010	0.150	0.102	0.193
			β_3	0.006	0.032	0.044	0.972	0.003	0.028	0.013	0.105	0.077	0.130
			β_4	-0.004	0.031	0.044	0.976	-0.001	0.027	-0.010	0.103	-0.056	0.122
		σ_3	β_1	0.002	0.016	0.023	0.962	0.003	0.016	-0.003	0.064	-0.020	0.070
			β_2	-0.007	0.024	0.035	0.954	-0.007	0.023	0.001	0.079	0.036	0.096
			β_3	-0.002	0.017	0.024	0.956	-0.002	0.017	0.005	0.055	0.028	0.064
			β_4	0.000	0.013	0.019	0.976	0.001	0.012	0.006	0.055	-0.013	0.057
200	40%	σ_1	β_1	-0.006	0.055	0.065	0.956	0.006	0.041	-0.023	0.194	-0.125	0.234
			β_2	0.016	0.087	0.104	0.952	-0.010	0.063	0.027	0.222	0.232	0.334
			β_3	0.009	0.056	0.066	0.962	-0.005	0.043	0.013	0.165	0.136	0.224
			β_4	-0.008	0.053	0.067	0.964	0.006	0.039	-0.013	0.166	-0.114	0.214
		σ_3	β_1	0.004	0.037	0.043	0.948	0.007	0.034	-0.009	0.107	-0.047	0.123
			β_2	-0.007	0.057	0.065	0.956	-0.012	0.052	0.009	0.122	0.084	0.168
			β_3	-0.002	0.038	0.044	0.962	-0.004	0.035	0.004	0.087	0.048	0.108
			β_4	0.001	0.031	0.039	0.968	0.003	0.028	0.028	0.099	-0.011	0.101
400	40%	σ_1	β_1	-0.007	0.039	0.053	0.966	-0.006	0.034	-0.024	0.143	-0.128	0.190
			β_2	0.012	0.064	0.092	0.978	0.009	0.053	0.024	0.159	0.233	0.286
			β_3	0.008	0.042	0.055	0.962	0.007	0.036	0.022	0.117	0.146	0.186
			β_4	-0.007	0.042	0.055	0.966	-0.005	0.036	-0.019	0.117	-0.122	0.172
		σ_3	β_1	0.000	0.023	0.030	0.950	0.001	0.022	-0.008	0.074	-0.046	0.090
			β_2	-0.002	0.034	0.046	0.966	-0.005	0.032	0.008	0.087	0.085	0.132
			β_3	0.001	0.024	0.031	0.950	-0.001	0.023	0.009	0.062	0.053	0.086
			β_4	0.001	0.019	0.026	0.970	0.002	0.019	0.022	0.069	-0.019	0.072