ANALYSIS OF MULTIVARIATE FAILURE TIME DATA USING MARGINAL PROPORTIONAL HAZARDS MODEL

Ying Chen, Kani Chen and Zhiliang Ying

SHUFE, HKUST and Columbia University

Supplementary Material

S1 Proof of the Theorem

Recall that β_0 and $\lambda_{k0}(\cdot)$ are the true values and functions of β and $\lambda_k(\cdot)$, and $M_{ik}(\cdot)$ are iid copies of $M_k(\cdot)$. The proof is essentially same as that of Andersen and Gill (1982). Write

$$U_{\mathbf{h}}(\beta_0) = \sum_{i=1}^{n}\sum_{k=1}^{K}\int_0^{\tau_k}[h_{ik}(t) - \mu_{k,\mathbf{h}}(t)]dM_{ik}(t) - \sum_{k=1}^{K}\sum_{i=1}^{n}\int_0^{\tau_k}[\bar{h}_k(t;\beta_0) - \mu_{k,\mathbf{h}}(t)]dM_{ik}(t) \tag{S1.1}$$

The first term of the right hand side is the sum of n random vectors which are iid copies of $\sum_{k=1}^{K}\xi_{k,\mathbf{h}}$ and is thus asymptotically normal at the rate $n^{1/2}$ with asymptotic variance $V_{\mathbf{h}}$. For every $1 \le k \le K$,

$$\begin{aligned} &E\Big(\textstyle\sum_{i=1}^{n}\int_0^{\tau_k}[\bar{h}_k(t;\beta_0) - \mu_{k,\mathbf{h}}(t)]dM_{ik}(t)\Big)^{\otimes 2} \\ &\le \textstyle\sum_{i=1}^{n} E\Big(\int_0^{\tau_k}[\bar{h}_k(t;\beta_0) - \mu_{k,\mathbf{h}}(t)]^{\otimes 2}dN_{ik}(t)\Big) = o(n). \end{aligned}$$

Therefore the second term of the right hand side of (S1.1) is $o_P(n^{1/2})$. As a result,

$$n^{-1/2}U_{\mathbf{h}}(\beta_0) \to N(0, V_{\mathbf{h}}). \tag{S1.2}$$

It follows from the law of large numbers that

$$\begin{aligned} \frac{1}{n}\frac{\partial}{\partial\beta}U_{\mathbf{h}}(\beta)\Big|_{\beta=\beta_0} &= -\frac{1}{n}\sum_{k=1}^{K}\sum_{i=1}^{n}\int_0^{\tau_k}\frac{\sum_{j=1}^{n}(h_{jk}(t) - \bar{h}_k(t;\beta_0))}{\sum_{j=1}^{n}e^{\beta_0' Z_{jk}}Y_{jk}(t)} \\ &\quad \times(Z_{jk}' - \bar{Z}_k'(t;\beta_0))e^{\beta_0' Z_{jk}}Y_{jk}(t)dN_{ik}(t) \\ &\to -\sum_{k=1}^{K}E(\xi_{k,\mathbf{h}}\xi_k') = -A_{\mathbf{h}} \end{aligned} \tag{S1.3}$$

in probability as $n \to \infty$. The above convergence can be shown to hold uniformly over $\{\beta : \|\beta - \beta_0\| \le \epsilon_n\}$ for any sequence of $\epsilon_n \downarrow 0$. Since $A_{\mathbf{h}}$ is assumed nondegenerate, let $a = \inf\{\|A_{\mathbf{h}} x\|/\|x\| : x \in R^p\}$. Then $a > 0$. Let B be the ball in R^p centered at β_0 with radius ϵ, where $\epsilon > 0$ is small but fixed, and let $D_n = \{(1/n)U_{\mathbf{h}}(x) : x \in B\}$ be the image of B for the continuous mapping $(1/n)U_{\mathbf{h}}(\cdot)$. With probability tending to 1, for any two p-vectors x_1 and x_2 in B,

$$1/n\|U_{\mathbf{h}}(x_1) - U_{\mathbf{h}}(x_2)\| > (a/2)\|x_1 - x_2\|.$$

It implies that, with probability tending to 1, $(1/n)U_{\mathbf{h}}(\beta)$ is a homeomorphism from B to D_n and, moreover, D_n contains a ball centered at $(1/n)U_{\mathbf{h}}(\beta_0)$ with radius $a/2$. Then, with probability tending to 1, this ball contains 0 since (S1.2) implies $(1/n)U_{\mathbf{h}}(\beta_0) = o_P(1)$. This proves that, with probability tending to 1, there exists a zero solution of the equation $U_{\mathbf{h}}(\beta) = 0$ in any small but fixed neighborhood of β_0. The consistency follows. Then, (S1.2) and (S1.3) together ensures asymptotic normality and the asymptotic variance is given by the sandwich formula. The proof is complete.

S2 Proof of the Proposition

The proof is divided into five steps. Step 1: Introducing some notations. Let L_0^2 be the space of all p-dimensional random vectors measurable to the σ-algebra generated by $\{(Y_k, \delta_k), k = 1, ..., K, \mathbf{Z}\}$ and with zero conditional mean given $\mathbf{Z}$ and finite variance. Define an inner product to be the sum of component-wise covariances so that L_0^2 can be verified to be a Hilbert space. Let $\mathcal{S}_k = \{\eta : \ \eta \in L_0^2, E(\eta|\mathbf{Z}) = 0, \eta \in \sigma(Y_k, \delta_k, \mathbf{Z})\}$. Denote by $\mathcal{M}_k$ the closure of $\{\eta : \ \eta = \int_0^{\tau_k} (h(t, \mathbf{Z}) - \mu_h(t)) dM_k(t)$, for all p-dimensional continuous functions h such that $\eta \in L_0^2\}$ where $\mu_h = E(h(t, \mathbf{Z})|Y_k = t, \delta_k = 1)$. It is seen that $\mathcal{M}_k$ and $\mathcal{S}_k$ are both closed linear subspaces of L_0^2 and that $\mathcal{M}_k \subseteq \mathcal{S}_k$. Set $\mathcal{M} = \mathcal{M}_1 + \cdots + \mathcal{M}_K$ and $\check{\mathcal{M}}_k = \mathcal{M}_1 + \cdots + \mathcal{M}_{k-1} + \mathcal{M}_{k+1} + \cdots + \mathcal{M}_K$. Then $\mathcal{S}$ and $\check{\mathcal{S}}_k$ are likewise defined. To avoid trivialities, we assume throughout the paper $\check{\mathcal{M}}_k$, $\check{\mathcal{S}}_k$, $\mathcal{M}_k + \check{\mathcal{M}}_k$ and $\mathcal{S}_k + \check{\mathcal{S}}_k$ are closed and $\mathcal{M}_k \cap \check{\mathcal{M}}_k = \{0\} = \mathcal{S}_k \cap \check{\mathcal{S}}_k$ for every $k = 1, ..., K$.

Step 2. Defining the score $\sum_{l=1}^K \xi_{l,\mathbf{h}^\star}$ through alternating projection. Denote the projection operator in L_0^2 by Π. Write

$$\Sigma_{\mathbf{h}} = \Big[\sum_{k=1}^K E(\xi_{k,\mathbf{h}}\xi_k')\Big]^{-1} \Big[\sum_{k=1}^K E\Big(\xi_{k,\mathbf{h}}\Big\{\Pi(\sum_{l=1}^K \xi_{l,\mathbf{h}}|\mathcal{M}_k)\Big\}'\Big)\Big] \Big[\sum_{k=1}^K E(\xi_k \xi_{k,\mathbf{h}}')\Big]^{-1}.$$

Let $\mathbf{h}^\star$ satisfy

$$\Pi(\sum_{l=1}^K \xi_{l,\mathbf{h}^\star}|\mathcal{M}_k) = \xi_k, \qquad k = 1, ..., K. \tag{S2.1}$$

Then,

$$\Sigma_{\mathbf{h}^\star} = \Big[\sum_{k=1}^K E(\xi_{k,\mathbf{h}^\star}\xi_k')\Big]^{-1} = \Big[\mathrm{var}(\sum_{k=1}^K \xi_{k,\mathbf{h}^\star})\Big]^{-1}.$$

The existence of the solution of (S2.1) can be argued as follows. Let $\Xi_k \equiv \xi_k - \Pi(\xi_k|\check{\mathcal{M}}_k) + \Pi(\Pi(\xi_k|\check{\mathcal{M}}_k)|\mathcal{M}_k) - \Pi(\Pi(\Pi(\xi_k|\check{\mathcal{M}}_k)|\mathcal{M}_k)|\check{\mathcal{M}}_k) + \cdots$. The convergence of the series follows from, e.g, Theorem 2 of Chapter A.4 of Bickel *et al.* (1993, pp.438) and Ξ_k is an element of $\mathcal{M}$. Furthermore, $\Pi(\Xi_k|\mathcal{M}_k) = \xi_k$, $\Pi(\Xi_k|\check{\mathcal{M}}_k) = 0$ and, therefore, $\Pi(\Xi_k|\mathcal{M}_l) = 0$ for $l \neq k$ since $\mathcal{M}_l \subseteq \check{\mathcal{M}}_k$ for $l \neq k$. Thus, $\Pi(\sum_{l=1}^K \Xi_l|\mathcal{M}_k) = \xi_k$, for $1 \leq k \leq K$. The existence is established. The uniqueness is argued as follows. Let $\sum_{k=1}^K \xi_{l,\mathbf{h}}$ be the difference of any two solutions. Then, $\Pi(\sum_{l=1}^K \xi_{l,\mathbf{h}}|\mathcal{M}_k) = 0$ for all $k = 1, ..., K$. This implies $\sum_{k=1}^K \xi_{k,\mathbf{h}} \perp \mathcal{M}$. Since $\sum_{k=1}^K \xi_{k,\mathbf{h}} \in \mathcal{M}$, it follows that $\sum_{k=1}^K \xi_{k,\mathbf{h}} = 0$.

Step 3. Martingale representations of the projections of $\sum_{l=1}^K \xi_{l,\mathbf{h}^\star}$. Let $M_k^\circ(t) = (1-\delta_k)I(Y_k \leq t) - \int_0^t \lambda_{C_k|\mathbf{Z}}(s, \mathbf{Z})Y_k(s)ds$ where $\lambda_{C_k|\mathbf{Z}}(\cdot, \mathbf{z})$ is the true conditional hazard of C_k given $\mathbf{Z} = \mathbf{z}$. It follows from the counting process martingale representation of random variables with zero mean and finite second moment that

$$\Pi(\sum_{l=1}^K \xi_{l,\mathbf{h}^\star}|\mathcal{S}_k) = E(\sum_{l=1}^K \xi_{l,\mathbf{h}^\star}|Y_k, \delta_k, \mathbf{Z})$$
$$= \int_0^{\tau_k} \tilde{h}_k(t, \mathbf{Z})dM_k(t) + \int_0^{\tau_k} \tilde{a}_k(t)dM_k(t) + \int_0^{\tau_k} \tilde{G}_k(t, \mathbf{Z})dM_k^\circ(t) \qquad \text{(S2.2)}$$

for some measurable functions $\tilde{h}_k$, $\tilde{a}_k$ and $\tilde{G}_k$, $1 \leq k \leq K$, where $\tilde{h}_k$ satisfies $E(\tilde{h}_k(t, \mathbf{Z})|Y_k = t, \delta_k = 1) = 0$ and $\tilde{a}_k$ is a non-random function. The last two terms of (S2.2) are orthogonal to each other and both are orthogonal to $\mathcal{M}_k$ while the first is an element of $\mathcal{M}_k$. Combining (S2.2) with (S2.1), it follows that $\tilde{h}_k(t, \mathbf{Z}) = Z_k(t) - \mu_k(t)$.

Step 4. Constructing a parametric submodel. Let β be in a small but fixed neighborhood of β_0. Let

$$\lambda_k(t; \beta) = \lambda_{k0}(t)e^{(\beta-\beta_0)'[-\mu_k(t)+\tilde{a}_k(t)]} \quad \text{and} \quad \lambda_{C_k|\mathbf{Z}}(t, \mathbf{z}; \beta) = \lambda_{C_k|\mathbf{Z}}(t, \mathbf{z})e^{(\beta-\beta_0)'\tilde{G}_k(t,\mathbf{z})}.$$

Define

$$f_k(y, d|\mathbf{z}; \beta) = e^{d\beta' z_k}\lambda_k^d(y; \beta)e^{-\int_0^{\tau_k \wedge y} e^{\beta' z_k}\lambda_k(t;\beta)dt} \times \lambda_{C_k|\mathbf{Z}}^{1-d}(y, \mathbf{z}; \beta)e^{-\int_0^{\tau_k \wedge y} \lambda_{C_k|\mathbf{Z}}(t,\mathbf{z};\beta)dt},$$

where $\mathbf{z} = (z_1, ..., z_K)$ and d takes value 0 or 1. If a parametric family, with parameter β, has (conditional) marginal densities as f_k, then the family is a parametric submodel since the expression of f_k fulfills the requirement of proportional hazards in (1). Such a family of densities is constructed in the following.

Let $u_k(\beta) = f_k(Y_k, \delta_k|\mathbf{Z}, \beta)/f_k(Y_k, \delta_k|\mathbf{Z}, \beta_0) - 1$. Then $u_k(\beta) \in \mathcal{S}_k$ and $u_k(\beta_0) = 0$. Let

$$v_k(\beta) = u_k(\beta) - \Pi(u_k(\beta)|\check{\mathcal{S}}_k) + \Pi(\Pi(u_k(\beta)|\check{\mathcal{S}}_k)|\mathcal{S}_k) - \Pi(\Pi(\Pi(u_k(\beta)|\check{\mathcal{S}}_k)|\mathcal{S}_k)|\check{\mathcal{S}}_k) + \cdots.$$

Theorem 2 of A.4 of Bickel *et al.* (1993) ensures the convergence of the series and that

$$v_k(\beta) \in \mathcal{S}, \quad \Pi(v_k(\beta)|\mathcal{S}_k) = u_k(\beta) \quad \text{and} \quad \Pi(v_k(\beta)|\check{\mathcal{S}}_k) = 0. \qquad \text{(S2.3)}$$

Let $v(\beta) = 1 + \sum_{k=1}^{K} v_k(\beta)$ and

$$f(y_1, \delta_1, ..., y_K, \delta_K|\mathbf{z}; \beta) = v(\beta) f_0(y_1, \delta_1, ..., y_K, \delta_K|\mathbf{z}; \beta_0)$$

where f_0 denotes the true conditional density of $(Y_1, \delta_1, ..., Y_K, \delta_K)$ given $\mathbf{Z}$. Notice that f is a (conditional) density since $E(v(\beta)|\mathbf{Z}) = 1$ and $v(\beta) \geq 0$ for β in a small neighborhood of β_0. Observe that $f_k(y_k, \delta_k|\mathbf{z}, \beta_0)$ are the true conditional marginal densities. Write

$$f(y_1, \delta_1, ..., y_K, \delta_K|\mathbf{z}; \beta) = f_k(y_k, \delta_k|\mathbf{z}; \beta_0) \times \frac{v(\beta) f_0(y_1, \delta_1, ..., y_K, \delta_K|\mathbf{z}; \beta_0)}{f_k(y_k, \delta_k|\mathbf{z}; \beta_0)}.$$

Then, the log of the marginal density of f is

$$\begin{aligned} & \log f_k(Y_k, \delta_k|\mathbf{Z}, \beta_0) + \log E(v(\beta)|Y_k, \delta_k, \mathbf{Z}) \\ = \; & \log f_k(Y_k, \delta_k|\mathbf{Z}, \beta_0) + \log[1 + \Pi(v(\beta) - 1|\mathcal{S}_k)] \\ = \; & \log f_k(Y_k, \delta_k|\mathbf{Z}, \beta_0) + \log[1 + \Pi(v_k(\beta)|\mathcal{S}_k)] \\ = \; & \log f_k(Y_k, \delta_k|\mathbf{Z}, \beta_0) + \log(1 + u_k(\beta)) \\ = \; & \log f_k(Y_k, \delta_k|\mathbf{Z}, \beta). \end{aligned}$$

Thus f as a parametric family of densities is indeed a parametric submodel with parameter β.

Step 5. Verifying that the score of the parametric submodel is $\sum_{l=1}^{K} \xi_{l,\mathbf{h}^\star}$. Observe that $v(\beta_0) = 1$ since $v_k(\beta_0) = u_k(\beta_0) = 0$. Moreover, $\frac{\partial}{\partial\beta} u_k(\beta)|_{\beta=\beta_0}$ is the same as (S2.2). The score of the parametric family f at $\beta = \beta_0$ is $\frac{\partial}{\partial\beta} \log v(\beta)|_{\beta=\beta_0} = \frac{\partial}{\partial\beta} v(\beta)|_{\beta=\beta_0}$. It follows from (S2.3) that

$$\Pi(\frac{\partial}{\partial\beta} v(\beta)|_{\beta=\beta_0}|\mathcal{S}_k) = \Pi(\frac{\partial}{\partial\beta} v_k(\beta)|_{\beta=\beta_0}|\mathcal{S}_k) = \frac{\partial}{\partial\beta} u_k(\beta)|_{\beta=\beta_0} = \Pi(\sum_{l=1}^{K} \xi_{l,\mathbf{h}^\star}|\mathcal{S}_k).$$

The uniqueness of the alternating projection solution then implies that the score of the parametric submodel f at $\beta = \beta_0$ is $\frac{\partial}{\partial\beta} v(\beta)|_{\beta=\beta_0} = \sum_{l=1}^{K} \xi_{l,\mathbf{h}^\star}$. The proof is complete.