Open Journal of Statistics, **2013, 3, 12-26**
http://dx.doi.org/10.4236/ojs.2013.31003 Published Online February 2013 (http://www.scirp.org/journal/ojs)

Quantile Regression Based on Semi-Competing Risks Data

Jin-Jian Hsieh[1], A. Adam Ding[2], Weijing Wang[3], Yu-Lin Chi[1]
[1]Department of Mathematics, National Chung Cheng University, Chia-Yi, Chinese Taipei
[2]Department of Mathematics, Northeastern University, Boston, USA
[3]Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Chinese Taipei
Email: jjhsieh@math.ccu.edu.tw, ding@neu.edu, wjwang@stat.nctu.edu.tw

Received December 7, 2012; revised January 10, 2013; accepted January 22, 2013

ABSTRACT

This paper considers quantile regression analysis based on semi-competing risks data in which a non-terminal event may be dependently censored by a terminal event. The major interest is the covariate effects on the quantile of the non-terminal event time. Dependent censoring is handled by assuming that the joint distribution of the two event times follows a parametric copula model with unspecified marginal distributions. The technique of inverse probability weighting (IPW) is adopted to adjust for the selection bias. Large-sample properties of the proposed estimator are derived and a model diagnostic procedure is developed to check the adequacy of the model assumption. Simulation results show that the proposed estimator performs well. For illustrative purposes, our method is applied to analyze the bone marrow transplant data in [1].

Keywords: Copula Model; Dependent Censoring; Quantile Regression; Semi-Competing Risks Data

1. Introduction

Quantile regression analysis has received increasing attentions in the recent literature of survival analysis. Compared with conventional regression models such as the proportional hazards (PH) model or the accelerated failure time (AFT) model, quantile regression models provide direct assessment of the covariate effect on different quantiles of the failure time variable. This model also allows covariates to affect both location and shape of the distribution. Let T be the failure time of interest, $\tilde{\boldsymbol{Z}}$ be a $p\times 1$ vector and $\boldsymbol{Z}=\left(1,\tilde{\boldsymbol{Z}}^{\mathrm{T}}\right)^{\mathrm{T}}$. Consider the following linear quantile regression model on $h(T)$, where $h(\cdot)$ is a known monotonic function, such that

$$\xi_\gamma\left(h(T)\middle|\boldsymbol{Z}\right)=\boldsymbol{\beta}_0^{\mathrm{T}}(\gamma)\boldsymbol{Z}, \tag{1}$$

where $0<\gamma<1$ and $\xi_\gamma\left(Y\middle|\boldsymbol{Z}\right)$ is the $(100\times\gamma)$ th quantile of Y conditional on Z. Note that when we set $\epsilon_\gamma=h(T)-\boldsymbol{\beta}_0^{\mathrm{T}}(\gamma)\boldsymbol{Z}$, model (1) is equivalent to $\Pr\left(\epsilon_\gamma\le 0\middle|\boldsymbol{Z}\right)=\gamma$. Many papers for estimating $\boldsymbol{\beta}_0(\gamma)$ without specifying the distribution of $T|Z$ or ϵ_γ have appeared in the literature. [2-5] considered quantile regression analysis under a fixed censoring mechanism in which all the censoring times are observed. Independent right censorship has been assumed by many papers including [6-11].

In this paper, we consider semi-competing risks data [12] in which the failure time of a non-terminal event T is subject to dependent censoring by a terminal event time D but not vice versa. Consider an example of bone marrow transplantation for leukemia patients described in [1] such that T is the time to leukemia relapse and D is the time to death. One important risk factor is the disease classification (*i.e.* ALL, AML low-risk, and AML high-risk) which was determined based on patient's status at the time of transplantation. Here we assume that T, the time to a non-terminal event, follows model (1). Note that [13,14] also considered quantile regression analysis for competing risks data and left-truncated semi-competing risks data respectively. They defined the quantiles based on the crude quantity, namely the cumulative incidence function $\Pr(T\le t, T\le D)$. In contrast, the proposed regression model (1) is defined based on the *net quantity* $\Pr(T\le t)$ which is not identifiable without extra assumption on the dependence structure. There has been some controversy over which quantity should be used in presence of dependent competing risks. We believe that both quantities are important and not mutually exclusive as they provide information on different aspects of the data. Here $\boldsymbol{\beta}_0(\gamma)$ measures the covariate effect on T after separating the potential influence from D. Such analysis is also useful in practical applications. For example, a covariate may prolong D so that increase $\Pr(T\le t, T\le D)$ but have no direct effect on the non-terminal event. The dependence between T and D complicates the estimation of $\boldsymbol{\beta}_0(\gamma)$. We will adopt a semi-parametric copula assumption to model their joint distribution and apply the technique of inverse probability weighting (IPW) to correct the bias due to dependent censoring in the estimation procedure. The association parameter in the copula model will also be estimated using existing methods.

The rest of this paper is organized as follows. In Section 2, we introduce the data structure and model assumptions. The proposed methodology for parameter estimation and model checking is presented in Section 3. The proofs of the asymptotic properties are given in the Appendix. Section 4 contains simulation results. In Section 5, we apply the proposed methods to analyze the bone marrow transplant data in [1] and in Section 6, we give some concluding remarks.

2. Data and Model Assumptions

Recall that T and D denote the time to a non-terminal event and the time to a terminal event respectively such that T is subject to censoring by D but not vice versa. In presence of additional external censoring due to drop-out or the end-of-study effect, one observes $(X, Y, \delta_X, \delta_Y)$ such that $X = T \wedge D \wedge C$, $Y = D \wedge C$, $\delta_X = I(T \le D \wedge C)$, $\delta_Y = I(D \le C)$, where $\wedge$ is the minimum operator and $I(\cdot)$ is the indicator function. The covariate vectors can be denoted as $\tilde{\boldsymbol{Z}}(p \times 1)$ and $\boldsymbol{Z} = \left(1, \tilde{\boldsymbol{Z}}^{\mathrm{T}}\right)^{\mathrm{T}}$. The sample Contains $\left\{\left(X_i, Y_i, \delta_{X_i}, \delta_{Y_i}, \boldsymbol{Z}_i\right)(i = 1, \cdots, n)\right\}$ which are random replications of $(X, Y, \delta_X, \delta_Y, \boldsymbol{Z})$. We will assume that (T, D) and C are independent given Z. The covariate effect on T is specified by model (1) and the major objective is to estimate $\boldsymbol{\beta}_0(\gamma)$ based on semi-competing risks data.

To handle dependent censoring, we have to make extra assumptions about the dependence structure between T and D in the upper wedge. According to [15] who extended Sklar's theorem to the regression setting, we consider the following copula model

$$\Pr\left(T > t, D > d \mid \boldsymbol{Z} = \boldsymbol{z}\right) = C_{\alpha(z)}\left\{S_{T|z}(t), S_{D|z}(d)\right\}, \quad (2)$$

where $0 \le t \le d \le \infty$, $S_{T|z}(t)$ and $S_{D|z}(d)$ are the marginal survival functions of T and D, given $\boldsymbol{Z} = \boldsymbol{z}$, and $C_\alpha(\cdot,\cdot)$ is a parametric copula function defined on the unit square. The association parameter α in (2) is related to Kendall's tau defined by

$$\tau = 4\int_0^1\int_0^1 C_\alpha(u, v) C_\alpha(\mathrm{d}u, \mathrm{d}v) - 1.$$

In particular, we will assume $(T, D) \mid Z$ in the upper wedge follows a popular subclass of copula models, namely Archimedean copula (AC), in which the copula function can be further expressed as

$$C_\alpha(u, v) = \phi_\alpha^{-1}\left\{\phi_\alpha(u) + \phi_\alpha(v)\right\}, 0 \le u, v \le 1, \quad (3)$$

where ϕ_α is a non-increasing convex function defined on $(0,1]$ with $\phi_\alpha(1) = 0$. Examples of Archimedean copula include Clayton's copula with

$$\phi_\alpha(s) = \left(s^{-\alpha} - 1\right)/\alpha$$

and

$$C_\alpha(u, v) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-1/\alpha};$$

and Frank's copula with

$$\phi_\alpha(s) = \log\{1 - \alpha\} - \log\{1 - \alpha^s\}$$

and

$$C_\alpha(u, v) = \log_\alpha\left\{1 + \left(\alpha^u - 1\right)\left(\alpha^v - 1\right)/(\alpha - 1)\right\}.$$

3. The Proposed Inference Methods

Our major objective is to develop an inference method for estimation $\beta_0(\gamma)$ but, in the mean time, employ existing methods for estimating α based on semi-competing risks data such as those proposed by [16] and [17].

3.1. Estimation of $\boldsymbol{\beta}(\gamma)$ for Discrete Covariates

In absence of censoring, one can estimate $\boldsymbol{\beta}_0(\gamma)$ by solving

$$n^{-1/2}\sum_{i=1}^n \boldsymbol{Z}_i\left(I\left\{h(T)_i \le \boldsymbol{\beta}^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\} - \gamma\right) = 0.$$

Since T_i is subject to censoring by $D_i \wedge C_i$, it follows that

$$\begin{aligned}
&E\left[\left.\frac{I\left\{h(X_i) \le \boldsymbol{\beta}_0^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\}\delta_{X_i}}{H_{\boldsymbol{Z}_i}(X_i)}\right|\boldsymbol{Z}_i\right] \\
&= E\left[\left. E\left\{\left.\frac{I\left\{h(T_i) \le \boldsymbol{\beta}_0^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\}\delta_{X_i}}{H_{\boldsymbol{Z}_i}(T_i)}\right| T_i, \boldsymbol{Z}_i\right\}\right|\boldsymbol{Z}_i\right] \\
&= E\left[I\left\{h(T_i) \le \boldsymbol{\beta}_0^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\}\middle|\boldsymbol{Z}_i\right] = \gamma,
\end{aligned}$$

where the reciprocal of the weight function is given by

$$\begin{aligned}
H_{\boldsymbol{Z}}(t) &= \Pr\left(\delta_X = 1 \mid T = t, \boldsymbol{Z}\right) \\
&= \Pr\left(D \wedge C > t \mid T = t, \boldsymbol{Z}\right) \\
&= \Pr\left(C > t \mid \boldsymbol{Z}\right) \times \Pr\left(D > t \mid T = t, \boldsymbol{Z}\right) \\
&= G_{\boldsymbol{Z}}(t) \times S_{D|T,\boldsymbol{Z}}(t).
\end{aligned}$$

The above derivations yield the following estimating function for $\boldsymbol{\beta}_0(\gamma)$

$$n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left(\frac{I\left\{h(X_i)\le\boldsymbol{\beta}^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\}\delta_{X_i}}{H_{z_i}(X_i)}-\gamma\right)=0.$$

This is the so called inverse probability weighting technique for bias correction. Since $H_{z_i}(X_i)$ needs to be estimated, it is natural to modify the estimating equation as

$$S_n\{\boldsymbol{\beta}(\gamma),\gamma\} = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left(\frac{I\left\{h(X_i)\le\boldsymbol{\beta}^{\mathrm{T}}(\gamma)\boldsymbol{Z}_i\right\}\delta_{X_i}}{\hat{H}_{z_i}(X_i)}-\gamma\right)=0, \quad (4)$$

where the estimated components in the weight can be denoted as

$$\hat{H}_{z_i}(X_i)=\hat{G}_{z_i}(X_i)\times\hat{S}_{D|T,z_i}(X_i) = \widehat{\Pr}\left(C_i>X_i\middle|\boldsymbol{Z}_i=z_i\right) \times\widehat{\Pr}\left(D_i>X_i\middle|T_i=X_i,\boldsymbol{Z}_i=z_i\right).$$

Now we discuss estimation of the weight components. We will first address the situation that $\boldsymbol{Z}$ takes discrete values, and then briefly discuss possible modification for continuous covariates. Since C is independent of T and D given $\boldsymbol{Z}$, $\Pr\left(C>x\middle|\boldsymbol{Z}=z\right)$ can be estimated by the Kaplan-Meier estimator based on data $\left\{\left(Y_i,1-\delta_{Y_i}\right)(i=1,\cdots,n)\right\}$ or $\left\{\left(X_i,1-\delta_{X_i}\delta_{Y_i}\right)(i=1,\cdots,n)\right\}$ with $\boldsymbol{Z}_i=z$. We will utilize some analytic properties of the chosen AC model to derive an explicit expression of $\Pr\left(D>x\middle|T=x,\boldsymbol{Z}_i\right)$. Denote $S_{T,z}(x)=\Pr\left(T>x\middle|\boldsymbol{Z}=\boldsymbol{z}\right)$, $S_{D,z}(x)=\Pr\left(D>x\middle|\boldsymbol{Z}=\boldsymbol{z}\right)$ and $S_{W,z}(x)=\Pr\left(W=T\wedge D>x\middle|\boldsymbol{Z}=\boldsymbol{z}\right)$. It follows that

$$\begin{aligned}S_{D|T,z}(x)&=\Pr\left(D>x\middle|T=x,\boldsymbol{Z}=\boldsymbol{z}\right)\\&=\frac{\partial}{\partial u}\phi_{\alpha(z)}^{-1}\left\{\phi_{\alpha(z)}(u)+\phi_{\alpha(z)}(v)\right\}\Big|_{u=S_{T,z}(x^-),v=S_{D,z}(x)}\\&=\frac{\phi'_{\alpha(z)}(u)}{\phi'_{\alpha(z)}\left[\phi_{\alpha(z)}^{-1}\left\{\phi_{\alpha(z)}(u)+\phi_{\alpha(z)}(v)\right\}\right]}\Bigg|_{u=S_{T,z}(x^-),v=S_{D,z}(x)}\\&=\frac{\phi'_{\alpha(z)}\left(S_{T,z}(x^-)\right)}{\phi'_{\alpha(z)}\left(S_{W,z}(x)\right)}.\end{aligned}$$

We suggest to estimate $S_{D|T,z}(x)$ by applying the estimators in [17] for quantities in the right-hand side of the above expression. Specifically $\hat{S}_{W,z}(x)$ is the Kaplan-Meier estimator of $\Pr\left(T\wedge D>x\middle|\boldsymbol{Z}=\boldsymbol{z}\right)$ based on $\left\{\left(X_i,\delta_{W_i}\right)(i=1,\cdots,n)\right\}$, where $\delta_{W_i}=I\left(T_i\wedge D_i<C_i\right)$, $\hat{S}_{T,z}(x)$ is the copula-graphic estimator

$$\hat{S}_{T,z}(x)=\phi_{\hat{\alpha}(z)}^{-1}\left\{\sum_{i=1}^{n}I\left(X_i\le x,\delta_{X_i}=1,\boldsymbol{Z}_i=\boldsymbol{z}\right)\times\left\{\phi_{\hat{\alpha}(z)}\left[\hat{S}_{W,z}(X_i)\right]-\phi_{\hat{\alpha}(z)}\left[\hat{S}_{W,z}(X_i^-)\right]\right\}\right\},$$

where the estimator $\hat{\alpha}(\boldsymbol{z})$ is the root of the following estimating equation,

$$\sum_{i<j,\boldsymbol{Z}_i=\boldsymbol{Z}_j=z}w\left(\tilde{X}_{ij},\tilde{Y}_{ij}\right)I\left(\tilde{T}_{ij}\le\tilde{D}_{ij}\le\tilde{C}_{ij}\right)\left\{I\left(\left(X_i-X_j\right)\left(Y_i-Y_j\right)>0\right)-\frac{\theta_{\alpha(z)}\left(\hat{\pi}_z\left(\tilde{X}_{ij},\tilde{Y}_{ij}\right)\right)}{\theta_{\alpha(z)}\left(\hat{\pi}_z\left(\tilde{X}_{ij},\tilde{Y}_{ij}\right)\right)+1}\right\}=0,$$

where $\tilde{X}_{ij}=X_i\wedge X_j$, $\tilde{Y}_{ij}=Y_i\wedge Y_j$, $\tilde{T}_{ij}=T_i\wedge T_j$, $\tilde{D}_{ij}=D_i\wedge D_j$, $\tilde{C}_{ij}=C_i\wedge C_j$, $w(\cdot,\cdot)$ is a weight function, $\theta_{\alpha(z)}(v)=-v\phi''_{\alpha(z)}(v)\big/\phi'_{\alpha(z)}(v)$, and

$$\hat{\pi}_z(s,t)=\Pr\left(T>s,D>t\middle|\boldsymbol{Z}=\boldsymbol{z}\right) =\sum_{i=1}^{n}I\left(X_i>x,Y_i>y,Z_i=z\right)\Big/\left\{n_z\hat{G}_z(y)\right\},$$

where $n_z=\sum_{j=1}^{n}I\left(Z_i=z\right)$. Then

$$\hat{S}_{D|T,z}(x)=\frac{\phi'_{\hat{\alpha}(z)}\left(\hat{S}_{T,z}(x^-)\right)}{\phi'_{\hat{\alpha}(z)}\left(\hat{S}_{W,z}(x)\right)}. \quad (5)$$

This estimator is then used in estimating Equation (4).

The Equation (4) may not be continuous so that an exact solution may not exist. Here we define $\hat{\boldsymbol{\beta}}(\gamma)$ as a generalized solution as in [13,18]. By the monotonic property of (4), the set of generalized solutions is convex. Using the arguments in [13], the solution of (4) can be reformulated as the minimizer of the following function,

$$U_n(\boldsymbol{b},\gamma)=\sum_{i=1}^{n}\delta_{X_i}\left|\frac{h(X_i)}{\hat{H}_{z_i}(X_i)}-\boldsymbol{b}^{\mathrm{T}}\frac{\boldsymbol{Z}_i}{\hat{H}_{z_i}(X_i)}\right|+\left|M-\boldsymbol{b}^{\mathrm{T}}\sum_{l=1}^{n}\frac{-\boldsymbol{Z}_l\delta_{X_l}}{\hat{H}_{z_l}(X_l)}\right|+\left|M-\boldsymbol{b}^{\mathrm{T}}\sum_{k=1}^{n}\left(2\boldsymbol{Z}_k\gamma\right)\right|,$$

where M is a large enough positive value to bound $\left|\boldsymbol{b}^{\mathrm{T}}\sum_{i=1}^{n}\frac{-\boldsymbol{Z}_l\delta_{X_l}}{\hat{H}_{z_l}(X_l)}\right|$ and $\left|\boldsymbol{b}^{\mathrm{T}}\sum_{k=1}^{n}(2\boldsymbol{Z}_k\gamma)\right|$ from above.

We suggest using a re-sampling approach for variance estimation since the analytic formula for the variance of $\hat{\boldsymbol{\beta}}(\gamma)$ is complicated to calculate. Based on the non-parametric bootstrap approach, we can sample replications $\left\{\left(X_i',Y_i',\delta_{x_i}',\delta_{y_i}'\right)(i=1,\cdots,n)\right\}$ from the original data. Given a bootstrap sample, we can compute $\hat{\boldsymbol{\beta}}'(\gamma)$. Repeating the re-sampling procedure B times, we obtain $\left\{\hat{\boldsymbol{\beta}}_b'(\gamma):b=1,\cdots,B\right\}$ and the variance of $\hat{\boldsymbol{\beta}}(\gamma)$ can be estimated by

$$V_{\hat{\beta}(\gamma)}=\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\boldsymbol{\beta}}_b'(\gamma)-\overline{\boldsymbol{\beta}}'(\gamma)\right)^2,$$

where $\overline{\boldsymbol{\beta}}'(\gamma)=\sum_{i=b}^{B}\hat{\boldsymbol{\beta}}_b'(\gamma)/B$. Furthermore, we can construct the $(1-\alpha)$ confidence interval for $\boldsymbol{\beta}(\gamma)$ as $\hat{\boldsymbol{\beta}}(\gamma)\pm V_{\hat{\beta}(\gamma)}^{1/2}z_{1-\alpha/2}$, where $z_{1-\alpha/2}=\Phi^{-1}(1-\alpha/2)$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. The bootstrap percentile method suggests another way of constructing a $(1-\alpha)$ confidence interval of $\boldsymbol{\beta}(\gamma)$ with the formula $\left[\hat{\boldsymbol{\beta}}_{(B\times\alpha/2)}'(\gamma),\hat{\boldsymbol{\beta}}_{(B\times(1-\alpha/2))}'(\gamma)\right]$, where $\hat{\boldsymbol{\beta}}_{(b)}'(\gamma)$, $b=1,\cdots,B$ are the order statistics of $\hat{\boldsymbol{\beta}}_b'(\gamma)$ for $b=1,\cdots,B$.

3.2. Asymptotic Properties for Discrete Covariates

We establish the uniform consistency and weak convergence of the proposed estimator $\hat{\boldsymbol{\beta}}(\gamma)$ for $\gamma\in[\gamma_L,\gamma_U]$, a region that $\beta_0(\gamma)$ is identifiable. We first state the regularity conditions.

(C1) Denote the set of possible covariate $\boldsymbol{Z}$ values as $\mathcal{Z}$ which is a compact set in $\mathcal{R}^{p+1}$. The probability density function $f_{\boldsymbol{Z}}(\boldsymbol{z})$ for covariate $\boldsymbol{Z}$ is uniformly bounded above and below on $\mathcal{Z}$.

(C2) There exists a compact set $\mathcal{A}$ in the parameter space for the copula parameter α such that all true values of $\alpha(\boldsymbol{z})$ are interior points of $\mathcal{A}$ for all $\boldsymbol{z}\in\mathcal{Z}$.

(C3) There exists $\nu>0$ such that $\Pr(C>\nu)=0$, $\Pr(C=\nu)>0$, $\inf_{z\in\mathcal{Z}}S_{T,z}(\nu)\wedge S_{D,z}(\nu)>0$ and $\sup_{z\in\mathcal{Z}}S_{T,z}(\nu)\vee S_{D,z}(\nu)<1$.

(C4) 1) $\boldsymbol{\beta}_0(\gamma)$ is Lipschitz continuous for $\gamma\in[\gamma_L,\gamma_U]$; 2) The density $f_{T,z}(t)=-\frac{\mathrm{d}}{\mathrm{d}t}S_{T,z}(t)$ is bounded above uniformly for $t\in[0,\nu]$ and $z\in\mathcal{Z}$; 3) The copula generator function $\phi_\alpha(u)$ has continuous derivatives $\phi_\alpha'(u)$, $\phi_\alpha''(u)$, $\phi_\alpha'''(u)$, $\dot{\phi}_\alpha(u)=\frac{\partial}{\partial\alpha}\phi_\alpha(u)$ and $\dot{\phi}_\alpha'(u)$ which do not equal 0 for all $\alpha\in\mathcal{A}$ and $u\in(0,1]$.

(C5) $\inf_{\boldsymbol{b}\in\mathcal{B}(\rho_0)}$ eigmin $\boldsymbol{A}(\boldsymbol{b})\ge c_0$, for some $\rho_0>0$ and $c_0>0$, where $\boldsymbol{A}(\boldsymbol{b})=E\left[\boldsymbol{Z}_{T,\boldsymbol{Z}}^{\otimes 2}\left\{h^{-1}\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}\right)\right\}\right]$,

$$B(\rho)=\left\{\boldsymbol{b}\in R^{p+1}:\inf_{\gamma\in[\gamma_L,\gamma_U]}\left\|\boldsymbol{b}-\boldsymbol{\beta}_0(\gamma)\right\|\le\rho\right\},$$

and $\boldsymbol{u}^{\otimes 2}=\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$ for a vector u.

Condition C1 assumes the boundedness of covariates and is satisfied for finite discrete covariates. This assumption is only used to derive the asymptotic properties of $\hat{S}_{D|T,z}$ for proving Theorem 1. Condition C2 assumes that the true value of α is an interior point in the parameter space which is a common regularity condition. Condition C3 is assumed to simplify theoretical arguments similar to condition C1 in [13], and generally ν is the study end time in practical applications. Conditions C4 1) and 2) assume the smoothness of coefficient processes, and the uniform boundedness on the density of T, which are standard for quantile regression methods. Condition C4 3) imposes the smoothness requirement on the copula generator function similar to the regularity conditions in [17,19]. Condition C5 is similar to condition C4 in [13] which ensures the identifiability of $\boldsymbol{\beta}_0(\gamma)$ and is needed for proving the consistency of $\hat{\boldsymbol{\beta}}(\gamma)$.

Therefore with finite $\mathcal{Z}$, we prove the following result.

Theorem 1 If conditions C1-C5 hold, then

$$\lim_{n\to\infty}\sup_{\gamma\in[\gamma_L,\gamma_U]}\left\|\hat{\boldsymbol{\beta}}(\gamma)-\boldsymbol{\beta}_0(\gamma)\right\|\xrightarrow{p}0,$$

and $n^{1/2}\left\{\hat{\boldsymbol{\beta}}(\gamma)-\boldsymbol{\beta}_0(\gamma)\right\}$ converges weakly to a mean-zero Gaussian process.

The detailed proofs are presented in the Appendix.

3.3. Model Checking and Model Diagnosis

Motivated by the work of [20-22] in which complete data are considered, we define the residual quantities as

$$e_i(\gamma)=I\left[h(X_i)\le\hat{\boldsymbol{\beta}}_\gamma^{\mathrm{T}}\boldsymbol{Z}_i\right]\delta_{X_i}/H_{z_i}(X_i)-\gamma$$

for $i=1,\cdots,n$ and consider

$$\ell_n(\gamma)=n^{-1/2}\sum_{i=1}^{n}q(\boldsymbol{Z}_i)e_i(\gamma),$$

where $q(\cdot)$ is a known bounded weight function. Similar to the arguments in [13,23], $\ell_n(\gamma)$ converges weakly to a zero-mean Gaussian process if model (1) is specified correctly and the covariate takes discrete values. Therefore we propose the following test statistic

$$T_n=n^{-1/2}\sum_{i=1}^{n}\frac{q(\boldsymbol{Z}_i)e_i(\gamma)}{\hat{\sigma}_e},$$

where $\hat{\sigma}_e$ is an estimator of the standard deviation of

$\ell_n(\gamma)$ which can be obtained by applying the bootstrap approach mentioned earlier. Thus, we have that T_n converges to the standard normal random variable asymptotically as the model is correct. On the other hand, when the model is mis-specified, T_n will deviate from zero. Accordingly we can reject the model assumption if $|T_n| > Z_{\kappa/2}$, where $Z_{\kappa/2}$ is the quantile of $N(0,1)$ and κ is the level of significance. If there are K candidate models under consideration, we compute the absolute value of T_n for each model for $k = 1,\cdots,K$ and choose the one with the smallest value.

3.4. Estimation for Continuous Covariates

We briefly discuss how to extend our estimation method for continuous covariates. One can apply a smoothing approach to estimate the probability functions conditional on z. Following [24], without loss of generality, assume that $Z \in [0,1]$ and $Z_1 < Z_2 < \cdots < Z_n$ are ordered. Let

$$w_{ni}(z,h_n) = \frac{1}{c_n(z,h_n)} \int_{t=Z_{i-1}}^{Z_i} \frac{1}{h_n} K\left(\frac{z-t}{h_n}\right) \mathrm{d}t, (i = 1,\cdots,n),$$

$$c_n(z,h_n) = \int_{t=Z_0}^{Z_n} \frac{1}{h_n} K\left(\frac{z-t}{h_n}\right) \mathrm{d}t,$$

where $Z_0 = 0$, $h_n \to 0$ is the bandwidth and K is the kernel. Then

$$\hat{S}_{W,z}(x) = \prod_{W_{(i)} \le x} \left(1 - \frac{w_{n(i)}(z,h_n)}{1 - \sum_{j=1}^{i-1} w_{n(i)}(z,h_n)}\right)^{\delta_{W(i)}},$$

where $\left\{\left(W_{(i)}, \delta_{W_{(i)}}, w_{n(i)}(z,h_n)\right)(i = 1,\cdots,n)\right\}$ are the rearrangement $\left\{\left(W_i, \delta_{W_i}, w_{ni}(z,h_n)\right)(i = 1,\cdots,n)\right\}$ sorted according to W_i, where $W_i = T_i \wedge D_i$ and $\delta_{W_i} = I(T_i \wedge D_i < C_i)$, and $\hat{S}_{T,z}(x)$ is the copula-graphic estimator in [24]

$$\hat{S}_{T,z}(x) = \phi^{-1}_{\hat{\alpha}(\mathbf{z})}\left\{-\sum_{i=1}^{n} I\left(X_i \le x, \delta_{X_i} = 1\right)\left\{\phi_{\hat{\alpha}(\mathbf{z})}\left[\hat{S}_{W,\mathbf{z}}\left(X_i^-\right)\right] - \phi_{\hat{\alpha}(\mathbf{z})}\left[\hat{S}_{W,\mathbf{z}}\left(X_i^-\right) - w_{ni}(z,h_n)\right]\right\}\right\}, \tag{6}$$

and $\hat{\alpha}(z)$ solves estimating equation

$$\sum_{i<j} K\left(\frac{z-Z_i}{h_n}\right) K\left(\frac{z-Z_j}{h_n}\right) w\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right) I\left(\tilde{T}_{ij} \le \tilde{D}_{ij} \le \tilde{C}_{ij}\right)\left\{I\left(\left(X_i - X_j\right)\left(Y_i - Y_j\right) > 0\right) - \frac{\theta_{\alpha(\mathbf{z})}\left(\hat{\pi}_z\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right)}{\theta_{\alpha(\mathbf{z})}\left(\hat{\pi}_z\left(\tilde{X}_{ij}, \tilde{Y}_{ij}\right)\right) + 1}\right\} = 0.$$

Special techniques are needed to derive the asymptotic properties for the case of continuous covariates. For example properties of the smoothed versions of $\hat{S}_{D|T,z}$ and $\hat{\alpha}(z)$ are not fully available yet. The $n^{1/2}$ convergence rate for the normality proof may not be directly extended since the smooth version of $\hat{S}_{D|T,z}$ may not be $n^{1/2}$ asymptotic normal. However the estimator for the quantile regression parameter may still be $n^{1/2}$ asymptotic normal even when some component converges at a slower rate.

4. Simulation Studies

We conduct simulation studies to examine the finite-sample performance of the proposed methods with R software. Here we consider two cases. For the first one, we consider the model,

$$\log(T) = \beta_0^{(1)}(\gamma) + \beta_0^{(2)}(\gamma) Z + \epsilon_\gamma, \tag{7}$$

where $Z \sim \mathrm{Ber}(0.5)$ and $\left(\beta_0^{(1)}(\gamma), \beta_0^{(2)}(\gamma)\right) = (-1,-1)$. We generate (ϵ_γ, D) which follow the Clayton copula and Frank copula with ϵ_γ marginally following $U(-0.5\gamma, 0.5-0.5\gamma)$ so that $\Pr(\varepsilon_\gamma \le 0) = \gamma$, and D marginally following exp(2). For the second case, we consider

$$\log(T) = b_0 + b_1(1+Z)\epsilon, \tag{8}$$

where $Z \sim \mathrm{Ber}(0.5)$, $(b_0, b_1) = (-1.5, 0.5)$ and (ϵ, D) generated from the Clayton copula and Frank copula with ϵ following $U(0,0.5)$ and $D \sim \exp(2)$. In this case, $\left(\beta_0^{(1)}(\gamma), \beta_0^{(2)}(\gamma)\right) = (b_0 + 0.5b_1\gamma, 0.5b_1\gamma)$. Three levels of association τ = 0.3, 0.5, 0.7 are considered. The censoring variable C follows a uniform distribution on $[0,12]$.

We evaluate the performances for γ = 0.1, 0.3, 0.5 and the sample size n = 100 based on 400 simulation runs. To obtain the standard error of the proposed estimator, we use the bootstrap method with B = 50. Based on the settings, we also present a naive estimator of $\boldsymbol{\beta}_0(\gamma)$, which is constructed under the wrong assumption that T is independently censored by $D \wedge C$. That is, we estimate $\boldsymbol{\beta}_0(\gamma)$ by solving the estimating Equation (4) with

$$\hat{H}_{z_i}(X_i) = \widehat{\Pr}\left(D \wedge C > X_i \mid Z = z_i\right) = \frac{\sum_{j=1}^{n} I\left(D_j \wedge C_j > X_i, Z_j = z_i\right)}{\sum_{j=1}^{n} I\left(Z_j = z_i\right)}.$$

Tables 1-4 report the average bias of the proposed

Table 1. Finite-sample results for estimating the quantile regression parameters under model (7) with Clayton copula.

τ	γ	Method	$\hat{\beta}^{(1)}(\gamma)$				$\hat{\beta}^{(2)}(\gamma)$			
			Bias	EmpSd	MSE	CP	Bias	EmpSd	MSE	CP
0.3	0.1	Proposed	0.0059	0.0263	0.0007	0.975	0.0011	0.0369	0.0013	0.980
		Naive	0.0129	0.0284	0.0009	0.975	−0.0019	0.0380	0.0014	0.975
	0.3	Proposed	0.0024	0.0383	0.0014	0.940	−0.0001	0.0498	0.0024	0.945
		Naive	0.0183	0.0379	0.0017	0.947	−0.0115	0.0487	0.0025	0.942
	0.5	Proposed	0.0001	0.0380	0.0014	0.935	0.0024	0.0546	0.0029	0.922
		Naive	0.0163	0.0386	0.0017	0.915	−0.0078	0.0554	0.0031	0.930
0.5	0.1	Proposed	0.0074	0.0300	0.0009	0.955	−0.0007	0.0422	0.0017	0.962
		Naive	0.0240	0.0318	0.0015	0.891	−0.0124	0.0404	0.0017	0.932
	0.3	Proposed	0.0064	0.0391	0.0015	0.925	−0.0024	0.0500	0.0025	0.962
		Naive	0.0369	0.0384	0.0028	0.817	−0.0232	0.0511	0.0031	0.922
	0.5	Proposed	−0.0001	0.0381	0.0014	0.912	0.0025	0.0497	0.0024	0.957
		Naive	0.0259	0.0357	0.0019	0.867	−0.0139	0.0495	0.0026	0.942
0.7	0.1	Proposed	0.0087	0.0323	0.0011	0.945	−0.0008	0.0440	0.0019	0.967
		Naive	0.0420	0.0335	0.0028	0.802	−0.0260	0.0432	0.0025	0.920
	0.3	Proposed	0.0073	0.0371	0.0014	0.925	0.0004	0.0519	0.0026	0.927
		Naive	0.0475	0.0358	0.0035	0.707	−0.0254	0.0528	0.0034	0.902
	0.5	Proposed	0.0065	0.0378	0.0014	0.937	−0.0029	0.0521	0.0027	0.945
		Naive	0.0314	0.0344	0.0021	0.845	−0.0159	0.0477	0.0025	0.942

The results are based on 400 simulation runs each with a sample size 100.

Table 2. Finite-sample results for estimating the quantile regression parameters under model (8) with Clayton copula.

τ	γ	Method	$\hat{\beta}^{(1)}(\gamma)$				$\hat{\beta}^{(2)}(\gamma)$			
			Bias	EmpSd	MSE	CP	Bias	EmpSd	MSE	CP
0.3	0.1	Proposed	−0.0025	0.0330	0.0011	0.922	−0.0094	0.1261	0.0160	0.957
		Naive	0.0032	0.0270	0.0007	0.957	−0.0076	0.1252	0.0157	0.965
	0.3	Proposed	0.0002	0.0194	0.0003	0.942	0.0005	0.0415	0.0017	0.932
		Naive	0.0071	0.0210	0.0005	0.927	0.0042	0.0584	0.0034	0.940
	0.5	Proposed	0.0004	0.0201	0.0004	0.920	0.0001	0.0442	0.0019	0.942
		Naive	0.0058	0.0199	0.0004	0.877	0.0070	0.0444	0.0020	0.915
0.5	0.1	Proposed	−0.0015	0.0326	0.0010	0.925	−0.0133	0.1291	0.0168	0.960
		Naive	0.0030	0.0288	0.0008	0.952	0.0011	0.1060	0.0112	0.977
	0.3	Proposed	0.0028	0.0196	0.0003	0.912	−0.0007	0.0425	0.0018	0.913
		Naive	0.0129	0.0189	0.0005	0.880	0.0098	0.0437	0.0020	0.915
	0.5	Proposed	0.0010	0.0199	0.0004	0.917	0.0001	0.0418	0.0017	0.935
		Naive	0.0098	0.0183	0.0004	0.891	0.0104	0.0387	0.0016	0.935
0.7	0.1	Proposed	−0.0011	0.0331	0.0011	0.927	−0.0087	0.1293	0.0167	0.957
		Naive	0.0086	0.0264	0.0007	0.952	−0.0103	0.1655	0.0275	0.942
	0.3	Proposed	0.0022	0.0166	0.0002	0.942	0.0032	0.0383	0.0014	0.932
		Naive	0.0166	0.0178	0.0005	0.877	0.0169	0.0413	0.0019	0.912
	0.5	Proposed	0.0022	0.0180	0.0003	0.917	0.0007	0.0367	0.0013	0.957
		Naive	0.0122	0.0174	0.0004	0.867	0.0108	0.0366	0.0014	0.932

The results are based on 400 simulation runs each with a sample size 100.

Table 3. Finite-sample results for estimating the quantile regression parameters under model (7) with Frank copula.

τ	γ	Method	$\hat{\beta}^{(1)}(\gamma)$				$\hat{\beta}^{(2)}(\gamma)$			
			Bias	EmpSd	MSE	CP	Bias	EmpSd	MSE	CP
0.3	0.1	Proposed	0.0077	0.0316	0.0010	0.937	0.0021	0.0428	0.0018	0.947
		Naive	0.0230	0.0314	0.0015	0.887	−0.0104	0.0411	0.0017	0.942
	0.3	Proposed	0.0045	0.0366	0.0013	0.945	0.0003	0.0517	0.0026	0.955
		Naive	0.0272	0.0376	0.0021	0.892	−0.0155	0.0500	0.0027	0.930
	0.5	Proposed	0.0001	0.0380	0.0014	0.957	−0.0005	0.0519	0.0027	0.947
		Naive	0.0203	0.0361	0.0017	0.932	−0.0127	0.0501	0.0026	0.930
0.5	0.1	Proposed	0.0109	0.0324	0.0011	0.932	−0.0014	0.0438	0.0019	0.935
		Naive	0.0399	0.0325	0.0026	0.770	−0.0213	0.0443	0.0024	0.925
	0.3	Proposed	0.0073	0.0388	0.0015	0.937	−0.0022	0.0524	0.0027	0.942
		Naive	0.0429	0.0360	0.0031	0.790	−0.0233	0.0507	0.0031	0.930
	0.5	Proposed	0.0050	0.0366	0.0013	0.947	−0.0033	0.0515	0.0026	0.962
		Naive	0.0296	0.0343	0.0020	0.852	−0.0171	0.0484	0.0026	0.952
0.7	0.1	Proposed	0.0261	0.0309	0.0016	0.885	−0.0130	0.0426	0.0019	0.937
		Naive	0.0586	0.0325	0.0044	0.572	−0.0311	0.0443	0.0029	0.907
	0.3	Proposed	0.0331	0.0323	0.0021	0.867	−0.0236	0.0452	0.0026	0.915
		Naive	0.0575	0.0329	0.0043	0.582	−0.0362	0.0449	0.0033	0.862
	0.5	Proposed	0.0210	0.0361	0.0017	0.902	−0.0158	0.0533	0.0030	0.947
		Naive	0.0312	0.0341	0.0021	0.830	−0.0195	0.0512	0.0030	0.955

The results are based on 400 simulation runs each with a sample size 100.

Table 4. Finite-sample results for estimating the quantile regression parameters under model (8) with Frank copula.

τ	γ	Method	$\hat{\beta}^{(1)}(\gamma)$				$\hat{\beta}^{(2)}(\gamma)$			
			Bias	EmpSd	MSE	CP	Bias	EmpSd	MSE	CP
0.3	0.1	Proposed	−0.0049	0.0331	0.0011	0.937	−0.0114	0.1031	0.0107	0.977
		Naive	0.0018	0.0288	0.0008	0.922	0.0028	0.1013	0.0102	0.987
	0.3	Proposed	0.0022	0.0179	0.0003	0.947	−0.0001	0.0400	0.0016	0.937
		Naive	0.0092	0.0196	0.0004	0.942	0.0082	0.0401	0.0016	0.955
	0.5	Proposed	0.0016	0.0187	0.0003	0.940	0.0012	0.0417	0.0017	0.950
		Naive	0.0084	0.0183	0.0004	0.920	0.0117	0.0396	0.0017	0.937
0.5	0.1	Proposed	0.0001	0.0313	0.0009	0.930	−0.0126	0.1317	0.0175	0.980
		Naive	0.0082	0.0316	0.0010	0.927	−0.0076	0.1440	0.0208	0.977
	0.3	Proposed	0.0041	0.0181	0.0003	0.950	0.0036	0.0394	0.0015	0.945
		Naive	0.0159	0.0178	0.0005	0.835	0.0167	0.0390	0.0018	0.927
	0.5	Proposed	0.0029	0.0167	0.0002	0.960	0.0024	0.0397	0.0015	0.950
		Naive	0.0126	0.0167	0.0004	0.872	0.0126	0.0387	0.0016	0.937
0.7	0.1	Proposed	−0.0002	0.0277	0.0007	0.920	−0.0177	0.1679	0.0285	0.967
		Naive	0.0141	0.0281	0.0009	0.932	−0.0075	0.1719	0.0296	0.967
	0.3	Proposed	0.0082	0.0158	0.0003	0.917	0.0123	0.0339	0.0013	0.930
		Naive	0.0178	0.0176	0.0006	0.857	0.0223	0.0355	0.0017	0.907
	0.5	Proposed	0.0075	0.0174	0.0003	0.920	0.0063	0.0390	0.0015	0.955
		Naive	0.0123	0.0167	0.0004	0.882	0.0102	0.0379	0.0015	0.942

The results are based on 400 simulation runs each with a sample size 100.

point estimator, $\sum_{i=1}^{400}\hat{\beta}_i^{(j)}(\gamma)/400-\beta_0^{(j)}(\gamma),(j=1,2)$, (Bias); the empirical standard deviation,

$$\sum_{i=1}^{400}\left(\hat{\beta}_i^{(j)}(\gamma)-\bar{\beta}^{(j)}(\gamma)\right)^2 \Big/ 399,$$

where $\bar{\beta}^{(j)}(\gamma)=\sum_{i=1}^{400}\hat{\beta}_i^{(j)}(\gamma)/400$, (EmpSd); the mean squared error, Bias^2 + EmpSd^2, (MSE); and the coverage probability of the 95% confidence intervals,

$$\sum_{i=1}^{400} I\left(\beta_0^{(j)}(\gamma)\in\hat{\beta}_i^{(j)}(\gamma)\pm 1.96\times Sd_i^{(j)}\right)\Big/400,$$

where $Sd_i^{(j)}$ is the estimated standard deviation of $\hat{\beta}_i^{(j)}(\gamma)$ by the bootstrap approach, (CP). From the results, we can see that our proposed estimator has much smaller bias and smaller mean squared error than the naive estimator. The confidence intervals coverage probabilities are close to the nominal level 95% in most cases while the naive estimator has the coverage rate far below the nominal level in many cases. Although the proposed estimator of $\beta_0^{(1)}$ has the coverage rate lower than 90% in the first case with Kendall's tau $\tau = 0.7$ but it still performs better than the naive estimator. As the sample size increases to $n = 200$ for that case (data omitted here), the coverage probabilities for proposed estimator become close to the nominal level while the coverage probabilities for the naive estimator get worse. This confirms that our estimator is asymptotically correct while the naive estimator is not.

Then we examine the proposed model diagnostic method when the true model is generated from

$$\log(T)=\beta_0(\gamma)Z+\epsilon_\gamma,$$

where $\beta_0(\gamma)=-1$, $Z\sim 1+Ber(0.5)$, and $\epsilon_\gamma\sim U(-0.5\gamma,0.5-0.5\gamma)$ so that $\xi_\gamma(\epsilon_\gamma)=0$ and (ϵ_γ,D) follow Clayton copula with $D\sim\exp(2)$. We consider $\tau = 0.3, 0.5, 0.7$ and $\gamma = 0.1, 0.3, 0.5$ under $n = 100$ based on 200 replications.

Three forms of transformation are fitted: 1) $h(t)=\log(t)$; 2) $h(t)=t$; 3) $h(t)=2(t^{1/2}-1)$. **Table 5** presents the rejection probability $\sum_{i=1}^{200} I\left(|T_{n,i}|>Z_{\alpha/2}\right)/200$, where $\alpha = 0.05$, and the probability that the fitted model is selected as the one which gives the smallest value of $|T_n|$ among the three candidates. From the results, we see that when $h(t)=\log(t)$, the rejection probability (type-I error rate) is close to the specified level of $\alpha = 0.05$. When the fitted model is wrong, the rejection probability (power of the test) is very high in most cases.

Table 5. Finite-sample results for the proposed model checking method.

Kendall's τ	Quantile γ		$h(t)=\log(t)$	$h(t)=t$	$h(t)=2(t^{1/2}-1)$
0.3	0.1	Power	0.07	0.385	0.965
		Selection rate	0.82	0.18	0
	0.3	Power	0.05	0.995	1
		Selection rate	0.995	0	0.005
	0.5	Power	0.045	1	0.97
		Selection rate	0.995	0	0.005
0.5	0.1	Power	0.06	0.415	0.99
		Selection rate	0.845	0.155	0
	0.3	Power	0.035	0.995	0.995
		Selection rate	0.995	0.005	0
	0.5	Power	0.025	1	0.99
		Selection rate	1	0	0
0.7	0.1	Power	0.06	0.425	0.965
		Selection rate	0.85	0.145	0.005
	0.3	Power	0.08	0.975	1
		Selection rate	0.995	0.005	0
	0.5	Power	0.045	1	0.985
		Selection rate	1	0	0

Note: The sample size is 100 and replications are 200. "Power" = $\sum_{i=1}^{200} I\left(|T_{n,i}|>Z_{\alpha/2}\right)/200$, where $\alpha=0.05$. "Selection rate" is the proportion that the fitted model is selected as the one giving the smallest value of $|T_n|$ among the three candidates.

Even for the case where the power is relatively low around 40% (the γ = 0.1 quantile for $h(t) = t$), the probabilities of selecting the correct model are still high.

5. Data Analysis

We apply the proposed methodology to analyze the bone marrow transplant data based on 137 leukemia patients provided by [1]. Patients were classified into three risk categories: ALL, AML low-risk, and AML high-risk based on their status at the time of transplantation. The covariates (Z_1, Z_2) are coded as ALL (Z_1 = 1, Z_2 = 0), AML low-risk (Z_1 = 1, Z_2 = 0), and AML high-risk (Z_1 = 0, Z_2 = 1). We want to investigate how the risk classification is related to the quantile of the relapse time. Specifically the fitted model is given by

$$\xi_\gamma\left(\log(T) \middle| Z_1, Z_2\right) = \beta_0(\gamma) + \beta_1(\gamma) Z_1 + \beta_2(\gamma) Z_2. \quad (9)$$

The results are summarized in the **Tables 6** and **7** based on B = 1000 bootstrap replications. **Table 6** contains the estimators and model checking tests with $q(Z_1, Z_2) = 1 / (Z_1 + Z_2 + 0.2)^2$. The p-value is the testing result by the model checking approach provided in Sub-Section 3.3. Since all the p-values are greater than 0.05, we adopt the model in (9) for further analysis.

From the analysis we see that patients of AML low-risk had longer relapse time than those in the other two groups and the difference is more obvious for those with earlier relapse. For example, the 10% quantile of the relapse time in the AML low-risk group is 3.2964 times of that in ALL group and 4.751 times of that in AML high-risk group. The group differences are statistically significant for the 10% and 30% quantiles. but no longer significant for the 50% quantile.

6. Concluding Remarks

In this paper, we consider quantile regression analysis for analyzing the failure-time of a non-terminal event under the semi-competing risks setting. The Archimedean copula assumption is adopted to specify the dependency between the two correlated events. This assumption is utilized to calculate the weight for bias correction in the estimation of quantile regression parameters. Here we focus on the case of discrete covariates and derive the asymptotic properties of the proposed estimators. The bootstrap method is suggested for variance estimation. For checking the adequacy of the fitted model, a model diagnostic approach is proposed. Simulation results confirm that the proposed methods have good performances in finite samples. In the data analysis, we see that the risk classification is particularly influential for earlier relapse. The methodology can be extended to allow for continuous covariates by employing some smoothing techniques but the corresponding theoretical analysis is beyond the scope of the paper.

Table 6. Estimation of quantile regression parameters and model checking test based on the bone marrow transplant data.

Quantile γ	$\boldsymbol{\beta}_0$				$\boldsymbol{\beta}_1$				$\boldsymbol{\beta}_2$				p-value
	$\hat{\boldsymbol{\beta}}_0$	Sd	95% CI		$\hat{\boldsymbol{\beta}}_1$	Sd	95% CI		$\hat{\boldsymbol{\beta}}_2$	Sd	95% CI		
0.1	4.587	0.262	4.109	5.080	1.193	0.301	0.700	1.691	−0.366	0.335	−1.137	0.191	0.581
0.3	5.571	0.198	5.222	6.169	1.278	0.331	0.532	1.894	−0.208	0.353	−0.936	0.380	0.897
0.5	6.129	0.409	5.498	6.819	1.155	0.514	−0.014	1.968	−0.374	0.608	−1.180	1.242	0.220

Table 7. Comparison of leukemia relapse time for the three risk groups.

Quantile γ	Disease Group	Diff.	Std. Err.	exp(Diff.)	95% CI of exp(Diff.)	
0.1	Low vs All $(\boldsymbol{\beta}_1)$	1.193	0.301	3.296	2.014	5.422
	High vs All $(\boldsymbol{\beta}_2)$	−0.366	0.335	0.694	0.321	1.210
	Low vs High $(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)$	1.558	0.254	4.751	2.902	8.446
0.3	Low vs All $(\boldsymbol{\beta}_1)$	1.278	0.331	3.588	1.702	6.647
	High vs All $(\boldsymbol{\beta}_2)$	−0.208	0.352	0.812	0.392	1.462
	Low vs High $(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)$	1.485	0.443	4.416	1.756	10.494
0.5	Low vs All $(\boldsymbol{\beta}_1)$	1.155	0.514	3.175	0.986	7.153
	High vs All $(\boldsymbol{\beta}_2)$	−0.374	0.608	0.687	0.307	3.462
	Low vs High $(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)$	1.529	0.626	4.616	0.829	8.877

7. Acknowledgement

This paper was financially supported by the National Science Council of Taiwan (NSC100-2118-M-194-003-MY2).

REFERENCES

[1] J. P. Klein and M. L. Moeschberger, "Survival Analysis: Techniques for Censored and Truncated Data," 2nd Edition, Springer, New York, 2003.

[2] J. Powell, "Least Absolute Deviations Estimation for the Censored Regression Model," *Journal of Econometrics*, Vol. 25, No. 3, 1984, pp. 303-325. doi:10.1016/0304-4076(84)90004-6

[3] J. Powell, "Censored Regression Qunatiles," *Journal of Econometrics*, Vol. 32, No. 1, 1986, pp. 143-155. doi:10.1016/0304-4076(86)90016-3

[4] B. Fitzenberger, "A Guide to Censored Quantile Regressions," In: G. S. Maddala and C. R. Rao, Eds., *Handbooks of Statistics*, North-Holland, Amsterdam, 1997, pp. 405-437. doi:10.1016/S0169-7161(97)15017-9

[5] M. Buchinsky and J. Hahn, "A Alternative Estimator for Censored Quantile Regression," *Econometrica*, Vol. 66, No. 3, 1998, pp. 653-671. doi:10.2307/2998578

[6] Z. Ying, S. H. Jung and L. J. Wei, "Survival Analysis with Median Regression Models," *Journal of the American Statistical Association*, Vol. 90, No. 429, 1995, pp. 178-184. doi:10.1080/01621459.1995.10476500

[7] S. Yang, "Censored Median Regression Using Weighted Empirical Survival and Hazard Functions," *Journal of the American Statistical Association*, Vol. 94, No. 445, 1999, pp. 137-145. doi:10.1080/01621459.1999.10473830

[8] S. Portnoy, "Censored Regression Quantiles," *Journal of the American Statistical Association*, Vol. 98, No. 464, 2003, pp. 1001-1012. doi:10.1198/016214503000000954

[9] L. Peng and Y. Huang, "Survival Analysis Based on Quantile Regression Models," *Journal of the American Statistical Association*, Vol. 103, No. 482, 2008, pp. 637-649. doi:10.1198/016214508000000355

[10] G. Yin, D. Zeng and H. Li, "Power Transformed Linear Quantile Regression with Censored Data," *Journal of the American Statistical Association*, Vol. 103, No. 483, 2008, pp. 1214-1224. doi:10.1198/016214508000000490

[11] S. Portnoy and G. Lin, "Asymptotics for Censored Regression Quantiles," *Journal of Nonparametric Statistics*, Vol. 22, No. 1, 2010, pp. 115-130. doi:10.1080/10485250903105009

[12] J. P. Fine, H. Jiang and R. Chappell, "On Semi-Competing Risks Data," *Biometrika*, Vol. 88, No. 4, 2001, pp. 907-919.

[13] L. Peng and J. P. Fine, "Competing Risks Quantile Regression," *Journal of the American Statistical Association*, Vol. 104, No. 488, 2009, pp. 1440-1453. doi:10.1198/jasa.2009.tm08228

[14] R. Li and L. Peng, "Quantile Regression for Left-Truncated Semicompeting Risks Data," *Biometrics*, Vol. 67 No. 3, 2011, pp. 701-710. doi:10.1111/j.1541-0420.2010.01521.x

[15] A. J. Patton, "Modeling Asymmetric Exchange Rate Dependence," *International Economic Review*, Vol. 47, No. 2, 2006, pp. 527-556. doi:10.1111/j.1468-2354.2006.00387.x

[16] W. Wang, "Estimating the Association Parameter for Copula Models under Dependent Censoring," *Journal of the Royal Statistical Society*, *Series B*, Vol. 65, No. 1, 2003, pp. 257-273. doi:10.1111/1467-9868.00385

[17] L. Lakhal, L.-P. Rivest and B. Abdous, "Estimating Survival and Association in a Semicompeting Risks Model," *Biometrics*, Vol. 64, No. 1, 2008, pp. 180-188. doi:10.1111/j.1541-0420.2007.00872.x

[18] M. Fygenson and Y. Ritov, "Monotone Estimating Equations for Censored Data," *The Annals of Statistics*, Vol. 22, No. 2, 1994, pp. 732-746. doi:10.1214/aos/1176325493

[19] L.-P. Rivest and M. Wells, "A Martingale Approach to the Copula-Graphic Estimator for the Survival Function under Dependent Censoring," *Journal of Multivariate Analysis*, Vol. 79, No. 1, 2001, pp. 138-155. doi:10.1006/jmva.2000.1959

[20] J. X. Zheng, "A Consistent Nonparametric Regression Models under Quantile Restrictions," *Econometric Theory*, Vol. 14, No. 1, 1998, pp. 123-138. doi:10.1017/S0266466698141051

[21] J. L. Horowitz and V. G. Spokoiny, "An Adaptive, Rate-Optimal Test of Linearity for Median Regression Models," *Journal of the American Statistical Association*, Vol. 97, No. 459, 2002, pp. 822-835. doi:10.1198/016214502388618627

[22] X. He and L. Zhu, "A Lack-of-Fit Test for Quantile Regression," *Journal of the American Statistical Association*, Vol. 98, No. 464, 2003, pp. 1013-1022. doi:10.1198/016214503000000963

[23] D. Y. Lin and L. J. Wei and Z. Ying, "Checking the Cox Model with Cumulative Sums of Martingle-Based Residuals," *Biometrika*, Vol. 80, No. 3, 1993, pp. 557-572. doi:10.1093/biomet/80.3.557

[24] R. Braekers and N. Veraverbeke, "A Copula-Graphic Estimator for the Conditional Survival Function under Dependent Censoring," *The Canadian Journal of Statistics*, Vol. 33, No. 3, 2005, pp. 429-447. doi:10.1002/cjs.5540330308

Appendix: Proofs of Theorem 1

The proof follow the outline similar to the proof in [13]. The technical details need to be adjusted for dependent censoring which make things harder.

Define

$$S_n^H(\boldsymbol{b},\gamma)=n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left(\frac{I\left(h(X_i)\le \boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}_i\right)\delta_{xi}}{H_{\boldsymbol{Z}_i}(X_i)}-\gamma\right),$$

$$\tilde{S}_n(\boldsymbol{b},\gamma)=n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left(F_{T,\boldsymbol{Z}_i}\left(h^{-1}(\boldsymbol{b}\boldsymbol{Z}_i)\right)-\gamma\right),$$

$$F_{T,\boldsymbol{Z}_i}(t)=\Pr\left(T\le t\,|\,\boldsymbol{Z}_i\right)=1-S_{T,\boldsymbol{Z}_i}(t),$$

$$\mu(\boldsymbol{b},\gamma)=E\left\{n^{-1/2}\tilde{S}_n(\boldsymbol{b},\gamma)\right\},\boldsymbol{b}\in R^{p+1},\gamma\in[\gamma_L,\gamma_U].$$

For simplicity, we use $\sup_{\boldsymbol{b}},\sup_{\gamma}$ and $\sup_{z}$ to denote supremum taken over $\boldsymbol{b}\in R^{p+1}$, $\gamma\in[\gamma_L,\gamma_U]$ and $\boldsymbol{z}\in\mathcal{Z}$ respectively.

First, we show that $S_n(\boldsymbol{b},\gamma)$ converges uniformly to $S_n^H(\boldsymbol{b},\gamma)$. Since $\mathcal{Z}$ is finite, $n_z=\sum_{i=1}^{n}I(\boldsymbol{Z}_i=\boldsymbol{z})\to\infty$ as $n\to\infty$. Hence by [17], $\hat{\alpha}(\boldsymbol{z})$ is consistent for $\alpha(\boldsymbol{z})$ for all $\boldsymbol{z}\in\mathcal{Z}$. Using this and conditions C2, it implies that $\hat{\alpha}(\boldsymbol{z})\in\mathcal{A}$ with probability 1 for large enough *n*. From condition C3, $\hat{G}_z(t)$ and $\hat{S}_{W,z}(t)$ converge to $G_z(t)$ and $S_{W,z}(t)$ uniformly for $t\in[0,\nu]$. Condition C4 (iii) together with conditions C1 and C3 ensure the uniform boundedness of the first two derivatives of $\phi_{\alpha(z)}(u)$ and $u\phi'_{\alpha(z)}(u)$ for $\boldsymbol{z}\in\mathcal{Z}$ and $u\in\left[\inf_{z\in\mathcal{Z}}S_{T,z}(\nu)\wedge S_{D,z}(\nu)>0,1\right]$, same as the regularity condition in [19]. Hence as $n_z\to\infty$, by [17] and [19], $\hat{S}_{T,z}(t)$ also converges to $S_{T,z}(t)$ uniformly for $t\in[0,\nu]$. Then there exists a number $s_0\in(0,1)$ such that $\hat{S}_{W,z}(t)$ and $\hat{S}_{T,z}(t)$ fall into $(s_0,1]$ with probability 1 for large enough *n* and all $t\in[0,\nu]$ by condition C3 and the uniform convergence of the two estimators. Denote $\Psi(\alpha,u,v)=\phi_{\alpha'}(v)/\phi_{\alpha'}(u)$. Condition C4 (iii) implies that $\left|\frac{\partial}{\partial\alpha}\Psi(\alpha,u,v)\right|$, $\left|\frac{\partial}{\partial u}\Psi(\alpha,u,v)\right|$ and $\left|\frac{\partial}{\partial v}\Psi(\alpha,u,v)\right|$ are all uniformly bounded above for $\alpha\in\mathcal{A},u\in[s_0,1]$ and $v\in[s_0,1]$. Hence $\hat{S}_{D|T,z}(t)=\Psi\left(\hat{S}_{T,z}(t),\hat{S}_{W,z}(t),\hat{\alpha}\right)$ converges to $S_{D|T,z}(t)$ uniformly for $t\in[0,\nu]$ and $\boldsymbol{z}\in\mathcal{Z}$. This result and the uniform convergence of $\hat{G}_z(t)$ imply the uniform convergence of $\hat{H}_z(t)$ for $t\in[0,\nu]$ and $\boldsymbol{z}\in\mathcal{Z}$. Hence we have $\sup_{\boldsymbol{b},\gamma}\left\|n^{-1/2}S_n(\boldsymbol{b},\gamma)-n^{-1/2}S_n^H(\boldsymbol{b},\gamma)\right\|=o_p(1)$.

The function class

$$\left\{\boldsymbol{Z}_i\left(\frac{I\left(X_i\le h^{-1}\left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}_i\right)\right)\delta_{xi}}{H_{\boldsymbol{Z}_i}(X_i)}-\gamma\right),\boldsymbol{b}\in R^{p+1},\gamma\in[\gamma_L,\gamma_U]\right\}$$

is Donsker because the class of indicator functions is Donsker and both $\boldsymbol{Z}_i$ and $1/H_{\boldsymbol{Z}_i}(X_i)$ are uniformly bounded by conditions C1 and C3. Therefore, by Glivenko-Cantelli theorem, $\sup_{\boldsymbol{b},\gamma}\left\|n^{-1/2}S_n^H(\boldsymbol{b},\gamma)-\mu(\boldsymbol{b},\gamma)\right\|=o_p(1)$. Also

$$\sup_{\boldsymbol{b},\gamma}\left\|n^{-1/2}S_n(\boldsymbol{b},\gamma)-\mu(\boldsymbol{b},\gamma)\right\|=o_p(1). \tag{1}$$

Then the consistency of $\hat{\beta}(\gamma)$ comes from the identifiability condition C5 using the arguments in the proof of Theorem 1 in [13].

Similar to [13] the following lemma holds with the uniform boundedness of $\boldsymbol{Z}$, $f_{T,z}(t)$ and $\mathcal{B}(\rho_0)$ that comes from conditions C1, C3 (i) and C3 (ii).

Lemma 1. For any positive sequence $\{d_n\}_{n=1}^{\infty}$ satisfying $d_n\to 0$,

$$\begin{aligned}\lim_{n\to\infty}\sup_{\boldsymbol{b},\boldsymbol{b}'\in\mathcal{B}(\rho_0)\|\boldsymbol{b}-\boldsymbol{b}'\|\le d_n}&\Bigg\|n^{-1/2}\sum_{i=1}^{n}\left[\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{xi}/H_{z_i}(X_i)-\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{b}'^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{xi}/H_{z_i}(X_i)\right]\\&-n^{-1/2}\left\{\mu(\boldsymbol{b},\gamma)-\mu(\boldsymbol{b}',\gamma)\right\}\Bigg\|=0\ a.s.\end{aligned}$$

Now, we provide the proof for the asymptotic normality of $\hat{\beta}(\gamma)$. One can write

$$\begin{aligned}S_n\left(\hat{\boldsymbol{\beta}}(\gamma),\gamma\right)-S_n\left(\boldsymbol{\beta}_0(\gamma),\gamma\right)&=S_n^H\left(\hat{\boldsymbol{\beta}}(\gamma),\gamma\right)-S_n^H\left(\boldsymbol{\beta}_0(\gamma),\gamma\right)\\&+n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left[I\left\{h(X_i)\le\hat{\boldsymbol{\beta}}(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}-I\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\right]\delta_{X_i}\left[\frac{1}{\hat{H}_{\boldsymbol{Z}_i}(X_i)}-\frac{1}{H_{\boldsymbol{Z}_i}(X_i)}\right].\end{aligned}$$

From the uniform convergence and asymptotic weak convergence of $1/\hat{H}_z(t)$ and condition C3, we have $\sup_i\left\{1/\hat{H}_{\boldsymbol{Z}_i}(X_i)-1/H_{\boldsymbol{Z}_i}(X_i)\right\}=o_p\left(n^{-1/2+r}\right)$ for any $r>0$. Hence the above quantity is dominated by the first term $S_n^H\left(\hat{\boldsymbol\beta}(\gamma),\gamma\right)-S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)$. By Lemma 1 and the uniform convergence of $\hat{\boldsymbol\beta}(\gamma)$ to $\boldsymbol\beta_0(\gamma)$,

$$S_n^H\left(\hat{\boldsymbol\beta}(\gamma),\gamma\right)-S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)$$
$$\approx n^{-1/2}\left[\mu\left\{\hat{\boldsymbol\beta}(\gamma),\gamma\right\}-\mu\left\{\boldsymbol\beta_0(\gamma),\gamma\right\}\right],$$

where $\approx$ denotes asymptotic equivalence uniformly in $\gamma\in[\gamma_L,\gamma_U]$. Applying Taylor expansions for $\mu(\boldsymbol b)$ at $\boldsymbol b=\boldsymbol\beta_0(\gamma)$, and using the uniform convergence of $\hat{\boldsymbol\beta}(\gamma)$ to $\boldsymbol\beta_0(\gamma)$, we have

$$S_n\left(\hat{\boldsymbol\beta}(\gamma),\gamma\right)-S_n\left(\boldsymbol\beta_0(\gamma),\gamma\right)$$
$$=A\left\{\boldsymbol\beta_0(\gamma)+\varepsilon_n(\gamma)\right\}\cdot n^{-1/2}\left[\hat{\boldsymbol\beta}(\gamma)-\boldsymbol\beta(\gamma)\right],$$

where $\varepsilon_n(\gamma)\approx 0$. Since $S_n\left(\hat{\boldsymbol\beta}(\gamma),\gamma\right)\approx 0$,

$$n^{-1/2}\left[\hat{\boldsymbol\beta}(\gamma)-\boldsymbol\beta_0(\gamma)\right]\approx -A\left\{\boldsymbol\beta_0(\gamma)\right\}^{-1}S_n\left(\boldsymbol\beta_0(\gamma),\gamma\right). \quad (2)$$

It remains to prove the weak convergence of $S_n\left(\boldsymbol\beta_0(\gamma),\gamma\right)$ to a zero-mean Gaussian process. It follows that

$$\begin{aligned}
S_n\left(\boldsymbol\beta_0(\gamma),\gamma\right)&=S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)+\left[S_n\left(\boldsymbol\beta_0(\gamma),\gamma\right)-S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)\right]\\
&=S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)-n^{-1/2}\sum_{i=1}^n \boldsymbol Z_i I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right\}\delta_{X_i}\frac{\hat H_{\boldsymbol Z_i}(X_i)-H_{\boldsymbol Z_i}(X_i)}{\hat H_{\boldsymbol Z_i}(X_i)H_{\boldsymbol Z_i}(X_i)}\\
&\approx S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)-n^{-1/2}\sum_{i=1}^n \boldsymbol Z_i I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right\}\delta_{X_i}\frac{\hat G_{\boldsymbol Z_i}(X_i)-G_{\boldsymbol Z_i}(X_i)}{H_{\boldsymbol Z_i}(X_i)G_{\boldsymbol Z_i}(X_i)}\\
&\quad -n^{-1/2}\sum_{i=1}^n \boldsymbol Z_i I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right\}\delta_{X_i}\frac{\hat S_{D|T,\boldsymbol Z_i}(X_i)-S_{D|T,\boldsymbol Z_i}(X_i)}{H_{\boldsymbol Z_i}(X_i)S_{D|T,\boldsymbol Z_i}(X_i)}\\
&=(I)-(II)-(III)
\end{aligned}$$

The first two terms (*I*) and (*II*) can be proved to converge to zero-mean Gaussian processes by applying the arguments in [13] as follows. The family

$$\left\{\boldsymbol Z_i\left(I\left(h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right)\delta_{X_i}\big/H_{\boldsymbol Z_i}(X_i)-\gamma\right),\ \gamma\in[\gamma_L,\gamma_U]\right\}$$

is Donsker by the Lipschitz continuity of $\boldsymbol\beta_0(\gamma)$ (condition C4 (i)) and uniformly boundedness of $\boldsymbol Z$ and $H_{\boldsymbol Z}(X)^{-1}$ (conditions C1 and C3). Thus, the first term

$$(I)=S_n^H\left(\boldsymbol\beta_0(\gamma),\gamma\right)$$
$$=n^{-1/2}\sum_{i=1}^n\boldsymbol Z_i\left(I\left(h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right)\delta_{xi}\big/H_{z_i}(X_i)-\gamma\right)$$

converges weakly to a zero-mean Gaussian process.

Denote $R_{X_i}(t)=I(X_i\ge t)$ and $R_{Y_i}(t)=I(Y_i\ge t)$ as the at-risk processes at time *t*. Let

$r_z^Y(t)=\Pr\left(Y\ge t\mid\boldsymbol Z=z\right)$,

$\lambda^{G_z}(t)=\lim_{\Delta\to 0}\Pr\left(C\in(t,t+\Delta)\mid C\ge t,\boldsymbol Z=z\right)\big/\Delta$,

$\Lambda^{G_z}(t)=\int_0^t\lambda^{G_z}(s)\mathrm ds$. $N_i^{G_z}(t)=I\left(Y_i\le t,\delta_{Y_i}=0,\boldsymbol Z_i=z\right)$.

Then $M_i^{G_z}(t)=N_i^{G_z}(t)-\int_0^t I(\boldsymbol Z_i=z)R_{Y_i}(s)\mathrm d\Lambda^{G_z}(s)$ is a martingale.

From martingale representation theory for univariate independent censoring,

$$n_z^{1/2}\left\{\hat G_z(t)-G_z(t)\right\}$$
$$\approx n_z^{-1/2}\sum_{i=1}^n I(\boldsymbol Z_i=z)G_z(t)\int_0^t r_z^Y(s)^{-1}\mathrm dM_i^{G_z}(s).$$

So the second term can be written as

$$\begin{aligned}
(II)&=n^{-1/2}\sum_{i=1}^n \boldsymbol Z_i I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right\}\delta_{X_i}\frac{\hat G_{\boldsymbol Z_i}(X_i)-G_{\boldsymbol Z_i}(X_i)}{H_{\boldsymbol Z_i}(X_i)G_{\boldsymbol Z_i}(X_i)}\\
&=n^{-1/2}\sum_{i=1}^n\frac{\boldsymbol Z_i I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_i\right\}\delta_{X_i}}{H_{\boldsymbol Z_i}(X_i)}\frac{1}{n_{\boldsymbol Z_i}}\sum_{j=1}^n I\left(\boldsymbol Z_j=\boldsymbol Z_i\right)\int_0^\infty\frac{R_{X_i}(s)}{r_{\boldsymbol Z_j}^Y(s)}\mathrm dM_j^{G_{\boldsymbol Z_j}}(s)\\
&=n^{-1/2}\sum_{j=1}^n\int_0^\infty\frac{1}{n_{\boldsymbol Z_j}}\sum_{i=1}^n I\left(\boldsymbol Z_j=\boldsymbol Z_i\right)\frac{\boldsymbol Z_j I\left\{h(X_i)\le\boldsymbol\beta_0(\gamma)^{\mathrm T}\boldsymbol Z_j\right\}\delta_{X_i}R_{X_i}(s)}{H_{\boldsymbol Z_i}(X_i)r_{\boldsymbol Z_j}^Y(s)}\mathrm dM_j^{G_{\boldsymbol Z_j}}(s)\\
&\approx n^{-1/2}\sum_{j=1}^n\int_0^\infty \boldsymbol w_1\left(\boldsymbol\beta_0(\gamma),\boldsymbol Z_j,s\right)\mathrm dM_j^{G_{\boldsymbol Z_j}}(s),
\end{aligned}$$

where $\boldsymbol{w}_1(\boldsymbol{b},\boldsymbol{z},t)=\boldsymbol{z}E\left[\frac{I\left\{h(T)\le \boldsymbol{b}^{\mathrm{T}}\boldsymbol{z}\right\}\delta_X R_X(t)}{H_z(X)r_z^Y(t)}\middle|\boldsymbol{Z}=\boldsymbol{z}\right]$.

From uniform boundedness of $\boldsymbol{Z}$, $H_{\boldsymbol{Z}}(X)^{-1}$ and $r_{\boldsymbol{Z}}^Y(X)^{-1}$, it is easy to show that $\int_0^\infty \boldsymbol{w}_1(\boldsymbol{b},\boldsymbol{Z}_j,s)\mathrm{d}M_j^{Gz_j}(s)$ is Lipschitz in $\boldsymbol{b}$. Then similarly $\left\{\int_0^\infty \boldsymbol{w}_1(\boldsymbol{\beta}_0(\gamma),\boldsymbol{Z}_j,s)\mathrm{d}M_j^{Gz_j}(s)\right\}$ can be shown to be Donsker, and the second term also converges weakly to a zero-mean Gaussian process.

For the third term (III), recall

$$\hat{S}_{D|T,z}(t)=\Psi\left(\hat{\alpha},\hat{S}_{W,z}(t),\hat{S}_{T,z}(t)\right)$$

with $\Psi(\alpha,u,v)=\phi_{\alpha'}(v)/\phi_{\alpha'}(u)$. Denote

$$\Psi_\alpha(\alpha,u,v)=\frac{\frac{\partial}{\partial\alpha}\Psi(\alpha,u,v)}{\Psi(\alpha,u,v)},$$

$$\Psi_u(\alpha,u,v)=\frac{\frac{\partial}{\partial u}\Psi(\alpha,u,v)}{\Psi(\alpha,u,v)}$$

and

$$\Psi_v(\alpha,u,v)=\frac{\frac{\partial}{\partial v}\Psi(\alpha,u,v)}{\Psi(\alpha,u,v)}.$$

So for $t\in[0,\nu], z\in\mathcal{Z}$,

$$\begin{aligned}\hat{S}_{D|T,z}(t)-S_{D|T,z}(t)\approx{}& S_{D|T,z}(t)\Big[\Psi_\alpha\left(\alpha(z),S_{W,z}(t),S_{T,z}(t)\right)\left\{\hat{\alpha}(z)-\alpha(z)\right\}\\&+\Psi_u\left(\alpha(z),S_{W,z}(t),S_{T,z}(t)\right)\left\{\hat{S}_{W,z}(t)-S_{W,z}(t)\right\}+\Psi_v\left(\alpha(z),S_{W,z}(t),S_{T,z}(t)\right)\left\{\hat{S}_{T,z}(t)-S_{T,z}(t)\right\}\Big].\end{aligned}$$

For notation brevity, denote

$$\Psi_{\alpha,i}=\Psi_\alpha\left(\alpha(\boldsymbol{Z}_i),S_{W,\boldsymbol{Z}_i}(X_i),S_{T,\boldsymbol{Z}_i}(X_i)\right),$$

$$\Psi_{u,i}=\Psi_u\left(\alpha(\boldsymbol{Z}_i),S_{W,\boldsymbol{Z}_i}(X_i),S_{T,\boldsymbol{Z}_i}(X_i)\right)$$

and

$$\Psi_{v,i}=\Psi_v\left(\alpha(\boldsymbol{Z}_i),S_{W,\boldsymbol{Z}_i}(X_i),S_{T,\boldsymbol{Z}_i}(X_i)\right).$$

The third term becomes

$$\begin{aligned}(III)&=n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\frac{\hat{S}_{D|T,\boldsymbol{Z}_i}(X_i)-S_{D|T,\boldsymbol{Z}_i}(X_i)}{H_{\boldsymbol{Z}_i}(X_i)S_{D|T,\boldsymbol{Z}_i}(X_i)}\\&=n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}H_{\boldsymbol{Z}_i}(X_i)^{-1}\Big[\Psi_{\alpha,i}\left\{\hat{\alpha}(\boldsymbol{Z}_i)-\alpha(\boldsymbol{Z}_i)\right\}\\&\quad+\Psi_{u,i}\left\{\hat{S}_{T,\boldsymbol{Z}_i}(X_i)-S_{T,\boldsymbol{Z}_i}(X_i)\right\}+\Psi_{v,i}\left\{\hat{S}_{W,\boldsymbol{Z}_i}(X_i)-S_{W,\boldsymbol{Z}_i}(X_i)\right\}\Big]=(A)+(B)+(C)\end{aligned}$$

Since $\mathcal{Z}$ is finite, $n_z\to\infty$ by condition C1 for all $z\in\mathcal{Z}$. So $n_z^{1/2}\left\{\hat{\alpha}(z)-\alpha(z)\right\}\approx M^z$ with M^z follows a Gaussian distribution. Let

$$\boldsymbol{w}_2(\boldsymbol{b},\boldsymbol{z})=\boldsymbol{z}E\left[\frac{I\left\{h(T)\le\boldsymbol{b}^{\mathrm{T}}\boldsymbol{z}\right\}\delta_X\Psi_\alpha\left(\alpha(z),S_{W,z}(X),S_{T,z}(X)\right)}{H_z(X)}\middle|\boldsymbol{Z}=\boldsymbol{z}\right].$$

Now term (A) is

$$\begin{aligned}&n^{-1/2}\sum_{i=1}^n\frac{\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\Psi_{\alpha,i}}{H_{\boldsymbol{Z}_i}(X_i)}\left\{\hat{\alpha}(\boldsymbol{Z}_i)-\alpha(\boldsymbol{Z}_i)\right\}\\&=n^{-1/2}\sum_{i=1}^n\frac{\boldsymbol{Z}_iI\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\Psi_{\alpha,i}}{H_{\boldsymbol{Z}_i}(X_i)}\frac{1}{n_{\boldsymbol{Z}_i}}\sum_{j=1}^nI\left(\boldsymbol{Z}_j=\boldsymbol{Z}_i\right)\left\{\hat{\alpha}(\boldsymbol{Z}_j)-\alpha(\boldsymbol{Z}_j)\right\}\\&=n^{-1/2}\sum_{j=1}^n\left\{\hat{\alpha}(\boldsymbol{Z}_j)-\alpha(\boldsymbol{Z}_j)\right\}\frac{1}{n_{\boldsymbol{Z}_j}}\sum_{i=1}^nI\left(\boldsymbol{Z}_j=\boldsymbol{Z}_i\right)\frac{\boldsymbol{Z}_jI\left\{h(X_i)\le\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_j\right\}\delta_{X_i}\Psi_{\alpha,i}}{H_{\boldsymbol{Z}_j}(X_i)}\\&\approx n^{-1/2}\sum_{j=1}^n\left\{\hat{\alpha}(\boldsymbol{Z}_j)-\alpha(\boldsymbol{Z}_j)\right\}\boldsymbol{w}_2\left(\boldsymbol{\beta}_0(\gamma),\boldsymbol{Z}_j\right)\approx\sum_{z\in\mathcal{Z}}f_{\boldsymbol{Z}}(\boldsymbol{z})^{1/2}M^z\boldsymbol{w}_2\left(\boldsymbol{\beta}_0(\gamma),\boldsymbol{z}\right)\end{aligned}$$

which follows a Gaussian distribution by the boundedness of $f_{\boldsymbol{Z}}(\boldsymbol{z})$ and $\boldsymbol{w}_2\left(\boldsymbol{\beta}_0(\gamma),\boldsymbol{z}\right)$.

Denote

$$\lambda^{S_W,z}(t)=\lim_{\Delta\to 0}\Pr\left(W\in(t,t+\Delta)\middle|W\geq t,\boldsymbol{Z}=\boldsymbol{z}\right)\Big/\Delta,$$

$$\Lambda^{S_W,z}(t)=\int_0^t\lambda^{S_W,z}(s).$$

Denote

$$N_i^{S_W,z}(t)=I\left(X_i\leq t,\delta_{W_i}=1,\boldsymbol{Z}_i=\boldsymbol{z}\right)$$

where

$$\delta_{W_i}=I\left(T_i\wedge D_i<C_i\right)=1-\left(1-\delta_{X_i}\right)\left(1-\delta_{Y_i}\right).$$

$$r_z^X(t)=\Pr\left(X\geq t\middle|\boldsymbol{Z}=\boldsymbol{z}\right).$$

Then

$$M_i^{S_W,z}(t)=N_i^{S_W,z}(t)-\int_0^t I\left(\boldsymbol{Z}_i=\boldsymbol{z}\right)R_{X_i}(s)\,\mathrm{d}\Lambda^{S_W,z}(s)$$

is a martingale. Then

$$n_z^{1/2}\left\{\hat{S}_{W,z}(t)-S_{W,z}(t)\right\}\approx n_z^{-1/2}\sum_{i=1}^n I\left(\boldsymbol{Z}_i=\boldsymbol{z}\right)S_{W,z}(t)\int_0^t r_z^X(s)^{-1}\,\mathrm{d}M_i^{S_W,z}(s).$$

Now similar to the martingale expression for the second term (II), we have term (B) as

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^n\frac{\boldsymbol{Z}_i I\left\{h(X_i)\leq\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\Psi_{u,i}}{H_{\boldsymbol{Z}_i}(X_i)}\left\{\hat{S}_{W,\boldsymbol{Z}_i}(t)-S_{W,\boldsymbol{Z}_i}(t)\right\}\\
&=n^{-1/2}\sum_{i=1}^n\frac{\boldsymbol{Z}_i I\left\{h(X_i)\leq\boldsymbol{\beta}_0(\gamma)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\Psi_{u,i}}{H_{\boldsymbol{Z}_i}(X_i)}\frac{1}{n_{\boldsymbol{Z}_i}}\sum_{j=1}^n I\left(\boldsymbol{Z}_j=\boldsymbol{Z}_i\right)\int_0^\infty\frac{R_{X_i}(s)}{r_{\boldsymbol{Z}_j}^X(s)}\,\mathrm{d}M_j^{S_W,\boldsymbol{z}_j}(s)\\
&\approx n^{-1/2}\sum_{j=1}^n\int_0^\infty\boldsymbol{w}_3\left(\boldsymbol{\beta}_0(\gamma),\boldsymbol{Z}_j,s\right)\mathrm{d}M_j^{S_W,\boldsymbol{z}_j}(s)
\end{aligned}$$

where

$$\boldsymbol{w}_3(\boldsymbol{b},\boldsymbol{z},t)=\boldsymbol{z}E\left[I\left\{h(T)\leq\boldsymbol{b}^{\mathrm{T}}\boldsymbol{z}\right\}\delta_X\Psi_u\left(\alpha(\boldsymbol{z}),S_{W,z}(X),S_{T,z}(X)\right)R_X(t)\Big/\left\{H_z(X)r_z^X(t)\right\}\middle|\boldsymbol{Z}=\boldsymbol{z}\right].$$

Similar as (*II*) above, $\int_0^\infty\boldsymbol{w}_3\left(\boldsymbol{b},\boldsymbol{Z}_j,s\right)\mathrm{d}M_j^{S_W,\boldsymbol{z}_j}(s)$ is Lipschitz in $\boldsymbol{b}$, and $\left\{\int_0^\infty\boldsymbol{w}_3\left(\boldsymbol{\beta}_0(\gamma),\boldsymbol{Z}_j,s\right)\mathrm{d}M_j^{S_W,\boldsymbol{z}_j}(s)\right\}$ is Donsker. Thus the term (*B*) converges weakly to a zero-mean Gaussian process.

Finally, for term (C), let

$$\lambda^{S_T,z}(t)=\lim_{\Delta\to 0}\Pr\left(T\in(t,t+\Delta)\middle|X\geq t,\boldsymbol{Z}=\boldsymbol{z}\right)\Big/\Delta$$

be the crude hazard function in [17],

$$\Lambda^{S_T,z}(t)=\int_0^t\lambda^{S_T,z}(s)\,\mathrm{d}s.$$

$$N_i^{S_T,z}(t)=I\left(X_i\leq t,\delta_{X_i}=1,\boldsymbol{Z}_i=\boldsymbol{z}\right).$$

Then

$$M_i^{S_T,z}(t)=N_i^{S_T,z}(t)-\int_0^t I\left(\boldsymbol{Z}_i=\boldsymbol{z}\right)R_{X_i}(s)\,\mathrm{d}\Lambda^{S_T,z}(s)$$

is a martingale. From [17], we have

$$\begin{aligned}
n_z^{1/2}\left\{\hat{S}_{T,z}(t)-S_{T,z}(t)\right\}\approx&-n_z^{-1/2}\sum_{i=1}^n I\left(\boldsymbol{Z}_i=\boldsymbol{z}\right)\frac{1}{\phi'_{\alpha(z)}\left\{S_{T,z}(t)\right\}}\int_0^t\frac{\phi'_{\alpha(z)}\left\{S_{W,z}(s)\right\}}{G_z(s)}\,\mathrm{d}M_i^{S_T,z}(s)\\
&+n_z^{-1/2}\sum_{i=1}^n I\left(\boldsymbol{Z}_i=\boldsymbol{z}\right)\frac{1}{\phi'_{\alpha(z)}\left\{S_{T,z}(t)\right\}}\int_0^t\frac{\boldsymbol{v}(\boldsymbol{z},t)-\boldsymbol{v}(\boldsymbol{z},s)}{S_{W,z}(s)G_z(s)}\,\mathrm{d}M_i^{S_W,z}(s)\\
&-n_z^{1/2}\left\{\hat{\alpha}(\boldsymbol{Z}_i)-\alpha(\boldsymbol{Z}_i)\right\}\left[\frac{\dot{\phi}_{\alpha(z)}\{S_{T,z}(t)\}}{\phi'_{\alpha(z)}\left\{S_{T,z}(t)\right\}}+\frac{\int_{s=0}^t S_{W,z}(s)\dot{\phi}'_{\alpha(z)}\left\{S_{W,z}(s)\right\}\lambda^{S_T,z}(s)\,\mathrm{d}s}{\phi'_{\alpha(z)}\left\{S_{T,z}(t)\right\}}\right]
\end{aligned}$$

where

$$\boldsymbol{v}(\boldsymbol{z},t)=-\int_{s=0}^t\left[\phi'_{\alpha(z)}\left\{S_{W,z}(s)\right\}+S_{W,z}(s)\phi''_{\alpha(z)}\left\{S_{W,z}(s)\right\}\right]S_{W,z}(s)\lambda^{S_T,z}(s)\,\mathrm{d}s.$$

Similar to terms (A) and (B), we have term (C) equals

$$n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i I\left\{h\left(X_i\right)\le\boldsymbol{\beta}_0\left(\gamma\right)^{\mathrm{T}}\boldsymbol{Z}_i\right\}\delta_{X_i}\times\frac{\Psi_v(\alpha\left(\boldsymbol{Z}_i\right),S_{W,\boldsymbol{Z}_i}\left(X_i\right),S_{T,\boldsymbol{Z}_i}\left(X_i\right))}{H_{\boldsymbol{Z}_i}\left(X_i\right)}\left(\hat{S}_{T,\boldsymbol{Z}_i}\left(X_i\right)-S_{T,\boldsymbol{Z}_i}\left(X_i\right)\right)$$

$$\approx n^{-1/2}\sum_{j=1}^{n}\int_0^\infty\boldsymbol{w}_4\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{Z}_j,s\right)\mathrm{d}M_j^{S_T,\boldsymbol{z}_j}\left(s\right)+n^{-1/2}\sum_{j=1}^{n}\int_0^\infty\boldsymbol{w}_5\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{Z}_j,s\right)\mathrm{d}M_j^{S_W,\boldsymbol{z}_j}\left(s\right)+\sum_{\boldsymbol{z}\in\mathcal{Z}}f_{\boldsymbol{Z}}\left(\boldsymbol{z}\right)^{1/2}M^{\boldsymbol{z}}\boldsymbol{w}_6\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{z}\right),$$

where

$$\boldsymbol{w}_4\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{z},t\right)=-E\left[I\left\{h\left(T\right)\le\boldsymbol{b}^{\mathrm{T}}\boldsymbol{z}\right\}\delta_X\frac{\Psi_v\left(\alpha\left(\boldsymbol{z}\right),S_{W,z}\left(X\right),S_{T,z}\left(X\right)\right)}{H_z\left(X\right)\phi'_{\alpha(z)}\left\{S_{T,z}\left(X\right)\right\}}R_X\left(t\right)\middle|\boldsymbol{Z}=\boldsymbol{z}\right]\boldsymbol{z}\frac{\phi'_{\alpha(z)}\left\{S_{W,z}\left(t\right)\right\}}{G_z\left(t\right)},$$

$$\boldsymbol{w}_5\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{z},t\right)=E\left[I\left\{h\left(T\right)\le\boldsymbol{b}^{\mathrm{T}}\boldsymbol{z}\right\}\delta_X\frac{\Psi_v\left(\alpha\left(\boldsymbol{z}\right),S_{W,z}\left(X\right),S_{T,z}\left(X\right)\right)}{H_z\left(X\right)\phi'_{\alpha(z)}\left\{S_{T,z}\left(X\right)\right\}}R_X\left(t\right)\frac{\boldsymbol{v}\left(\boldsymbol{z},X\right)-\boldsymbol{v}\left(\boldsymbol{z},t\right)}{S_{W,z}\left(t\right)G_z\left(t\right)}\middle|\boldsymbol{Z}=\boldsymbol{z}\right]\boldsymbol{z},$$

$$\boldsymbol{w}_6\left(\boldsymbol{\beta}_0\left(\gamma\right),\boldsymbol{z}\right)=-E\left[\frac{\dot{\phi}_{\alpha(z)}\left\{S_{T,z}\left(X\right)\right\}}{\phi'_{\alpha(z)}\left\{S_{T,z}\left(X\right)\right\}}+\frac{\int_{s=0}^{X}S_{W,z}\left(s\right)\dot{\phi}'_{\alpha(z)}\left\{S_{W,z}\left(s\right)\right\}\lambda^{S_{T,z}}\left(s\right)\mathrm{d}s}{\phi'_{\alpha(z)}\left\{S_{T,z}\left(X\right)\right\}}\middle|\boldsymbol{Z}=\boldsymbol{z}\right].$$

Then each of the three terms in (C) can be shown to converge to zero-mean Gaussian process similarly as above.

Summarizing the results above, $S_n\left\{\boldsymbol{\beta}_0\left(\gamma\right),\gamma\right\}$ converges weakly to a zero-mean Gaussian process. Hence (2) implies that $n^{-1/2}\left[\hat{\boldsymbol{\beta}}\left(\gamma\right)-\boldsymbol{\beta}_0\left(\gamma\right)\right]$ converges weakly to a zero-mean Gaussian process for $\gamma\in\left[\gamma_L,\gamma_U\right]$.

