Smoothed quantile regression for censored residual life

Kyuhyun Kim[a,b], Sangwook Kang[a,b,*]

[a] *Department of Statistics and Data Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, Korea*
[b] *Department of Applied Statistics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, Korea*

Abstract

We consider a regression modeling of the quantiles of residual life, remaining lifetime at a specific time. For estimation of regression parameters, we propose an induced smoothed version of the existing non-smooth estimating equations approaches. The proposed estimating equations are smooth in regression parameters, so solutions can be readily obtained via standard numerical algorithms. Moreover, the smoothness in the proposed estimating equations enables one to obtain a robust sandwich-type covariance estimator of regression estimators aided by an efficient resampling method. To handle data subject to right censoring, inverse probabilities of censoring are incorporated as weights. The consistency and asymptotic normality of the proposed estimator are established. Extensive simulation studies are conducted to verify performances of the proposed estimator under various finite samples settings. We apply the proposed method to dental study data evaluating the longevity of dental restorations.

Keywords: Induced smoothing, Inverse of censoring weighting, Median regression, Resampling, Sandwich estimator, Survival analysis

1. Introduction

In many clinical and epidemiological studies, remaining lifetimes at a specific time are often of interest. Unlike the usual failure time defined as the elapsed time from the baseline until an event occurs, residual life can be defined at any time t after the baseline given that the subject has not experienced the event of interest by t. Since survival data are often collected through a series of follow-up visits after an initial visit at baseline, modeling residual life at a specific visiting time t could provide more dynamic and meaningful information. For example, in a dental study, longevity of a tooth treated with a restoration procedure might be of interest. Subjects who has received dental restoration of a tooth typically visit the clinic on a regular basis for a check-up. At each visit, it would be of great interest to assess the effects of factors that might be related to residual life of the treated tooth and predict its longevity. While my tooth restored is still alive 1 year after restoration, how long does it is expected to be alive?

To investigate the effects of various factors on remaining lifetimes, regression modeling have typically been on the mean and quantiles of residual life. Modeling the mean residual life has mostly focused on a proportional mean residual life model, a counterpart of Cox proportional hazard model (Maguluri and Zhang, 1994; Oakes and Dasu, 1990, 2003; Chen et al., 2005). Statistical methods also have been proposed under alternative models including an additive mean residual life model (Chen and Cheng, 2006; Chen, 2007; Zhang et al., 2010), proportional scaled mean residual life model (Liu and Ghosh, 2008) which can be considered as the accelerated failure time (AFT) model, and semiparametric transformation models (Sun and Zhang, 2009; Sun et al., 2012). Although a mean has been a popular quantity representing the central tendency in data, it might not be a suitable summary measure for survival times that are typically skewed, possibly with a heavy tail and extreme observations. Moreover, the identifiability could be an issue for the mean residual life when the follow-up time is not long enough (Li et al., 2016). In such cases, a median, a special case of quantiles, could serve as a nice alternative; the median is a less sensitive measure to outliers and thus offer a more meaningful summary for skewed survival data (Ying and Sit, 2017).

[*] Corresponding author
Email addresses: kyuhyunkim07@yonsei.ac.kr (Kyuhyun Kim), kanggi1@yonsei.ac.kr (Sangwook Kang)

Quantile regression models, originally proposed for modeling a continuous response (Koenker and Bassett Jr, 1978), have been adapted to modeling failure time data without considering censoring (Jung, 1996; Portnoy et al., 1997; Wei et al., 2006; Whang, 2006) and expanded to accommodating censored failure time data (Ying et al., 1995; Bang and Tsiatis, 2002; Gelfand and Kottas, 2003; Portnoy, 2003; Peng and Huang, 2008; Wang and Wang, 2009; Huang, 2010; Portnoy and Lin, 2010). For residual life, Jung et al. (2009) considered a semiparametric regression model and proposed estimating equations approach. Kim et al. (2012) proposed an alternative estimating equations approach, for which an empirical likelihood approach is taken. Li et al. (2016) considered a more general setting that allows repeated measurements of covariates. They proposed an estimating equations approach that employs the inverse probability of censoring weighted (IPCW) principle to handle right-censored observations. Asymptotic properties of the proposed estimators were rigorously established in all these approaches.

Note that all the estimating equations considered in these proposed works are non-smooth in regression parameters. Thus, in calculating the proposed estimators, the standard numerical algorithms such as Newton-Raphson cannot be directly applied, and solutions might not be uniquely defined. Nevertheless, recent developments in computing algorithms have alleviated these issues substantially. Efficient and reliable calculation of regression parameters estimates have been shown to be feasible. Regression coefficients estimators were calculated via a grid search method (Jung et al., 2009) or L_1-minimization algorithm based on the linear programming technique (Kim et al., 2012; Li et al., 2016). Despite these progresses in point estimation, variance estimation still could be problematic. The proposed estimating functions are not differentiable with respect to regression parameters, so a well-known robust sandwich-type estimator cannot be directly calculated. Furthermore, a direct estimation of variance involve nonparametric estimation of unspecified conditional error density, which could be computationally unstable. For these reasons, a direct estimation of variance was either avoided (Jung et al., 2009; Kim et al., 2012) or done with a computationally intensive multiplier bootstrap method that requires solving perturbed estimating equations multiple times. One way to tackle this issue is to apply an induced smoothing procedure that modifies non-smooth estimating equations into smooth ones.

The induced smoothing approach (Brown and Wang, 2005) has been frequently employed in survival analysis especially for the rank-based approach in fitting semiparametric AFT models (Brown and Wang, 2007; Johnson and Strawderman, 2009; Fu et al., 2010; Pang et al., 2012; Chiou et al., 2014a,b, 2015a,b; Kang, 2017) and quantile regression models (Choi et al., 2018). Based on the asymptotic normality of the non-smooth estimator, the non-smooth estimating functions are smoothed by averaging out the random perturbation generated by adding a scaled mean-zero normal random variable to the regression parameters. The resulting estimating functions are smooth in regression parameters and thus the standard numerical algorithms such as Newton-Raphson can be readily applied. Furthermore, due to the smoothness, estimation of standard errors is straightforward by using the robust sandwich-type estimator. To our knowledge, this induced smoothing approach has not yet been applied in the quantile residual life regression context. Thus, we propose to take advantage of the induced smoothing approach in fitting a semiprametric quantile residual life regression model.

The rest of the article is organized as follows. In Section 2, a semiparametric regression model for quantiles of residual life and the proposed estimation methods are provided. In Section 3, two computing algorithms are introduced. Finite sample properties of the proposed estimators are investigated by simulation experiments in Section 4. In Section 5, the proposed methods are illustrated with a dental restoration longevity study data for older adults (Caplan et al., 2019). Concluding remarks are provided with some discussions in Section 6.

2. Model and Estimation

2.1. Semiparametric quantile regression model

Suppose T and C denote the potential failure time and censoring time, respectively. In the presence of right censorship, we observe $Z = \min(T, C)$, which T and C are independent. Let $\delta = I[T \leq C]$ be the failure indicator where $I[\cdot]$ is an indicator function. Then, with a $p \times 1$ vector of covariate X, we observe n independent copies of (Z, δ, X), $(Z_i, \delta_i, X_i), i = 1, \ldots, n$, where n and i denote the sample size and subject, respectively. At time t_0, the τ-th quantile of the residual life is defined as $\theta_\tau(t_0)$ that satisfies

$P(T_i - t_0 \geq \theta_\tau(t_0) \mid T_i \geq t_0) = 1 - \tau$. As the underlying model for $\theta_\tau(t_0)$, we consider the following regression model with an exponential link:

$$\begin{aligned}\theta_\tau &= \exp\left\{X_i^\top \beta_0(\tau, t_0)\right\}, \text{ or equivalently,} \\ \log(T_i - t_0) &= X_i^\top \beta_0(\tau, t_0) + \epsilon_i \qquad (1)\end{aligned}$$

where $\beta_0(\tau, t_0)$ is a $(p+1) \times 1$ vector of regression coefficients and ϵ_i is a random variable taking zero at the τ-th quantile. Note that when $t_0 = 0$, (1) reduces to the quantile regression model for continuous responses (Koenker and Bassett Jr, 1978). Hereafter, we use β and θ instead of $\beta(\tau, t)$ and $\theta_\tau(t)$ for notational convenience.

2.2. Estimation via non-smooth functions

In the absence of censoring, the regression coefficients β_0 in (1) could be estimated by solving the following estimating equations (Kim et al., 2012)

$$n^{-1} \sum_{i=0}^{n} I[T_i \geq t_0] X_i \Big\{ I\big[\log(T_i - t_0) \leq X_i^\top \beta\big] - \tau \Big\} = 0 \qquad (2)$$

Note that when $t_0 = 0$, (2) reduces to those developed for estimating the regression parameters in the quantile regression model for continuous responses (Koenker, 2005).

In the presence of right-censorship, not all T_is are observable. Thus, (2) cannot directly be evaluated. To account for this incompleteness in right-censored observations, weighted estimating equations in which a complete observation is weighted by the IPCW have been proposed (Li et al., 2016). The corresponding weighted estimating equations are

$$U_{t_0}(\beta, \tau) = n^{-1} \sum_{i=1}^{n} I[Z_i \geq t_0] X_i \left\{ I\big[\log(Z_i - t_0) \leq X_i^\top \beta\big] \frac{\delta_i}{\hat{G}(Z_i)/\hat{G}(t_0)} - \tau \right\} \qquad (3)$$

where $\hat{G}(\cdot)$ is the Kaplan-Meier estimate of the survival function of the censoring time C. The estimator for β_0 in (1), $\hat{\beta}_{NS}$ is defined as the solution to (3).

Note that (3) are non-smooth step functions in β, whose exact solutions might not exist. Thus, instead of directly solving (3), $\hat{\beta}_{NS}$ can equivalently be obtained via minimizing $L_{t_0}(\beta, \tau)$, a L_1-objective function with the following form (Li et al., 2016):

$$\begin{aligned} L_{t_0}(\beta, \tau) = n^{-1} \sum_{i=1}^{n} \frac{\delta_i I[Z_i > t_0]}{\hat{G}(Z_i)/\hat{G}(t_0)} \left|\log(Z_i - t_0) - X_i^\top \beta\right| + \left| M - \beta^\top \sum_{l=1}^{n} -X_l \frac{\delta_l I[Z_l > t_0]}{\hat{G}(Z_l)/\hat{G}(t_0)} \right| \\ + \left| M - \beta^\top \sum_{l=1}^{n} 2\tau X_l I[Z_l > t_0] \right|. \qquad (4)\end{aligned}$$

where M is an extremely large positive constant (for example, $M = 10^6$) that bound $\left|\beta^\top \sum_{i=1}^{n} -X_i \frac{\delta_i I[Z_i > t_0]}{\hat{G}(Z_i)/\hat{G}(t_0)}\right|$ and $\left|\beta^\top \sum_{i=1}^{n} 2\tau X_i I[Z_i > t_0]\right|$ from above. A straightforward calculation shows that the first derivative $L_{t_0}(\beta, \tau)$ with respect to β is proportional to $U_{t_0}(\beta, \tau)$. $\hat{\beta}_{NS}$ can be easily obtained using some existing software that can implement L_1-minimization algorithm such as the `rq()` function in the `quantreg` package in R (Koenker et al., 2012).

2.3. Estimation via induced smoothed functions

The estimating functions (3) are non-smooth in regression coefficients. In this subsection, we propose to apply the induced smoothing procedure (Brown and Wang, 2005) to (3). The proposed induced smoothed estimating functions are constructed by adding a scaled mean-zero random noise to the regression parameters in (3) and averaging it out. Specifically,

$$\begin{aligned}\tilde{U}_{t_0}(\beta, \tau, H) &= E_w\{U_{t_0}(\beta + H^{1/2} W, \tau)\} \\ &= n^{-1} \sum_{i=1}^{n} I[Z_i \geq t_0] X_i \left\{ \Phi\Big(\frac{X_i^\top \beta - \log(Z_i - t_0)}{\sqrt{X_i^\top H X_i}} \Big) \frac{\delta_i}{\hat{G}(Z_i)/\hat{G}(t_0)} - \tau \right\} \qquad (5)\end{aligned}$$

where $H = O(n^{-1})$, $W \sim N(0, I_p)$, I_p is the $p \times p$ identity matrix, and $\Phi(\cdot)$ is the standard normal cumulative distribution function. The estimator for β_0 in (1), $\hat{\beta}_{IS}$, is defined as the solution to $\tilde{U}_{t_0}(\beta, \tau, H) = 0$.

Note that (5) is smooth in β, so calculation of $\hat{\beta}_{IS}$ can be readily done via the standard numerical algorithms such as the Newton-Raphson method. Moreover, since (5) is differentiable with respect to β, the robust sandwich-type estimator, a typical approach employed in estimating equation-based approaches for variance estimation, can be directly applied; a slope matrix can be directly estimated. The resulting estimator is consistent and asymptotically normally distributed. In addition, as with other induced smoothed estimators for semiparametric AFT models (Johnson and Strawderman, 2009; Pang et al., 2012; Chiou et al., 2014a, 2015b) and quantile regression models (Choi et al., 2018), $n^{1/2}(\hat{\beta}_{IS} - \beta_0)$ and $n^{1/2}(\hat{\beta}_{NS} - \beta_0)$ are shown to be asymptotically equivalent, another important and useful feature of the induced smoothing method. These asymptotic properties are provided in Section 4.

2.4. Variance estimation

To estimate the variance-covariance function of $\hat{\beta}_{IS}$, we use the robust sandwich-form estimator, i.e., $\hat{\text{Var}}(\hat{\beta}_{IS}, \tau) = \hat{A}(\hat{\beta}_{IS})^\top \hat{V}(\hat{\beta}_{IS}) \hat{A}(\hat{\beta}_{IS})$. For obtaining the slope matrix part, $\hat{A}(\beta)$, we take advantage of the smoothness of $\tilde{U}_{t_0}(\beta, \tau, H)$ in β - $\hat{A}(\hat{\beta}_{IS})$ is the first derivative of $\tilde{U}_{t_0}(\beta, \tau, H)$ with respect to β evaluated at $\hat{\beta}_{IS}$. Specifically,

$$\begin{aligned} \hat{A}(\hat{\beta}_{IS}) &= \frac{\partial \tilde{U}_{t_0}(\hat{\beta}_{IS}, \tau, H)}{\partial \beta} \\ &= n^{-1} \sum_{i=1}^{n} I[Z_i > t_0] X_i \frac{\hat{G}(t_0)\delta_i}{\hat{G}(Z_i)} \phi\left(\frac{X_i{}^\top \hat{\beta}_{IS} - \log(Z_i - t_0)}{\sqrt{X_i{}^\top H X_i}} \right) \left(\frac{-X_i}{\sqrt{X_i{}^\top H X_i}} \right) \end{aligned} \tag{6}$$

where $\phi(\cdot)$ is the density function of a standard normal distribution.

For calculating $\hat{V}(\beta)$, we propose to employ a computationally efficient resampling method. Similar procedures have been proposed for fitting semiparametric AFT models using the induced smoothing methods (Chiou et al., 2014a). In this procedure, we first generate n independently and identically distributed (*i.i.d.*) positive multiplier random variables η_i's, $i = 1, ..., n$, with both mean and variance being 1, independent of the observed data. Given data with a realization of $(\eta_1, \dots, \eta_n)$, we obtain $\tilde{U}^*_{t_0}(\hat{\beta}_{IS}, \tau, H)$, a perturbed version of $\tilde{U}_{t_0}(\beta, \tau, H)$ where

$$\tilde{U}^*_{t_0}(\beta, \tau, H) = n^{-1} \sum_{i=1}^{n} I[Z_i \geq t_0] X_i \eta_i \left\{ \Phi\left(\frac{X_i^\top \beta - \log(Z_i - t_0)}{\sqrt{X_i^\top H X_i}} \right) \frac{\hat{G}(t_0)\delta_i}{\hat{G}(Z_i)} - \tau \right\} \tag{7}$$

We repeat this procedure m times. Let $\tilde{U}^{*(l)}_{t_0}(\hat{\beta}_{IS}, \tau, H)$ denote the kth perturbed version of $\tilde{U}_{t_0}(\beta, \tau, H)(l = 1, \dots, m)$. Then, for given data, $\{\tilde{U}^{*(1)}_{t_0}(\hat{\beta}_{IS}, \tau, H), \dots, \tilde{U}^{*(m)}_{t_0}(\hat{\beta}_{IS}, \tau, H)\}$ can be generated. $\hat{V}(\hat{\beta}_{IS})$ can be obtained by the sample variance of $\tilde{U}^*_{t_0}(\hat{\beta}_{IS}, \tau, H)$.

3. Computation

For calculating $\hat{\beta}_{IS}$ and its estimated variance $\hat{\text{Var}}(\hat{\beta}_{IS})$, we propose an iterative algorithm similar to those considered previously for the induced smoothing approach (Johnson and Strawderman, 2009; Chiou et al., 2014a, 2015b; Choi et al., 2018). The iterative algorithm uses the Newton-Raphson method while sequentially updating $\hat{\beta}_{IS}$ and $\hat{\text{Var}}(\hat{\beta}_{IS})$ until convergence. This algorithm can be summarized as follows:

Step 1: Set the initial value as the non-smooth estimator for β_0 that minimizes (4), $\hat{\beta}^{(0)} = \hat{\beta}_{NS}, \hat{\Sigma}^{(0)} = I_p$ and $H^{(0)} = n^{-1}\hat{\Sigma}^{(0)}$.

Step 2: Given $\hat{\beta}^{(k)}$ and $H^{(k)}$ at the k-th step, update $\hat{\beta}^{(k)}$ by

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} - \hat{A}(\hat{\beta}^{(k)})^{-1} \tilde{U}_{t_0}(\hat{\beta}^{(k)}, \tau, H^{(k)})$$

Step 3: Given $\hat{\beta}^{(k+1)}$ and $\hat{\Sigma}^{(k)}$, update $\hat{\Sigma}^{(k)}$ by

$$\hat{\Sigma}^{(k+1)} = \hat{A}(\hat{\beta}^{(k+1)})^{-1} \hat{V}(\hat{\beta}^{(k+1)}, \tau) \hat{A}(\hat{\beta}^{(k+1)})^{-1}$$

Step 3: Set $H^{(k+1)} = n^{-1}\hat{\Sigma}^{(k+1)}$. Repeat Steps 2 and 3 until $\hat{\beta}^{(k)}$ and $\hat{\Sigma}^{(k)}$ converge.

$\hat{\beta}_{IS}$ and $\hat{\Sigma}$ are the values of $\hat{\beta}^{(k)}$ and $\hat{\Sigma}^{(k)}$ at convergence. $\hat{\text{Var}}(\hat{\beta}_{IS}) = n^{-1}\hat{\Sigma}$.

In practice, a simpler version can be considered. Instead of updating $\hat{\Sigma}^{(k)}$ when calculating $\hat{\beta}^{(k)}$, we calculate $\hat{\beta}_{IS}$ at a fixed H, say $H = n^{-1}I_p$. Evaluating at $\hat{\beta}_{IS}$, $\hat{\Sigma}$ can be calculated using the variance estimation procedure described in Section 2.4. As shown in Appendix, as long as $H = O(n^{-1})$, the choice of H does not change the asymptotic properties of the resulting estimator. In addition, for induced smoothed estimators under semiparametric AFT models, the iterative algorithm and its simpler version were reported to produce similar estimates (Chiou et al., 2014a, 2015b). In the simulation studies in Section 5, our findings were also similar, so only the results from the simpler version are reported.

4. Asymptotic properties

In this section, we present the asymptotic properties of the proposed estimator. We assume some regularity conditions similar to those specified in Li et al. (2016) and Pang et al. (2012). These conditions are required to establish the asymptotic properties of the proposed estimator. Under Conditions C1 - C3, the proposed estimator can be shown to be consistent and asymptotically normally distributed:

C1 For any $t_0 \in \mathcal{T}$, the conditional density of $T - t_0$ given $T \geq t_0$, $g_{T-t_0}(s)$ is uniformly bounded from above and away from 0, and $g'_{T-t_0}(s)$ exists and is uniformly bounded on the real line.

C2 For each $i = 1, \ldots, n$, X_i satisfies the following conditions:

(a) $n^{-1}\Sigma_{i=1}^{n} X_i X_i^\top g_{T_i-t_0}(0)$ converges to a positive definite matrix A;
(b) There is a finite positive constant M_c such that $\sup_i \|X_i\| \leq M_c$, where $\|\cdot\|$ denotes the Euclidean norm.

C3 There exists positive constant $\nu > 0$ such that

(a) $P(C > \nu) = 0$ and $P(C = \nu) \geq c_0$, where c_0 is some positive constant and
(b) $\sup_{X, t\in\mathcal{T}}[t + \exp(X^\top \beta)] \leq \nu$.

The following theorem summarizes the asymptotic properties of the proposed estimator.

Theorem 1. *Under Conditions C1 - C3, $\hat{\beta}_{IS}$ solving* (5) *is consistent for β_0 and $n^{1/2}(\hat{\beta}_{IS} - \beta_0)$ is asymptotically normally distributed with mean zero and a finite covariance function matrix. Moreover, $n^{1/2}(\hat{\beta}_{IS} - \beta_0)$ has the same asymptotic distribution as that of $n^{1/2}(\hat{\beta}_{NS} - \beta_0)$ where $\hat{\beta}_{NS}$ minimizes* (4).

A sketch of the proofs are provided in Appendix A.

5. Simulation

To evaluate the performance of the proposed estimators in finite samples, we conduct extensive simulation experiments. We use the simulation settings similar to those in Jung et al. (2009). We assume the proposed model (1) with a single binary covariate with success probability 0.5. We generate T from a Weibull distribution with the survival function $S(t) = \exp\{-(\rho t)^\kappa\}$. We set the scale parameter κ to 2. For $t_0 = 0$, we set the intercept $\beta^{(0)}(\tau, t_0) = \log(5)$. We consider two settings for the regression coefficient for X, $\beta^{(1)}$: $\beta^{(1)}(\tau, t_0) = 0$ and $\beta^{(1)}(\tau, t_0) \neq 0$. For $\beta^{(1)}(\tau, t_0) \neq 0$ with $t_0 = 0$, we set $\beta^{(1)}(\tau, t_0) = \log(2)$. For a given τ, $\beta^{(0)}(\tau, t_0)$ and $\beta^{(1)}(\tau, t_0)$ at $t_0 = 0$, the shape parameter ρ can be obtained by solving

$$\rho^{-1}\{(\rho t_0)^\kappa - \log(1-\tau)\}^{(1/\kappa)} - t_0 = \exp\{\beta^{(0)}(\tau, t_0) + \beta^{(1)}(\tau, t_0)\}. \tag{8}$$

Let $\rho_0(\tau)$ and $\rho_1(\tau)$ denote ρs at a given τ for $\beta^{(1)}(\tau, t_0) = 0$ and $\beta^{(1)}(\tau, t_0) \neq 0$, respectively. When $t_0 > 0$, for a given τ and κ, $\beta^{(0)}(\tau, t_0)$s and $\beta^{(1)}(\tau, t_0)$s can be obtained by solving (8) sequentially for $\beta^{(0)}(\tau, t_0)$ by setting $\beta^{(1)}(\tau, t_0) = 0$ and then for $\beta^{(1)}(\tau, t_0)$ at the given $\beta^{(0)}(\tau, t_0)$. We consider $t_0 = 0, 1, 2$ and 3 for t_0 and $\tau = 0.25$ and 0.5 for τ. For $\tau = 0.5$, the corresponding $\beta^{(0)}(\tau, t_0)$s and $\beta^{(1)}(\tau, t_0)$s are $1.61 (= \log(5)), 1.41, 1.22, 1.04$ and $0.69 (= \log(2)), 0.80, 0.91, 1.02$, respectively.

Potential censoring times, C_is, are generated, independently from T_is, from $unif(0, c)$ where c is determined by the desired censoring proportions. We consider censoring proportions of $0\%, 10\%, 30\%$ and 50%.

Table 1: Simulation results of fitting quantile regression model for residual lifetimes using the proposed induced smoothing method at $\tau = 0.5$. $\beta^{(0)}(0.5, t_0) = 1.61, 1.41, 1.22, 1.04$ at $t_0 = 0, 1, 2, 3$ and $\beta^{(1)}(0.5, t_0) = 0$. PE is the mean of point estimates for regression parameters, $\beta^{(0)}(\tau, t_0)$ and $\beta^{(1)}(\tau, t_0)$. ESE is the mean of estimated standard error of regression parameter. SD is the sample standard deviation of point estimates. CP is coverage proportion of the nominal 95% confidence intervals. Cens is the censoring proportions. The sample size is set to 200. The number of repetition is 2000.

t_0	Cens	$\beta^{(0)}(0.5, t_0)$				$\beta^{(1)}(0.5, t_0)$			
		PE	ESE	SD	CP	PE	ESE	SD	CP
0	0	1.609	0.068	0.069	0.925	-0.004	0.097	0.095	0.941
	10	1.610	0.073	0.073	0.936	-0.002	0.104	0.101	0.940
	30	1.612	0.082	0.081	0.930	-0.002	0.118	0.116	0.935
	50	1.614	0.096	0.091	0.930	-0.003	0.139	0.135	0.936
1	0	1.408	0.084	0.084	0.926	-0.003	0.120	0.115	0.947
	10	1.409	0.089	0.088	0.929	-0.002	0.128	0.123	0.942
	30	1.412	0.102	0.099	0.925	-0.003	0.147	0.143	0.935
	50	1.415	0.121	0.117	0.921	-0.002	0.177	0.170	0.932
2	0	1.215	0.100	0.101	0.914	0.000	0.144	0.139	0.941
	10	1.216	0.109	0.107	0.915	0.002	0.156	0.150	0.942
	30	1.219	0.126	0.124	0.914	0.002	0.182	0.178	0.933
	50	1.225	0.151	0.150	0.892	-0.002	0.221	0.216	0.918
3	0	1.035	0.121	0.120	0.916	0.001	0.172	0.168	0.938
	10	1.035	0.132	0.131	0.909	0.003	0.190	0.185	0.931
	30	1.039	0.157	0.153	0.890	0.003	0.229	0.221	0.921
	50	1.046	0.197	0.189	0.864	-0.005	0.293	0.277	0.906

The sample size is set to $n = 200$. For variance estimation, the resampling size for estimating $\hat{V}$ is set to 200. The estimates obtained for each configuration are based on 2000 repetitions. All the analyses were conducted by R 4.02 (Koenker, 2005). We used the `dfsane()` function in the R package `BB` to solve the proposed induced smoothed estimating equations (Varadhan and Gilbert, 2009).

Simulation results for $\tau = 0.5$ when $\beta^{(1)} = 0$ are summarized in Table 1.

Overall, our proposed estimators work reasonably well in most cases considered. The point estimates are all close to the true regression parameters and their standard errors estimates are virtually identical to their empirical counterparts. The coverage rates of the nominal 95% confidence intervals are in the range of 93% to 95% except for $t_0 = 2$ and 3, and especially when the censoring proportion is 50%. This is mainly due to the decreased effective sample sizes. When $t_0 = 2$ and 3, the number of effective sample sizes decrease to, on average, 168 and 146, respectively. When we increase the sample size to 400, the coverage rates get closer to the nominal level of 95% (results not shown).

The simulation results for $\tau = 0.5$ when $\beta^{(1)} \neq 0$ are presented in Table 2. The overall findings are similar to those under $\beta^{(1)} = 0$. The coverage rates of the nominal 95% confidence intervals are relatively low when $t_0 = 2$ and 3 with the 50% censoring proportion. Again, by increasing the sample size to 400, the coverage rates become closer to the 95% level (results not shown).

We also consider a lower quantile with $\tau = 0.25$ when $\beta^{(1)} \neq 0$. Simulation results under $\tau = 0.25$ are presented in Table 3.

In general, the proposed estimator also seem to perform well under this setting by showing negligible biases in the point estimates and little discrepancies between the proposed standard errors estimates and their empirical counterparts. The coverage proportions are, however, low in the range of 88% to 91% for $\beta^{(0)}(0.25, t_0)$ when $t_0 \neq 0$. Again, increasing the sample size to 400 shows an improvement in the coverage rates, which get closer to 95%. These results are available in a separate Table 3 of Supplementary material. Simulation results for a higher quantile with $\tau = 0.75$ are also available in Tables 6 and 7 of Supplementary material.

We also compare the proposed induced smoothed estimator with its non-smooth counterpart (Li et al., 2016). Under the setting previously considered for $\tau = 0.5$ with nonzero $\beta^{(1)}(\tau, t_0)$, we calculated the non-smooth estimates. The non-smoothed estimator is obtained via the L_1-minimization method in (4) using the `rq()` function in the `quantreg` package (Koenker et al., 2012). To assess the degree to which the two

Table 2: Simulation results of fitting quantile regression model for residual lifetimes using the proposed induced smoothing method at $\tau = 0.5$. $\beta^{(0)}(0.25, t_0) = 1.61, 1.41, 1.22, 1.04$ and $\beta^{(1)}(0.25, t_0) = 0.69, 0.80, 0.91, 1.02$ at $t_0 = 0, 1, 2, 3$. PE is the mean of point estimates for regression parameters, $\beta^{(0)}(\tau, t_0)$ and $\beta^{(1)}(\tau, t_0)$. ESE is the mean of estimated standard error of regression parameter. SD is the sample standard deviation of point estimates. CP is coverage proportion of the nominal 95% confidence intervals. Cens is the censoring proportions. The sample size is set to 200. The number of repetition is 2000.

t_0	Cens	$\beta^{(0)}(0.5, t_0)$				$\beta^{(1)}(0.5, t_0)$			
		PE	ESE	SD	CP	PE	ESE	SD	CP
0	0	1.608	0.069	0.069	0.928	0.690	0.098	0.097	0.939
	10	1.611	0.071	0.072	0.921	0.690	0.103	0.103	0.938
	30	1.611	0.077	0.076	0.940	0.693	0.119	0.116	0.944
	50	1.613	0.085	0.083	0.941	0.694	0.143	0.135	0.939
1	0	1.408	0.083	0.084	0.927	0.789	0.114	0.110	0.952
	10	1.410	0.087	0.087	0.929	0.789	0.121	0.118	0.941
	30	1.411	0.094	0.093	0.930	0.792	0.140	0.135	0.942
	50	1.412	0.103	0.101	0.927	0.792	0.163	0.158	0.918
2	0	1.215	0.100	0.101	0.913	0.883	0.133	0.127	0.941
	10	1.216	0.105	0.106	0.913	0.883	0.141	0.137	0.933
	30	1.218	0.115	0.115	0.915	0.885	0.163	0.158	0.933
	50	1.220	0.130	0.126	0.908	0.881	0.192	0.188	0.914
3	0	1.035	0.121	0.120	0.918	0.966	0.155	0.147	0.929
	10	1.036	0.125	0.126	0.902	0.966	0.165	0.159	0.925
	30	1.038	0.140	0.139	0.905	0.968	0.192	0.183	0.924
	50	1.038	0.157	0.150	0.891	0.964	0.222	0.206	0.914

methods coincide with each other, we construct scatter plots comparing the proposed induced-smoothed estimates with the non-smooth estimates for $\beta^{(0)}$ and $\beta^{(1)}$s at each combination of different t_0s ($t_0 = 0$ and 2) and censoring proportions (0% and 30%). In Figure 1, scatter plots of comparing the proposed induced smoothed and non-smooth estimates for these combinations are provided.

Each plot shows that pairs of the two estimates (circles) are scattered around the straight line with the 45 degree angle (red line), which means that, overall, the two methods produce similar estimates. The empirical standard errors based on 2000 estimates are close to each other but those from the proposed estimator are consistently smaller than those from the non-smooth estimator by 3 to 5% (Table 5 of Supplementary material). Similar findings can also be found that compares the non-smooth and induced smoothed estimators for quantile regression models with $t_0 = 0$, a special case of the setting we consider, but for competing risks data (Choi et al., 2018).

The simulation results presented are based on the non-iterative algorithm using $H = I_2$. The results under the iterative algorithm are slightly better but, overall, the performances of the non-iterative and iterative methods are similar under the settings we considered.

6. Real data analysis

We illustrate our proposed methods by analyzing the dental restoration longevity study data for older adults (Caplan et al., 2019). Dental caries was known to be the most frequently encountered health condition worldwide, including older adult population, in 2010 (Kassebaum et al., 2015). As the older adult population grows, a larger burden and impact on society and health care system are expected. A recent study addresses the importance of dental restoration in older adults population and evaluated the longevity of dental restoration and factors related (Caplan et al., 2019).

Dental restoration refers to a general term of dental cavity treatment. Data are composed of patients who have visited Geriatric and Special Needs Dentistry Clinic at the University of Iowa Collage of Dentistry (COD). For a more detailed description of the data set, see Caplan et al. (2019). We select a sample of the patients who visited the clinics at year 2000 or later, and are 65 years or older at their times of visits. Since a patient could have multiple teeth restored, we randomly choose the first restored tooth for each patient. The resulting sample has $1,551$ unique patients who contribute one restored tooth. The failure time for a

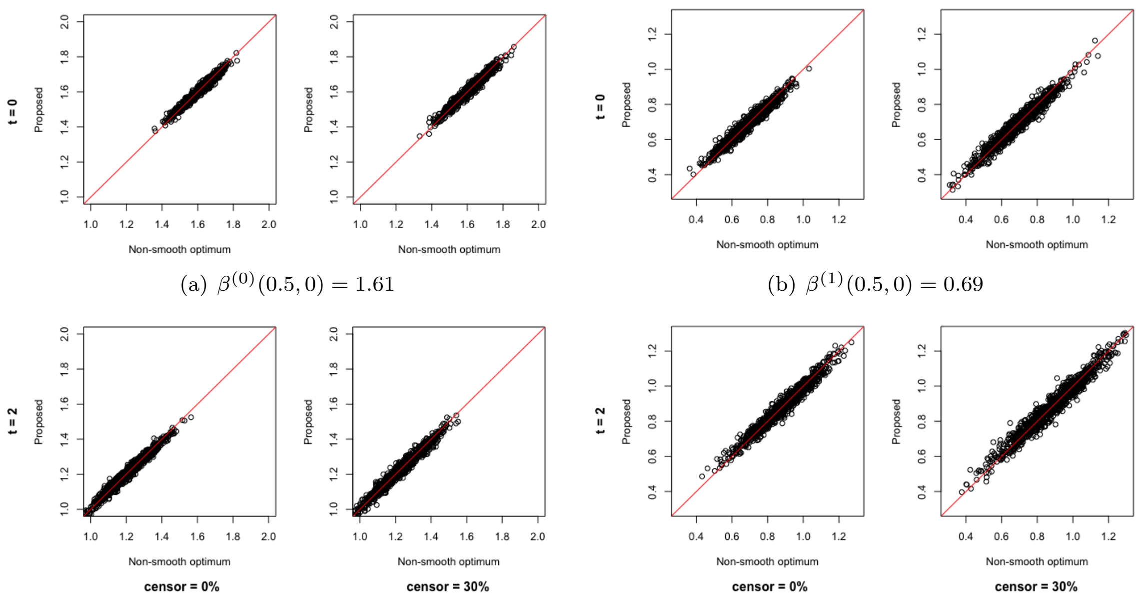
Table 3: Simulation results of fitting quantile regression model for residual lifetimes using the proposed induced smoothing method at $\tau = 0.25$. $\beta^{(0)}(0.25, t_0) = 1.61, 1.41, 1.22, 1.04$ and $\beta^{(1)}(0.25, t_0) = 0.69, 0.80, 0.91, 1.02$ at $t_0 = 0, 1, 2, 3$. PE is the mean of point estimates for regression parameters, $\beta^{(0)}(\tau, t_0)$ and $\beta^{(1)}(\tau, t_0)$. ESE is the mean of estimated standard error of regression parameter. SD is the sample standard deviation of point estimates. CP is coverage proportion of the nominal 95% confidence intervals. Cens is the censoring proportions. The sample size is set to 200. The number of repetition is 2000.

t_0	Cens	$\beta^{(0)}(0.25, t_0)$				$\beta^{(1)}(0.25, t_0)$			
		PE	ESE	SD	CP	PE	ESE	SD	CP
0	0	1.598	0.097	0.094	0.923	0.701	0.138	0.135	0.937
	10	1.603	0.097	0.099	0.913	0.696	0.141	0.140	0.942
	30	1.609	0.100	0.102	0.901	0.693	0.149	0.143	0.940
	50	1.611	0.106	0.106	0.908	0.687	0.163	0.154	0.932
1	0	1.400	0.117	0.119	0.901	0.794	0.161	0.155	0.931
	10	1.407	0.119	0.121	0.899	0.786	0.165	0.159	0.939
	30	1.411	0.123	0.124	0.894	0.784	0.175	0.169	0.941
	50	1.412	0.129	0.130	0.894	0.787	0.191	0.181	0.933
2	0	1.217	0.138	0.138	0.900	0.878	0.185	0.176	0.927
	10	1.222	0.139	0.141	0.887	0.870	0.189	0.185	0.925
	30	1.220	0.146	0.145	0.892	0.875	0.201	0.192	0.930
	50	1.221	0.154	0.153	0.881	0.879	0.222	0.209	0.922
3	0	1.036	0.158	0.159	0.888	0.957	0.208	0.199	0.910
	10	1.037	0.159	0.161	0.889	0.958	0.212	0.201	0.916
	30	1.038	0.166	0.168	0.882	0.959	0.228	0.217	0.907
	50	1.038	0.176	0.178	0.877	0.965	0.250	0.235	0.911

tooth is defined as the longevity of the initial restoration of the tooth, i.e., time from the initial restoration to the subsequent restoration or extraction of the tooth. If a tooth has not received any subsequent restoration nor extracted at the last visit, the failure time of the tooth is considered as censored at the last visit. The corresponding censoring proportion is 56.6%. For factors that might be related to the longevity of the initial restoration of the tooth, we consider the following 7 variables as covariates: gender(male/female), age at baseline, cohort effects (4, 5 and 6 for group of patients who took treatment between 2000 and 2004, between 2005 and 2009, and between 2010 and 2014, respectively), provider type (predoctoral student / graduate student or faculty), payment (private / Medicaid(XIX) / self-pay), tooth type (molar/pre-molar/anterior), and restoration type (amalgam / composite / GIC / crown or bridge).

We consider a semiparametric quantile regression model to assess the effects of aforementioned factors on quantiles of residual longevity after dental restoration at different times representing the baseline ($t_0 = 0$) and subsequent follow-up visits ($t_0 = 1$ and 2 years). We set female patients who received the initial restoration between 2010 and 2014 (cohort = 6) in an anterior tooth with crown and bridges by graduate student or faculty and made a self-pay as the reference group.

Model fitting results at $\tau = 0.1$ and 0.2 quantiles and various follow-up times ($t_0 = 0, 1$ and 2 years) are summarized in Table 4. Note that, due to the identifiability issue related to the high censoring rate, we focus on lower quantiles that can be reliably estimated. Regression coefficients are estimated by the proposed induced smoothing approach. Standard errors of the estimators are estimated by the proposed robust sandwich-type estimator aided by a resampling-based method.



(a) $\beta^{(0)}(0.5, 0) = 1.61$ (b) $\beta^{(1)}(0.5, 0) = 0.69$

(c) $\beta^{(0)}(0.5, 2) = 1.215$ (d) $\beta^{(1)}(0.5, 2) = 0.883$

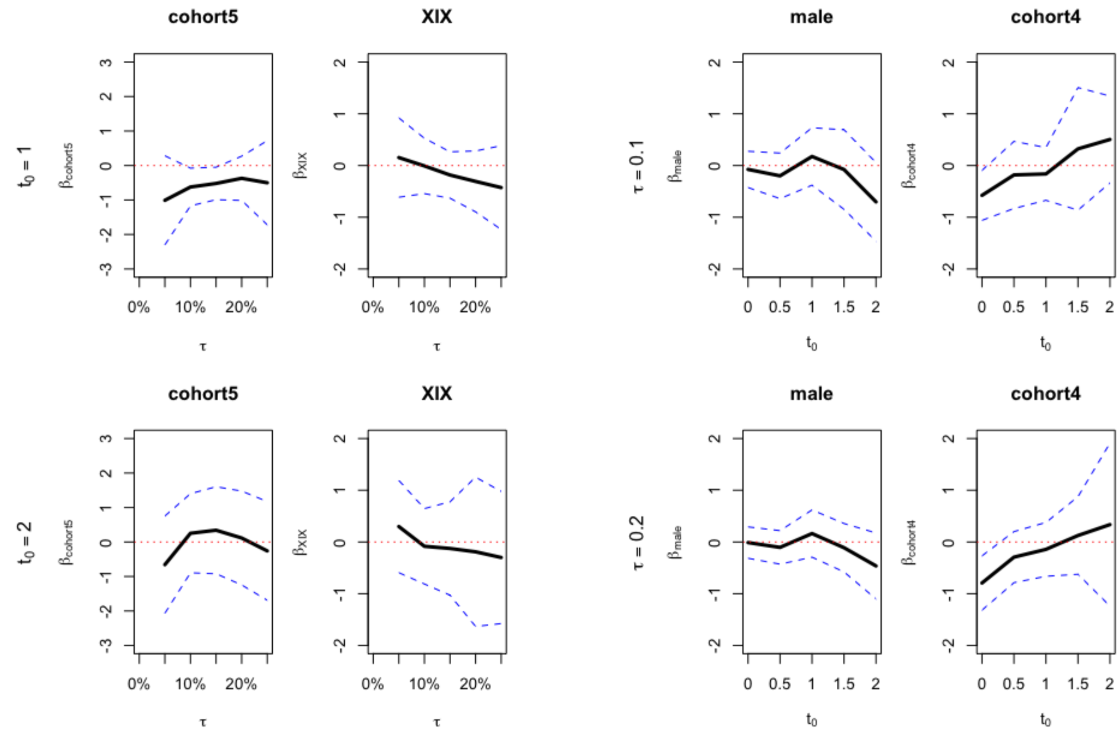
Figure 1: Scatter plots of regression coefficients estimates. The horizontal and vertical axes represent the estimates from the non-smooth and proposed methods, respectively. The circle represents a pair of the two point estimates ($\hat{\beta}_{NS}$, $\hat{\beta}_{IS}$) and The solid line represents the line with the 45 degree angle going through the origin.

Table 4: Results of analyzing the dental restoration data for $\tau = 0.1$ and 0.2 quantiles of residual longevity after dental restoration at $t_0 = 0, 1,$ and 2 (years), respectively. PE is point estimate of the regression parameter. SE is the estimated standard error of the regression parameter estimator.

Covariate	τ											
	0.1						0.2					
	$t_0 = 0$		$t_0 = 1$		$t_0 = 2$		$t_0 = 0$		$t_0 = 1$		$t_0 = 2$	
	PE	SE	PE	SE	PE	SE	PE	SE	PE	SE	PE	SE
Male	-0.076	0.177	0.175	0.261	-0.704	0.339	-0.011	0.147	0.163	0.236	-0.463	0.324
Age	0.014	0.010	-0.022	0.019	-0.035	0.028	0.030	0.011	-0.004	0.016	-0.026	0.025
Cohort4	-0.577	0.241	-0.163	0.276	0.503	0.433	-0.793	0.226	-0.142	0.274	0.336	0.824
Cohort5	-0.606	0.311	-0.624	0.275	0.253	0.494	-0.673	0.176	-0.371	0.299	0.116	0.773
Predoc	0.170	0.230	0.482	0.227	-0.042	0.384	0.381	0.133	0.036	0.395	-0.145	0.590
Private	0.082	0.223	0.057	0.292	-0.132	0.510	0.390	0.194	-0.051	0.253	-0.254	0.552
XIX	-0.132	0.263	-0.008	0.256	-0.083	0.346	0.243	0.215	-0.308	0.321	-0.190	0.640
Molar	0.205	0.212	0.012	0.372	0.372	0.766	0.093	0.197	0.010	0.348	-0.038	0.388
Pre-molar	0.082	0.173	0.910	0.208	-0.324	0.395	0.126	0.180	0.669	0.244	-0.214	0.454
Amalgam	-2.101	1.317	-1.975	0.406	-0.198	1.027	-2.525	0.291	-1.656	0.641	-0.373	0.544
Composite	-2.312	1.288	-2.160	0.406	-0.055	1.140	-2.626	0.286	-1.743	0.713	-0.470	0.770
GIC	-2.001	1.318	-2.167	0.450	-1.015	1.065	-2.303	0.323	-1.948	0.650	-0.858	0.549

The results show varying effects of factors by quantiles considered or times for defining residual life. For example, when $\tau = 0.1$, i.e., the 10^{th} percentile of the residual longevity of the restored teeth, At the 5% significance level, the residual longevity of the restored tooth at baseline or 1 year after the initial restoration for a male does not seem to be statistically significantly different from that for a female while holding the remaining covariates fixed. When evaluating at 2 years after the initial restoration, however, the residual longevity of the restored tooth a male is estimated to be 0.7 years shorter in the logarithm scale than that for a female, a statistically significant effect (p-value $= 0.038$). A similar pattern can be found for $\tau = 0.2$ but none of the effects seem to be statistically significantly different.

To display these varying effects more clearly, we visualize the estimated regression coefficients for some covariates by different quantiles and times along with the associated 95% Wald-type pointwise confidence intervals in Figure 2. For example, at $t_0 = 1$, the estimated effects of Medicaid (XIX) compared to self-pay show a decreasing trend as quantiles get increased. At $\tau = 0.2$, a similar pattern can be seen but with a slower decreasing trend.



(a) (b)

Figure 2: (a) Estimated effects of cohort5 (group of patients who received treatment between 2005 and 2009) and XIX (Medicaid) with the associated 95% pointwise confidence intervals for quantiles ranging from 0.05 to 0.25 at $t_0 = 1$ and 2. Black solid lines are the point estimates of regression parameters for cohort5 and XIX, $\beta_{cohort5}$ and β_{XIX}. Blue dotted lines are the upper and lower bounds of 95% pointwise confidence intervals for $\beta_{cohort5}$ and β_{XIX}. (b) Estimated effects of male and cohort4 (group of patients who received treatment between 200 and 2004) with the associated 95% pointwise confidence intervals for times ranging from 0 to 2 at $\tau = 0.1$ and 0.2. Black solid lines are the point estimates of regression parameters for male and cohort4, β_{male} and $\beta_{cohort4}$. Blue dotted lines are the upper and lower bounds of 95% pointwise confidence intervals for β_{male} and $\beta_{cohort4}$

7. Discussion

In this paper, we propose a statistical inference procedure of fitting a semiparametric quantile regression model for residual life subject to right-censoring. Existing estimation methods (Jung et al., 2009; Kim et al., 2012; Li et al., 2016) are based on estimating equations non-smooth in regression parameters. Our proposed estimation procedure for regression parameters adapts the induced smoothing procedure that have been shown to be computationally efficient and reliable under semiparametric AFT model settings or quantile regression settings (Chiou et al., 2014b, 2015b; Choi et al., 2018). The standard error of the proposed estimator can be estimated by the robust sandwich-type estimator with the application of an efficient resampling method. The proposed estimator is shown to have desirable asymptotic properties: consistent and asymptotically normal. Through extensive simulation experiments, the proposed estimator is also shown to perform reasonably well for finite samples.

Kim et al. (2012) also considered a similar problems and proposed estimating equations with the following form:

$$U_n(\beta) = n^{-1}\sum_{i=1}^{n} I[Z_i \geq t_0]X_i \frac{\delta_i}{\hat{G}(Z_i)}\left(I[\log(Z_i - t_0) \leq X_i^\top \beta] - \tau\right) \tag{9}$$

(9) is the same as (3) on which our proposed estimating equations are based except for the position of the IPCW. An alternative estimation procedure can also be considered by applying the induced smoothing approach to (9). The proposed computing algorithm and standard error estimation procedure can be directly applied to this alternative version. The asymptotic properties can be obtained using the arguments in Appendix with a minimal modification. We also considered this alternative version and compared this with our proposed estimator through some simulation experiments. Overall, the performances are similar. But as the censoring proportion gets higher, the performances of the alternative version become inferior to those of the proposed estimator. This phenomenon requires a further investigation.

Appendix A. Proof of Theorem 1

In Appendix A, we provide a proof of Theorem 1: consistency and asymptotic normality of the proposed induced smoothed estimator.

First, we establish the consistency of the proposed estimator $\hat{\beta}_{IS}$. The consistency of the non-smooth counterpart, $\hat{\beta}_{NS}$, is shown in Li et al. (2016). Based on this consistency result, it suffices to we prove that, as $n \to \infty$, the difference between $\tilde{U}_{t_0}(\beta, \tau, H)$ and $U_{t_0}(\beta, \tau)$ scaled by $n^{1/2}$ converges uniformly to zero in probability for β in the compact neighborhood of β_0.

Let $\sigma_i = (X_i^\top H X_i)^{1/2}$, $\epsilon_i(\beta) = X_i\beta - \log(Z_i - t_0)$ and $d_i(\beta) = \text{sign}(\epsilon_i^\beta)\Phi(-|\epsilon_i^\beta/\sigma_i|)$. Then,

$$\begin{aligned}
&n^{1/2}\{\tilde{U}_{t_0}(\beta,\tau,H) - U_{t_0}(\beta,\tau))\} \\
&= n^{-1/2}\sum_{i=1}^{n} I[Z_i \geq t_0]X_i\delta_i \frac{\hat{G}(t_0)}{\hat{G}(Z_i)}\left\{\Phi\left(\frac{-\epsilon_i(\beta)}{\sigma_i}\right) - I[\epsilon_i(\beta) < 0]\right\} \\
&= n^{-1/2}\sum_{i=1}^{n} I[Z_i \geq t_0]X_i\delta_i \frac{G(t_0)}{G(Z_i)} d_i(\beta) + n^{-1/2}\sum_{i=1}^{n} I[Z_i \geq t_0]X_i\delta_i \left\{\frac{\hat{G}(t_0)}{\hat{G}(Z_i)} - \frac{G(t_0)}{G(Z_i)}\right\} d_i(\beta) \\
&= D_n^{(1)}(\beta) + D_n^{(2)}(\beta)
\end{aligned}$$

To show $\|D_n^{(1)}(\beta)\| \xrightarrow{p} 0$ as $n \to \infty$, we first note that

$$\mathrm{E}\{D_n^{(1)}(\beta)\} = \mathrm{E}\left\{n^{-1/2}\sum_{i=1}^{n} I[Z_i \geq t_0]X_i\delta_i\frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\} = n^{-1/2}\sum_{i=1}^{n} X_i\, \mathrm{E}\{d_i(\beta)|T_i \geq t_0\}.$$

Let ω_{1i}^* be the line segment lying between $X_i^\top(\beta - \beta_0)$ and $X_i^\top(\beta - \beta_0) + \sigma_i t$. Then,

$$\begin{aligned}
\mathrm{E}\{d_i(\beta)|T_i \geq t_0\} &= \int_{-\infty}^{\infty} d_i(\beta) g_{T_i - t_0}\{\epsilon_i(\beta) + X_i^\top(\beta - \beta_0)|T_i \geq t_0\} d\epsilon_i(\beta) \\
&= \sigma_i \int_{-\infty}^{\infty} \Phi(-|t|)\{2I[t > 0] - 1\}\left[g_{T_i - t_0}\{\sigma_i t + X_i^\top(\beta - \beta_0)\} + g'_{T_i - t_0}\{\omega_i^*(t)\}\sigma_i t\right] dt
\end{aligned}$$

It follows from Conditions C1 and C3 that $\sup_i g_{T_i - t_0}\{\sigma_i t + X_i^\top(\beta - \beta_0)\} < \infty$. Since $\int_{-\infty}^{\infty} \Phi(-|t|)\{2I[t > 0] - 1\}dt = 0$, we have $\int_{-\infty}^{\infty} \Phi(-|t|)\{2I[t > 0] - 1\}g_{T_i - t_0}\{X_i^{\star\top}(\beta - \beta_0)\}dt = 0$. Again, by Condition C1, $\exists M > 0$ such that $\sup_i |g_i'\{\omega_i^*(t)\}| < M$. Thus, $|E\{d_i(\beta)\}| \leq \int_{-\infty}^{\infty} |t|\Phi(|t|)|g_i'\{\omega_i^*(t)\}|dt \leq M\sigma_i^2/2$. Note that $\sum_{i=1}^n \sigma_i^2 = \mathrm{tr}(XHX^\top) = \mathrm{tr}(HX^\top X)$ is bounded by $H = O(n^{-1})$ and Condition C2. Then, $\sum_{i=1}^n |E\{d_i(\beta)\}| \leq M\sum_{i=1}^n \sigma_i^2/2$ is also bounded. Therefore,

$$\|\mathrm{E}\{D_n^{(1)}(\beta)\}\| \leq n^{-1/2}\sqrt{p}\sup_{i,j}|X_{ij}|\sum_{i=1}^{n}|\mathrm{E}\{d_i(\beta)|T_i \geq t_0\}| \to 0 \quad \text{as } n \to 0.$$

By applying Condition C3, we have

$$\mathrm{Var}\{D_n^{(1)}(\beta)\} = \mathrm{Var}\left\{n^{-1/2}\sum_{i=1}^n X_iX_i^\top I[Z_i \ge t_0]\delta_i \frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\} \le n^{-1}\sum_{i=1}^n \frac{X_iX_i^\top}{c_0}\,\mathrm{E}\{d_i^2(\beta)|T_i \ge t_0\}.$$

It follows from the arguments similar to evaluating $\mathrm{E}\{d_i(\beta)|T_i \ge t_0\}$ combining with Conditions C1 and C2, we have, as $n \to \infty$, $\|\mathrm{E}\{d_i^2(\beta)|T_i \ge t_0\}\| \to 0$. This implies $\|\mathrm{Var}\{D_n^{(1)}(\beta)\}\| \to 0$. Then, by the Weak Law of Large Numbers,

$$\|D_n^{(1)}(\beta)\| \xrightarrow{p} 0, \quad \text{as } n \to \infty. \tag{10}$$

for β in a compact neighborhood of β_0.

To show $\|D_n^{(2)}(\beta)\| \xrightarrow{p} 0$ as $n \to \infty$, we use the martingale representation of the Kaplan-Meier estimator (Fleming and Harrington, 2011)). Specifically, $\hat{G}(t)$ can be represented as

$$\frac{\hat{G}(t) - G(t)}{G(t)} = -\sum_{i=1}^n \int_0^t \left\{\frac{\hat{G}(u^-)}{G(u)}\right\}\frac{dM_i^C(u)}{Y(u)}$$

where $M_i^C(u) = N_i^C(u) - \int_0^t Y_i(u)d\Lambda^C(s)$, $N_i^C(u) = (1-\delta_i)I[Z_i \le u], \Lambda^C(u) = -\log\{G(u)\}$, $Y(u) = \sum_{i=1}^n Y_i(u)$, and $Y_i(u) = I[Z_i \ge u]$. By combining this with an application of the functional delta method and the uniform convergence result of $\hat{G}(\cdot)$ to $G(\cdot)$, we have

$$\begin{aligned} D_n^{(2)}(\beta) &= n^{-1/2}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_i n^{-1}\sum_{j=1}^n\left\{\frac{h_j(t_0)}{G(Z_i)} - \frac{h_j(Z_i)G(t_0)}{G^2(Z_i)}\right\}d_i(\beta) + o_p(1) \\ &= n^{-1/2}\sum_{j=1}^n\int_{t_0}^{\nu}\left\{n^{-1}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_iY_i(u)\frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\}\frac{dM_j^C(u)}{y(u)} + o_p(1) \end{aligned}$$

where

$$h_j(t) = G(t)\int_0^t \frac{dM_j^C(u)}{Y(u)} \text{ and } y(t) = \lim_{n\to\infty} n^{-1}Y(t).$$

Using the arguments similar to those used to establish $\|D_n^{(1)}(\beta)\| \xrightarrow{p} 0$, as $n \to \infty$, it can be shown that $\mathrm{E}\left\{|Y_i(u)d_i(\beta)| \mid T_i \ge t_0\right\} = O(n^{-1/2})$. Thus,

$$\begin{aligned} &\left\|\mathrm{E}\left\{n^{-1}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_iY_i(u)\frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\}\right\| = \left\|n^{-1}\sum_{i=1}^n X_i\,\mathrm{E}\left\{Y_i(u)d_i(\beta)|T_i \ge t_0\right\}\right\| \\ &\le \sqrt{p}\sup_{i,j}|X_{ij}|n^{-1}\sum_{i=1}^n \mathrm{E}\left\{|Y_i(u)d_i(\beta)| \mid T_i \ge t_0\right\} \to 0. \end{aligned}$$

It then follows that, as $n \to \infty$

$$\left\|n^{-1}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_iY_i(u)\frac{G(t_0)}{G(Z_i)}d_i(\beta) - \mathrm{E}\left\{n^{-1}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_iY_i(u)\frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\}\right\| \xrightarrow{p} 0$$

uniformly in β for β in the compact neighborhood of β_0. By applying the martingale central limit theorem and the Kolmogorov-Centsov Theorem (Karatzas and Shreve, 1988, p53),

$$n^{-1/2}\sum_{j=1}^n \frac{dM_j^C(u)}{y(u)} \text{ converges weakly to a zero-mean Gaussian process with continuous sample paths.}$$

By combining these results, it follows from Lemma 1 in Lin (2000) that

$$\left\|n^{-1/2}\sum_{j=1}^n\int_{t_0}^{\nu}\left\{n^{-1}\sum_{i=1}^n I[Z_i \ge t_0]X_i\delta_iY_i(u)\frac{G(t_0)}{G(Z_i)}d_i(\beta)\right\}\frac{dM_j^c(u)}{y(u)}\right\| \xrightarrow{p} 0. \tag{11}$$

By combining (10) and (11), we have

$$\|n^{1/2}\{\tilde{U}_n(\beta,\tau,H) - U_{t_0}(\beta,\tau)\}\| \xrightarrow{p} 0. \tag{12}$$

Note that both $\tilde{U}_{t_0}(\beta,\tau,H)$ and $U_{t_0}(\beta,\tau)$ are monotone functions, thus the point-wise convergence could be strengthened to uniform convergence (Shorack and Wellner, 2009).

To establish the asymptotic equivalence of $n^{1/2}(\hat{\beta}_{IS} - \beta_0)$ and $n^{1/2}(\hat{\beta}_{NS} - \beta_0)$, it suffices to show that the following two convergence results hold: As $n \to \infty$,

$$\text{(i)} \ \left\|\hat{A}(\beta_0, H) - A\right\| \to 0 \text{ and (ii) } \left\|n^{1/2}\left\{\tilde{U}_{t_0}(\beta_0,\tau,H) - U_{t_0}(\beta_0,\tau)\right\}\right\| \to 0.$$

Note that (12) implies (ii). Thus, we prove (i). For any vectors $a, b \in R^p$,

$$\begin{aligned}
\mathrm{E}\left[a^\top \hat{A}(\beta_0,H)b\right] &= a^\top \mathrm{E}\left[n^{-1}\sum_{i=1}^{n} I[Z_i > t_0]X_iX_i^\top \frac{\hat{G}(t_0)\delta_i}{\hat{G}(Z_i)}\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\right] b \\
&= a^\top \left[n^{-1}\sum_{i=1}^{n} X_iX_i^\top \mathrm{E}\left\{\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\Big|\ T_i > t_0, X_i\right\}\right] b
\end{aligned}$$

It follows from the variable transformation $t = \epsilon(\beta_0)/\sigma_i$ and the Taylor expansion at 0 that

$$\begin{aligned}
\mathrm{E}\left\{\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\Big|T_i > t_0, X_i\right\} &= \int_{-\infty}^{\infty} \phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big) g_{T_i-t_0}(\epsilon_i)d\epsilon_i \\
&= \int_{-\infty}^{\infty} \phi(-t)\, g_{T_i-t_0}(0)dt + \sigma_i \int_{-\infty}^{\infty} t\phi(-t)\, g'_{T_i-t_0}(\omega^*_{2i})dt
\end{aligned}$$

where ω^*_{2i} is some value lying between 0 and $\sigma_i t$.

By Condition C1, we have $\sigma_i \int_{-\infty}^{\infty} t\phi(-t)\, g'_{T_i-t_0}(\omega^*_{2i})dt \le M\sigma_i \int_{-\infty}^{\infty} |t|\phi(-|t|)\,dt \to 0$. Since $\int_{-\infty}^{\infty} \phi(-t)\, g_{T_i-t_0}(0)dt = 0$, we have $\mathrm{E}\left\{\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\Big|T_i > t_0, X_i\right\} \to g_{T_i-t_0}(0)$ and, therefore,

$$\lim_{n\to\infty} \mathrm{E}\left[a^\top \hat{A}(\beta_0,H)b\right] = a^\top \left\{\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} X_iX_i^\top g_{T_i-t_0}(0)\right\} b = a^\top Ab. \tag{13}$$

By Condition C1 and applying the arguments in Pang et al. (2012, p795, Appendix), it can be shown that

$$\mathrm{E}\left[\left\{\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\right\}^2 \Big|\ T_i > t_0, X_i\right] = O(n^{1/2}).$$

Then,

$$\begin{aligned}
&\mathrm{Var}\left[a^\top \left\{n^{-1}\sum_{i=1}^{n} I[Z_i > t_0]X_iX_i^\top \frac{\hat{G}(t_0)\delta_i}{\hat{G}(Z_i)}\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\right\} b\right] \\
&\le \frac{1}{n^2 c_0}\sum_{i=1}^{n}(a^\top X_iX_i^\top)^2\, \mathrm{E}\left[\left\{\phi\Big(-\frac{\epsilon_i(\beta_0)}{\sigma_i}\Big)\Big(\frac{1}{\sigma_i}\Big)\right\}^2 \Big|\ T_i > t_0, X_i\right] \to 0.
\end{aligned} \tag{14}$$

By combining the results in (13) and (14), we have $\left\|\hat{A}(\beta_0,H) - A\right\| \to 0$ as $n \to \infty$. This completes the proof of (i).

Appendix B. Supplementary material

Supplementary material related to this article can be found online

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (2020R1A2C1A0101313911)

Conflict of interest

The authors declare that they have no conflict of interest.

References

G. Maguluri, C.-H. Zhang, Estimation in the mean residual life regression model, Journal of the Royal Statistical Society: Series B (Methodological) 56 (1994) 477–489.

D. Oakes, T. Dasu, A note on residual life, Biometrika 77 (1990) 409–410.

D. Oakes, T. Dasu, Inference for the proportional mean residual life model, Lecture Notes-Monograph Series (2003) 105–116.

Y. Chen, N. Jewell, X. Lei, S. Cheng, Semiparametric estimation of proportional mean residual life model in presence of censoring, Biometrics 61 (2005) 170–178.

Y. Q. Chen, S. Cheng, Linear life expectancy regression with censored data, Biometrika 93 (2006) 303–313.

Y. Q. Chen, Additive expectancy regression, Journal of the American Statistical Association 102 (2007) 153–166.

Z. Zhang, X. Zhao, L. Sun, Goodness-of-fit tests for additive mean residual life model under right censoring, Lifetime data analysis 16 (2010) 385–408.

S. Liu, S. K. Ghosh, Regression analysis of mean residual life function, Technical Report, North Carolina State University. Dept. of Statistics, 2008.

L. Sun, Z. Zhang, A class of transformed mean residual life models with censored survival data, Journal of the American Statistical Association 104 (2009) 803–815.

L. Sun, X. Song, Z. Zhang, Mean residual life models with time-dependent coefficients under right censoring, Biometrika 99 (2012) 185–197.

R. Li, X. Huang, J. E. Cortes, Quantile residual life regression with longitudinal biomarker measurements for dynamic prediction, Journal of the Royal Statistical Society. Series C: Applied Statistics 65 (2016) 755–773.

Z. Ying, T. Sit, Survival analysis: a quantile perspective, in: Handbook of Quantile Regression, Chapman and Hall/CRC, 2017, pp. 89–108.

R. Koenker, G. Bassett Jr, Regression quantiles, Econometrica: journal of the Econometric Society (1978) 33–50.

S.-H. Jung, Quasi-likelihood for median regression models, Journal of the american statistical association 91 (1996) 251–257.

S. Portnoy, R. Koenker, et al., The gaussian hare and the laplacian tortoise: computability of squared-error versus absolute-error estimators, Statistical Science 12 (1997) 279–300.

Y. Wei, A. Pere, R. Koenker, X. He, Quantile regression methods for reference growth charts, Statistics in medicine 25 (2006) 1369–1382.

Y.-J. Whang, Smoothed empirical likelihood methods for quantile regression models, Econometric Theory (2006) 173–205.

Z. Ying, S.-H. Jung, L.-J. Wei, Survival analysis with median regression models, Journal of the American Statistical Association 90 (1995) 178–184.

H. Bang, A. A. Tsiatis, Median regression with censored cost data, Biometrics 58 (2002) 643–649.

A. E. Gelfand, A. Kottas, Bayesian semiparametric regression for median residual life, Scandinavian Journal of Statistics 30 (2003) 651–665.

S. Portnoy, Censored regression quantiles, Journal of the American Statistical Association 98 (2003) 1001–1012.

L. Peng, Y. Huang, Survival analysis with quantile regression models, Journal of the American Statistical Association 103 (2008) 637–649.

H. J. Wang, L. Wang, Locally weighted censored quantile regression, Journal of the American Statistical Association 104 (2009) 1117–1128.

Y. Huang, Quantile calculus and censored regression, Annals of statistics 38 (2010) 1607.

S. Portnoy, G. Lin, Asymptotics for censored regression quantiles, Journal of Nonparametric Statistics 22 (2010) 115–130.

S.-H. Jung, J.-H. Jeong, H. Bandos, Regression on quantile residual life, Biometrics 65 (2009) 1203–1212.

M.-O. Kim, M. Zhou, J.-H. Jeong, Censored quantile regression for residual lifetimes, Lifetime data analysis 18 (2012) 177–194.

B. Brown, Y.-G. Wang, Standard errors and covariance matrices for smoothed rank estimators, Biometrika 92 (2005) 149–158.

B. Brown, Y.-G. Wang, Induced smoothing for rank regression with censored survival times, Statistics in medicine 26 (2007) 828–836.

L. M. Johnson, R. L. Strawderman, Induced smoothing for the semiparametric accelerated failure time model: asymptotics and extensions to clustered data, Biometrika 96 (2009) 577–590.

L. Fu, Y.-G. Wang, Z. Bai, Rank regression for analysis of clustered data: a natural induced smoothing approach, Computational statistics & data analysis 54 (2010) 1036–1050.

L. Pang, W. Lu, H. J. Wang, Variance estimation in censored quantile regression via induced smoothing, Computational statistics & data analysis 56 (2012) 785–796.

S. H. Chiou, S. Kang, J. Yan, Fast accelerated failure time modeling for case-cohort data, Statistics and Computing 24 (2014a) 559–568.

S. H. Chiou, S. Kang, J. Yan, et al., Fitting accelerated failure time models in routine survival analysis with r package aftgee, Journal of Statistical Software 61 (2014b) 1–23.

S. Chiou, S. Kang, J. Yan, Rank-based estimating equations with general weight for accelerated failure time models: an induced smoothing approach, Statistics in Medicine 34 (2015a) 1495–1510.

S. H. Chiou, S. Kang, J. Yan, Semiparametric accelerated failure time modeling for clustered failure times from stratified sampling, Journal of the American Statistical Association 110 (2015b) 621–629.

S. Kang, Fitting semiparametric accelerated failure time models for nested case–control data, Journal of Statistical Computation and Simulation 87 (2017) 652–663.

S. Choi, S. Kang, X. Huang, Smoothed quantile regression analysis of competing risks, Biometrical Journal 60 (2018) 934–946.

D. Caplan, Y. Li, W. Wang, S. Kang, L. Marchini, H. Cowen, J. Yan, Dental restoration longevity among geriatric and special needs patients, JDR Clinical & Translational Research 4 (2019) 41–48.

R. Koenker, Quantile Regression, Econometric Society Monographs, Cambridge University Press, 2005. URL: `http://www.amazon.de/Quantile-Regression-Econometric-Society-Monographs/dp/0521608279/ref=sr`$_{11}?ie = UTF8qid = 1312553603sr = 8{-}1. R.\ Koenker, S.\ Portnoy, P.\ T.Ng, A.\ Zeileis, P.\ Grosjean, B.\ D.Ripley, Package`qu$

R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2020. URL: https://www.R-project.org/.

R. Varadhan, P. Gilbert, Bb: An r package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function, Journal of statistical software 32 (2009) 1–26.

N. Kassebaum, E. Bernabé, M. Dahiya, B. Bhandari, C. Murray, W. Marcenes, Global burden of untreated caries: a systematic review and metaregression, Journal of dental research 94 (2015) 650–658.

T. R. Fleming, D. P. Harrington, Counting processes and survival analysis, volume 169, John Wiley & Sons, 2011.

I. Karatzas, S. Shreve, Brownian motion and stochastic calculus, New York:Springer, 1988.

D. Lin, On fitting cox's proportional hazards models to survey data, Biometrika 87 (2000) 37–47.

G. R. Shorack, J. A. Wellner, Empirical processes with applications to statistics, SIAM, 2009.