Published in final edited form as:

Comput Stat Data Anal. 2012 April 1; 56(4): 785–796. doi:10.1016/j.csda.2010.10.018.

Variance Estimation in Censored Quantile Regression via Induced Smoothing

Lei Panga^a, Wenbin Lu^a, and Huixia Judy Wang^{a,*}

^a Department of Statistics, North Carolina State University, Raleigh, NC 27606, U.S.A.

Abstract

Statistical inference in censored quantile regression is challenging, partly due to the unsmoothness of the quantile score function. A new procedure is developed to estimate the variance of Bang and Tsiatis's inverse-censoring-probability weighted estimator for censored quantile regression by employing the idea of induced smoothing. The proposed variance estimator is shown to be asymptotically consistent. In addition, numerical study suggests that the proposed procedure performs well in finite samples, and it is computationally more efficient than the commonly used bootstrap method.

Keywords

censored quantile regression; smoothing; survival analysis; variance estimation

1. Introduction

Consider a random sample subject to random right censoring: $\{Y_i = \min(T_i, C_i), \Delta_i, Z_i\}_{i=1}^n$, where T_i and C_i denote the failure time and the censoring time (or some monotone transformations thereof), respectively, $\Delta_i = I(T_i \leq C_i)$ is the censoring indicator, and Z_i is the p -dimensional covariate vector with 1 as the first component corresponding to the intercept. We assume the following quantile regression model at a fixed quantile level $\tau \in (0, 1)$,

$$T_i = Z_i' \beta_0(\tau) + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where $\beta_0(\tau)$ is the p -dimensional quantile coefficient at the τ th quantile, and ε_i is the random error whose τ th conditional quantile given Z_i equals 0. In the rest of the article, we will suppress τ in $\beta_0(\tau)$ for notational simplicity.

The quantile regression model, first introduced by Koenker and Bassett (1978), is a valuable alternative to the Cox proportional hazards model (Cox, 1972) and the accelerated failure time (AFT) model (Cox and Oakes, 1984) in survival analysis. Compared to the Cox and the AFT models, quantile regression offers an automatic and flexible way to capture the heterogeneity in the data. Moreover, the relations between the conditional quantiles of the failure time and the covariates are directly interpretable.

*Corresponding author. judywang@ncsu.edu, Phone: (919) 513-1661, Fax: (919) 515-7591.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Considerable research efforts have been devoted to the estimation of regression parameters in censored quantile regression. Early works of Powell (1984) and Powell (1986) focus on fixed censoring where the censoring variable C_i is always observable. For quantile regression with random censoring, recent developments include Ying et al. (1995), Bang and Tsiatis (2002), Honor et al. (2002), Portnoy (2003), Peng and Huang (2008), Wang and Wang (2009), Huang (2010) and Portnoy and Lin (2010). However, relatively fewer methods are available for estimating the variance of censored quantile regression estimators and as such the associated inference becomes more difficult. In general, a plug-in estimator of the variance can be obtained based on the asymptotical normality of the corresponding estimator, while the asymptotical variance-covariance matrix usually involves $f_X(0|Z_i)$, the unknown conditional error density function $f_X(\cdot|Z_i)$ evaluated at zero. Nonparametric approaches can be used to estimate $f_X(0|Z_i)$, but most of them require moderate to large sample size and choosing proper tuning parameters, such as the bandwidth parameter in kernel smoothing. Previous research show that inference based on nonparametric density estimation is very sensitive to the choice of tuning parameters in finite samples; see Chen and Wei (2005) and Kocherginsky et al. (2005) for comparison of different inference methods in quantile regression without censoring. Some other researchers considered computationally intensive methods for variance estimation, e.g. Portnoy (2003) and Wang and Wang (2009) used bootstrap, and Peng and Huang (2008) employed the resampling approach of Jin et al. (2001). To bypass the variance estimation in the context of censored quantile regression, an alternative inference approach via the inversion of score tests was also considered by Ying et al. (1995) and Bang and Tsiatis (2002). However, the construction of confidence intervals is cumbersome, because one needs to perform hypothesis testing repeatedly on a fine grid in the p -dimensional parameter space.

The difficulty in variance estimation for censored quantile regression arises partly from the unsmoothness of the corresponding estimating function that involves indicator functions. The unsmoothness issue can be overcome by approximating the indicator functions with some suitable smooth functions. A common smoothing method is the kernel smoothing; see Chen and Hall (1993), Horowitz (1998), Heller (2007) and Whang (2006). Using kernel smoothing, if the bandwidth parameter converges to zero at a proper rate, the smoothed estimating function is asymptotically equivalent to the original estimating function, and the covariance can be consistently estimated by the sandwich formula derived from the smoothed estimating equation. Though conceptually simple, kernel smoothing usually needs a well-chosen bandwidth parameter to achieve good finite sample performance.

In this paper, we employ the induced smoothing idea of Brown and Wang (2005) and develop a variance estimation procedure for the censored quantile coefficient estimator of Bang and Tsiatis (2002). The induced smoothing method smoothes the estimating function by taking its expectation under the distribution of a random perturbation of the unknown parameter. Induced smoothing is practically convenient, as it does not require selecting any tuning parameters. Recently, Wang et al. (2009) demonstrated that induced smoothing provides reliable variance estimation for quantile regression with uncensored data. Other notable applications of induced smoothing in different contexts include Brown and Wang (2006), Johnson and Strawderman (2009), Fu et al. (2010) and Wang and Fu (2010). In this paper, we will show that for censored quantile regression, our proposed variance estimator via induced smoothing is asymptotically consistent. In addition, it has good finite sample performance and is computationally efficient, requiring only a fraction of the computational cost as compared to the bootstrap method.

The rest of this article is organized as follows. In Section 2, we present the variance estimation procedure. In Section 3, we establish the asymptotic property of the variance estimator. The finite sample performance of the proposed procedure is investigated through

a simulation study in Section 4, and an analysis of a multiple myeloma data set in Section 5. We conclude the paper with some discussions in Section 6. Theoretical proofs are relegated to the Appendix.

2. Induced Smoothing and Computational Algorithm

When there is no censoring, the quantile coefficient β_0 in model (1) can be consistently estimated by the solution of the following estimating equation

$$n^{-1} \sum_{i=1}^n Z_i \{I(Y_i - \beta' Z_i < 0) - \tau\} = 0.$$

In the presence of random right censoring, a number of estimation methods have been proposed. One popular estimator is the inverse-censoring-probability-weighted (ICPW) estimator proposed by Bang and Tsiatis (2002). Specifically, the ICPW estimator $\hat{\beta}$ is defined as the solution of the estimating equation

$$U_n(\beta) = n^{-1} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} \{I(Y_i - \beta' Z_i < 0) - \tau\} = 0, \quad (2)$$

where $\widehat{G}(\cdot)$ is the Kaplan-Meier estimate of the survival function of the censoring variable C_i . Note that solving (2) is equivalent to minimizing a weighted objective function

$$n^{-1} \sum_{i=1}^n \frac{\Delta_i}{\widehat{G}(Y_i)} \rho_{\tau}(Y_i - \beta' Z_i), \quad (3)$$

where $\rho_{\tau}(t) = t\tau - tI(t < 0)$. The minimization can be solved efficiently by using existing linear programming algorithms, for instance, the *rq* function in R package *quantreg*.

In Bang and Tsiatis's method, censoring times are assumed to be independent and identically distributed, and independent of covariates, which we also assume through the remaining of the paper. Bang and Tsiatis (2002) proved that under some regularity conditions, $\hat{\beta}$ is consistent and asymptotically normal:

$$n^{1/2}(\widehat{\beta} - \beta_0) \xrightarrow{d} \mathcal{N}(0, \Gamma),$$

where $\Gamma = \mathbf{A}^{-1} \Sigma \mathbf{A}^{-1}$, $\mathbf{A} = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n Z_i Z_i' f_i(0|Z_i)$, and $\Sigma = \lim_{n \rightarrow \infty} \text{Var}\{n^{1/2} U_n(\beta_0)\}$. It is well-known that direct estimation of Γ is difficult, since the matrix $\mathbf{A}$ depends on the unknown error densities $f_i(\cdot|Z_i)$, $i = 1, \dots, n$.

By the asymptotic normality of $\hat{\beta}$, we can write $\hat{\beta} = \beta_0 + \mathbf{H}^{1/2} V$, where $\mathbf{H} = n^{-1} \Gamma$, $V \sim \mathcal{N}(0, I_p)$, and I_p is the $p \times p$ identity matrix. In this way, $\hat{\beta}$ can be regarded as a random perturbation of β_0 . Motivated by the induced smoothing idea of Brown and Wang (2005), we consider the smoothed estimating function $\tilde{U}_n(\beta, \mathbf{H}) = E\{U_n(\beta + \mathbf{H}^{1/2} V)\}$, the expectation of the unsmoothed estimating function with respect to V . We have

$$\tilde{U}_n(\beta, \mathbf{H}) = E_v\{U_n(\beta + \mathbf{H}^{1/2}V)\} = \frac{1}{n} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} \left\{ \Phi \left(\frac{\beta' Z_i - Y_i}{\sqrt{Z_i' \mathbf{H} Z_i}} \right) - \tau \right\}, \quad (4)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. When $\mathbf{H}$ is given, the smoothed quantile coefficient estimator $\tilde{\beta}$ can be obtained by solving $\tilde{U}_n(\tilde{\beta}, \mathbf{H}) = 0$. However, since $\mathbf{H}$ is unknown and has to be estimated, we propose to use a positive definite $p \times p$ matrix $\tilde{\mathbf{H}}_0 = O(n^{-1})$ as an initial value in the smoothing, and then update through iterations to estimate $\mathbf{H}$ and hence $\mathbf{\Gamma}$. In fact, we will show, as long as the matrix used for initial smoothing is of order $O(n^{-1})$ and positive definite, the corresponding smoothed coefficient estimator $\tilde{\beta}$ is asymptotically equivalent to $\hat{\beta}$.

For any positive definite matrix $\mathbf{W} = O(n^{-1})$, we define

$$\tilde{\mathbf{A}}_n(\beta, \mathbf{W}) = \frac{\partial \tilde{U}_n(\beta, \mathbf{W})}{\partial \beta} = n^{-1} \sum_{i=1}^n \frac{\Delta_i}{\widehat{G}(Y_i)} \varphi \left(\frac{\beta' Z_i - Y_i}{\sqrt{Z_i' \mathbf{W} Z_i}} \right) \frac{Z_i Z_i'}{\sqrt{Z_i' \mathbf{W} Z_i}}.$$

The estimation algorithm can be summarized as follows.

Step 1. Let $\tilde{\beta}_0 = \hat{\beta}$, the Bang-Tsiatis estimator, and $\tilde{\mathbf{H}}_0 = n^{-1} \mathbf{I}_p$.

Step 2. Given $\tilde{\beta}_{k-1}$ and $\tilde{\mathbf{H}}_{k-1}$ from the $(k-1)$ th step, update $\tilde{\beta}_k$ and $\tilde{\mathbf{H}}_k$ as:

$$\begin{aligned} \tilde{\beta}_k &= \tilde{\beta}_{k-1} + \{-\tilde{\mathbf{A}}_n(\tilde{\beta}_{k-1}, \tilde{\mathbf{H}}_{k-1})\}^{-1} \tilde{U}_n(\tilde{\beta}_{k-1}, \tilde{\mathbf{H}}_{k-1}), \\ \text{and } \tilde{\mathbf{H}}_k &= n^{-1} \tilde{\mathbf{A}}_n^{-1}(\tilde{\beta}_k, \tilde{\mathbf{H}}_{k-1}) \widehat{\Sigma}(\tilde{\beta}_k) \tilde{\mathbf{A}}_n^{-1}(\tilde{\beta}_k, \tilde{\mathbf{H}}_{k-1}), \end{aligned}$$

where $\widehat{\Sigma}(\tilde{\beta}_k)$ is the consistent estimator of the covariance matrix of $\sqrt{n}U_n(\tilde{\beta}_k)$, as defined in (10) in the Appendix.

Step 3. Repeat step 2 till convergence. Denote the coefficient estimate and the covariance estimate at convergence as $\tilde{\beta}$ and $\tilde{\mathbf{H}}$, respectively, and let $\tilde{\mathbf{\Gamma}} = n\tilde{\mathbf{H}}$.

In addition, we denote the covariance estimator of $n^{1/2}\tilde{\beta}_k$ at the k th iteration as $\tilde{\mathbf{\Gamma}}_k = n\tilde{\mathbf{H}}_k$. Thereafter, we will refer to this procedure as IS (Induced Smoothing).

3. Asymptotic Results

In this section, we establish the consistency and asymptotic normality of $\tilde{\beta}_k$, and the consistency of $\tilde{\mathbf{\Gamma}}_k$. We impose the following regularity conditions.

A1. The conditional error distribution functions, $F_i(\cdot|Z_i)_{i=1}^n$, are absolutely continuous with continuous densities $f_i(\cdot|Z_i)$ uniformly bounded away from 0 and ∞ in a neighborhood of 0, and $f_i'(\cdot|Z_i)$ exists and is uniformly bounded on the real line.

A2. For each $i = 1, \dots, n$, Z_i satisfies the following conditions:

- a. $n^{-1} \sum_{i=1}^n Z_i Z_i' f_i(0|Z_i)$ converges to a positive definite matrix $\mathbf{A}$;
- b. $\sup_i \|Z_i\| < \infty$, where $\|\cdot\|$ denotes the Euclidean norm.

A3. There exists $L > 0$ such that $P(C > L) = 0$ and $P(C = L) = \nu$, where ν is some positive constant.

Theorem 1

Assume conditions A1–A3 hold, and the initial smoothing matrix $\mathbf{H}_0$ is positive definite and $O(n^{-1})$, as $n \rightarrow \infty$, we have

- i. $n^{1/2}(\tilde{\beta}_k - \beta_0) \xrightarrow{d} \mathcal{N}(0, \Gamma)$, for any $k \geq 1$;
- ii. $\tilde{\Gamma}_k \xrightarrow{p} \Gamma$, for any $k \geq 1$.

Remark

Assumption A1 and A2 impose some conditions on the conditional error distributions and the covariates, both of which will hold in most practical situations. Assumption A3 will be satisfied in many clinical settings with administrative censoring. Theorem 1 (i) suggests that the smoothed coefficient estimator, $\tilde{\beta}_k$, has the same asymptotic distribution as the unsmoothed estimator $\hat{\beta}$. Theorem 1 (ii) provides the theoretical justification for IS as a variance estimation procedure for the censored quantile regression model (1). In addition, Theorem 1 suggest that $\tilde{\beta}$ and $\tilde{\Gamma}$, the coefficient and variance estimators at convergence, have the same asymptotic properties as the one-step estimators $\tilde{\beta}_1$ and $\tilde{\Gamma}_1$, respectively. However, our numerical studies show that the IS variance estimator at convergence has smaller bias than the one-step estimator in finite samples. The proposed algorithm converges quite fast, usually within 4 iterations for the simulation study in Section 4 and within 4–8 iterations for the data analysis in Section 5.

4. Simulation

In this section, we conduct a simulation study to compare three variance estimation methods: the proposed IS method, the bootstrap method (BT-boot) and the kernel smoothing method with prespecified bandwidths. For BT-boot, the bootstrap coefficient estimates are obtained by minimizing (3) using bootstrap samples generated by resampling the triplets $\{Y_i, \Delta_i, Z_i\}$ with replacement. The bootstrap covariance matrix is calculated as the sample covariance of the bootstrap coefficient estimates. For kernel smoothing, the indicator functions $I(Y_i - \beta'Z_i < 0)$ are approximated by $\Phi\{(Y_i - \beta'Z_i)/h\}$, where h is the bandwidth, and the covariance matrix is estimated by the sandwich formula using M-estimation technique. We consider three bandwidths in the kernel smoothing $h = n^{-0.1}$, $n^{-0.3}$ and $n^{-0.7}$, and denote them as S_1 , S_2 and S_3 , respectively. Note that the induced smoothing approach can be viewed as a variation of the kernel smoothing method: in induced smoothing, rather than a fixed bandwidth h , each subject receives a subject-specific bandwidth $h_i = \sqrt{Z_i' \mathbf{H} Z_i}$ that is determined automatically through iteration.

We generate data from the following model.

$$\log(T_i) = Z_i' \beta + \varepsilon_i, \quad i = 1, \dots, n,$$

where $Z_i = (1, Z_{i1}, Z_{i2})'$, $Z_{i1} \sim \text{Unif}(0, 4)$, $Z_{i2} \sim \text{Bernoulli}(0.5)$ and $\beta = (\beta_1, \beta_2, \beta_3)' = (1, 1, 1)'$. We consider 3 different scenarios for generating the random error ε_i .

Scenario 1: homoscedastic normal, $\varepsilon_i \sim N(0, 1)$;

Scenario 2: heteroscedastic normal, $\varepsilon_i \sim N(0, (1 + 0.3Z_{i1})^2)$;

Scenario 3: heteroscedastic logistic, $\varepsilon_i \sim \text{Logistic}(0, 1 + 0.3Z_{i1})$.

The censoring times C_i are generated by $\log(C_i) \sim \text{Unif}(0, k)$, where k is set as 23, 12 and 7, corresponding to 15%, 30% and 50% of censoring, respectively. For each scenario, 2000 simulated samples of size $n = 200$ are generated. We focus on two quantiles, $\tau = 0.25$ and $\tau = 0.5$.

Tables 1, 2 and 3 summarize the simulation results for 15%, 30% and 50% censoring, respectively. For each method, we report the mean bias and the mean squared error (MSE) of each estimate of β_i , $i = 1, 2, 3$, the empirical coverage probability of 95% Wald-type confidence intervals, and SE/SD, where SD is the Monte Carlo standard deviation and SE is the average of estimated standard errors across 2000 simulations.

In cases with moderate censoring (15% and 30%), the coefficient estimates from IS and BT-boot are essentially unbiased, while those from the kernel smoothing methods tend to be biased when the bandwidth is selected too large. Theorem 1 suggests that the smoothed coefficient estimator from IS, $\tilde{\beta}_k$, is asymptotically equivalent to the BT estimator. However, simulation results show that the ultimate point estimator from IS has smaller mean squared errors than the BT estimator and thus has some finite sample advantages. In terms of variance estimation, IS performs similarly as BT-boot, and both provide reasonable variance estimates and confidence intervals with coverage probabilities close to the nominal level 95%. In contrast, the kernel smoothing is very sensitive to the choice of bandwidth h . Generally speaking, kernel smoothing tends to produce inflated variance estimation when more smoothing (larger h) is applied.

Estimation and inference are more challenging in cases with 50% censoring. When there is heavy right censoring, the upper quantiles may not be identifiable. Table 3 shows that for data sets with 50% censoring, all methods yield point estimates with larger bias especially at median, which is possibly due to the identifiability issue. Consequently, the confidence intervals from IS and BT-boot have coverage probabilities lower than 95% at $\tau = 0.5$. However, two methods still provide reasonable confidence intervals at the lower quantile $\tau = 0.25$.

In summary, the IS method is comparable to bootstrap in terms of variance estimation. However, IS is computationally more efficient than BT-boot. For example, in Scenario 1 with $n = 200$ and 15% censoring, it took IS 4.27 seconds and BT-boot 621 seconds to finish the computation for 100 data sets by using R 2.9.2 on a PC with an Intel i7 920 CPU clocked at 3.6Ghz. In this simulation study, we focused on data sets with 2 covariates. Additional simulation studies suggest that the proposed method also work well for data sets with more covariates at the cost of more computational time. For instance, the IS method required 134 seconds for analyzing 2000 simulated data sets with 6 covariates, compared to 94 seconds for analyzing 2000 datasets with 2 covariates.

5. Analysis of a Multiple Myeloma Data

In this section, we apply the proposed method to analyze the survival times of multiple myeloma patients. Multiple myeloma is a plasma cell cancer that happens in 4 per 10 million in the human population. Almost all multiple myeloma patients have genetic lesions in their cancer cells. Recently, to identify specific genetic lesions associated with prognosis, a genetic analysis of malignant plasma cells from 192 multiple myeloma patients was done using high density SNP arrays (Avet-Loiseau et al., 2009). After discarding those observations with missing values in the covariates, we have 170 observations with 54% censoring. We fit the quantile regression model to study the relation between the log survival times and 4 covariates, including serum Beta-2 microglobulin (B_2M) level, genetic

markers del(13) (deletion on chromosome 13), t(4; 14) (translocation between chromosome 4 and 14) and t(11; 14) (translocation between chromosome 11 and 14). We focus on two quantile levels $\tau = 0.25$ and 0.5 . In cancer studies, lower quantiles are of particular interest, as they correspond to high-risk cases with shorter survivals.

Table 3 summarizes the quantile coefficient estimates, the corresponding estimated standard errors, 95% confidence intervals and p -values obtained from three methods: IS, BT-boot and kernel smoothing with $h = n^{-0.3}$. In accordance with our simulation study, all three methods give similar coefficient estimates. All three methods suggest that del(13) and t(11;14) have no significant effect, while t(4;14) has a significant adverse effect on the two quantiles of the survival times, which is consistent with the existing results (Keats et al., 2003). In addition, the quantile coefficient of t(4; 14) is estimated as -1.062 at $\tau = 0.5$ and -1.037 at $\tau = 0.25$, indicating that the adverse effects of the genetic lesion t(4; 14) on patients survival are about the same at both quantiles.

The main discrepancy of three methods lies in the assessment of the B_2M effect. At $\tau = 0.5$, both IS and BT-boot suggest a marginal significance of B_2M , while the kernel smoothing method gives a much larger variance estimate, leading to an insignificance. Note that this data has heavy right censoring (54%). Simulation studies in Section 4 suggest that analysis at median could be unstable under such heavy censoring, and this may partly explain the different variance estimation of the kernel smoothing method. At the lower quantile $\tau = 0.25$, all three methods show that B_2M has a significantly adverse effect. In biomedical literatures, B_2M has been shown to be an important factor in multiple myeloma prognosis; see Avet-Loiseau et al. (2009), Greipp et al. (1993) and Facon et al. (2001). By studying at different quantiles, our analysis reveals a richer picture. The results of IS show that the estimated quantile coefficient of B_2M is -0.083 at $\tau = 0.5$ and is -0.134 at $\tau = 0.25$, suggesting that high B_2M level has an adverse effect on patients' survival times, and this adverse effect tends to be larger for those with shorter survivals, i.e. at the lower quantiles. Further biological investigations are needed to provide a more definite explanation of the differential impacts of B_2M on different locations of the survival distribution.

6. Conclusion and Remarks

We developed a new inference procedure for censored quantile regression based on induced smoothing, which is computationally fast and theoretically valid. In this paper, we focused on Bang-Tsatis estimator because of its computational convenience. The proposed idea can be extended to other censored quantile estimators. For example, for regression at the τ th quantile, we can extend the median regression approach of Ying et al. (1995) and consider the following estimating equation

$$n^{-1} \sum_{i=1}^n Z_i \left\{ \frac{I(Y_i - Z_i' \beta \geq 0)}{\widehat{G}(Z_i' \beta)} - 1 + \tau \right\} = 0.$$

Note that in the above equation, β is involved in both the numerator and the denominator, and thus the induced smoothing cannot be directly applied. However, the difficulty can be overcome if there exists a good initial estimator of β . For instance, we can use $\hat{\beta}$, the estimate obtained based on equation (2) as an initial estimate. Then an iterative estimation can proceed as follows: at each step, we plug the β estimate from the previous step into $\widehat{G}(Z_i' \beta)$, apply induced smoothing to $I(Y_i - Z_i' \beta \geq 0)$ and update both the coefficient and variance estimates. Further investigation is needed in this direction.

In this paper, we focused on survival data by making the unconditional independence assumption of survival and censoring times as in Ying et al. (1995). The method can be directly extended to analyze the censored medical cost data with informative censoring, the original context studied in Bang and Tsiatis (2002). In survival analysis, it is more practical to assume the conditional independence of survival and censoring times given covariates. To incorporate covariate-dependent censoring, we can adopt the local Kaplan-Meier estimates $\hat{G}(Y_i|Z_i)$ in constructing the inverse censoring probability weights in (2); see Wang and Wang (2009) for related discussions of the local Kaplan-Meier estimates. The corresponding variance estimation can be carried out following the same line as in the current paper.

The R implementation of the proposed method is available at <http://www4.stat.ncsu.edu/~wang/software.html>.

Acknowledgments

Lu's research was partly supported by NIH/NCI grant R01 CA-140632, and Wang's Research was partly supported by NSF grants DMS-0706963 and DMS-1007420. The authors would like to thank the editor, an associate editor, and two referees for their valuable suggestions.

References

- Avet-Loiseau H, Li C, Magrangeas F, Gouraud W, Charbonnel C, Harousseau JL, Attal M, Marit G, Mathiot C, Facon T, Moreau P, Anderson KC, Campion L, Munshi NC, Minvielle S. Prognostic significance of copy-number alterations in multiple myeloma. *Journal of Clinical Oncology*. 2009; 27:4585–4590. [PubMed: 19687334]
- Bang H, Tsiatis A. Median regression with censored cost data. *Biometrics*. 2002; 58:643–649. [PubMed: 12229999]
- Brown BM, Wang YG. Standard errors and covariance matrices for smoothed rank estimators. *Biometrika*. 2005; 92:149–158.
- Brown BM, Wang YG. Induced smoothing for rank regression with censored survival times. *Statistics in Medicine*. 2006; 26:828–836. [PubMed: 16646005]
- Chen C, Wei Y. Computational issues for quantile regression. *Journal of Computational and Graphical Statistics*. 2005; 67:399–417.
- Chen SX, Hall P. Smoothed empirical likelihood confidence intervals for quantiles. *The Annals of Statistics*. 1993; 21:1166–1181.
- Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972; 34:187–220.
- Cox DR, Oakes D. Analysis of survival data. *Monographs on Statistics and Applied Probability*. 1984:21.
- Facon T, Avet-Loiseau H, Guillermin G, Moreau P, Genevieve F, Zandecki M, Lai JL, Leleu X, Jouet JP, Bauters F, Harousseau JL, Bataille R, Mary JY. Chromosome 13 abnormalities identified by FISH analysis and serum beta2-microglobulin produce a powerful myeloma staging system for patients receiving high-dose therapy. *Blood*. 2001; 97:1566–1571. [PubMed: 11238092]
- Fleming, RT.; Harrington, PD. *Counting Processes and Survival Analysis*. Wiley-Interscience; 1991.
- Fu L, Wang YG, Bai ZD. Rank regression for analysis of clustered data: A natural induced smoothing approach. *Computational Statistics and data Analysis*. 2010; 54:1036–1050.
- Greipp P, Lust J, O'Fallon W, Katzmann J, Witzig T, Kyle R. Plasma cell labeling index and beta 2-microglobulin predict survival independent of thymidine kinase and C-reactive protein in multiple myeloma [see comments]. *Blood*. 1993; 81:3382–3387. [PubMed: 8507875]
- Heller G. Smoothed rank regression with censored data. *Journal of the American Statistical Association*. 2007; 102:552–559.
- Honor B, Khan S, Powell JL. Quantile regression under random censoring. *Journal of Econometrics*. 2002; 109:67–105.

- Horowitz JL. Bootstrap methods for covariance structures. *The Journal of Human Resources*. 1998; 33:39–61.
- Huang Y. Quantile calculus and censored regression. *Annals of Statistics*. 2010; 38:1607–1637. [PubMed: 20592942]
- Jin Z, Ying Z, Wei LJ. A simple resampling method by perturbing the minimand. *Biometrika*. 2001; 88:381–390.
- Johnson LM, Strawderman RL. Induced smoothing for the semiparametric accelerated failure time model: asymptotics and extensions to clustered data. *Biometrika*. 2009; 96:577–590. [PubMed: 23049117]
- Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, Belch AR, Pilarski LM. In multiple myeloma, t(4;14) (p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood*. 2003; 101:1520–1529. [PubMed: 12393535]
- Kocherginsky M, He X, Mu Y. Practical confidence intervals for regression quantiles. *Journal of Computational and Graphical Statistics*. 2005; 14:41–55.
- Koenker R, Bassett G. Regression quantiles. *Econometrica*. 1978; 46:33–50.
- Peng L, Huang Y. Survival analysis with quantile regression models. *Journal of the American Statistical Association*. 2008; 103:637–649.
- Portnoy S. Censored regression quantiles. *Journal of the American Statistical Association*. 2003; 98:1001–1012.
- Portnoy S, Lin G. Asymptotics for censored regression quantiles. *Journal of Nonparametric Statistics*. 2010; 22:115–130.
- Powell JL. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*. 1984; 25:303–325.
- Powell JL. Censored regression quantiles. *Journal of Econometrics*. 1986; 32:143–155.
- Shorack, GA.; Wellner, JA. *Empirical Processes With Applications to Statistics*. New York: Wiley; 1986.
- van der Vaart, A.; Wellner, JA. *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York: Springer-Verlag; 1996.
- Wang H, Wang L. Locally weighted censored quantile regression. *Journal of the American Statistical Association*. 2009; 104:1117–1128.
- Wang, YG.; Fu, L. Accounting for within cluster correlations in rank regression for the accelerated failure time model with censored data. 2010. unpublished manuscript
- Wang YG, Shao Q, Zhu M. Quantile regression without the curse of unsmoothness. *Computational Statistics and Data Analysis*. 2009; 53:3696–3705.
- Whang YJ. Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory*. 2006; 22:173–205.
- Ying Z, Jung SH, Wei LJ. Survival analysis with median regression models. *Journal of the American Statistical Association*. 1995; 90:178–184.

8. Appendix

To prove Theorem 1, we need the following Lemma.

Lemma 1

Let $\mathbf{W} = O(n^{-1})$ be any positive definite matrix, and define

$$\tilde{U}_n(\beta, \mathbf{W}) = n^{-1} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} \Phi\left(\frac{\beta_0' Z_i - Y_i}{\sqrt{Z_i' \mathbf{W} Z_i}}\right)$$

as the smoothed estimating function. Under Assumptions A1–A3, we have

$$\sup_{\|\beta - \beta_0\| \leq \varepsilon_n} \|n^{1/2}\{\tilde{U}_n(\beta, \mathbf{W}) - U_n(\beta)\}\| \xrightarrow{p} 0, \text{ as } n \rightarrow \infty,$$

where $\{\varepsilon_n\}$ is a positive sequence that converges to 0.

Proof of Lemma 1

Let $\sigma_i = (Z_i' \mathbf{W} Z_i)^{1/2}$, $\varepsilon_i^\beta = Y_i - Z_i' \beta$ and $d_i(\beta) = \text{sgn}(\varepsilon_i^\beta) \Phi(-|\varepsilon_i^\beta|/\sigma_i)$, where $\text{sgn}(\cdot)$ is the sign function. We have

$$\begin{aligned} n^{1/2}\{\tilde{U}_n(\beta, \mathbf{W}) - U_n(\beta)\} &= n^{-1/2} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} \left\{ \Phi\left(-\frac{\varepsilon_i^\beta}{\sigma_i}\right) - I(\varepsilon_i^\beta < 0) \right\} \\ &= n^{-1/2} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} d_i(\beta). \end{aligned}$$

Denote $D_n(\beta) = n^{-1/2} \sum_{i=1}^n \frac{Z_i \Delta_i}{\widehat{G}(Y_i)} d_i(\beta)$ and $D_n^G(\beta) = n^{-1/2} \sum_{i=1}^n \frac{Z_i \Delta_i}{G(Y_i)} d_i(\beta)$. It follows that

$$D_n(\beta) = D_n^G(\beta) - n^{-1/2} \sum_{i=1}^n \frac{Z_i \Delta_i \{\widehat{G}(Y_i) - G(Y_i)\}}{G^2(Y_i)} d_i(\beta) + o_p(1). \quad (5)$$

Taking the expectation of $D_n^G(\beta)$, we have

$$E\{D_n^G(\beta)\} = n^{-1/2} \sum_{i=1}^n Z_i E\{d_i(\beta)\},$$

where

$$\begin{aligned} E\{d_i(\beta)\} &= \int_{-\infty}^{+\infty} \text{sgn}(\varepsilon_i^\beta) \Phi(-|\varepsilon_i^\beta|/\sigma_i) f_i\{\varepsilon_i^\beta + Z_i'(\beta - \beta_0)\} d\varepsilon_i^\beta \\ &= \sigma_i \int_{-\infty}^{+\infty} \Phi(-|t|) \{2I(t > 0) - 1\} f_i\{\sigma_i t + Z_i'(\beta - \beta_0)\} dt \\ &= \sigma_i \int_{-\infty}^{+\infty} \Phi(-|t|) \{2I(t > 0) - 1\} [f_i\{Z_i'(\beta - \beta_0)\} + f_i'\{\omega_i^*(t)\} \sigma_i t] dt, \end{aligned}$$

where f_i is the density of $\varepsilon_i = \varepsilon_i^{\beta_0}$, and $\omega_i^*(t)$ is between $Z_i'(\beta - \beta_0)$ and $Z_i'(\beta - \beta_0) + \sigma_i t$. Notice that for β that satisfies $\|\beta - \beta_0\| \leq \varepsilon_n$, where $\varepsilon_n \rightarrow 0$, we have $|Z_i'(\beta - \beta_0)| \rightarrow 0$. It follows from Assumption A1 that $\sup_i f_i\{Z_i'(\beta - \beta_0)\} < \infty$, and since $\int_{-\infty}^{+\infty} \Phi(-|t|) \{2I(t > 0) - 1\} dt = 0$, we have $\int_{-\infty}^{+\infty} \Phi(-|t|) \{2I(t > 0) - 1\} f_i\{Z_i'(\beta - \beta_0)\} dt = 0$. In addition, by Assumption A1, we can find $M > 0$ such that $\sup_i |f_i'(\omega_i^*(t))| < M$. Thus, it follows that

$$|E\{d_i(\beta)\}| \leq \sigma_i^2 \int_{-\infty}^{+\infty} |t| \Phi(-|t|) |f'_{\beta,i}(\omega_i^*(t))| dt \leq M \sigma_i^2 / 2,$$

where the last equality holds because $\int_{-\infty}^{+\infty} |t| \Phi(-|t|) dt = 1/2$.

By Assumption A2 and the fact that $\mathbf{W} = O(n^{-1})$, $\sum_{i=1}^n \sigma_i^2 = \text{tr}(\mathbf{Z}\mathbf{W}\mathbf{Z}') = \text{tr}(\mathbf{W}\mathbf{Z}'\mathbf{Z})$ is bounded, and $\sum_{i=1}^n |E\{d_i(\beta)\}| \leq M \sum_{i=1}^n \sigma_i^2 / 2$ is also bounded. Therefore,

$$\|E\{D_n^G(\beta)\}\| \leq n^{-1/2} \sqrt{p} \sup_{i,j} |Z_{i,j}| \sum_{i=1}^n |E\{d_i(\beta)\}| \rightarrow 0, \text{ as } n \rightarrow \infty.$$

In addition, by Assumption A3,

$$\text{Var}\{D_n^G(\beta)\} = \frac{1}{n} \sum_{i=1}^n Z_i Z_i' \text{Var}\left\{\frac{\Delta_i}{G(Y_i)} d_i(\beta)\right\} \leq \frac{1}{n} \sum_{i=1}^n \frac{Z_i Z_i'}{v} E\{d_i^2(\beta)\},$$

where

$$\begin{aligned} E\{d_i^2(\beta)\} &= \int_{-\infty}^{+\infty} \Phi^2(-|s|) f_i\{\sigma_i s + Z_i'(\beta - \beta_0)\} d(\sigma_i s) \\ &= \int_{|s| > \Delta} \Phi^2(-|s|) f_i\{\sigma_i s + Z_i'(\beta - \beta_0)\} d(\sigma_i s) + \int_{|s| \leq \Delta} \Phi^2(-|s|) f_i\{\sigma_i s + Z_i'(\beta - \beta_0)\} d(\sigma_i s) \leq \Phi^2(-\Delta) + \sigma_i \int_{|s| \leq \Delta} f_i\{\sigma_i s + Z_i'(\beta - \beta_0)\} ds \\ &= \Phi^2(-\Delta) + 2\sigma_i \Delta f_i(\omega_i^*). \end{aligned}$$

Note that $\omega_i^* \in (Z_i'(\beta - \beta_0) - \sigma_i \Delta, Z_i'(\beta - \beta_0) + \sigma_i \Delta)$. Let $\Delta = n^{1/4}$ and since $\sigma_i = O(n^{-1/2})$, both $\sigma_i \Delta$ and ω_i^* go to 0 as n increases. As $f_i(\cdot)$ is uniformly bounded around zero, both $\Phi^2(-\Delta)$ and $\sigma_i \Delta f_i(\omega_i^*)$ go to 0 as $n \rightarrow \infty$.

Thus, it follows that $\lim_{n \rightarrow \infty} E\{d_i^2(\beta)\} = 0$, and $\lim_{n \rightarrow \infty} \text{Var}\{D_n^G(\beta)\} = 0$. By the Weak Law of Large Numbers, for β that satisfies $\|\beta - \beta_0\| = o_p(1)$, we have

$$\|D_n^G(\beta)\| \xrightarrow{p} 0, \text{ as } n \rightarrow \infty. \quad (6)$$

The second term on the right side of (5) can be written as

$$n^{-1/2} \sum_{j=1}^n \int_0^L \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_i Z_i d_i(\beta) Y_i(u)}{G(Y_i)} \right\} \frac{dM_j^c(u)}{y(u)} + o_p(1),$$

where $M_j^c(u) = N_j^c(u) - \int_0^u I(Y_i \geq s) d\Lambda^c(s)$, $N_j^c(u) = (1 - \delta_i) I(Y_i \leq u)$, $\Lambda^c(u) = -\log\{G(u)\}$ is the censoring cumulative hazard, $Y_{i\lambda}(u) = I(Y_i \geq u)$ is the λ th at-risk process, and

$y(u) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n Y_i(u)$ is bounded from below in $(0, L]$ by Assumption A3.

Define $I_n(u, \beta) = n^{-1} \sum_{i=1}^n \frac{\Delta_i Z_i d_i(\beta) Y_i(u)}{G(Y_i)}$, and $I(u, \beta) = E\{I_n(u, \beta)\}$. We have

$$I(u, \beta) = n^{-1} \sum_{i=1}^n Z_i E\{d_i(\beta) Y_i(u)\},$$

where $|E\{d_i(\beta) Y_i(u)\}| \leq E|d_i(\beta)|$, and

$$\begin{aligned} E|d_i(\beta)| &= \int_{-\infty}^{+\infty} \Phi\left(-\left|\frac{t}{\sigma_i}\right|\right) f_i\{\sigma_i t + Z_i'(\beta - \beta_0)\} d\mathcal{E}_i^\beta \\ &= \sigma_i \int_{-\infty}^{+\infty} \Phi(-|t|) f_i\{\sigma_i t + Z_i'(\beta - \beta_0)\} dt \\ &= \sigma_i f_i\{Z_i'(\beta - \beta_0)\} \int_{-\infty}^{+\infty} \Phi(-|t|) dt + \sigma_i^2 \int_{-\infty}^{+\infty} t \Phi(-|t|) f_i'\{\omega_i^*(t)\} dt. \end{aligned}$$

By Assumption A1, we have $f_i\{Z_i'(\beta - \beta_0)\} \int_{-\infty}^{+\infty} \Phi(-|t|) dt < \infty$, and $\int_{-\infty}^{+\infty} t \Phi(-|t|) f_i'\{\omega_i^*(t)\} dt < \infty$. Thus, it follows that $E|d_i(\beta)| = O(n^{-1/2})$, and

$$\|I(u, \beta)\| \leq \sqrt{p} \sup_{i,j} |Z_{ij}| n^{-1} \sum_{i=1}^n E|d_i(\beta)| = O(n^{-1/2}) \rightarrow 0.$$

Define

$$\mathcal{F} = \frac{\Delta_i Z_i d_i(\beta) Y_i(u)}{G(Y_i)}$$

, $\|\beta - \beta_0\| \leq \varepsilon_n$ and $u \in (0, \infty)$. The function class $\mathcal{F}$ is Glivenko-Cantelli (van der Vaart and Wellner, 1996) because the class of indicator functions is Glivenko-Cantelli, and Z_i , $d_i(\beta)$, and $1/G(Y_i)$ are uniformly bounded. It follows that

$$\sup_{\|\beta - \beta_0\| \leq \varepsilon_n, u \in (0, \infty)} \|I_n(u, \beta) - I(u, \beta)\| \xrightarrow{a.s.} 0, \text{ and we have}$$

$$n^{-1/2} \sum_{j=1}^n \int_0^L I_n(u, \beta) \frac{dM_j^c(u)}{y(u)} = n^{-1/2} \sum_{j=1}^n \int_0^L I(u, \beta) \frac{dM_j^c(u)}{y(u)} + o_p(1).$$

By the Martingale Central Limit Theorem (Fleming and Harrington, 1991),

$n^{-1/2} \sum_{j=1}^n \int_0^L I(u, \beta) \frac{dM_j^c(u)}{y(u)}$ is $o_p(1)$ as n goes to infinity. It follows that, for β that satisfies $\|\beta - \beta_0\| \leq \varepsilon_n$

$$\left| n^{-1/2} \sum_{j=1}^n \int_0^L I_n(u, \beta) \frac{dM_j^c(u)}{y(u)} \right| \xrightarrow{p} 0. \quad (7)$$

Collating (6) and (7), we have

$$\left\| n^{1/2} \{ \tilde{U}_n(\beta, \mathbf{W}) - U_n(\beta) \} \right\| \xrightarrow{p} 0,$$

for any β such that $\|\beta - \beta_0\| \leq \varepsilon_n$. Lemma 1 is thus proven by the fact that both $\tilde{U}_n(\beta, \mathbf{W})$ and $U_n(\beta)$ are monotone functions, thus the point-wise convergence could be strengthened to uniform convergence (Shorack and Wellner, 1986).

Proof of Theorem 1 (i)

Consider the one-step estimator of β_0 ,

$$\tilde{\beta}_1 = \tilde{\beta}_0 + \{-\tilde{\mathbf{A}}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0)\}^{-1} \tilde{U}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0),$$

where $\tilde{\beta}_0 = \hat{\beta}$ is the unsmoothed estimator.

To show the asymptotic equivalency of $\tilde{\beta}_1$ and $\tilde{\beta}_0$, we first prove that for any positive

definite matrix $\mathbf{W} = O(n^{-1})$, $\mathbf{W} = O(n^{-1})$, $\tilde{\mathbf{A}}_n(\beta_0, \mathbf{W}) \xrightarrow{p} \mathbf{A}$. Define $\sigma_i = \sqrt{Z_i' \mathbf{W} Z_i}$, $\varepsilon_i = Y_i - Z_i' \beta_0$, and $\eta_i = \varphi(-\varepsilon_i / \sigma_i) / \sigma_i$. Then

$$\tilde{\mathbf{A}}_n(\beta_0, \mathbf{W}) = n^{-1} \sum_{i=1}^n Z_i Z_i' \frac{\Delta_i}{\widehat{G}(Y_i)} \eta_i.$$

For any vectors $a, b \in R^p$,

$$a' \tilde{\mathbf{A}}_n(\beta_0, \mathbf{W}) b = a' \left\{ n^{-1} \sum_{i=1}^n Z_i Z_i' \frac{\Delta_i}{\widehat{G}(Y_i)} \eta_i \right\} b + o_p(1),$$

where

$$E \left[a' \left\{ n^{-1} \sum_{i=1}^n Z_i Z_i' \frac{\Delta_i}{\widehat{G}(Y_i)} \eta_i \right\} b \right] = a' \left\{ n^{-1} \sum_{i=1}^n Z_i Z_i' E(\eta_i) \right\} b,$$

and

$$\begin{aligned} E(\eta_i) &= \frac{1}{\sigma_i} \int \varphi\left(-\frac{\varepsilon_i}{\sigma_i}\right) f_i(\varepsilon_i) d\varepsilon_i = \int \varphi(-t) f_i(\sigma_i t) dt \\ &= \int \varphi(-t) f_i(0) dt + \sigma_i \int t \varphi(-t) f_i'(\omega_i^*) dt, \end{aligned}$$

where $\int \varphi(-t) f_i(0) dt = f_i(0)$, and $\sigma_i \int t \varphi(-t) f_i'(\omega_i^*) dt \leq M \sigma_i \int |t| \varphi(-t) dt \rightarrow 0$.

It follows that $E(\eta_i) \rightarrow f_i(0)$ and

$$\lim_{n \rightarrow \infty} E \left[a' \left\{ n^{-1} \sum_{i=1}^n Z_i Z_i' \frac{\Delta_i}{G(Y_i)} \eta_i \right\} b \right] = a' \left\{ \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n Z_i Z_i' f_i(0) \right\} b = a' \mathbf{A} b. \quad (8)$$

By Assumption A1,

$$\begin{aligned} E(\eta_i^2) &= \frac{1}{\sigma_i^2} \int \varphi^2(-\frac{\varepsilon_i}{\sigma_i}) f_i(\varepsilon_i) d\varepsilon_i = \frac{1}{\sigma_i^2} \int \varphi^2(-t) f_i(\sigma_i t) dt \\ &= \frac{1}{\sigma_i^2} \int \varphi^2(-t) \{f_i(0) + f_i'(\omega_i^*) \sigma_i t\} dt \\ &= \frac{1}{\sigma_i^2} \int \varphi^2(-t) f_i(0) dt + \int t \varphi^2(-t) f_i'(\omega_i^*) dt \\ &= O(n^{1/2}). \end{aligned}$$

Consequently,

$$\text{Var} \left[a' \left\{ n^{-1} \sum_{i=1}^n Z_i Z_i' \frac{\Delta_i}{G(Y_i)} \eta_i \right\} b \right] \leq \frac{1}{n^2 v} \sum_{i=1}^n (a' Z_i Z_i' b)^2 E(\eta_i^2) \rightarrow 0. \quad (9)$$

By (8) and (9), we have $\tilde{\mathbf{A}}_n(\beta_0, \mathbf{W}) \xrightarrow{p} \mathbf{A}$. Thus, $\tilde{\mathbf{A}}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0) \xrightarrow{p} \mathbf{A}$ because $\tilde{\beta}_0 \xrightarrow{p} \beta_0$ (Bang and Tsiatis, 2002). By lemma 1, $n^{1/2} \tilde{U}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0) = n^{1/2} U_n(\tilde{\beta}_0) + o_p(1) = o_p(1)$.

Now consider

$$n^{1/2}(\tilde{\beta}_1 - \beta_0) = n^{1/2}(\tilde{\beta}_0 - \beta_0) + \{-\tilde{\mathbf{A}}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0)\}^{-1} \{n^{1/2} \tilde{U}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0)\},$$

where $\tilde{\mathbf{A}}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0) \xrightarrow{p} \mathbf{A}$ and $n^{1/2} \tilde{U}_n(\tilde{\beta}_0, \tilde{\mathbf{H}}_0) \xrightarrow{p} 0$. It follows that $n^{1/2}(\tilde{\beta}_1 - \beta_0) \xrightarrow{p} n^{1/2}(\tilde{\beta}_0 - \beta_0)$, which together with the asymptotic normality of $\tilde{\beta}_0$ (Bang and Tsiatis, 2002) gives

$$n^{1/2}(\tilde{\beta}_1 - \beta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Gamma}).$$

The proof for the asymptotic properties of $\tilde{\beta}_k$ with $k > 1$ follows with similar arguments.

Proof of Theorem 1 (ii)

We define

$$\tilde{\mathbf{\Gamma}}_k = \{\tilde{\mathbf{A}}_n(\tilde{\beta}_k, \tilde{\mathbf{H}}_{k-1})\}^{-1} \sum \widehat{(\tilde{\beta}_k)} \{\tilde{\mathbf{A}}_n(\tilde{\beta}_k, \tilde{\mathbf{H}}_{k-1})\}^{-1},$$

where

$$\begin{aligned}
\widehat{\Sigma}(\beta) = & \frac{1}{4n} \sum_{i=1}^n \frac{\Delta_i}{\widehat{G}(Y_i)} Z_i Z_i' \\
& - \left[n^{-1} \sum_{i=1}^n \frac{\Delta_i}{\widehat{G}(Y_i)} \{I(Y_i - Z_i' \beta \leq 0) - \tau\} Z_i \right]^{\otimes 2} \\
& + n^{-1} \int \frac{dN^c(u)}{\widehat{G}^2(Y_i)} [\widehat{K}(\{I(Y_i - Z_i' \beta \leq 0) - \tau\} Z_i)^{\otimes 2}, u] - \widehat{K}^{\otimes 2}(\{I(Y_i - Z_i' \beta \leq 0) - \tau\} Z_i, u).
\end{aligned} \tag{10}$$

Here $\widehat{K}(\mathbf{W}, u) = n^{-1} \widehat{S}^{-1}(u) \sum_{i=1}^n \frac{\Delta_i \mathbf{W}_i I(Y_i \geq u)}{\widehat{G}(Y_i)}$, $a \otimes 2 = aa'$ for vector a , and $S(t)$ is the survival function for the failure time T . Since $\tilde{\beta}_k \xrightarrow{p} \beta_0$ for any k , and $\widehat{\Sigma}(\beta_0) \xrightarrow{p} \Sigma$ by Bang and Tsiatis (2002), it follows that $\widehat{\Sigma}(\tilde{\beta}_k) \xrightarrow{p} \Sigma$. In addition, $\tilde{\mathbf{H}}_{k-1} = O(n^{-1})$, thus we have $\tilde{\mathbf{A}}_n(\tilde{\beta}_k, \tilde{\mathbf{H}}_{k-1}) \xrightarrow{p} \mathbf{A}$, and consequently, $\tilde{\Gamma}_k \xrightarrow{p} \mathbf{A}^{-1} \Sigma \mathbf{A}^{-1} = \Gamma$.

Table 1

Simulation results for 15% censoring, S_1 , S_2 and S_3 are three kernel smoothing methods with bandwidth $h = n^{-0.1}$, $n^{-0.3}$ and $n^{-0.7}$, respectively. Bias and MSE are reported as the original numbers times 1000. SE/SD: SD is the Monte Carlo standard deviation and SE is the mean of estimated standard errors. CP is the empirical coverage probability of the 95% Wald-type confidence intervals. The standard errors of MSEs range from 1 to 5, and the standard errors of SE/SD range from 0.002 to 0.006.

	β_1				β_2				β_3			
	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
$\tau = 0.25$, Scenario 1												
IS	-11	42	1.08	0.947	3	7	1.08	0.951	3	35	1.08	0.953
BT-boot	-4	48	1.08	0.953	4	8	1.07	0.952	4	40	1.10	0.960
S_1	-106	46	1.34	0.980	3	5	1.33	0.986	4	29	1.34	0.991
S_2	-16	42	1.13	0.957	4	7	1.12	0.958	4	34	1.14	0.969
S_3	-4	47	0.98	0.844	4	8	1.00	0.847	3	39	1.01	0.873
$\tau = 0.25$, Scenario 2												
IS	-24	73	1.09	0.952	9	16	1.07	0.950	6	80	1.08	0.955
BT-boot	-12	83	1.10	0.955	10	19	1.07	0.955	6	91	1.09	0.955
S_1	-105	75	1.26	0.975	21	15	1.20	0.970	7	71	1.23	0.984
S_2	-22	74	1.12	0.952	11	17	1.08	0.943	7	81	1.11	0.957
S_3	-12	82	0.90	0.807	10	19	0.90	0.789	6	90	0.90	0.830
$\tau = 0.25$, Scenario 3												
IS	-23	66	1.08	0.946	1	15	1.06	0.951	1	77	1.04	0.948
BT-boot	-2	73	1.11	0.958	-1	17	1.08	0.956	-3	87	1.06	0.955
S_1	-126	74	1.25	0.973	16	14	1.20	0.978	4	69	1.19	0.981
S_2	-20	67	1.12	0.951	2	15	1.09	0.942	1	78	1.08	0.952
S_3	-3	73	0.94	0.814	-1	17	0.93	0.793	-2	86	0.90	0.832
$\tau = 0.5$, Scenario 1												
IS	-6	38	1.06	0.945	4	6	1.08	0.955	3	33	1.06	0.947
BT-boot	-5	45	1.04	0.946	4	7	1.06	0.956	3	39	1.04	0.949
S_1	-5	31	1.38	0.992	4	5	1.37	0.992	1	27	1.36	0.994

	β_1				β_2				β_3			
	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
S_2	-5	38	1.11	0.959	3	6	1.12	0.961	2	33	1.11	0.958
S_3	-5	44	0.92	0.842	4	7	1.00	0.856	4	38	0.97	0.868
$\tau = 0.5$, Scenario 2												
IS	-8	67	1.08	0.946	5	15	1.07	0.949	5	77	1.06	0.950
BT-boot	-8	78	1.06	0.949	5	18	1.06	0.954	8	88	1.06	0.951
S_1	-7	60	1.27	0.984	5	14	1.21	0.976	2	69	1.22	0.981
S_2	-8	69	1.09	0.945	5	16	1.09	0.947	5	79	1.09	0.953
S_3	-7	77	0.91	0.811	5	17	0.93	0.798	7	87	0.92	0.840
$\tau = 0.5$, Scenario 3												
IS	-8	54	1.07	0.950	1	12	1.06	0.946	10	63	1.04	0.943
BT-boot	-11	61	1.07	0.954	3	14	1.06	0.956	10	72	1.04	0.948
S_1	-6	49	1.29	0.987	1	11	1.24	0.981	7	59	1.21	0.982
S_2	-8	55	1.10	0.954	1	13	1.08	0.942	9	64	1.07	0.951
S_3	-11	61	0.95	0.831	2	14	0.94	0.807	10	71	0.93	0.843

Table 2

Simulation results for 30% censoring, S_1 , S_2 and S_3 are three kernel smoothing methods with bandwidth $h = n^{-0.1}$, $n^{-0.3}$ and $n^{-0.7}$, respectively. Bias and MSE are reported as the original numbers times 1000. SE/SD: SD is the Monte Carlo standard deviation and SE is the mean of estimated standard errors. CP is the empirical coverage probability of the 95% Wald-type confidence intervals. The standard errors of MSEs range from 1 to 6, and the standard errors of SE/SD range from 0.003 to 0.008.

	β_1				β_2				β_3			
	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
$\tau = 0.25$, Scenario 1												
IS	-9	47	1.07	0.947	2	8	1.07	0.950	0	41	1.08	0.953
BT-boot	-2	55	1.08	0.955	4	9	1.06	0.956	1	48	1.09	0.954
S_1	-102	49	1.32	0.979	2	6	1.32	0.985	1	34	1.33	0.990
S_2	-13	47	1.12	0.957	3	8	1.11	0.956	1	41	1.13	0.965
S_3	-2	54	0.95	0.826	4	9	0.97	0.816	1	47	0.97	0.854
$\tau = 0.25$, Scenario 2												
IS	-14	80	1.08	0.955	2	19	1.05	0.940	-2	92	1.08	0.954
BT-boot	1	94	1.09	0.959	1	23	1.05	0.948	2	105	1.09	0.960
S_1	-92	81	1.24	0.977	13	18	1.17	0.968	-2	84	1.22	0.980
S_2	-11	83	1.10	0.953	3	21	1.04	0.932	0	95	1.10	0.953
S_3	0	93	0.88	0.791	2	23	0.85	0.737	2	104	0.88	0.802
$\tau = 0.25$, Scenario 3												
IS	-19	73	1.08	0.948	-6	17	1.07	0.943	-1	89	1.04	0.946
BT-boot	1	82	1.11	0.961	-7	20	1.09	0.952	-4	103	1.06	0.950
S_1	-119	80	1.24	0.978	9	16	1.20	0.974	1	81	1.18	0.980
S_2	-14	74	1.11	0.949	-5	18	1.09	0.932	-1	91	1.07	0.947
S_3	1	81	0.92	0.801	-7	19	0.90	0.760	-4	101	0.85	0.795
$\tau = 0.5$, Scenario 1												
IS	1	44	1.07	0.946	1	7	1.08	0.953	-3	42	1.06	0.949
BT-boot	2	51	1.06	0.947	1	8	1.08	0.956	-1	48	1.05	0.950
S_1	2	37	1.37	0.991	1	6	1.37	0.991	-5	35	1.34	0.992

	β_1					β_2					β_3				
	Bias	MSE	SE/SD	CovP		Bias	MSE	SE/SD	CovP		Bias	MSE	SE/SD	CovP	
S_2	2	44	1.11	0.957	1	7	1.12	0.962	-3	41	1.10	0.959			
S_3	2	50	0.96	0.832	1	8	1.00	0.837	-1	47	0.94	0.854			
$\tau = 0.5$, Scenario 2															
IS	14	80	1.08	0.947	-8	19	1.06	0.944	-10	101	1.05	0.946			
BT-boot	12	93	1.08	0.954	-7	22	1.07	0.953	-8	115	1.05	0.957			
S_1	15	74	1.25	0.980	-9	18	1.18	0.971	-13	93	1.19	0.980			
S_2	13	83	1.09	0.948	-7	20	1.07	0.940	-9	105	1.07	0.946			
S_3	12	92	0.94	0.795	-7	22	0.93	0.763	-8	114	0.87	0.816			
$\tau = 0.5$, Scenario 3															
IS	5	66	1.07	0.949	-10	16	1.04	0.939	4	82	1.03	0.943			
BT-boot	3	75	1.07	0.954	-9	18	1.05	0.946	4	92	1.04	0.954			
S_1	7	62	1.26	0.986	-11	15	1.20	0.978	1	78	1.19	0.979			
S_2	5	68	1.08	0.947	-10	17	1.05	0.930	4	84	1.05	0.946			
S_3	3	74	0.95	0.817	-9	18	0.92	0.769	5	91	0.91	0.830			

Table 3

Simulation results for 50% censoring, S_1 , S_2 and S_3 are three kernel smoothing methods with bandwidth $h = n^{-0.1}$, $n^{-0.3}$ and $n^{-0.7}$, respectively. Bias and MSE are reported as the original numbers times 1000. SE/SD: SD is the Monte Carlo standard deviation and SE is the mean of estimated standard errors. CP is the empirical coverage probability of the 95% Wald-type confidence intervals. The standard errors of MSEs range from 1 to 6, and the standard errors of SE/SD range from 0.002 to 0.005.

	β_1				β_2				β_3			
	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
$\tau = 0.25$, Scenario 1												
IS	21	61	1.05	0.937	-18	12	1.04	0.933	-28	62	1.05	0.945
BT-boot	29	73	1.06	0.954	-16	14	1.06	0.947	-26	73	1.07	0.954
S_1	-63	54	1.29	0.984	-22	10	1.26	0.978	-30	54	1.27	0.986
S_2	20	62	1.08	0.948	-17	12	1.06	0.934	-26	62	1.09	0.951
S_3	29	72	0.91	0.805	-15	14	0.90	0.757	-26	71	0.86	0.809
$\tau = 0.25$, Scenario 2												
IS	95	112	1.04	0.921	-98	37	1.00	0.880	-61	137	1.02	0.938
BT-boot	110	131	1.07	0.934	-95	42	1.04	0.899	-55	154	1.06	0.957
S_1	25	97	1.17	0.971	-90	35	1.09	0.924	-63	132	1.13	0.969
S_2	101	119	1.05	0.917	-95	39	1.00	0.869	-58	143	1.04	0.934
S_3	110	130	0.82	0.753	-95	42	0.80	0.646	-56	153	0.79	0.768
$\tau = 0.25$, Scenario 3												
IS	75	96	1.05	0.923	-93	33	1.03	0.883	-57	130	1.00	0.933
BT-boot	97	114	1.10	0.946	-91	36	1.09	0.914	-58	144	1.06	0.952
S_1	-14	85	1.19	0.976	-82	30	1.14	0.932	-58	124	1.11	0.965
S_2	81	100	1.08	0.922	-89	33	1.04	0.872	-55	133	1.03	0.937
S_3	97	112	0.84	0.739	-91	36	0.82	0.668	-57	142	0.76	0.750
$\tau = 0.5$, Scenario 1												
IS	60	64	1.06	0.943	-34	13	1.03	0.928	-45	71	1.03	0.940
BT-boot	59	77	1.06	0.954	-32	16	1.03	0.937	-44	85	1.03	0.948
S_1	67	57	1.35	0.991	-40	12	1.30	0.975	-53	65	1.27	0.984

	β_1					β_2					β_3					
	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
S_2	60	67	1.08	0.952	-34	14	1.04	0.927	-46	75	1.04	0.948	-46	75	1.04	0.948
S_3	59	76	0.92	0.802	-32	16	0.88	0.744	-44	84	0.84	0.799	-44	84	0.84	0.799
$\tau = 0.5$, Scenario 2																
IS	201	161	1.00	0.907	-167	61	0.95	0.776	-101	183	0.98	0.931	-101	183	0.98	0.931
BT-boot	191	179	1.03	0.921	-160	66	0.98	0.830	-99	207	1.01	0.945	-99	207	1.01	0.945
S_1	210	158	1.14	0.948	-174	62	1.05	0.830	-110	181	1.07	0.960	-110	181	1.07	0.960
S_2	195	166	1.00	0.898	-162	63	0.93	0.766	-101	193	0.99	0.923	-101	193	0.99	0.923
S_3	191	177	0.79	0.708	-160	65	0.71	0.546	-99	205	0.72	0.724	-99	205	0.72	0.724
$\tau = 0.5$, Scenario 3																
IS	164	126	1.00	0.911	-143	47	0.96	0.800	-78	147	0.96	0.927	-78	147	0.96	0.927
BT-boot	157	142	1.02	0.924	-136	51	1.00	0.845	-78	164	1.00	0.938	-78	164	1.00	0.938
S_1	176	125	1.18	0.962	-152	49	1.10	0.858	-89	146	1.09	0.964	-89	146	1.09	0.964
S_2	159	130	1.00	0.899	-139	49	0.95	0.777	-78	153	0.98	0.922	-78	153	0.98	0.922
S_3	157	140	0.82	0.729	-137	50	0.75	0.574	-78	163	0.75	0.728	-78	163	0.75	0.728

Table 4

Results of the multiple myeloma data analysis. For each method, $\hat{\beta}$ denotes the quantile coefficient estimate, lb and ub denote the lower and the upper bounds of a 95% Wald-type confidence interval, and SE is the estimated standard error.

	$\tau=0.25$				$\tau=0.5$					
	$\hat{\beta}$	lb	ub	SE	p -value	$\hat{\beta}$	lb	ub	SE	p -value
IS										
intercept	3.805	3.346	4.264	0.234	0.000	3.980	3.462	4.498	0.264	0.000
$B_2 M$	-0.134	-0.238	-0.030	0.053	0.011	-0.083	-0.163	-0.002	0.041	0.045
del(13)	0.087	-0.552	0.726	0.326	0.789	0.256	-0.133	0.646	0.199	0.197
$\kappa(4; 14)$	-1.037	-1.646	-0.429	0.311	0.001	-1.062	-1.655	-0.470	0.302	0.000
$\kappa(11; 14)$	-0.497	-1.421	0.426	0.471	0.291	-0.153	-0.680	0.374	0.269	0.569
BT-boot										
intercept	3.867	3.268	4.467	0.306	0.000	4.049	3.475	4.623	0.293	0.000
$B_2 M$	-0.137	-0.220	-0.055	0.042	0.001	-0.081	-0.170	0.007	0.045	0.072
del(13)	0.327	-0.451	1.105	0.397	0.410	0.252	-0.174	0.678	0.218	0.247
$\kappa(4; 14)$	-1.176	-1.921	-0.430	0.380	0.002	-1.349	-2.046	-0.653	0.355	0.000
$\kappa(11; 14)$	-0.503	-1.648	0.641	0.584	0.389	-0.233	-0.876	0.411	0.328	0.478
Kernel smoothing with $h= n^{-0.3}$										
intercept	3.803	3.416	4.190	0.198	0.000	3.918	3.093	4.743	0.421	0.000
$B_2 M$	-0.137	-0.158	-0.115	0.011	0.000	-0.071	-0.247	0.105	0.090	0.431
del(13)	0.269	-0.339	0.877	0.310	0.386	0.241	-0.165	0.647	0.207	0.245
$\kappa(4; 14)$	-1.139	-1.704	-0.573	0.289	0.000	-1.138	-1.697	-0.579	0.285	0.000
$\kappa(11; 14)$	-0.447	-1.540	0.646	0.558	0.423	-0.119	-0.643	0.405	0.267	0.657

Table 5

Simulation results at median for models with 6 covariates and 15% or 30% censoring. The “Cen” stands for the censoring proportion. The “See” stands for the error scenario. The β_k , $k = 1, \dots$ are the 7 median coefficients.

Cen	See	β_1				β_2				β_3			
		Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
15%	1	16	887	1.11	0.960	-2	93	1.11	0.961	-9	288	1.09	0.960
	2	16	887	1.11	0.960	-2	93	1.11	0.961	-9	288	1.09	0.960
	3	26	1638	1.12	0.964	-7	196	1.10	0.960	-10	577	1.09	0.962
30%	1	140	2257	1.09	0.958	-50	267	1.09	0.957	-36	794	1.08	0.957
	2	6	1390	1.08	0.957	-11	164	1.08	0.956	-15	470	1.09	0.963
	3	97	1852	1.07	0.965	-62	221	1.08	0.958	-12	649	1.06	0.961
		β_4				β_5				β_6			
		Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP	Bias	MSE	SE/SD	CovP
15%	1	10	94	1.11	0.960	-18	306	1.08	0.957	-1	94	1.11	0.961
	2	10	94	1.11	0.960	-18	306	1.08	0.957	-1	94	1.11	0.961
	3	16	188	1.11	0.962	-25	618	1.08	0.960	-4	191	1.11	0.966
30%	1	-10	270	1.07	0.950	-49	854	1.06	0.956	-20	260	1.09	0.958
	2	-5	156	1.09	0.966	5	493	1.08	0.960	7	158	1.08	0.958
	3	-24	216	1.07	0.961	-8	647	1.08	0.963	-16	219	1.06	0.951
		β_7											
		Bias	MSE	SE/SD	CovP								
15%	1	-14	290	1.11	0.970								
	2	-14	290	1.11	0.970								
	3	-22	583	1.11	0.967								
30%	1	-30	767	1.13	0.964								
	2	38	492	1.09	0.959								
	3	35	706	1.05	0.963								