Statistics and Its Interface Volume 14 (2021) 183–195

Estimating equation estimators of quantile differences for one sample with length-biased and right-censored data

Li Xun, Guangchao Zhang, Dehui Wang, and Yong Zhou*

This paper estimates quantile differences for one sample with length-biased and right-censored (LBRC) data. To ensure the asymptotic unbiasedness of the estimator, the estimating equation method is adopted. To improve the efficiency of the estimator, in the sense of having a lower mean squared error, the kernel-smoothed approach is employed. To make full use of the features of LBRC data, the augmented inverse probability complete case weight is investigated in detail. Moreover, the consistency and asymptotic normality of the proposed estimators are established. The numerical simulations are conducted to examine the performance of the estimators.

Keywords and phrases: Quantile difference, Length bias, Informative censoring, Estimating equation, Kernel function.

1. INTRODUCTION

For one population, its quantile difference, denoted by θ_{pq}, is the difference between its two quantiles ξ_p and ξ_q for $0 < q < p < 1$. The quantile difference describes a range of the values of the population corresponding to the probability pair (p, q). Thanks to the complete flexibility of the pair (p, q), quantile differences are versatile in the sense that it encompasses a variant of probability levels. For example, the quantile differences express characteristics of the right tail of the distribution with q large, and those of the left tail of the distribution with p small.

For one sample, its quantile difference, denoted by $\hat{\theta}_{pq}$, can be interpreted as a measure of the statistical dispersion of the observations. For example, taking (p, q) to be (0.75, 0.25), it retrieves the interquartile range (25% trimmed range). Taking (p, q) to be (0.90, 0.10), it becomes the interdecile range (10% trimmed range). Taking (p, q) to be (0.93, 0.07), it generalizes the 7% trimmed range. Taking p close to 1 and q close to 0, $\hat{\theta}_{pq}$ represents the sample range.

*Corresponding author. Yong Zhou, Key Laboratory of Advanced Theory and Application in Statistics and Data Science, Ministry of Education, and Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai 200062, China. E-mail address: yzhou@amss.ac.cn.

It is worthy noting that quantile differences for one sample are more robust than the sample standard deviation which is greatly influenced by outliers. For one population, quantile differences always exist regardless of the types of distributions.

As a robust measure of scale, quantile differences for one population deserve an investigation in great detail. From a statistical viewpoint, a natural estimator of the quantile difference is the difference between two sample quantiles, like spacing, see for example, Chapters 6 and 8 in [16]. In particular, if ξ_q denotes an observing time point of a life population, then θ_{pq} is said to be the quantile residual life. As a special case of quantile differences, quantile residual lifetime has been studied widely, see for instance, Chapter 3 in [6]. Actually, quantile differences are receiving increasing attention in the recent literature, including [26, 34, 19, 7, 33, 11, 27, 31, 30] among others.

The majority of the works above were motivated by different data types, such as the complete data, right-censored (RC) data, and left-truncated and right-censored (LTRC) data. To date, however, little scholarly work has been done on the estimators of quantile differences for one sample with length-biased and right-censored (LBRC) data.

LBRC data are popular in many practical applications, such as the studies of prevalent cohort, cancer screening trials, and labor economics. These studies can be found in the literature, see for instance, [20, 23, 24, 25, 2, 3, 32], and so on.

For LBRC data, there are some especially noteworthy features. On the one hand, the probability of being observed for a variable of interest is proportional to its length, such as the duration of unemployment. That means LBRC data type is a subtype of LTRC data. It is worth mentioning that the truncation variable under LBRC data follows a uniform distribution because of an assumption of stationarity, see for example [28, 5]. Meanwhile, the residual lifetime has just the same distribution as the truncation variable. That provides more information to the statistical inference. On the other hand, LBRC data type is a subtype of RC data. Particularly, the censoring under LBRC data is informative because the censoring variable shares the common truncation variable with the failure time. Therefore, the traditional nonparametric methods for the independent censoring aren't suitable to LBRC data. In a word, it is obvious that making full use of the aforementioned characteristics should contribute significantly to the statistical inference.

For these reasons, it is important to utilize the features of LBRC data to estimate quantile differences. Because of the complexity of LBRC data, it is also significant to efficiently and effectively estimate quantile differences for one population. Motivated by these, we pursue the estimating equation estimators of quantile differences and establish the consistency and asymptotic normality of these estimators under LBRC data.

The numerical simulations are conducted to examine the performance of these estimators. In the sense of having a lower mean squared error, the smoothed estimating equation estimator is more efficient than the nonsmooth estimating equation estimators. That profits from the contribution of a kernel smoother. In the sense of having a lower asymptotic variance, the augmented inverse probability weighted complete-case estimating equation estimator is the most effective. It benefits from the features of LBRC data.

The rest of the paper is organized as follows. In Section 2, we introduce the notations, describe the estimating methods, and establish the asymptotic properties of the estimators of quantile differences. In Section 3, we conduct the simulations to examine the performance of these estimators, and apply the proposed methods to a real example. Finally, we delay the proofs of the theorems in Appendix A.

2. MODEL AND ESTIMATION

Let $\tilde{X}$ be a variable of interest with the unknown distribution function $F(\cdot)$, the density function $f(\cdot)$, and the survival function $S(\cdot)$. Denote by ξ_p the p-quantile of the distribution function $F(\cdot)$, then $\xi_p = F^{-1}(p) = \inf\{x : F(x) \geq p\}$ for $0 < p < 1$, and the quantile difference is

$$\theta_{pq} = F^{-1}(p) - F^{-1}(q) = \xi_p - \xi_q,$$

for $0 < q < p < 1$.

In a prevalent cohort study, $\tilde{X}$ represents the lifetime of the subjects. In fact, the failure events before the examination time are not observed. Meanwhile, at the end of the experiment some subjects are still alive, or in the process of the experiment some individuals drop out because of other unrelated events. Therefore, the observed sample is subject to a left truncation variable A and a right censoring variable C. Moreover, if the disease incidence is stationary over time, then A follows a uniform distribution, and the sampling generates a length-biased data set.

Denote by T the residual lifetime. The length-biased time $X^0 = A+T$ is obtained only when $\tilde{X} > A$. Then the density function of the length-biased variable X^0 is given by

$$f_{X^0}(x) = \frac{xf(x)}{\mu}, \tag{1}$$

where $\mu = \int_0^\infty xf(x)dx$.

Let $\delta = I(T \leq C)$ be the censoring indicator, and $X = \min\{X^0, A+C\}$. Assume that C is independent of (A, T), and the observed sample is independent and identically distributed triples $(X_i,\ A_i,\ \delta_i)$, $i = 1, \cdots, n$, where $X_i = \min\{X_i^0,\ A_i + C_i\}$ and $\delta_i = I(X_i^0 \leq A_i + C_i)$.

Under the above-mentioned LBRC data, we estimate θ_{pq} by constructing the following systems of estimating equations.

2.1 Estimating equation

Denote by $S_T(\cdot)$ and $S_C(\cdot)$ the survival functions of T and C, respectively, by $\Lambda_C(\cdot)$ the cumulative hazard function of C, and by $\lambda_C(\cdot)$ the derivative of $\Lambda_C(\cdot)$.

To construct estimating equations of the quantile difference θ_{pq}, we use the probability of observing the length-biased failure time during the period $(t, t+dt)$ as presented by [14], that is,

$$\begin{aligned} & Pr\{X \in (t, t+dt),\ A \in (a, a+da),\ \delta = 1\} \\ = & \frac{f(t)S_C(t-a)}{\mu} dadt. \end{aligned}$$

For the true value $\mathbf{\Delta}_0 = (\xi_p, \theta_0)^T$ and an arbitrary function $g(\cdot,\cdot)$ satisfying $E_F[g(\tilde{X}, \mathbf{\Delta}_0)] = 0$, it can be deduced that

$$\begin{aligned} & E\left[\frac{\delta}{XS_C(X-A)} g(X, \mathbf{\Delta}_0)\right] \\ = & \int_0^\infty \int_0^t \frac{g(t, \mathbf{\Delta}_0)f(t)S_C(t-a)}{tS_C(t-a)\mu} dadt \\ = & \frac{1}{\mu}\int_0^\infty g(t, \mathbf{\Delta}_0)f(t)dt \\ = & 0. \end{aligned}$$

Therefore, based on the function $\frac{\delta}{XS_C(X-A)} g(X, \mathbf{\Delta})$ and $\mathbf{\Delta} = (\xi, \theta)^T$, one system of estimating equations is constructed as follows.

$$\frac{1}{n}\sum_{i=1}^n \frac{\delta_i}{X_i S_C(X_i - A_i)} \boldsymbol{v}(X_i, \mathbf{\Delta}) = 0, \tag{2}$$

where

$$\boldsymbol{v}(X, \mathbf{\Delta}) = \begin{pmatrix} I(X \leq \xi) - p \\ I(X \leq \xi - \theta) - q \end{pmatrix}, \quad 0 < q < p < 1.$$

Then the solution of the system of equations (2) is one estimator of $\mathbf{\Delta}_0$.

Unfortunately, the survival function $S_C(\cdot)$ is unknown. A preferred approach is to replace $S_C(\cdot)$ with its Kaplan–Meier estimator (see [9]). That is,

$$\hat{S}_C(t) = \prod_{s \leq t}\left(1 - \frac{dN_C(s)}{Y(s)}\right),$$

where $N_C(t)=\sum_{i=1}^n N_i^C(t),\ N_i^C(t)=I(X_i-A_i\le t,\ \delta_i=0),\ Y(t)=\sum_{i=1}^n Y_i(t),\ Y_i(t)=I(X_i-A_i\ge t)$. Thus the system of equations (2) can be replaced with

$$(3)\quad \hat{\boldsymbol\psi}(\boldsymbol\Delta):=\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat S_C(X_i-A_i)}\boldsymbol v(X_i,\boldsymbol\Delta)=0.$$

That is because

$$\begin{aligned}E\left[\hat{\boldsymbol\psi}(\boldsymbol\Delta_0)\right]&=\frac{1}{\mu}E_F[\boldsymbol v(\tilde X,\boldsymbol\Delta_0)]+o_p(1)\\&=\frac{1}{\mu}\begin{pmatrix}F(\xi_p)-p\\F(\xi_p-\theta_0)-q\end{pmatrix}+o_p(1)\\&=o_p(1).\end{aligned}$$

For simplicity, for any differentiable function $l(x,y)$, let $\dot l_x(x,y)=\partial l(x,y)/\partial x$. Denote by

$$\boldsymbol\psi(\boldsymbol\Delta)=\frac{1}{\mu}E_F[\boldsymbol v(\tilde X,\boldsymbol\Delta)],$$

then

$$\dot{\boldsymbol\psi}_{\boldsymbol\Delta}(\boldsymbol\Delta)=\frac{\partial\boldsymbol\psi(\boldsymbol\Delta)}{\partial\boldsymbol\Delta}=\frac{1}{\mu}\begin{pmatrix}f(\xi)&0\\f(\xi-\theta)&-f(\xi-\theta)\end{pmatrix}.$$

Denote by $\hat{\boldsymbol\Delta}=(\hat\xi,\hat\theta)^T$ the solution of the system of equations (3), we have the following theorem.

Theorem 2.1. *Assume that the conditions (C1)-(C3) in Appendix A hold, then $\hat{\boldsymbol\Delta}$ is consistent and asymptotically normal with*

$$\sqrt{n}(\hat{\boldsymbol\Delta}-\boldsymbol\Delta_0)\xrightarrow{\mathcal D}N(0,\Sigma),$$

where $\Sigma=\Gamma_0^{-1}\Gamma(\Gamma_0^{-1})^T$, and

$$\Gamma_0=\frac{1}{\mu}\begin{pmatrix}f(\xi_p)&0\\f(\xi_q)&-f(\xi_q)\end{pmatrix},\quad\Gamma=\begin{pmatrix}\gamma_1&\gamma_3\\\gamma_3&\gamma_2\end{pmatrix},$$

$$\begin{aligned}\gamma_1&=E\left(\frac{\delta}{XS_C(X-A)}\left[I(X\le\xi_p)-p\right]\right)^2\\&\quad+\int_0^{\tau_C}\frac{b_1^2(s)}{\phi^2(s)}I(X-A\ge s)d\Lambda_C(s),\\\gamma_2&=E\left(\frac{\delta}{XS_C(X-A)}\left[I(X\le\xi_q)-q\right]\right)^2\\&\quad+\int_0^{\tau_C}\frac{b_2^2(s)}{\phi^2(s)}I(X-A\ge s)d\Lambda_C(s),\\\gamma_3&=E\left(\frac{\delta^2}{X^2S_C^2(X-A)}\left[I(X\le\xi_p)-p\right]\left[I(X\le\xi_q)-q\right]\right)\\&\quad+\int_0^{\tau_C}\frac{b_1(s)b_2(s)}{\phi^2(s)}I(X-A\ge s)d\Lambda_C(s),\end{aligned}$$

$$b_1(t)=\frac{1}{\mu}E_F\left[\frac{\tilde X-t}{\tilde X}\left[I\left(\tilde X\le\xi_p\right)-p\right]I\left(\tilde X\ge t\right)\right],$$

$$b_2(t)=\frac{1}{\mu}E_F\left[\frac{\tilde X-t}{\tilde X}\left[I\left(\tilde X\le\xi_q\right)-q\right]I\left(\tilde X\ge t\right)\right],$$

$$\phi(t)=S_C(t)S_T(t).$$

The conclusion of Theorem 1 can be used to construct a confidence region for $\boldsymbol\Delta_0$. For this purpose, denote by

$$\hat\Gamma_0(\boldsymbol\Delta)=\frac{1}{nh_f}\sum_{i=1}^n\frac{\delta_i}{X_i\hat S_C(X_i-A_i)}v_f(X_i,\boldsymbol\Delta),$$

$$\begin{aligned}\hat\Gamma(\boldsymbol\Delta)&=\frac{1}{n}\sum_{i=1}^n\left[\frac{\delta_i}{X_i\hat S_C(X_i-A_i)}\boldsymbol v(X_i,\boldsymbol\Delta)\right.\\&\qquad\left.+\int_0^{\tau_C}\frac{\hat{\boldsymbol B}(t,\boldsymbol\Delta)}{\hat\phi(t)}d\hat M_i^C(t)\right]\\&\qquad\times\left[\frac{\delta_i}{X_i\hat S_C(X_i-A_i)}\boldsymbol v(X_i,\boldsymbol\Delta)+\int_0^{\tau_C}\frac{\hat{\boldsymbol B}(t,\boldsymbol\Delta)}{\hat\phi(t)}d\hat M_i^C(t)\right]^T,\end{aligned}$$

$$v_f(X_i,\boldsymbol\Delta)=\begin{pmatrix}u\left(\frac{\xi-X_i}{h_f}\right)&0\\u\left(\frac{\xi-\theta-X_i}{h_f}\right)&-u\left(\frac{\xi-\theta-X_i}{h_f}\right)\end{pmatrix},$$

$$\hat{\boldsymbol B}(t,\boldsymbol\Delta)=\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat S_C(X_i-A_i)}\boldsymbol v(X_i,\boldsymbol\Delta)I(X_i-A_i\ge t),$$

$\hat\phi(t)=\frac{1}{n}\sum_{i=1}^n I(X_i-A_i\ge t)$, $\hat M_i^C(t)=I(X_i-A_i\le t,\delta_i=0)-\int_0^t I(X_i-A_i\ge s)d\hat\Lambda_C(s)$, and $\hat\Lambda_C(\cdot)$ the Nelson–Aalen estimator of $\Lambda_C(\cdot)$. Therefore, one of the consistent estimators of the covariance matrix Σ is $\hat\Sigma=\hat\Gamma_0(\hat{\boldsymbol\Delta})^{-1}\hat\Gamma(\hat{\boldsymbol\Delta})(\hat\Gamma_0(\hat{\boldsymbol\Delta})^{-1})^T$, where $\hat\Gamma_0(\hat{\boldsymbol\Delta})$ and $\hat\Gamma(\hat{\boldsymbol\Delta})$ are the consistent estimators of Γ_0 and Γ, respectively. Here, the estimator of the density function f is obtained by a kernel smoother with the kernel function $u(\cdot)$ and the bandwidth sequence $\{h_f\}$.

Corollary 2.1. *Assume that the conditions (C1)-(C3) in Appendix A hold, then $\hat\theta$ is consistent and asymptotically normal with*

$$\sqrt{n}(\hat\theta-\theta_0)\xrightarrow{\mathcal D}N(0,\sigma^2),$$

where $\sigma^2=\mu^2\gamma_1/f^2(\xi_p)$.

2.2 Smooth method

Another method, to estimate θ_0, is to construct one system of smooth estimating equations. Here the non-differentiable indicator functions are smoothed by the kernel function $K(\cdot)$ with a bandwidth sequence $\{h\}$. Denote by

$$\boldsymbol\varphi(X,\boldsymbol\Delta)=\begin{pmatrix}K\left(\frac{\xi-X}{h}\right)-p\\K\left(\frac{\xi-\theta-X}{h}\right)-q\end{pmatrix},$$

where $K(t)=\int_{-\infty}^t k(u)du$, and the function $k(\cdot)$ is a kernel density. Assume that the kernel $k(\cdot)$ is bounded and compactly supported, $k^{(2)}(\cdot)$ exists and is bounded, and

$$\int u^j k(u)du = \begin{cases} 1, & j = 0, \\ 0, & 1 \le j \le \gamma - 1, \\ c_0, & j = \gamma, \end{cases}$$

where $\gamma \ge 2$ is an integer, c_0 is some finite constant, and $c_0 \neq 0$. Then the smooth estimating vector function is

$$\frac{\delta}{X\hat{S}_C(X-A)}\boldsymbol{\varphi}(X, \boldsymbol{\Delta}).$$

It can be verified that

$$E\left[\frac{\delta}{X\hat{S}_C(X-A)}\boldsymbol{\varphi}(X, \boldsymbol{\Delta})\right] \to 0.$$

Then the smooth system of estimating equations is

$$\text{(4)} \qquad \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i - A_i)}\boldsymbol{\varphi}(X_i, \boldsymbol{\Delta}) = 0.$$

The solution $\tilde{\boldsymbol{\Delta}} = (\tilde{\xi}, \tilde{\theta})^T$ of (4) is one of the smoothed estimators of $\boldsymbol{\Delta}_0$. We derive that $\tilde{\boldsymbol{\Delta}}$ is consistent and asymptotically normal as follows.

Theorem 2.2. *Assume that the conditions (C1)-(C5) in Appendix A hold, then $\tilde{\boldsymbol{\Delta}}$ is consistent and asymptotically normal with*

$$\sqrt{n}(\tilde{\boldsymbol{\Delta}} - \boldsymbol{\Delta}_0) \xrightarrow{\mathcal{D}} N(0, \Sigma).$$

Note that, by the conclusions of Theorems 2.1 and 2.2, $\tilde{\boldsymbol{\Delta}}$ is asymptotically equivalent to $\hat{\boldsymbol{\Delta}}$. For a large sample size, they have similar performance. For a small sample size, however, the estimator $\tilde{\boldsymbol{\Delta}}$ with an appropriate bandwidth is more efficient than $\hat{\boldsymbol{\Delta}}$ in the sense of having a lower mean squared error. That is an advantage of using the kernel method.

Because the choice of the smoothness parameter h can affect the values of the mean squared error of $\tilde{\boldsymbol{\Delta}}$, the optimal bandwidth is required. In data analysis, the optimal bandwidth can be derived by minimizing the asymptotic mean squared error of the estimator. It can also be obtained by minimizing the asymptotic mean integrated squared error of the estimator. As the optimal bandwidth of the kernel distribution function introduced by [4], $h = O(n^{-1/3})$.

One confidence region for $\boldsymbol{\Delta}_0$ can also be constructed by using the result of Theorem 2. Here, $\tilde{\Sigma} = \hat{\Gamma}_0(\tilde{\boldsymbol{\Delta}})^{-1}\hat{\Gamma}(\tilde{\boldsymbol{\Delta}})(\hat{\Gamma}_0(\tilde{\boldsymbol{\Delta}})^{-1})^T$.

Corollary 2.2. *Assume that the conditions (C1)-(C5) in Appendix A hold, then $\tilde{\theta}$ is consistent and asymptotically normal with*

$$\sqrt{n}(\tilde{\theta} - \theta_0) \xrightarrow{\mathcal{D}} N(0, \sigma^2),$$

where σ^2 is defined in Corollary 2.1.

2.3 Efficiency improvement

The aforementioned estimators $\hat{\boldsymbol{\Delta}}$ and $\tilde{\boldsymbol{\Delta}}$ are all obtained by using the incomplete information of the observed data. To make the most of the features of the data and gain much more efficiency, an augmented inverse probability weighted complete-case (AIPWCC) system of estimating equations is constructed.

As we know, LBRC data type is a special type of RC data because the truncation variable A can be observed. Note that the censored data are monotone coarsening at random (CAR). For the monotone CAR data, there exists the explicit form of the most efficient estimator.

Moreover, it is easy to see that the estimator $\hat{\boldsymbol{\Delta}}$ satisfies the conditions of Theorem 2.2 in [12], hence it is a regular estimator. By the proof of Theorem 2.1, $\hat{\boldsymbol{\Delta}}$ is also an asymptotically linear estimator. That is,

$$\begin{aligned} &\sqrt{n}(\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}_0) \\ &= -\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\delta_i}{X_i S_C(X_i - A_i)}\boldsymbol{v}(X_i, \boldsymbol{\Delta}_0)\right. \\ &\qquad \left. + \int_0^{\tau_C}\frac{\boldsymbol{B}(s, \boldsymbol{\Delta}_0)}{\phi(s)}dM_i^C(s)\right] + o_p(1), \end{aligned}$$

where

$$\boldsymbol{B}(t, \boldsymbol{\Delta}) = E\left[\frac{\delta}{XS_C(X-A)}\boldsymbol{v}(X, \boldsymbol{\Delta})I(X - A \ge t)\right].$$

Therefore, the influence function of ith observation (X_i, δ_i, A_i) of the estimator $\hat{\boldsymbol{\Delta}}$ is

$$\begin{aligned} h_i &= -\Gamma_0^{-1}\left[\frac{\delta_i}{X_i S_C(X_i - A_i)}\boldsymbol{v}(X_i, \boldsymbol{\Delta}_0)\right. \\ &\qquad \left. + \int_0^{\tau_C}\frac{\boldsymbol{B}(s, \boldsymbol{\Delta}_0)}{\phi(s)}dM_i^C(s)\right], \quad i = 1, 2, \cdots, n. \end{aligned}$$

Based on the result of [17], it is true that

$$\frac{\delta_i}{S_C(X_i - A_i)} = 1 - \int_0^{\tau_C}\frac{dM_i^C(s)}{S_C(s)}.$$

Denote by

$$\boldsymbol{V}_i(\boldsymbol{\Delta}) = \frac{\boldsymbol{v}(X_i^0, \boldsymbol{\Delta})}{X_i^0}, \quad \boldsymbol{U}(s, \boldsymbol{V}) = \frac{E[\boldsymbol{V}(\boldsymbol{\Delta}_0)I(X^0 - A \ge s)]}{S_T(s)}.$$

According to Theorem 10.1 of [22], the efficient influence function is in the class of influence functions

$$\begin{aligned} \text{(5)} \qquad & -\Gamma_0^{-1}\left[\boldsymbol{V}_i(\boldsymbol{\Delta}_0) - \int_0^{\tau_C}[\boldsymbol{V}_i(\boldsymbol{\Delta}_0) - \boldsymbol{U}(s, \boldsymbol{V})]\frac{dM_i^C(s)}{S_C(s)}\right. \\ &\qquad \left. + \int_0^{\tau_C}[\boldsymbol{L}_i(s) - \boldsymbol{U}(s, \boldsymbol{L})]\frac{dM_i^C(s)}{S_C(s)}\right], \end{aligned}$$

where $\boldsymbol{L}_i(s)$ is an arbitrary function of the observed censoring data $(X_i - A_i = s, \delta_i = 0, A_i)$, $i = 1, \cdots, n$. When

$\boldsymbol{L}(s) = E[\boldsymbol{V}(\boldsymbol{\Delta}_0)|X^0 - A \geq s, A]$, the variance of the influence function is the smallest. The corresponding estimator is the most efficient estimator.

Nevertheless, we still do not know the explicit form of the conditional expectation $E[\boldsymbol{V}(\boldsymbol{\Delta}_0)|X^0 - A \geq s, A]$. To deal with this problem, one of the methods is to posit a model for the distribution of the full data X^0. Note that the estimator may not be the most efficient if the model is misspecified.

For simplicity, we use another method. Multiplying the augmented term in (5) by a 2×2 coefficient matrix D leads to the following class of influence functions

$$(6) \qquad -\Gamma_0^{-1}\left[\boldsymbol{V}_i(\boldsymbol{\Delta}_0) - \int_0^{\tau_C}[\boldsymbol{V}_i(\boldsymbol{\Delta}_0) - \boldsymbol{U}(s,\boldsymbol{V})]\frac{dM_i^C(s)}{S_C(s)} + D\int_0^{\tau_C}[\boldsymbol{L}_i(s) - \boldsymbol{U}(s,\boldsymbol{L})]\frac{dM_i^C(s)}{S_C(s)}\right].$$

The problem is transformed to find the optimal D that makes the variance matrix of the influence functions in (6) smallest. The optimal choice D_{opt} of D can be derived by using the ordinary least squares. After a little algebra, we have $D_{opt} = D_1 D_2^{-1}$, where

$$D_1 = E\left[\int_0^{\tau_C}[\boldsymbol{V}(\boldsymbol{\Delta}_0) - \boldsymbol{U}(s,\boldsymbol{V})][\boldsymbol{L}(s) - \boldsymbol{U}(s,\boldsymbol{L})]^T \times I(X - A \geq s)\frac{d\Lambda_C(s)}{S_C^2(s)}\right],$$

$$D_2 = E\left[\int_0^{\tau_C}[\boldsymbol{L}(s) - \boldsymbol{U}(s,\boldsymbol{L})][\boldsymbol{L}(s) - \boldsymbol{U}(s,\boldsymbol{L})]^T \times I(X - A \geq s)\frac{d\Lambda_C(s)}{S_C^2(s)}\right].$$

That is because the first term $\boldsymbol{V}_i(\boldsymbol{\Delta}_0)$ is independent of the third term in the square brackets of (6). Then the smallest variance of the influence functions in the class (6) is $\Sigma - \Gamma_0^{-1}D_1D_2^{-1}D_1^T(\Gamma_0^{-1})^T$ which is smaller than Σ when $\Gamma_0^{-1}D_1D_2^{-1}D_1^T(\Gamma_0^{-1})^T \neq 0$.

Finally, the AIPWCC system of estimating equations can be constructed as the form

$$(7) \qquad \sum_{i=1}^n \frac{\delta_i}{\hat{S}_C(X_i - A_i)}\boldsymbol{V}_i(\boldsymbol{\Delta}) + \sum_{i=1}^n \hat{D}\int_0^{\tau_C}[\boldsymbol{L}_i(s) - \hat{\boldsymbol{U}}(s,\boldsymbol{L})]\frac{dM_i^C(s)}{\hat{S}_C(s)} = 0,$$

where

$$\hat{D} = \hat{D}_1\hat{D}_2^{-1}, \quad \hat{\boldsymbol{U}}(s,\boldsymbol{L}) = \frac{\sum_{i=1}^n \boldsymbol{L}_i(s)I(X_i - A_i \geq s)}{\sum_{i=1}^n I(X_i - A_i \geq s)},$$

$$\hat{D}_1 = \frac{1}{n}\sum_{i=1}^n\left[\int_0^{\tau_C}[\boldsymbol{V}_i(\hat{\boldsymbol{\Delta}}) - \hat{\boldsymbol{U}}(s,\hat{\boldsymbol{V}})][\boldsymbol{L}_i(s) - \hat{\boldsymbol{U}}(s,\boldsymbol{L})]^T \times I(X_i - A_i \geq s)\frac{d\hat{\Lambda}_C(s)}{\hat{S}_C^2(s)}\right],$$

$$\hat{D}_2 = \frac{1}{n}\sum_{i=1}^n\left[\int_0^{\tau_C}[\boldsymbol{L}_i(s) - \hat{\boldsymbol{U}}(s,\boldsymbol{L})][\boldsymbol{L}_i(s) - \hat{\boldsymbol{U}}(s,\boldsymbol{L})]^T \times I(X_i - A_i \geq s)\frac{d\hat{\Lambda}_C(s)}{\hat{S}_C^2(s)}\right],$$

$$\hat{\boldsymbol{U}}(s,\hat{\boldsymbol{V}}) = \frac{\sum_{i=1}^n \boldsymbol{V}_i(\hat{\boldsymbol{\Delta}})I(X_i - A_i \geq s)}{\sum_{i=1}^n I(X_i - A_i \geq s)}.$$

Let $L(s) = A$, then the solution $\boldsymbol{\Delta}^* = (\xi^*, \theta^*)^T$ of the improved system of estimating equations (7) is more efficient than the estimators $\hat{\boldsymbol{\Delta}}$ and $\tilde{\boldsymbol{\Delta}}$ in the sense of having a lower asymptotic variance. Here, the asymptotic variance of $\boldsymbol{\Delta}^*$ is estimated by $\hat{\Gamma}_0(\boldsymbol{\Delta}^*)^{-1}\left[\hat{\Gamma}(\boldsymbol{\Delta}^*) - \hat{D}_1\hat{D}_2^{-1}\hat{D}_1^T\right](\hat{\Gamma}_0(\boldsymbol{\Delta}^*)^{-1})^T$.

3. SIMULATION STUDIES AND REAL DATA EXAMPLE

3.1 Simulation studies

In this subsection, we conduct simulations to assess the performance of the proposed methods.

Assume that the random variable $\tilde{X}$ follows Gamma distribution $\Gamma(\alpha, \beta)$, $\alpha = 2$, $\beta = 1/2$, the probability pair $(p, q) = (0.75, 0.25)$, then the true value of $\boldsymbol{\Delta}$ is $(5.385, 3.463)$. Let the sample size $n = 200,\ 400,\ 800$, and the censoring rate C% $= 15\%,\ 30\%$. According to [10], a LBRC sample $\{(X_i, A_i, \delta_i),\ i = 1, 2, \cdots, n\}$ is generated by the steps as follows.

Step 1. Generate $\{X_i^0,\ i = 1, 2, \cdots, n\}$ from the density function $f_{X^0}(x)$ which is introduced in equation (1).

Step 2. Generate A_i from the continuous uniform distribution $U(0, X_i^0)$, $i = 1, 2, \cdots, n$.

Step 3. Generate $\{C_i,\ i = 1, 2, \cdots, n\}$ from the survival distribution $S_C(\cdot)$. Here, we take an exponential distribution with the parameter corresponding to the censoring rate.

Step 4. Obtain X_i and δ_i from $X_i = \min\{X_i^0,\ A_i + C_i\}$ and $\delta_i = I(X_i^0 \leq A_i + C_i)$.

Let the value of the confidence level be 0.95, the asymptotic variance be estimated by using the plug-in method, and the density function be estimated by using a Gaussian kernel function with the bandwidth sequence $\{h_f\}$. As the optimal bandwidth introduced by [18, 8], the order of the optimal bandwidth is $O(n^{-1/5})$. Here, we observe the results with different bandwidths $h_f = cn^{-1/5}$, $c = 2.0, 2.1, 2.2, \cdots, 3.6, 3.7, 3.8$. Then we find that the result is better when $h_f = 2.8n^{-1/5}$. In addition, for $\boldsymbol{\varphi}(x, \boldsymbol{\Delta})$, assume that the kernel function is Gaussian with the bandwidth sequence $\{h\}$. Similar to the above-mentioned steps, we select the bandwidth $h = c_h n^{-1/3}$, $c_h = 2.0$. Each study consists of 500 replications. The corresponding results are summarized in Table 1. Similarly, the results of the case $(p, q) = (0.90, 0.10)$ are presented in Table 2.

Tables 1 and 2 show that the censoring rate does not affect the performance of the estimators. As we see, for each estimator, the standard errors (SE), mean squared er-

Table 1. The comparison of the estimates for $(p,q) = (0.75, 0.25)$, $h_f = 2.8n^{-1/5}$, and $h = 2.0n^{-1/3}$

C%	n	Est	Bias	SE	MSE	SD	CP	Length
15%	200	$(\hat{\xi}, \hat{\theta})$	(0.014, -0.006)	(0.336, 0.329)	(0.113, 0.109)	(0.337, 0.347)	(95.2, 95.6)	(1.319, 1.359)
		$(\tilde{\xi}, \tilde{\theta})$	(0.034, 0.015)	(0.328, 0.302)	(0.109, 0.091)	(0.338, 0.347)	(95.2, 97.0)	(1.325, 1.361)
		(ξ^*, θ^*)	(0.014, -0.006)	(0.335, 0.330)	(0.112, 0.109)	(0.335, 0.346)	(94.8, 95.4)	(1.315, 1.355)
	400	$(\hat{\xi}, \hat{\theta})$	(0.008, -0.002)	(0.246, 0.237)	(0.061, 0.056)	(0.241, 0.240)	(94.2, 95.4)	(0.945, 0.942)
		$(\tilde{\xi}, \tilde{\theta})$	(0.019, 0.009)	(0.241, 0.221)	(0.059, 0.049)	(0.242, 0.241)	(95.0, 96.8)	(0.948, 0.943)
		(ξ^*, θ^*)	(0.009, -0.001)	(0.245, 0.237)	(0.060, 0.056)	(0.240, 0.240)	(93.2, 94.8)	(0.942, 0.939)
	800	$(\hat{\xi}, \hat{\theta})$	(0.001, -0.013)	(0.178, 0.167)	(0.032, 0.028)	(0.172, 0.168)	(94.4, 94.0)	(0.676, 0.660)
		$(\tilde{\xi}, \tilde{\theta})$	(0.009, -0.004)	(0.172, 0.155)	(0.030, 0.024)	(0.173, 0.168)	(94.8, 95.8)	(0.677, 0.660)
		(ξ^*, θ^*)	(0.001, -0.012)	(0.178, 0.167)	(0.032, 0.028)	(0.172, 0.168)	(94.4, 94.8)	(0.674, 0.658)
30%	200	$(\hat{\xi}, \hat{\theta})$	(0.022, -0.003)	(0.350, 0.346)	(0.123, 0.120)	(0.359, 0.372)	(95.0, 95.6)	(1.407, 1.457)
		$(\tilde{\xi}, \tilde{\theta})$	(0.041, 0.015)	(0.338, 0.317)	(0.116, 0.100)	(0.360, 0.372)	(95.6, 96.8)	(1.413, 1.458)
		(ξ^*, θ^*)	(0.028, 0.001)	(0.347, 0.343)	(0.122, 0.117)	(0.356, 0.369)	(94.8, 96.0)	(1.396, 1.445)
	400	$(\hat{\xi}, \hat{\theta})$	(0.001, 0.003)	(0.256, 0.251)	(0.066, 0.063)	(0.258, 0.261)	(95.0, 95.0)	(1.012, 1.023)
		$(\tilde{\xi}, \tilde{\theta})$	(0.013, 0.010)	(0.245, 0.231)	(0.060, 0.053)	(0.259, 0.261)	(95.8, 97.0)	(1.015, 1.023)
		(ξ^*, θ^*)	(0.003, 0.004)	(0.255, 0.250)	(0.065, 0.063)	(0.257, 0.259)	(94.8, 95.2)	(1.007, 1.016)
	800	$(\hat{\xi}, \hat{\theta})$	(0.010, -0.003)	(0.192, 0.187)	(0.037, 0.035)	(0.182, 0.180)	(94.4, 94.6)	(0.712, 0.705)
		$(\tilde{\xi}, \tilde{\theta})$	(0.017, 0.004)	(0.187, 0.174)	(0.035, 0.030)	(0.182, 0.180)	(95.0, 95.6)	(0.714, 0.706)
		(ξ^*, θ^*)	(0.013, 0.000)	(0.191, 0.185)	(0.037, 0.034)	(0.181, 0.179)	(94.2, 94.6)	(0.709, 0.701)

(p,q): the probability pair; h_f: the bandwidth of the Gaussian kernel density function; h: the bandwidth of the Gaussian kernel distribution function; C%: the censoring rate; n: sample size; Est: the type of the estimators; $\hat{\xi}$: the estimator of pth quantile based on the nonsmooth estimating equation; $\tilde{\xi}$: the estimator of pth quantile based on the smooth estimating equation; ξ^*: the estimator of pth quantile based on the AIPWCC estimating equation; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation; Bias: the biases of the estimator; SE: the standard errors of the estimator; MSE: the mean squared errors of the estimator; SD: the average of the estimators of the standard deviation; CP: the empirical 95% coverage probability; Length: the average length of the confidence intervals.

Table 2. The comparison of the estimates for $(p,q) = (0.90, 0.10)$, $h_f = 2.8n^{-1/5}$, and $h = 2.0n^{-1/3}$

C%	n	Est	Bias	SE	MSE	SD	CP	Length
15%	200	$(\hat{\xi}, \hat{\theta})$	(0.031, 0.002)	(0.428, 0.479)	(0.184, 0.229)	(0.415, 0.507)	(95.4, 95.4)	(1.627, 1.988)
		$(\tilde{\xi}, \tilde{\theta})$	(0.049, 0.035)	(0.417, 0.435)	(0.176, 0.190)	(0.417, 0.506)	(95.6, 97.4)	(1.634, 1.984)
		(ξ^*, θ^*)	(0.026, -0.003)	(0.422, 0.474)	(0.179, 0.225)	(0.408, 0.501)	(94.6, 95.4)	(1.599, 1.963)
	400	$(\hat{\xi}, \hat{\theta})$	(0.005, -0.014)	(0.308, 0.330)	(0.095, 0.109)	(0.295, 0.355)	(94.6, 95.4)	(1.156, 1.391)
		$(\tilde{\xi}, \tilde{\theta})$	(0.018, 0.010)	(0.300, 0.308)	(0.090, 0.095)	(0.296, 0.355)	(94.8, 97.2)	(1.159, 1.393)
		(ξ^*, θ^*)	(0.001, -0.018)	(0.302, 0.327)	(0.091, 0.107)	(0.290, 0.351)	(93.8, 96.4)	(1.138, 1.375)
	800	$(\hat{\xi}, \hat{\theta})$	(0.009, 0.014)	(0.209, 0.243)	(0.044, 0.059)	(0.210, 0.250)	(95.2, 95.6)	(0.824, 0.979)
		$(\tilde{\xi}, \tilde{\theta})$	(0.016, 0.027)	(0.205, 0.229)	(0.042, 0.053)	(0.211, 0.250)	(95.2, 96.8)	(0.825, 0.979)
		(ξ^*, θ^*)	(0.006, 0.011)	(0.207, 0.241)	(0.043, 0.058)	(0.207, 0.247)	(95.4, 95.2)	(0.812, 0.969)
30%	200	$(\hat{\xi}, \hat{\theta})$	(0.001, -0.040)	(0.492, 0.526)	(0.242, 0.278)	(0.459, 0.551)	(93.0, 95.2)	(1.798, 2.160)
		$(\tilde{\xi}, \tilde{\theta})$	(0.023, 0.002)	(0.476, 0.481)	(0.227, 0.231)	(0.460, 0.551)	(93.0, 95.6)	(1.805, 2.161)
		(ξ^*, θ^*)	(-0.002, -0.045)	(0.492, 0.518)	(0.242, 0.270)	(0.433, 0.528)	(91.4, 93.6)	(1.699, 2.071)
	400	$(\hat{\xi}, \hat{\theta})$	(-0.009, -0.032)	(0.344, 0.374)	(0.118, 0.141)	(0.330, 0.388)	(93.4, 95.8)	(1.295, 1.522)
		$(\tilde{\xi}, \tilde{\theta})$	(0.008, -0.007)	(0.331, 0.347)	(0.110, 0.121)	(0.332, 0.389)	(95.0, 96.8)	(1.301, 1.524)
		(ξ^*, θ^*)	(-0.009, -0.033)	(0.337, 0.369)	(0.114, 0.137)	(0.317, 0.376)	(92.8, 94.8)	(1.243, 1.474)
	800	$(\hat{\xi}, \hat{\theta})$	(-0.013, -0.015)	(0.249, 0.259)	(0.062, 0.067)	(0.236, 0.275)	(94.0, 94.6)	(0.924, 1.076)
		$(\tilde{\xi}, \tilde{\theta})$	(-0.004, 0.001)	(0.241, 0.243)	(0.058, 0.059)	(0.236, 0.275)	(94.2, 96.8)	(0.927, 1.077)
		(ξ^*, θ^*)	(-0.021, -0.024)	(0.241, 0.256)	(0.058, 0.066)	(0.228, 0.267)	(92.6, 94.8)	(0.892, 1.045)

(p,q): the probability pair; h_f: the bandwidth of the Gaussian kernel density function; h: the bandwidth of the Gaussian kernel distribution function; C%: the censoring rate; n: sample size; Est: the type of the estimators; $\hat{\xi}$: the estimator of pth quantile based on the nonsmooth estimating equation; $\tilde{\xi}$: the estimator of pth quantile based on the smooth estimating equation; ξ^*: the estimator of pth quantile based on the AIPWCC estimating equation; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation; Bias: the biases of the estimator; SE: the standard errors of the estimator; MSE: the mean squared errors of the estimator; SD: the average of the estimators of the standard deviation; CP: the empirical 95% coverage probability; Length: the average length of the confidence intervals.

rors (MSE), standard deviations (SD), and average length (Length) of the confidence intervals all decrease as the sample size increases from 200, 400 to 800. Meanwhile, the confidence intervals of $\mathbf{\Delta}_0$ have good coverage probabilities (CP) which are very close to the nominal level 0.95. The SD of $\mathbf{\Delta}^*$ are smaller than those of $\hat{\mathbf{\Delta}}$ and $\tilde{\mathbf{\Delta}}$. That is to say, $\mathbf{\Delta}^*$ is more efficient than $\hat{\mathbf{\Delta}}$ and $\tilde{\mathbf{\Delta}}$ in the sense of having a lower asymptotic variance.

Moreover, the biases (Bias) of $\tilde{\mathbf{\Delta}}$ may be larger than those of $\hat{\mathbf{\Delta}}$ and $\mathbf{\Delta}^*$. That is because the advantage of the unbiasedness is sacrificed by using the smooth method. Nevertheless, the MSE of the smoothed estimator are smaller than those of the estimators $\hat{\mathbf{\Delta}}$ and $\mathbf{\Delta}^*$. It means that the estimator $\tilde{\mathbf{\Delta}}$ is more efficient than the estimators $\hat{\mathbf{\Delta}}$ and $\mathbf{\Delta}^*$ in the sense of having a lower mean squared error. It is also confirmed by the results of the improved efficiency shown in Tables 3 and 4. For simplicity, denote by

$$R1_\xi = 100 * \left(\sqrt{\mathrm{MSE}(\hat{\xi})} - \sqrt{\mathrm{MSE}(\tilde{\xi})}\right) \Big/ \sqrt{\mathrm{MSE}(\hat{\xi})}\,,$$

$$R1_\theta = 100 * \left(\sqrt{\mathrm{MSE}(\hat{\theta})} - \sqrt{\mathrm{MSE}(\tilde{\theta})}\right) \Big/ \sqrt{\mathrm{MSE}(\hat{\theta})}\,,$$

$$R2_\xi = 100 * \left(\sqrt{\mathrm{MSE}(\xi^*)} - \sqrt{\mathrm{MSE}(\tilde{\xi})}\right) \Big/ \sqrt{\mathrm{MSE}(\xi^*)}\,,$$

$$R2_\theta = 100 * \left(\sqrt{\mathrm{MSE}(\theta^*)} - \sqrt{\mathrm{MSE}(\tilde{\theta})}\right) \Big/ \sqrt{\mathrm{MSE}(\theta^*)}\,.$$

Furthermore, to analyze the characteristics of a distribution, some comparisons of the quantile differences from the same distribution are of great significance. Assume that the censoring rate is 57.31% and $n = 766$ as those of the real data which will be introduced in the next subsection. The bandwidths are chosen as $h_f = 2.8n^{-1/5}$ and $h = 0.3n^{-1/3}$, respectively. The results for (0.75, 0.25), (0.75, 0.50), and (0.50, 0.25) are summarized in Table 5.

Table 5 shows that the Bias of the estimators of $\theta_{(0.75,0.50)}$ are larger than those of $\theta_{(0.50,0.25)}$. The SE and SD are very close to each other. Meanwhile, the CP are close to the nominal level. It is confirmed that $h_f = 2.8n^{-1/5}$ and $h = 0.3n^{-1/3}$ are appropriate to the cases with probability pairs (0.75, 0.25) and (0.75, 0.50). For the case with probability pair (0.50, 0.25), the performance is a bit worse than that for the other cases. That is because measuring a less dispersion requires a smaller bandwidth. The MSE of $\tilde{\theta}$ are all smaller than those of $\hat{\theta}$ and θ^*. The SD of θ^* are all smaller than those of $\hat{\theta}$ and $\tilde{\theta}$. It is indicated that θ^* is more efficient in the sense of having a lower asymptotic variance.

In addition, we compare one of the proposed method with a method that fails to adjust left truncation, in order to consider how much gain we have by using the proposed method. For one sample with RC data, the estimator $\hat{\theta}_w$ of the quantile difference is the solution of the estimating equation

$$\sum_{i=1}^{n} \frac{\delta_i}{\hat{S}_{A+C}(X_i)} \boldsymbol{v}(X_i, \mathbf{\Delta}) = 0,$$

where $\hat{S}_{A+C}(\cdot)$ is the Kaplan–Meier estimator of the survival function of the right censoring variable $A + C$. In the sense of having a lower mean squared error, we calculate the ratio

$$R_w = 100 * \left(\sqrt{\mathrm{MSE}(\hat{\theta}_w)} - \sqrt{\mathrm{MSE}(\hat{\theta})}\right) \Big/ \sqrt{\mathrm{MSE}(\hat{\theta}_w)}\,.$$

The comparison of the results is exhibited in Table 6.

Table 6 illustrates that much more information, like the distribution type of the truncation variable, indeed contributes to the statistical inference.

3.2 Real data example

In this subsection, the proposed methods are used to analyze Oscar data (see for instance, [15, 21]).

Oscar data include 1670 performers, who were identified as nominees and non-nominees at the year from 1928 to 2001. Here, we focus on the lifetime distribution of the nominees. In the data, two nominees (ID number 1075 and 1430) are excluded because of the wrong information. Then there are 766 nominees left, in which the censoring rate is 57.31%. According to [29], the observed lifetime of the nominees is a length-biased data set if we regard the age at nomination as the truncation time. Moreover, according to the formal test proposed by [1], the lifetime data of the nominees satisfy the stationarity assumption. It is a precondition to use our methods.

Specifically, let the variable of interest $\tilde{X}$ be the lifetime of the nominees, and the truncation variable A be the age at nomination. The observed sample is the triples (X_i, A_i, δ_i), $i = 1, \cdots, 766$. To derive the asymptotic variance, the plug-in method is used. Let $(p, q) = (0.75, 0.25)$, the kernel function all be Gaussian, $h_f = 2.8n^{-1/5}$, and $h = 0.3n^{-1/3}$. The corresponding results are presented in Table 7. Table 7 shows that the SD of $\mathbf{\Delta}^*$ are smaller than those of $\hat{\mathbf{\Delta}}$ and $\tilde{\mathbf{\Delta}}$ which implies that $\mathbf{\Delta}^*$ is more efficient than $\hat{\mathbf{\Delta}}$ and $\tilde{\mathbf{\Delta}}$ in the sense of having a lower asymptotic variance. All estimates illustrate that the difference between the 75%th quantile and the 25%th quantile is about 21 years.

APPENDIX A. APPENDIX SECTION

To establish the asymptotic properties of the estimators of quantile differences, the following conditions are required.

- $(C1)$ The parameter space $\mathbf{\Omega}$ of $\mathbf{\Delta}$ is compact, and the true value $\mathbf{\Delta}_0$ is an interior point of $\mathbf{\Omega}$.
- $(C2)$ $\mathbf{\Delta}_0$ is the unique value such that $\boldsymbol{\psi}(\mathbf{\Delta}) = 0$.
- $(C3)$ Let $\tau_1 = \sup\{t : S(t)S_C(t) > 0\}$, $\tau_T = \sup\{t : S_T(t) > 0\} \le \sup\{t : S_C(t) > 0\} = \tau_C$, and $\Pr(\delta = 1) > 0$.
- $(C4)$ The kernel function $K(\cdot)$ is continuous, differentiable and bounded in the neighborhood of $\mathbf{\Delta}_0$. The function $f(\cdot)$ is continuous, and $f(\xi_p)f(\xi_q) > 0$.
- $(C5)$ For some integer $\gamma \ge 2$, $f^{(\gamma-1)}(x)$ exists in a neighborhood of ξ_p and ξ_q, and is continuous at ξ_p and ξ_q, respectively.

Table 3. The comparison of the efficiencies for $n = 200$ and $h_f = 2.8n^{-1/5}$

		$p=0.75$	$q=0.25$			$p=0.90$	$q=0.10$		
C%	c_h	$R1_\xi$	$R1_\theta$	$R2_\xi$	$R2_\theta$	$R1_\xi$	$R1_\theta$	$R2_\xi$	$R2_\theta$
15%	1.6	2.449	7.022	2.182	6.827	3.151	6.802	2.806	5.802
	1.7	2.405	7.014	3.009	6.375	3.501	7.392	1.578	6.043
	1.8	2.723	8.093	2.284	7.901	2.804	6.191	1.041	6.186
	1.9	2.527	8.643	1.714	8.656	3.850	10.045	4.393	9.845
	2.0	4.609	10.763	3.604	9.721	4.528	11.372	4.551	11.185
	2.1	5.002	9.549	4.382	9.610	4.706	10.078	5.200	10.175
	2.2	2.735	12.170	1.958	11.359	4.864	11.073	3.350	10.702
	2.3	5.076	11.030	5.147	10.479	5.490	11.268	3.522	10.215
	2.4	1.306	11.299	0.376	11.015	3.627	9.310	1.924	8.195
30%	1.6	2.520	8.929	2.294	8.766	3.495	7.873	6.884	11.351
	1.7	4.069	8.532	3.531	6.001	3.973	8.402	4.936	9.368
	1.8	2.517	9.102	2.175	9.865	4.442	8.325	4.037	7.950
	1.9	3.656	10.319	3.244	9.924	4.888	9.908	11.990	14.810
	2.0	4.072	6.607	4.365	7.686	5.647	9.169	8.676	12.026
	2.1	4.176	12.314	3.849	11.567	3.406	8.558	3.189	8.064
	2.2	3.452	13.197	3.024	10.385	6.493	10.769	9.134	11.432
	2.3	2.512	10.772	2.241	10.804	5.602	9.513	6.561	10.374
	2.4	6.926	13.529	17.279	21.009	7.720	11.357	8.851	12.558

n: sample size; h_f: the bandwidth of the Gaussian kernel density function; (p,q): the probability pair; C%: the censoring rate; c_h: the coefficient of the bandwidth sequence of the Gaussian kernel distribution function; $R1_\xi$: the improved efficiency of $\tilde{\xi}$ with respect to $\hat{\xi}$; $R1_\theta$: the improved efficiency of $\tilde{\theta}$ with respect to $\hat{\theta}$; $R2_\xi$: the improved efficiency of $\tilde{\xi}$ with respect to ξ^*; $R2_\theta$: the improved efficiency of $\tilde{\theta}$ with respect to θ^*; $\hat{\xi}$: the estimator of pth quantile based on the nonsmooth estimating equation; $\tilde{\xi}$: the estimator of pth quantile based on the smooth estimating equation; ξ^*: the estimator of pth quantile based on the AIPWCC estimating equation; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation.

Table 4. The comparison of the efficiencies for $n = 400$ and $h_f = 2.8n^{-1/5}$

		$p=0.75$	$q=0.25$			$p=0.90$	$q=0.10$		
C%	c_h	$R1_\xi$	$R1_\theta$	$R2_\xi$	$R2_\theta$	$R1_\xi$	$R1_\theta$	$R2_\xi$	$R2_\theta$
15%	1.6	2.807	6.293	2.343	6.795	1.712	6.982	0.106	6.885
	1.7	2.941	6.620	2.478	7.121	1.777	7.291	0.172	7.196
	1.8	3.072	6.944	2.609	7.443	1.839	7.579	0.234	7.484
	1.9	3.195	7.266	2.733	7.764	1.893	7.844	0.289	7.749
	2.0	3.308	7.580	2.846	8.076	1.944	8.088	0.342	7.993
	2.1	3.411	7.886	2.950	8.380	1.996	8.311	0.394	8.216
	2.2	3.508	8.186	3.047	8.678	2.039	8.506	0.438	8.412
	2.3	3.599	8.484	3.139	8.975	2.078	8.678	0.478	8.583
	2.4	3.680	8.770	3.220	9.260	2.113	8.822	0.513	8.728
30%	1.6	0.011	0.059	0.266	0.304	0.004	0.031	0.312	0.344
	1.7	0.011	0.062	0.267	0.307	0.006	0.022	0.254	0.268
	1.8	0.012	0.064	0.267	0.310	0.007	0.023	0.254	0.269
	1.9	0.012	0.067	0.268	0.312	0.007	0.025	0.255	0.271
	2.0	0.013	0.069	0.268	0.315	0.008	0.026	0.255	0.272
	2.1	0.013	0.072	0.269	0.317	0.008	0.027	0.256	0.273
	2.2	0.014	0.074	0.269	0.319	0.009	0.028	0.256	0.274
	2.3	0.014	0.076	0.270	0.322	0.009	0.029	0.257	0.275
	2.4	0.015	0.079	0.270	0.324	0.010	0.030	0.257	0.276

n: sample size; h_f: the bandwidth of the Gaussian kernel density function; (p,q): the probability pair; C%: the censoring rate; c_h: the coefficient of the bandwidth sequence of the Gaussian kernel distribution function; $R1_\xi$: the improved efficiency of $\tilde{\xi}$ with respect to $\hat{\xi}$; $R1_\theta$: the improved efficiency of $\tilde{\theta}$ with respect to $\hat{\theta}$; $R2_\xi$: the improved efficiency of $\tilde{\xi}$ with respect to ξ^*; $R2_\theta$: the improved efficiency of $\tilde{\theta}$ with respect to θ^*; $\hat{\xi}$: the estimator of pth quantile based on the nonsmooth estimating equation; $\tilde{\xi}$: the estimator of pth quantile based on the smooth estimating equation; ξ^*: the estimator of pth quantile based on the AIPWCC estimating equation; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation.

Table 5. The comparison of the estimates for $C\% = 57.31\%$, $n = 766$, $h_f = 2.8n^{-1/5}$, and $h = 0.3n^{-1/3}$

$(p,\, q)$	Est	Bias	SE	MSE	SD	CP	Length
(0.75, 0.25)	$\hat{\theta}$	-0.049	0.246	0.063	0.251	95.2	0.986
	$\tilde{\theta}$	-0.049	0.244	0.062	0.251	94.8	0.985
	θ^*	-0.036	0.250	0.064	0.247	94.4	0.967
(0.75, 0.50)	$\hat{\theta}$	-0.034	0.214	0.047	0.215	93.6	0.842
	$\tilde{\theta}$	-0.033	0.210	0.045	0.215	95.0	0.842
	θ^*	-0.023	0.218	0.048	0.212	92.8	0.830
(0.50, 0.25)	$\hat{\theta}$	-0.015	0.152	0.023	0.163	95.6	0.638
	$\tilde{\theta}$	-0.016	0.150	0.023	0.163	96.6	0.638
	θ^*	-0.018	0.156	0.025	0.161	95.2	0.633

C%: the censoring rate; n: sample size; h_f: the bandwidth of the Gaussian kernel density function; h: the bandwidth of the Gaussian kernel distribution function; (p, q): the probability pair; Est: the type of the estimators; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation; Bias: the biases of the estimator; SE: the standard errors of the estimator; MSE: the mean squared errors of the estimator; SD: the average of the estimators of the standard deviation; CP: the empirical 95% coverage probability; Length: the average length of the confidence intervals.

Table 6. The comparison of the estimates for $C\% = 57.31\%$, $n = 766$, and $h_f = 2.8n^{-1/5}$

	$\hat{\theta}_w$			$\hat{\theta}$			
$(p,\, q)$	Bias	SE	MSE	Bias	SE	MSE	R_w
(0.75, 0.25)	0.095	0.376	0.150	-0.049	0.246	0.063	35.3
(0.80, 0.20)	0.130	0.442	0.212	-0.064	0.281	0.083	37.4
(0.90, 0.10)	0.215	0.656	0.477	-0.151	0.493	0.266	25.4

C%: the censoring rate; n: sample size; h_f: the bandwidth of the Gaussian kernel density function; $\hat{\theta}_w$: the estimator of quantile difference when reviewing LBRC data only as RC data; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; (p, q): the probability pair; Bias: the biases of the estimator; SE: the standard errors of the estimator; MSE: the mean squared errors of the estimator; R_w: the improved efficiency of $\hat{\theta}$ with respect to $\hat{\theta}_w$.

Table 7. The estimates on Oscar data for $(p, q) = (0.75, 0.25)$, $h_f = 2.8n^{-1/5}$, and $h = 0.3n^{-1/3}$

Est	$(\hat{\xi}, \hat{\theta})$	$(\tilde{\xi}, \tilde{\theta})$	(ξ^*, θ^*)
est	(85.0, 21.0)	(84.9, 21.0)	(85.0, 21.0)
SD	(3.1, 2.4)	(3.1, 2.4)	(0.2, 0.1)

(p, q): the probability pair; h_f: the bandwidth of the Gaussian kernel density function; h: the bandwidth of the Gaussian kernel distribution function; Est: the type of the estimators; $\hat{\xi}$: the estimator of pth quantile based on the nonsmooth estimating equation; $\tilde{\xi}$: the estimator of pth quantile based on the smooth estimating equation; ξ^*: the estimator of pth quantile based on the AIPWCC estimating equation; $\hat{\theta}$: the estimator of quantile difference based on the nonsmooth estimating equation; $\tilde{\theta}$: the estimator of quantile difference based on the smooth estimating equation; θ^*: the estimator of quantile difference based on the AIPWCC estimating equation; est: the estimates of the quantile difference; SD: the average of the estimators of the standard deviation.

The condition $(C4)$ implies that the functions $\boldsymbol{\varphi}(x, \boldsymbol{\Delta})$ and $\dot{\boldsymbol{\varphi}}_{\boldsymbol{\Delta}}(x, \boldsymbol{\Delta})$ are continuous and bounded by some integrable function $\boldsymbol{g}(x)$ in a neighborhood of the true value $\boldsymbol{\Delta}_0$ with $\int \boldsymbol{g}(u)dF(u) < \infty$, and $E_F[\dot{\boldsymbol{\varphi}}_{\boldsymbol{\Delta}}(\tilde{X}, \boldsymbol{\Delta}_0)] \neq 0$.

Proof of Theorem 2.1. Firstly, we prove the consistency of the estimator $\hat{\boldsymbol{\Delta}}$. Based on the condition $(C2)$, to prove the consistency of $\hat{\boldsymbol{\Delta}}$, it is required to prove that

$$\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}} \left\| \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i}{X_i \hat{S}_C(X_i - A_i)} \boldsymbol{v}(X_i, \boldsymbol{\Delta}) - \boldsymbol{\psi}(\boldsymbol{\Delta}) \right\| = o_p(1). \tag{8}$$

In actuality,

$$\begin{aligned} &\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}} \left\| \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i}{X_i \hat{S}_C(X_i - A_i)} \boldsymbol{v}(X_i, \boldsymbol{\Delta}) - \boldsymbol{\psi}(\boldsymbol{\Delta}) \right\| \\ \leq\ &\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}} \left\| \frac{1}{n}\sum_{i=1}^{n} \left[\frac{\delta_i}{X_i \hat{S}_C(X_i - A_i)} - \frac{\delta_i}{X_i S_C(X_i - A_i)} \right] \right. \\ &\qquad \left. \times \boldsymbol{v}(X_i, \boldsymbol{\Delta}) \right\| \\ &+ \sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}} \left\| \frac{1}{n}\sum_{i=1}^{n} \frac{\delta_i}{X_i S_C(X_i - A_i)} \boldsymbol{v}(X_i, \boldsymbol{\Delta}) - \boldsymbol{\psi}(\boldsymbol{\Delta}) \right\| \\ :=\ & I_1 + I_2. \end{aligned}$$

From the condition $(C3)$, we have

$$\sup_{0\leq t\leq\tau_1} \left| \frac{S_C(t)}{\hat{S}_C(t)} \right| \leq 1 + \sup_{0\leq t\leq\tau_1} \left| \frac{S_C(t) - \hat{S}_C(t)}{S_C(t)} \right| \sup_{0\leq t\leq\tau_1} \left| \frac{S_C(t)}{\hat{S}_C(t)} \right|.$$

The uniform consistency of the Kaplan–Meier estimator $\hat{S}_C(t)$ implies that $\sup_{0\le t\le\tau_1}\left|\frac{S_C(t)}{\hat{S}_C(t)}\right| \le M$ for some constant M. Combining the law of large numbers with the condition ($C1$), we get

$$I_1 \le \frac{2}{\mu}\sup_{0\le t\le\tau_1}\left|\frac{S_C(t)-\hat{S}_C(t)}{S_C(t)}\right|\sup_{0\le t\le\tau_1}\left|\frac{S_C(t)}{\hat{S}_C(t)}\right| = o_p(1).$$

From the uniform law of large numbers, it is easy to see that $I_2 \xrightarrow{\mathcal{P}} 0$. Hence (8) holds. That leads to, for any $r>0$ and neighborhood $\mathcal{N}(\boldsymbol{\Delta}_0, r)$ of $\boldsymbol{\Delta}_0$,

$$\begin{aligned}&\sup_{\boldsymbol{\Delta}\notin\mathcal{N}(\boldsymbol{\Delta}_0,r)}\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{v}(X_i,\boldsymbol{\Delta})-\boldsymbol{\psi}(\boldsymbol{\Delta})\right\|\\ &= o_p(1).\end{aligned}$$

It follows that

$$\begin{aligned}(9)\qquad &\inf_{\boldsymbol{\Delta}\notin\mathcal{N}(\boldsymbol{\Delta}_0,r)}\|\boldsymbol{\psi}(\boldsymbol{\Delta})\|\\ \ge\ &\inf_{\boldsymbol{\Delta}\notin\mathcal{N}(\boldsymbol{\Delta}_0,r)}\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{v}(X_i,\boldsymbol{\Delta})\right\| - o_p(1)\\ >\ & 0 = \|\boldsymbol{\psi}(\boldsymbol{\Delta}_0)\|.\end{aligned}$$

By the conditions ($C1$) and ($C2$), and the conclusions (8) and (9), the estimator $\hat{\boldsymbol{\Delta}}$ is consistent.

To prove the asymptotic normality, denote by $\mathcal{F}(t) = \sigma\{I(X_i - A_i \le u,\ \delta_i = 0),\ I(X_i - A_i \ge u),\ i = 1, 2, \cdots, n,\ 0 \le u \le t\}$, then $M_i^C(t) = N_i^C(t) - \int_0^t Y_i(s)d\Lambda_C(s)$ is a martingale with respect to $\mathcal{F}(t)$, $i = 1, 2, \cdots, n$. Meanwhile,

$$\sqrt{n}\frac{S_C(t)-\hat{S}_C(t)}{S_C(t)} = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^{\tau_C}\frac{I(s\le t)}{\phi(s)}dM_j^C(s) + o_p(1),$$

where $\phi(s) = S_T(s)S_C(s)$.

By the consistency of $\hat{\boldsymbol{\Delta}}$ and the mean value theorem, we have

$$\begin{aligned}&\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Delta}}) - \hat{\boldsymbol{\psi}}(\boldsymbol{\Delta}_0)\\ =\ &\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Delta}}) - \boldsymbol{\psi}(\hat{\boldsymbol{\Delta}}) + \boldsymbol{\psi}(\hat{\boldsymbol{\Delta}}) - \hat{\boldsymbol{\psi}}(\boldsymbol{\Delta}_0) + \boldsymbol{\psi}(\boldsymbol{\Delta}_0) - \boldsymbol{\psi}(\boldsymbol{\Delta}_0)\\ =\ &\boldsymbol{\beta}_n + \boldsymbol{\psi}(\hat{\boldsymbol{\Delta}}) - \boldsymbol{\psi}(\boldsymbol{\Delta}_0)\\ =\ &\boldsymbol{\beta}_n + [\dot{\boldsymbol{\psi}}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta}_0) + o_p(1)](\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}_0),\end{aligned}$$

where $\boldsymbol{\beta}_n = \boldsymbol{\alpha}_n(\hat{\boldsymbol{\Delta}}) - \boldsymbol{\alpha}_n(\boldsymbol{\Delta}_0)$, $\boldsymbol{\alpha}_n(\boldsymbol{\Delta}) = \hat{\boldsymbol{\psi}}(\boldsymbol{\Delta}) - \boldsymbol{\psi}(\boldsymbol{\Delta})$. The function class $\{\boldsymbol{v}(\cdot,\boldsymbol{\Delta}), \boldsymbol{\Delta}\in\boldsymbol{\Omega}\}$ is Euclidean with a square-integrable envelope, and it is $\mathcal{L}^2(p)$ continuous at $\boldsymbol{\Delta}_0$. Hence, by Lemma 2.17 of [13], for each sequence of positive numbers $\{\varepsilon_n\}$ converging to zero, we get

$$\sup_{\|\hat{\boldsymbol{\Delta}}-\boldsymbol{\Delta}_0\|<\varepsilon_n}\|\boldsymbol{\alpha}_n(\hat{\boldsymbol{\Delta}}) - \boldsymbol{\alpha}_n(\boldsymbol{\Delta}_0)\| = o_p(n^{-\frac{1}{2}}),$$

that is to say,

$$\sup_{\|\hat{\boldsymbol{\Delta}}-\boldsymbol{\Delta}_0\|<\varepsilon_n}\|\boldsymbol{\beta}_n\| = o_p(n^{-\frac{1}{2}}).$$

Let

$$\Gamma_0 = \frac{1}{\mu}\begin{pmatrix} f(\xi_p) & 0\\ f(\xi_q) & -f(\xi_q)\end{pmatrix}.$$

It is derived that

$$\begin{aligned}&\sqrt{n}(\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}_0)\\ =\ & -\sqrt{n}[\dot{\boldsymbol{\psi}}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta}_0) + o_p(1)]^{-1}[\hat{\boldsymbol{\psi}}(\boldsymbol{\Delta}_0) + \boldsymbol{\beta}_n]\\ =\ & -\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{v}(X_i,\boldsymbol{\Delta}_0) + o_p(1)\\ =\ & -\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{v}(X_i,\boldsymbol{\Delta}_0)\\ &\times\left(1 + \frac{S_C(X_i-A_i)-\hat{S}_C(X_i-A_i)}{S_C(X_i-A_i)}\right) + o_p(1)\\ =\ & -\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{v}(X_i,\boldsymbol{\Delta}_0)\\ & -\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau_C}\frac{\boldsymbol{B}(s,\boldsymbol{\Delta}_0)}{\phi(s)}dM_i^C(s) + o_p(1),\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{B}(t,\boldsymbol{\Delta}_0) &= E\left[\frac{\delta}{XS_C(X-A)}\boldsymbol{v}(X,\boldsymbol{\Delta}_0)I(t\le X-A)\right]\\ &= \frac{1}{\mu}E_F\left[\frac{\tilde{X}-t}{\tilde{X}}\boldsymbol{v}(\tilde{X},\boldsymbol{\Delta}_0)I(t\le\tilde{X})\right].\end{aligned}$$

By the central limit theorems of the martingale and the sum of i.i.d. random variables, we have

$$\sqrt{n}(\hat{\boldsymbol{\Delta}} - \boldsymbol{\Delta}_0) \xrightarrow{\mathcal{D}} N(0,\Sigma),$$

where $\Sigma = \Gamma_0^{-1}\Gamma(\Gamma_0^{-1})^T$, and

$$\begin{aligned}\Gamma &= E\Bigg(\left[\frac{\delta}{XS_C(X-A)}\boldsymbol{v}(X,\boldsymbol{\Delta}_0) + \int_0^{\tau_C}\frac{\boldsymbol{B}(s,\boldsymbol{\Delta}_0)}{\phi(s)}dM^C(s)\right]\\ &\quad\times\left[\frac{\delta}{XS_C(X-A)}\boldsymbol{v}(X,\boldsymbol{\Delta}_0) + \int_0^{\tau_C}\frac{\boldsymbol{B}(s,\boldsymbol{\Delta}_0)}{\phi(s)}dM^C(s)\right]^T\Bigg)\\ &= \begin{pmatrix}\gamma_1 & \gamma_3\\ \gamma_3 & \gamma_2\end{pmatrix},\end{aligned}$$

$$\begin{aligned}\gamma_1 &= E\left(\frac{\delta}{XS_C(X-A)}[I(X\le\xi_p)-p]\right)^2\\ &\quad+\int_0^{\tau_C}\frac{b_1^2(s)}{\phi^2(s)}I(X-A\ge s)d\Lambda_C(s),\\ \gamma_2 &= E\left(\frac{\delta}{XS_C(X-A)}[I(X\le\xi_q)-q]\right)^2\end{aligned}$$

$$+\int_0^{\tau_C}\frac{b_2^2(s)}{\phi^2(s)}I(X-A\geq s)d\Lambda_C(s),$$

$$\gamma_3=E\left(\frac{\delta^2}{X^2S_C^2(X-A)}[I(X\leq\xi_p)-p][I(X\leq\xi_q)-q]\right)$$
$$+\int_0^{\tau_C}\frac{b_1(s)b_2(s)}{\phi^2(s)}I(X-A\geq s)d\Lambda_C(s),$$

$$b_1(t)=\frac{1}{\mu}E_F\left[\frac{\tilde{X}-t}{\tilde{X}}\left[I\left(\tilde{X}\leq\xi_p\right)-p\right]I\left(t\leq\tilde{X}\right)\right],$$

$$b_2(t)=\frac{1}{\mu}E_F\left[\frac{\tilde{X}-t}{\tilde{X}}\left[I\left(\tilde{X}\leq\xi_q\right)-q\right]I\left(t\leq\tilde{X}\right)\right],$$

$$\phi(t)=S_C(t)S_T(t).$$

The proof is completed. □

Proof of Theorem 2.2. Firstly, we prove the consistency of the estimator $\tilde{\boldsymbol{\Delta}}$. Based on the condition $(C2)$, it is required to prove that

$$\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\boldsymbol{\psi}(\boldsymbol{\Delta})\right\|=o_p(1).\tag{10}$$

Actually,

$$\begin{aligned}&\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\boldsymbol{\psi}(\boldsymbol{\Delta})\right\|\\
\leq&\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})\right\|\\
&+\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\boldsymbol{\psi}(\boldsymbol{\Delta})\right\|\\
:=&J_1+J_2.\end{aligned}$$

By the uniform consistency of the Kaplan–Meier estimator $\hat{S}_C(t)$, it is easy to deduce that

$$\begin{aligned}J_1=&\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})\right\|\\
=&\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\left(\frac{1}{n}\sum_{i=1}^n\frac{S_C(X_i-A_i)-\hat{S}_C(X_i-A_i)}{[S_C(X_i-A_i)]^2}+\frac{1}{n}\sum_{i=1}^n\frac{[S_C(X_i-A_i)-\hat{S}_C(X_i-A_i)]^2}{\hat{S}_C(X_i-A_i)[S_C(X_i-A_i)]^2}\right)\times\frac{\delta_i}{X_i}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})\right\|\\
\leq&\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{1}{n}\sum_{i=1}^n\frac{S_C(X_i-A_i)-\hat{S}_C(X_i-A_i)}{[S_C(X_i-A_i)]^2}\times\frac{\delta_i}{X_i}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})\right\|+o_p(1)\\
=&o_p(1).\end{aligned}$$

From the uniform law of large numbers and the condition $(C5)$, it is easy to get

$$J_2\leq\sup_{\boldsymbol{\Delta}\in\boldsymbol{\Omega}}\left\|\frac{\boldsymbol{F}^{(\gamma)}(\boldsymbol{\Delta}')}{\gamma!}c_0h^\gamma\right\|+o_p(1)\xrightarrow{\mathcal{P}}0,$$

where

$$\boldsymbol{F}^{(\gamma)}(\boldsymbol{\Delta}')=\begin{pmatrix}F^{(\gamma)}(\xi')\\F^{(\gamma)}(\xi''-\theta'')\end{pmatrix},$$

for any $t>0$ and fixed h, the point ξ' is between ξ and $\xi-ht$, and $\xi''-\theta''$ is between $\xi-\theta$ and $\xi-\theta-ht$. Hence (10) holds. Then for any $r>0$ and the neighborhood $\mathcal{N}^*(\boldsymbol{\Delta}_0,r)$ of $\boldsymbol{\Delta}_0$, we have

$$\sup_{\boldsymbol{\Delta}\notin\mathcal{N}^*(\boldsymbol{\Delta}_0,r)}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})-\boldsymbol{\psi}(\boldsymbol{\Delta})\right\|=o_p(1).$$

It follows that

$$\begin{aligned}&\inf_{\boldsymbol{\Delta}\notin\mathcal{N}^*(\boldsymbol{\Delta}_0,r)}\|\boldsymbol{\psi}(\boldsymbol{\Delta})\|\\
\geq&\inf_{\boldsymbol{\Delta}\notin\mathcal{N}^*(\boldsymbol{\Delta}_0,r)}\left\|\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta})\right\|-o_p(1)\\
>&0=\|\boldsymbol{\psi}(\boldsymbol{\Delta}_0)\|.\end{aligned}\tag{11}$$

By the conditions $(C1)$ and $(C2)$, and the conclusions (10) and (11), the estimator $\hat{\boldsymbol{\Delta}}$ is consistent.

The asymptotic normality is proved as follows, and some notations are identical to those in the proof of Theorem 2.1. Combining the consistency of $\tilde{\boldsymbol{\Delta}}$ and the condition $(C4)$ with Taylor's formula, we have

$$\begin{aligned}0=&\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\tilde{\boldsymbol{\Delta}})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta}_0)\\
&+\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}[\dot{\boldsymbol{\varphi}}_{\boldsymbol{\Delta}}(X_i,\boldsymbol{\Delta}_0)+o_p(1)]\times\sqrt{n}(\tilde{\boldsymbol{\Delta}}-\boldsymbol{\Delta}_0).\end{aligned}$$

By Bernstein inequality and Borel–Cantelli Lemma, we have

$$\frac{1}{nh}\sum_{i=1}^n\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}k\left(\frac{\xi_p-X_i}{h}\right)\to\frac{f(\xi_p)}{\mu},\quad a.s..$$

Therefore,

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\dot{\boldsymbol{\varphi}}_{\boldsymbol{\Delta}}(X_i,\boldsymbol{\Delta}_0)\\
\to\ &\frac{1}{\mu}\begin{pmatrix} f(\xi_p) & 0\\ f(\xi_q) & -f(\xi_q)\end{pmatrix}=\Gamma_0 \quad a.s.,
\end{aligned}
$$

then we get

$$
\begin{aligned}
&\sqrt{n}(\tilde{\boldsymbol{\Delta}}-\boldsymbol{\Delta}_0)\\
= &-\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}[\dot{\boldsymbol{\varphi}}_{\boldsymbol{\Delta}}(X_i,\boldsymbol{\Delta}_0)+o_p(1)]\right)^{-1}\\
&\times\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta}_0)\\
= &-\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_i\hat{S}_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta}_0)+o_p(1)\\
= &-\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta}_0)\\
&\times\left(1+\frac{S_C(X_i-A_i)-\hat{S}_C(X_i-A_i)}{S_C(X_i-A_i)}\right)+o_p(1)\\
= &-\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{X_iS_C(X_i-A_i)}\boldsymbol{\varphi}(X_i,\boldsymbol{\Delta}_0)\\
&-\Gamma_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\tau_C}\frac{\boldsymbol{B}^{\boldsymbol{\varphi}}(s,\boldsymbol{\Delta}_0)}{\phi(s)}dM_i^C(s)+o_p(1),
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{B}^{\boldsymbol{\varphi}}(t,\boldsymbol{\Delta}_0) &= E\left[\frac{\delta}{XS_C(X-A)}\boldsymbol{\varphi}(X,\boldsymbol{\Delta}_0)I(t\le X-A)\right]\\
&\to \frac{1}{\mu}E_F\left[\frac{\tilde{X}-t}{\tilde{X}}\boldsymbol{v}(\tilde{X},\boldsymbol{\Delta}_0)I(t\le\tilde{X})\right],\quad h\to 0.
\end{aligned}
$$

By the central limit theorems of the martingale and the sum of i.i.d. random variables, we have

$$\sqrt{n}(\tilde{\boldsymbol{\Delta}}-\boldsymbol{\Delta}_0)\xrightarrow{\mathcal{D}}N(0,\Sigma).$$

The proof is completed. □

DISCUSSION

Theoretically and numerically, the three estimating equation estimators of quantile differences perform well. In the sense of having a lower mean squared error, the smoothed estimator is more efficient than the others. Meanwhile, in the sense of having a lower asymptotic variance, the AIPWCC estimating equation estimator is more efficient than the others.

Note that, in the sense of having a lower mean squared error, the smoothed estimator is more efficient than the difference between the estimators of quantiles based on the nonparametric maximum likelihood estimator of the survival function. That benefits from the smoother but sacrifices a little unbiasedness.

It is also worthy noting that the proposed methods are appropriate to LBRC data. The stationarity assumption is a precondition to use these estimators. If the stationarity assumption fails, the observations may be LTRC data or RC data. Some approaches appropriate to these data types have been explored in the literature mentioned in Section 1. Note that neglecting the features of LBRC data will lead to lower efficiency of the estimation under LBRC data.

ACKNOWLEDGEMENTS

We are very grateful to the reviewers for the valuable comments on this paper. Xun's work is partially supported by National Natural Science Foundation of China (11701043), Science and Technology Program of Jilin Educational Department during the "13th Five-Year" Plan Period (JJKH20191299KJ), and Scholarship from China Scholarship Council (201808220176). Wang's work is supported by National Natural Science Foundation of China (No. 11871028, 11731015, 11571051, 11501241), Natural Science Foundation of Jilin Province (No. 20180101216JC, 20170101057JC, 20150520053JH), and Program for Changbaishan Scholars of Jilin Province (2015010). Zhou's work is supported by the State Key Program of National Natural Science Foundation of China (71331006, 71931004), the State Key Program in the Major Research Plan of National Natural Science Foundation of China (91546202).

Received 28 July 2018

REFERENCES

[1] Addona, V., Wolfson, D. B. (2006). A formal test for the stationarity of the incidence rate using data from a prevalent cohort study with follow-up. *Lifetime Data Analysis* **12** 267–284. MR2328577

[2] Asgharian, M., M'Lan, C., Wolfson, D. (2002). Length-biased sampling with right censoring. *Journal of the American Statistical Association* **97** 201–209. MR1947280

[3] Asgharian, M., Wolfson, D. (2005). Asymptotic behavior of the unconditional NPMLE of the length-biased survivor function from right censored prevalent cohort data. *Annals of statistics* **33** 2109–2131. MR2211081

[4] Bowman, A., Hall, P., Prvan, T. (1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika* **85**(4) 799–808. MR1666695

[5] de Uña-Álvarez, J. (2004). Nonparametric estimation under length-biased sampling and Type I censoring: a moment based approach. *Annals of the Institute of Statistical Mathematics* **56**(4) 667–681. MR2126804

[6] Jeong, J. H. (2014). *Statistical inference on residual life*, New York: Springer. MR3157047

[7] Jeong, J. H., Jung, S. H., Costantino, J. P. (2008). Nonparametric inference on median residual life function. *Biometrics* **64**(1) 157–163. MR2422830

[8] Jones, M. C., Marron, J. S., Sheather, S. J. (1996). A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association* **91**(433) 401–407. MR1394097

[9] Kaplan, E. L., Meier, P. (1958). Nonparametric estimation from incomplete observation. *Journal of the American Statistical Association* **53** 457–481. MR0093867
[10] Lin, C., Zhou, Y. (2014). Inference for the treatment effects in two sample problems with right-censored and length-biased data. *Statistics and Probability Letters* **90** 17–24. MR3196852
[11] Liu, P., Wang, Y., Zhou, Y. (2015). Quantile residual lifetime with right-censored and length-biased data. *Annals of the Institute of Statistical Mathematics* **67**(5) 999–1028. MR3390176
[12] Newey, W. K. (1990). Semiparametric efficiency bounds. *Journal of Applied Econometrics* **5** 99–135.
[13] Pakes, A., Pollard, D. (1989). Simulation and the Asymptotics of Optimization Estimators. *Econometrica* **57**(5) 1027–1057. MR1014540
[14] Qin, J., Shen, Y. (2010). Statistical methods for analyzing right-censored length-biased data under Cox model. *Biometrics* **66** 382–392. MR2758818
[15] Redelmeier, D. A., Singh, S. M. (2001). Survival in academy award-winning actors and actresses. *Annals of Internal Medicine* **134** 955–962.
[16] Reiss, R. D. (1989). *Approximate distributions of order statistics*, Springer-Verlag, New York. MR0988164
[17] Robins, J. M., Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology-Methodological Issues* N. Jewell, K. Dietz and V. Farewell (eds), 297–331, Birkhauser, Boston.
[18] Sheather, S. J., Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)* **53**(3) 683–690. MR1125725
[19] Shen, J., He, S. (2007). Empirical likelihood for the difference of quantiles under censorship. *Statistical Papers* **48** 437–457. MR2391028
[20] Simon, R. (1980). Length-biased sampling in etiologic studies. *American Journal of Epidemiology* **111** 444–452.
[21] Sylvestre, M., Hustzi, E., Hanley, J. (2006). Do Oscar winners live longer than less successful peers? A reanalysis of the evidence. *Annals of Internal Medicine* **145** 361–363.
[22] Tsiatis, A. A. (2006). *Semiparametric Theory and Missing Data*, Springer-Verlag, New York. MR2233926
[23] Vardi, Y. (1982). Nonparametric estimation in the presence of length bias. *The Annals of Statistics* **10** 616–620. MR0653536
[24] Vardi, Y. (1985). Empirical distributions in selection bias models (with discussion). *The Annals of Statistics* **13** 178–203. MR0773161
[25] Vardi, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika* **76** 751–761. MR1041420
[26] Veraverbeke, N. (2001). Estimation of the quantiles of the duration of old age. *Journal of Statistical Planning and Inference* **98** 101–106. MR1860228
[27] Wang, Y., Liu, P., Zhou, Y. (2015). Quantile residual lifetime for left-truncated and right-censored data. *Science China Mathematics* **58**(6) 1217–1234. MR3344056
[28] Winter, B. B., Földes, A. (1988). A product-limit estimator for use with length-biased data. *The Canadian Journal of Statistics* **16**(4) 337–355. MR1007165
[29] Wolkewitz, M., Allignol, A., Schumacher, M. and Beyersmann, J. (2010). Two pitfalls in survival analysis of time-dependent exposure: a case study in a cohort of Oscar nominees. *The American Statistician* **64** 205–211. MR2757164
[30] Xun, L., Shao, L., Zhou, Y. (2017). Efficiency of estimators for quantile differences with left truncated and right censored data. *Statistics and Probability Letters* **121** 29–36. MR3575406
[31] Xun, L., Zhou, Y. (2017). Estimators and their asymptotic properties for quantile difference with left truncated and right censored data. *Acta Mathematica Sinica (in Chinese)* **60**(3) 451–464. MR3700774
[32] Zelen, M. (2005). Forward and backward recurrence times and length biased sampling: age specific models. *Lifetime Data Analysis* **10** 325–334. MR2125419
[33] Zhang, L., Liu, P., Zhou, Y. (2015). Smoothed estimator of quantile residual lifetime for right censored data. *Journal of Systems Science and Complexity* **28**(6) 1374–1388. MR3428722
[34] Zhou, W., Jing, B (2003). Smoothed empirical likelihood confidence intervals for the difference of quantiles. *Statistica Sinica* **13** 83–95. MR1963921

Li Xun
School of Mathematics and Statistics
Changchun University of Technology
Jilin
China
E-mail address: xunli@ccut.edu.cn

Guangchao Zhang
School of Mathematics and Statistics
Changchun University of Technology
Jilin
China
E-mail address: zgc890@gmail.com

Dehui Wang
School of Mathematics
Jilin University
Jilin
China
E-mail address: wangdh@jlu.edu.cn

Yong Zhou
Faculty of Economics and Management
East China Normal University
Shanghai
China
E-mail address: yzhou@amss.ac.cn