

Research Article

Rank Estimation for Mean Residual Life Transformation Model

Xiaoping Chen 

College of Mathematics and Informatics & FJKLMAA, Fujian Normal University, Fuzhou 350117, China

Correspondence should be addressed to Xiaoping Chen; xpchen@fjnu.edu.cn

Received 20 August 2020; Revised 5 December 2020; Accepted 9 December 2020; Published 21 December 2020

Academic Editor: Chenxi Huang

Copyright © 2020 Xiaoping Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a new and important class of mean residual life regression model, which is called the mean residual life transformation model. The link function is assumed to be unknown and increasing in its second argument, but it is permitted to be not differentiable. The mean residual life transformation model encompasses the proportional mean residual life model, the additive mean residual life model, and so on. Under maximum rank correlation estimation, we present the estimation procedures, whose asymptotic and finite sample properties are established. The consistent variance can be estimated by a resampling method via perturbing the U -statistics objective function repeatedly which avoids the usual sandwich choice. Monte Carlo simulations reveal good finite sample performance and the estimators are illustrated with the Oscar data set.

1. Introduction

The mean residual life function at time t is the remaining life expectancy for those given survival up to t . It is defined as

$$m(t) = E(T - t | T > t) = \frac{\int_t^\infty S(u) du}{S(t)}, \quad (1)$$

where T is the time from the onset to the failure event of interest and $S(t) = P(T > t)$ is the survival function of T . The mean residual life function is often of interest and obviously important in many fields of application, such as industrial reliability study, actuarial research, and survival analysis. For example, when T is the patient's remaining life from cancer to death, $m(t)$ is the remaining life expectancy for those given survival up to t . The patients and physicians both care about how long cancer patients can survive. When $t = 300$ is the minimum score of Graduate Record Examination foreign language test for a school graduate program, $m(t)$ is the average excess score for those more than 300 points. From model (1), $m(0) = E(T)$, and it is easy to induce that the conditional survival function of T given covariate X :

$$S(t|X) = \frac{m(0|X)}{m(t|X)} \exp\left(-\int_0^t \frac{1}{m(u|X)} du\right). \quad (2)$$

How to describe the effect of the factors on the remaining residual life expectancy has aroused much interest (Song et al., Wang et al., Lin et al., and therein [1–3]). For example, they are interested in exploring the relation between the remaining survival time of cancer patients and age, smoking, and various treatments. All kinds of semiparametric statistical models have been widely used for assessing such relation. For instance, in order to concern with the correlation between $m(t|X)$ and $X'\beta_0$, Oakes and Dasu [4] studied the proportional mean residual life model:

$$m(t|X) = m_0(t) \exp(X'\beta_0), \quad (3)$$

where $m_0(t)$ is the unknown baseline hazard function, covariate X is p -dimensional vector, and β_0 is the true parameter. The proportional mean residual life model can compare the ratio when the covariate is a binary indicator. For example,

$$\frac{m(t|Xt = n1)}{m(t|Xt = n0)} = \exp(\beta_0), \quad (4)$$

where $X = 1$, combined treatment, and $X = 0$, no treatment. For instance, $\beta_0 = \log 3$ means that the combined treatment would be 3 times the mean remaining life as compared to no treatment.

In model (3) without censoring, Maguluri and Zhang [5] introduced the estimation method for β_0 . The regression methods for model (3) with right censoring data have been studied by Chen and Cheng [6]. Chen and Cheng [6] proposed the estimating equations which can be used to estimate $m_0(\cdot)$ and β_0 :

$$\begin{aligned} \sum_i m_0(t) (dN_i(t) - Z_i(t) d\Lambda_i(t; \beta, m_0)) &= 0; \\ \sum_i \int X_i(m_0(t) [dN_i(t) - Z_i(t) (\exp(X_i' \beta_0) dt + dm_0(t))]) &= 0. \end{aligned} \quad (5)$$

Sun and Zhao [7] presented a class of mean residual life regression model:

$$m(t|X) = m_0(t)g(X'\beta_0), \quad (6)$$

where $g(\cdot) \geq 0$ is a prespecified link function and is assumed to be continuous almost everywhere and twice differentiable. In model (6) with censoring data, Sun and Zhao [7] proposed the estimating equations which can be used to estimate $m_0(\cdot)$ and β_0 :

$$\begin{aligned} \sum_i m_0(t) dN_i(t) - Z_i(t) (g(X_i' \beta_0) dt + dm_0(t)) &= 0, \\ \sum_i \int \frac{g^{(1)}(X_i' \beta)}{g(X_i' \beta)} X_i \left[\hat{m}_0(t, \beta) dN_i(t) - Z_i(t) \left\{ \frac{1}{g(X_i' \beta) dt + d\hat{m}_0(t, \beta)} \right\} \right] &= 0, \end{aligned} \quad (7)$$

where $g^{(1)}(x) = dg(x)/dx$. They also developed estimation procedures and a goodness-of-fit test for model (6) with right censoring data.

Just like the additive hazards regression model being an attractive alternative to the Cox proportional hazards model, Chen [8] considered the additive mean residual life model:

$$m(t|X) = m_0(t) + X'\beta_0. \quad (8)$$

Chen [8] adapted estimating equations for estimating parameter β_0 . The model (8) explains the relative difference for mean residual life:

$$m(t|X = 1) - m(t|X = 0) = \beta_0. \quad (9)$$

In this paper, we consider a new and important class of mean residual life regression model given by

$$m(t|X) = h(t, X'\beta_0), \quad (10)$$

where link function $h(\cdot, \cdot) > 0$ is assumed to be an unknown and increasing function in its second argument. The function $h(\cdot, \cdot)$ is permitted to be not differentiable. Model (10) is called the mean residual life transformation model. Obviously, the mean residual life transformation model includes the proportional mean residual life model (3) and the additive mean residual life model (8) and partially extends model (6). Model (10) is robust, but traditional methods like the counting process approach cannot be used to estimate the unknown regression parameter β_0 because the estimating equations depend on the differentiation of link function.

In this paper, under maximum rank correlation estimation, we propose the smoothed objective function to estimate unknown regression parameter β_0 for model (10) with complete data and right censoring data. We motivate this paper in three ways. First, the consistencies and asymptotic normalities of such approximated estimators for the general mean residual life model have not been established yet in the literature. Second, the smoothed objective functions make it feasible to adapt the rank-based approach, without loss of asymptotic efficiency. Third, the censoring variable is correlated, not being independent with a covariate.

The rest of this article is organized as follows. In the next section, we give estimation procedures for model (10) with complete data. Section 3 considers the extension for right censoring data. A numerical study of Monte Carlo simulations and an application to the Oscar data are conducted in Section 4 to investigate the finite sample properties. Section 5 discusses future research.

2. Estimation

Without loss of generality, we suppose the survival time T is a positive random variable and T is absolutely continuous with a density $f(t)$ with respect to the Lebesgue measure. In the mean residual life transformation model (10), the link function $h(\cdot, \cdot)$ is unknown; thus, the approach of estimation equations cannot estimate parameter β_0 for model (10). Under mean residual life transformation model (10) with complete data, we present the following smoothed objective function to estimate parameter β_0 :

$$\beta_n = \operatorname{argmax}_{\beta \in \mathcal{B}} \frac{1}{n(n-1)} \sum_{i \neq j} I(T_i > T_j) \times \left(1 + \exp \left\{ -\frac{X_i' \beta - X_j' \beta}{\sigma_n} \right\} \right)^{-1}, \quad (11)$$

where \mathcal{B} is a parameter subset of \mathbb{R}^p , σ_n is a strictly positive and decreasing number satisfying $\lim_{n \rightarrow \infty} \sigma_n = 0$, and I is the indicator function.

To obtain the consistency of β_n , the sufficient and necessary condition of the following proposition is important.

We list four regularity assumptions for rank estimation below. All assumptions are ordinary, which are also used in other literature.

Assumption 1. (Covariates): (a) the support of the regressor $X = (X_1, X_2, \dots, X_p)$ is not contained in a proper linear subspace of \mathbb{R}^p . (b) X 's first component X_1 has an everywhere positive Lebesgue density, conditional on (X_2, \dots, X_p) .

Assumption 2. (Normalization): the parameter space \mathcal{B} is a compact subset of $\{\beta = (\beta_1, \beta_2, \dots, \beta_p) \in \mathbb{R}^p: \beta_1 = 1\}$.

To identify the parameter β_0 , Assumptions 1 and 2 are sufficient conditions. To obtain the asymptotic normality, we list Assumptions 3 and 4, which were presented by Sherman [9].

Assumption 3. Let \mathcal{N} be a neighborhood of true parameter β_0 .

- (i) All mixed second partial derivatives of $\tau(z, \cdot)$ exist on \mathcal{N} , for any z .
- (ii) There exists an integrable function $M(z)$. For any z and $\beta \in \mathcal{N}$, $M(z)$ satisfies

$$\|\nabla_2 \tau(z, \beta) - \nabla_2 \tau(z, \beta_0)\| \leq M(z) |\beta - \beta_0|, \quad (12)$$

where $\|(c_{ij})\| = (\sum_{i,j} c_{ij}^2)^{1/2}$.

- (iii) $E\|\nabla_1 \tau(z, \beta_0)\|^2 < \infty$.
- (iv) $E|\nabla_2 \tau(z, \beta_0)| < \infty$, where $|\nabla_2 \tau(z, \beta) = \sum_{i_1, i_2} |\partial^2 / \partial \beta_{i_1} \partial \beta_{i_2} \tau(z, \beta)|$.
- (v) The $(p-1) * (p-1)$ dimensional matrix $E[\nabla_2 \tau(z, \beta_0)]$ is negative definite.

Assumption 4. Let \mathcal{N} be a neighborhood of true parameter β_0 . Then,

- (i) All mixed second partial derivatives of $\tau_c(z, \cdot)$ exist on \mathcal{N} , for any z .
- (ii) There exists an integrable function $M_c(z)$. For any z and $\beta \in \mathcal{N}$, $M_c(z)$ satisfies

$$\|\nabla_2 \tau_c(z, \beta) - \nabla_2 \tau_c(z, \beta_0)\| \leq M_c(z) |\beta - \beta_0|. \quad (13)$$

- (iii) $E\|\nabla_1 \tau_c(z, \beta_0)\|^2 < \infty$.
- (iv) $E|\nabla_2 \tau_c(z, \beta_0)| < \infty$.
- (v) The $(p-1) * (p-1)$ dimensional matrix $E[\nabla_2 \tau_c(z, \beta_0)]$ is negative definite.

Proposition 1. Under mean residual life transformation model (10) with complete data,

$$P(T_i > T_j | X_i, X_j) > P(T_i < T_j | X_i, X_j), \quad \text{iff } X'_i \beta_0 > X'_j \beta_0. \quad (14)$$

It is known that the interest parameter β_0 is not unique, that is, not identifiable. To avoid identifiability problem and without loss of generality, the first component of β_0 is assumed to be 1, that is, $\beta_0 = (1, \theta'_0)'$. Recall that $Z = (T, X)$ denotes an observation from the distribution P on the set $S \subseteq \mathbb{R} \otimes \mathbb{R}^p$. For each z in S , define

$$\begin{aligned} \tau(z, \beta) = & E[I(t > T)I(x' \beta > X' \beta)] \\ & + E[I(T > t)I(X' \beta > x' \beta)]. \end{aligned} \quad (15)$$

The \sqrt{n} consistency and asymptotic normality of β_n are given by the following Theorem 1.

Proof of Proposition 1

$$\begin{aligned} P(T_i > T_j | X_i, X_j) &= \int_0^\infty P(T_i > t_j) dF_j(t_j) \\ &= \int_0^\infty S_i(t_j) dF_j(t_j) \\ &= \int_0^\infty S_i(t_j) f_j(t_j) dt_j \\ &= \int_0^\infty S_i(t_j) \lambda_j(t_j) S_j(t_j) dt_j, \end{aligned} \quad (16)$$

where the fourth equality holds since the hazard function $\lambda_j(t) = f_j(t)/S_j(t)$. Then,

$$\begin{aligned} P(T_i > T_j | X_i, X_j) - P(T_j > T_i | X_i, X_j) &= \int_0^\infty S_i(t_j) \lambda_j(t_j) S_j(t_j) dt_j - \int_0^\infty S_j(t_i) \lambda_i(t_i) S_i(t_i) dt_i \\ &= \int_0^\infty S_i(t) S_j(t) (\lambda_j(t) - \lambda_i(t)) dt. \end{aligned} \quad (17)$$

Because

$$S(t|X) = \frac{m(0|X)}{m(t|X)} \exp\left(-\int_0^t \frac{1}{m(u|X)} du\right), \quad (18)$$

the derivative of $S(t|X)$ is

$$\begin{aligned} \frac{dS(t|X)}{dX' \beta_0} = & \frac{1}{m^2(t|X)} \left\{ \frac{dm(t|X)}{dX' \beta_0} m(X' X) \exp\left(-\int_0^t \frac{1}{m(u|X)} du\right) + m(t|X) \right. \\ & \cdot \left. \left[\frac{dm(0|X)}{dX' \beta_0} \exp\left(-\int_0^t \frac{1}{m(u|X)} du\right) + m(0|X) \exp\left(-\int_0^t \frac{1}{m(u|X)} du\right) \left(-\int_0^t \frac{1}{m^2(u|X)} \frac{dm(t|X)}{dX' \beta_0} du\right) \right] \right\}. \end{aligned} \quad (19)$$

Because $m(t|X)$ is an increasing function in $X' \beta_0$, that is, $dm(t|X)/dX' \beta_0 \geq 0$, then

$$\frac{dS(t|X)}{dX' \beta_0} \geq 0. \quad (20)$$

Thus, $S(t|X)$ is an increasing function of $X' \beta_0$. Because $S_i(t) = \exp\{-\Lambda_i(t)\}$, where $\Lambda_i(t) = \int_0^t \lambda_i(u) du$ is a cumulative hazard function. Hence, $\lambda_i(t)$ is a decreasing function of $X'_i \beta_0$, that is,

$$\lambda_j(t) > \lambda_i(t) \quad \text{if and only if } X'_i\beta_0 > X'_j\beta_0. \quad (21)$$

Therefore,

$$P(T_i > T_j | X_i, X_j) > P(T_j > T_i | X_i, X_j), \quad \text{if and only if } X'_i\beta_0 > X'_j\beta_0. \quad (22)$$

□

Theorem 1. Under Assumptions 1 and 2,

$$\beta_n \longrightarrow p\beta_0, \quad (23)$$

where $\longrightarrow p$ denotes convergence in probability. Under Assumptions 1-3,

$$\sqrt{n}(\beta_n - \beta_0) \longrightarrow D \begin{pmatrix} 0 \\ W \end{pmatrix}, \quad (24)$$

where $\longrightarrow D$ denotes convergence in distribution and W has a normal distribution $N(0, V^{-1}\Delta V^{-1})$ with

$$\begin{aligned} \Delta &= E\nabla_1\tau(\cdot, \beta_0)[\nabla_1\tau(\cdot, \beta_0)]', \\ 2V &= E\nabla_2\tau(\cdot, \beta_0). \end{aligned} \quad (25)$$

From theorem 1, we find that the asymptotic variance of β_n is $V^{-1}\Delta V^{-1}$, which is equivalent to nonsmooth. Also, the smooth estimation β_n has another advantage, which avoids the situation that the estimation of β_0 involves maximizing indicator function. The asymptotic variance of β_n can be estimated by a resampling method, which was studied by Cai et al. [10]. For each z in S , let the mean and variance of the positive random variable U both be 1. Define

$$\begin{aligned} \beta_n^* &= \arg \max_{\beta \in \mathcal{B}} \frac{1}{n(n-1)} \sum_{i \neq j} U_i U_j I(t_i > t_j) \\ &\times \left(1 + \exp \left\{ -\frac{x'_i\beta - x'_j\beta}{\sigma_n} \right\} \right)^{-1}. \end{aligned} \quad (26)$$

Theorem 2. Under Assumptions 1-3,

$$\sqrt{n}(\beta_n^* - \beta_n) \longrightarrow D \begin{pmatrix} 0 \\ W^* \end{pmatrix}, \quad (27)$$

where W^* follows $N(0, V^{-1}\Delta V^{-1})$ and V and Δ are given in Theorem 1.

From theorem 2, we find that the asymptotic variance of β_n can be approximated by the sample variance of β_n^* . The proof of Theorem 2 is omitted because we can use similar proof as Cai et al. [10].

3. Extension

For right censoring data, we observe the data $Z_i = (V_i, \delta_i, X_i)$, where $\delta_i = I(T_i \leq C_i)$, $V_i = T_i I(\delta_i = 1) + C_i I(\delta_i = 0)$. C_i are right censoring random variables, $i = 1, \dots, n$.

Under model (10) with right censoring data, we present the following smoothed objective function to estimate parameter β_0 :

$$\tilde{\beta} = \arg \max_{\beta \in \mathcal{B}} \tilde{G}_n(\beta), \quad (28)$$

where

$$\tilde{G}_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} \delta_j I(V_i > V_j) \times \left(1 + \exp \left\{ -\frac{X'_i\beta - X'_j\beta}{\sigma_n} \right\} \right)^{-1}. \quad (29)$$

\mathcal{B} is a parameter subset of \mathbb{R}^p , and σ_n is a strictly positive and decreasing number satisfying $\lim_{n \rightarrow \infty} \sigma_n = 0$. The objective function (28) is inspired by Khan and Tamer [11] and Song et al. [12].

For obtaining the consistency of $\tilde{\beta}$, the identification (i.e., uniquely maximum) is an important condition. As Khan and Tamer [11] proposed, define

$$\begin{aligned} V_{1i} &= \delta_i V_i + (1 - \delta_i)(+\infty), \\ V_{0i} &= V_i. \end{aligned} \quad (30)$$

Proposition 2. Under mean residual life transformation model (10) with censoring data,

$$P(V_{0i} > V_{1j} | X_i, X_j) > P(V_{1i} < V_{0j} | X_i, X_j), \quad \text{iff } X'_i\beta_0 > X'_j\beta_0. \quad (31)$$

Recall that $Z = (V, \delta, X)$ denotes an observation from the distribution P on the set $S \subseteq \mathbb{R} \otimes \mathbb{R} \otimes \mathbb{R}^p$. For each z in S , define

$$\begin{aligned} \tau_c(z, \beta) &= E[\delta I(v > V)I(x'\beta > x'\beta)] \\ &+ E[\delta I(V > v)I(x'\beta > x'\beta)]. \end{aligned} \quad (32)$$

The \sqrt{n} consistency and asymptotic normality of $\tilde{\beta}$ are given by Theorem 3.

Proof of Proposition 2

First of all, we calculate the conditional expectation of $V_{0j} > V_{1i}$ given (X_i, X_j, C_i, C_j) , which can be decomposed as two expressions:

$$\begin{aligned}
 &P(V_{0j} > V_{1i} | X_i = x_i, X_j = x_j, C_i = c_i, C_j = c_j) \\
 &= P(V_{0j} > V_{1i}, \delta_i = 1 | x_i, x_j, c_i, c_j) + P(V_{0j} > V_{1i}, \delta_i = 0 | x_i, x_j, c_i, c_j) \\
 &= P(V_{0j} > V_{1i}, \delta_i = 1 | x_i, x_j, c_i, c_j) \\
 &= P(V_{0j} > V_{1i}, \delta_i = 1, \delta_j = 1 | x_i, x_j, c_i, c_j) + P(V_{0j} > V_{1i}, \delta_i = 1, \delta_j = 0 | x_i, x_j, c_i, c_j) \\
 &\stackrel{\text{def}}{=} I_1 + I_2.
 \end{aligned} \tag{33}$$

The two above expressions are calculated as follows:

$$\begin{aligned}
 I_1 &= P(T_j > T_i, T_i \leq C_i, T_j \leq C_j | x_i, x_j, c_i, c_j) \\
 &= \int_0^{c_i \wedge c_j} \int_{t_i}^{c_j} dF_j(t_j) dF_i(t_i) \\
 &= \int_0^{c_i \wedge c_j} (F_j(c_j) - F_j(t_i)) dF_i(t_i) \\
 &= F_j(c_j) F_i(c_i \wedge c_j) - \int_0^{c_i \wedge c_j} F_j(t_i) dF_i(t_i),
 \end{aligned} \tag{34}$$

$$\begin{aligned}
 I_2 &= P(C_j > T_i, T_i \leq C_i, T_j > C_j | x_i, x_j, c_i, c_j) \\
 &= P(T_i < C_i \wedge C_j, T_j > C_j | x_i, x_j, c_i, c_j) \\
 &= F_i(c_i \wedge c_j) (1 - F_j(c_j)).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 I_1 + I_2 &= F_i(c_i \wedge c_j) - \int_0^{c_i \wedge c_j} F_j(t_i) dF_i(t_i) \\
 &= \int_0^{c_i \wedge c_j} (1 - F_j(t_i)) dF_i(t_i) \\
 &= \int_0^{c_i \wedge c_j} S_j(t_i) dF_i(t_i).
 \end{aligned} \tag{35}$$

Similarly, the conditional expectation of $V_{0i} > V_{1j}$ given (X_i, X_j, C_i, C_j) is

$$P(V_{0i} > V_{1j} | x_i, x_j, c_i, c_j) = \int_0^{c_i \wedge c_j} S_i(t_j) dF_j(t_j). \tag{36}$$

The rest of the proof of Proposition 2 is the same as Proposition 1. \square

Theorem 3. Under Assumptions 1 and 2,

$$\tilde{\beta} \longrightarrow p\beta_0, \tag{37}$$

where $\longrightarrow p$ denotes convergence in probability. Under Assumptions 1, 2, and 4,

$$\sqrt{n}(\tilde{\beta} - t\beta_0) \longrightarrow D \begin{pmatrix} 0 \\ \tilde{W} \end{pmatrix}, \tag{38}$$

where $\longrightarrow D$ denotes convergence in distribution and W has a normal distribution $N(0, V_c^{-1} \Delta_c V_c^{-1})$ with

$$\begin{aligned}
 \Delta_c &= EV_1 \tau_c(\cdot, \beta_0) [\nabla_1 \tau_c(\cdot, \beta_0)]', \\
 2V_c &= EV_2 \tau_c(\cdot, \beta_0).
 \end{aligned} \tag{39}$$

From Theorem 3, we find that the asymptotic variance of $\tilde{\beta}$ is $V_c^{-1} \Delta_c V_c^{-1}$, which is equivalent to nonsmooth. Also, the smooth estimation $\tilde{\beta}$ has another advantage, which avoids the situation that the estimation of β_0 involves maximizing indicator function. The asymptotic variance of $\tilde{\beta}$ can be estimated by a resampling method, which was studied by Cai et al. [10]. For each z in S , let the mean and variance of the positive random variable U both be 1. Define

$$\tilde{\beta}^* = \arg \max_{\beta \in \mathcal{B}} \tilde{G}_n^*(\beta), \tag{40}$$

where

$$\begin{aligned}
 \tilde{G}_n^*(\beta) &= \frac{1}{n(n-1)} \sum_{i \neq j} U_i U_j \delta_j I(v_i > v_j) \\
 &\times \left(1 + \exp \left\{ -\frac{x'_i \beta - x'_j \beta}{\sigma_n} \right\} \right)^{-1}.
 \end{aligned} \tag{41}$$

Proof of Theorem 3

We consider the nonsmoothly objective function of β :

$$\begin{aligned}
 \tilde{\beta}_{\text{non}} &= \arg \max_{\beta \in \mathcal{B}} \frac{1}{n(n-1)} \sum_{i \neq j} I(V_{0i} > V_{1j}) I(X'_i \beta > X'_j \beta) \\
 &= \arg \max_{\beta \in \mathcal{B}} \frac{1}{n(n-1)} \sum_{i \neq j} \delta_j I(T_i > T_j) I(X'_i \beta > X'_j \beta).
 \end{aligned} \tag{42}$$

Define

$$G_n(\beta) = \frac{1}{n(n-1)} \sum_{i \neq j} \delta_j I(T_i > T_j) I(X'_i \beta > X'_j \beta); \tag{43}$$

$$G(\beta) = E[G_n(\beta)].$$

Then,

$$\begin{aligned}
 G_n(\beta) - G(\beta) &= \frac{1}{n(n-1)} \sum_{i \neq j} \delta_j I(T_i > T_j) I(X'_i \beta > X'_j \beta) \\
 &\quad - E(\delta_2 I(T_1 > T_2) I(X'_1 \beta > X'_2 \beta)) \\
 &\equiv \frac{1}{n(n-1)} \sum_{i \neq j} f(z_i, z_j, \beta),
 \end{aligned} \tag{44}$$

that is,

$$f(z_i, z_j, \beta) = \delta_j I(T_i > T_j) I(X'_i \beta > X'_j \beta) - E(\delta_2 I(T_1 > T_2) I(X'_1 \beta > X'_2 \beta)). \quad (45)$$

First of all, we prove the consistency of $\tilde{\beta}_{\text{non}}$.

Identification ($G(\beta)$ has unique maximum at β_0) follows from Proposition 2 and Assumption 1. We omit the details, which follow from identical steps used in proving Theorem 2.1 (Khan and Tamer [11]).

Obviously, $G_n(\beta) - G(\beta)$ is a degenerated U -statistic of order 2. Because the objective function is composed of an indicator, $\{f(\cdot, \cdot, \beta), \beta \in \mathcal{B}\}$ is Euclidean with an envelope. We can derive from Corollary 7 of Sherman [9] that

$$\sup_{\beta \in \mathcal{B}} |G_n(\beta) - G(\beta)| = O_p\left(\frac{1}{\sqrt{n}}\right). \quad (46)$$

Note that the conclusion is stronger than $\sup_{\beta \in \mathcal{B}} |G_n(\beta) - G(\beta)| = o_p(1)$.

By Assumption 1, it is easy to obtain that $P(X'_1 \beta = X'_2 \beta) = 0$, where $X'_1 \neq X'_2$. Then,

$$\lim_{\tilde{\beta}_{\text{non}} \rightarrow \beta} \{I(X'_1 \tilde{\beta}_{\text{non}} > X'_2 \tilde{\beta}_{\text{non}}) - I(X'_1 \beta > X'_2 \beta)\} = 0, \quad (47)$$

with probability one. Thus, continuity ($G(\beta)$ is continuous about β) follows directly from the dominated convergence theorem.

Because of Assumption 2 (compact parameter space), identification, uniform convergence, and continuity, we obtain $\tilde{\beta}_{\text{non}} \rightarrow p\beta_0$.

In order to obtain $\tilde{\beta} \rightarrow p\beta_0$, it suffices to show that $\sup_{\beta \in \mathcal{B}} |\tilde{G}_n(\beta) - G_n(\beta)| \rightarrow p0$. For any $\epsilon > 0$,

$$\begin{aligned} |\tilde{G}_n(\beta) - G_n(\beta)| &= \frac{1}{n(n-1)} \left| \sum_{i \neq j} \delta_j I(T_i > T_j) \left(I(X'_i \beta > X'_j \beta) - \left(1 + \exp\left\{-\frac{X'_i \beta - X'_j \beta}{\sigma_n}\right\} \right)^{-1} \right) \right| \\ &\leq \frac{1}{n(n-1)} \sum_{i \neq j} I(X'_i \beta - X'_j \beta \geq \epsilon) \left| \delta_j I(T_i > T_j) \left(I(X'_i \beta > X'_j \beta) - \left(1 + \exp\left\{-\frac{X'_i \beta - X'_j \beta}{\sigma_n}\right\} \right)^{-1} \right) \right| \\ &\quad + \frac{1}{n(n-1)} \sum_{i \neq j} I(X'_i \beta - X'_j \beta < \epsilon) \left| \delta_j I(T_i > T_j) \left(I(X'_i \beta > X'_j \beta) - \left(1 + \exp\left\{-\frac{X'_i \beta - X'_j \beta}{\sigma_n}\right\} \right)^{-1} \right) \right|. \end{aligned} \quad (48)$$

Note that

$$\begin{aligned} &\left| I(X'_i \beta > X'_j \beta) - \left(1 + \exp\left\{-\frac{X'_i \beta - X'_j \beta}{\sigma_n}\right\} \right)^{-1} \right| I(X'_i \beta - X'_j \beta \geq \epsilon) \\ &\leq \exp\left\{-\frac{|X'_i \beta - X'_j \beta|}{\sigma_n}\right\} I(X'_i \beta - X'_j \beta \geq \epsilon) \\ &\leq \exp\left\{-\frac{\epsilon}{\sigma_n}\right\}. \end{aligned} \quad (49)$$

Under Theorem 7 of Nolan and Pollard [13], it is easy to obtain that

$$\frac{1}{n(n-1)} \sum_{i \neq j} I(X'_i \beta - X'_j \beta < \epsilon) \rightarrow E(I(X'_1 \beta - X'_2 \beta < \epsilon)) a.s. \quad (50)$$

According to Assumption 1, $E(I(X'_1 \beta - X'_2 \beta < \epsilon)) \rightarrow 0$ as $\epsilon \rightarrow 0$. Hence,

$$\sup_{\beta \in \mathcal{B}} |\tilde{G}_n(\beta) - G_n(\beta)| \rightarrow p0. \quad (51)$$

Define

$$\begin{aligned} \tau_s(z, \beta) &= E \left[\delta I(t > T) \left(1 + \exp\left\{-\frac{x' \beta - X' \beta}{\sigma_n}\right\} \right)^{-1} \right] \\ &\quad + E \left[\delta I(T > t) \left(1 + \exp\left\{-\frac{X' \beta - x' \beta}{\sigma_n}\right\} \right)^{-1} \right]. \end{aligned} \quad (52)$$

Because $\tilde{G}_n(\beta)$ is an U -statistic of order 2, it is easy to obtain that by the Hoeffding decomposition:

$$\tilde{G}_n(\beta) - \tilde{G}_n(\beta_0) = \Gamma_n(\beta) + P_n f_{1n}(\cdot, \beta) + U_n^2 f_{2n}(\cdot, \cdot, \beta). \quad (53)$$

Similarly to the proof of Theorem 4 of Sherman [9], we obtain

$$\begin{aligned}\Gamma_n(\beta) &= \frac{1}{2}(\beta - \beta_0)' \frac{E\nabla_2\tau_s(z, \beta_0)}{2}(\beta - \beta_0) + o_p(|\beta - \beta_0|^2); \\ P_n f_{1n}(\cdot, \beta) &= \frac{1}{\sqrt{n}}(\beta - \beta_0)' \sqrt{n}P_n \nabla_1 \tau_s(\cdot, \beta_0) + o_p(|\beta - \beta_0|^2); \\ U_n^2 f_{2n}(\cdot, \beta) &= o_p\left(\frac{1}{n}\right), \text{ uniformly in } o_p(1) \text{ neighborhoods of } \beta_0.\end{aligned}\quad (54)$$

According to Assumption 4, we obtain

$$\begin{aligned}\nabla_1 \tau_s(z, \beta) &= \nabla_1 \tau_c(z, \beta) + O_p(\sigma_n); \\ \nabla_2 \tau_s(z, \beta) &= \nabla_2 \tau_c(z, \beta) + O_p(\sigma_n).\end{aligned}\quad (55)$$

Hence, we obtain

$$\begin{aligned}\tilde{G}_n(\beta) - \tilde{G}_n(\beta_0) &= \frac{1}{2}(\beta - \beta_0)' V_c(\beta - \beta_0) + \frac{1}{\sqrt{n}}(\beta - \beta_0)' W_n \\ &\quad + o_p(|\beta - \beta_0|^2) + o_p\left(\frac{1}{n}\right),\end{aligned}\quad (56)$$

uniformly in $o_p(1)$ neighborhoods of β_0 , where $W_n \rightarrow dN(0, \Delta_c)$, $V_c = E\nabla_2\tau_c(z, \beta_0)/2$,

$$\Delta_c = E\nabla_1\tau_c(z, \beta_0)[\nabla_1\tau_c(z, \beta_0)]'. \quad (57)$$

Then, by Theorem 2 of Sherman [9],

$$\sqrt{n}(\tilde{\beta} - t\beta_0) \rightarrow D\begin{pmatrix} 0 \\ \tilde{W} \end{pmatrix}, \quad (58)$$

where $\rightarrow D$ denotes convergence in distribution and W has a normal distribution $N(0, V_c^{-1}\Delta_c V_c^{-1})$. \square

Theorem 4. Under Assumptions 1, 2, and 4,

$$\sqrt{n}(\tilde{\beta}^* - \tilde{\beta}) \rightarrow D\begin{pmatrix} 0 \\ W_c^* \end{pmatrix}, \quad (59)$$

where W_c^* follows $N(0, V_c^{-1}\Delta_c V_c^{-1})$ and V_c and Δ_c are given in Theorem 3.

4. Numerical Examples

4.1. Monte Carlo Results. This section aims to conduct the finite sample properties of parameter estimation β_n for model (10) with complete data and $\tilde{\beta}$ for model (10) with right censoring data by a number of simulation studies. Simulation results are reported for $n=100$ and 400 observations, with 500 replications. We choose the smoothed parameter $\sigma_n = o(n^{-1/2})$. Weighted random variables U follow a standard exponential distribution with 500 replications. The empirical bias (BIAS), empirical standard deviation (SD), an average of the estimated standard errors (SE), and empirical coverage probability of 95% confidence interval (CP) are reported in Tables 1–4.

TABLE 1: Simulation results of β_n and β_{com} for the additive model.

Size	$m_0(t)$	Method	β			
			BIAS	SD	SE	CP
$n = 100$	$t + 1$	β_n	0.1094	1.5864	1.5091	92.8
	$t + 1$	β_{com}	0.2474	1.4522	1.3145	85.2
$n = 100$	1	β_n	0.0861	0.6231	0.6576	97.0
	1	β_{com}	0.1762	0.5614	0.4183	81.2
$n = 400$	$t + 1$	β_n	0.0634	0.4736	0.4581	93.0
	$t + 1$	β_{com}	0.1504	0.3967	0.3842	93.2
$n = 400$	1	β_n	0.0281	0.1651	0.1872	96.4
	1	β_{com}	0.0794	0.1147	0.1711	90.4

Our design for model (10) is specified by

$$m(t|X) = h(t, X_1 + X_2\beta), \quad (60)$$

where $\beta = 1$ and $h(t, X_1 + X_2\beta)$ is assumed to be an increasing function in $X_1 + X_2\beta$. The following subsections consider three functional forms for link function $h(\cdot, \cdot)$ and different covariates $X = (X_1, X_2)$.

4.1.1. Complete Data. First, we want to evaluate the performance of parameter estimation β_n for model (10) with complete data. The first link function $h(\cdot, \cdot)$ and different covariates $X = (X_1, X_2)$ which we consider are as follows:

$$\text{Additive model } m(t|X) = h(t, X_1 + X_2\beta) = m_0(t) + X_1 + X_2\beta, \quad (61)$$

where $\beta = 1$, $m_0(t) = t + 1$ or 1, X_1 is generated from a uniform $U(0, 1)$, X_2 follows a chi-square distribution with 1 degree of freedom, and X_1 is independent of X_2 . According to the conditional distribution function $F(t|X) = 1 - S(t|X)$, equation (2) and model (53), then $T = F^{-1}(U)$.

Let β_{com} denote the parameter estimation for additive mean residual life model (53) with complete data proposed by Chen [8]. Table 1 shows the simulation results of β_n and β_{com} for the additive model (53) with complete data. The biases and standard errors of parameter estimation β_n and β_{com} decrease as sample size from 100 to 400. The biases and standard errors of parameter estimation β_n perform well as sample size $n = 400$. The average of estimated standard errors of β_n are close to empirical standard deviations, and the empirical coverage probabilities are approximately 95%. The biases of parameter estimation β_n are smaller than β_{com} . The standard errors of β_n are larger than β_{com} ; it is probably because that mean residual life transformation model (11) is more general than the additive model (8). The results of Table 1 demonstrate that the parameter estimation β_n for the mean residual life transformation model (10) with complete data is robust and reasonable.

4.1.2. Censoring Data. Second, we want to evaluate the performance of parameter estimation $\tilde{\beta}$ for model (10) with right censoring data. The second link function $h(\cdot, \cdot)$ and

different covariates $X = (X_1, X_2)$ which we consider are as follows:

$$\begin{aligned} \text{proportional model } m(t|X) &= h(t, X_1 + X_2\beta) \\ &= m_0(t)\exp(X_1 + X_2\beta), \end{aligned} \quad (62)$$

where $\beta = 1$, $m_0(t) = 1$, X_1 follows a binary distribution of 0 or 1 with equal probabilities of 0.5, X_2 follows a standard normal distribution, and X_1 is independent of X_2 . Similar to the above section, we calculate survival time T for model (61). The right censoring variable C follows an exponential distribution, where the exponential parameter is selected to yield the censoring rate which is approximately 10%, 30%, and 60%, respectively.

Let β_{rc} be the parameter estimation for the proportional mean residual life model (3) with right censoring data being studied by Chen and Cheng [6]. The simulation results of $\tilde{\beta}$ and β_{rc} for the proportional model with right censoring data are shown in Table 2. The mean biases and standard errors of $\tilde{\beta}$ and β_{rc} increase when the censoring rate increases or sample size decreases. The average of estimated standard errors of $\tilde{\beta}$ is close to empirical standard deviations, and the empirical coverage probabilities are approximately 95%. The biases of parameter estimation $\tilde{\beta}$ are smaller than β_{rc} . The standard errors of $\tilde{\beta}$ are larger than β_{rc} ; it is probably because that mean residual life transformation model (10) is more general than the proportional mean residual life model (3). The results of Table 2 demonstrate that the proposed estimator $\tilde{\beta}$ is comparable with β_{rc} .

4.1.3. Nondifferential Link Function. Third, we want to evaluate the performance of parameter estimation $\tilde{\beta}$ for model (10) with right censoring data and nondifferential link function. The third link function $h(\cdot, \cdot)$ and different covariates $X = (X_1, X_2)$ which we consider are as follows:

$$m(t|X) = h(t, X_1 + X_2\beta) = m_0(t)q(X_1 + X_2\beta), \quad (63)$$

where $\beta = 1$, $m_0(t) = 1$, $q(s) = \log(1 + e^s)I(s < 3) + e^s I(s \geq 3)$, X_1 follows a chi-square distribution with 1 degree of freedom, X_2 is generated from a uniform $U(0, 4)$, and X_1 is independent of X_2 .

The simulation results for model (62) with right censoring data and nondifferential link function are summarized in Table 3. Table 3 shows that $\tilde{\beta}$ performs well in finite sample size $n = 100$ and 400. The results are in accordance with the theory; that is, $\tilde{\beta}$ is asymptotically unbiased and normal. The mean biases and standard errors of $\tilde{\beta}$ decrease as sample size n increases. The 95% confidence intervals of $\tilde{\beta}$ are close to their nominal levels. With the censoring rate increasing, the mean biases and standard errors of $\tilde{\beta}$ increase. From Table 3, we find that $\tilde{\beta}$ is suitable for the link function $h(\cdot, \cdot)$ which is not differentiable. Table 3 demonstrates that $\tilde{\beta}$ is reasonable for model (10) with right censoring data and nondifferential link function.

4.2. Application to Oscar Data. Redelmeier and Singh [14] compiled the Oscar Awards from 1929 to 2000 which list 766

TABLE 2: Simulation results of $\tilde{\beta}$ and β_{rc} for the proportional model with right censoring data.

Size	Rate	Method	β			
			BIAS	SD	SE	CP
$n = 100$	10%	$\tilde{\beta}$	0.0183	0.1850	0.1946	96.2
	30%	$\tilde{\beta}$	0.0274	0.2239	0.2118	93.8
	60%	$\tilde{\beta}$	0.0382	0.2708	0.2438	93.4
$n = 100$	10%	β_{rc}	-0.0151	0.1252	0.1145	88.2
	30%	β_{rc}	-0.0671	0.1310	0.1225	86.2
	60%	β_{rc}	-0.2322	0.1797	0.1359	61.4
$n = 400$	10%	$\tilde{\beta}$	-0.0029	0.1037	0.1017	94.6
	30%	$\tilde{\beta}$	0.0093	0.1218	0.1283	95.4
	60%	$\tilde{\beta}$	0.0208	0.1641	0.1691	95.8
$n = 400$	10%	β_{rc}	-0.0085	0.0577	0.0535	91.6
	30%	β_{rc}	-0.0404	0.0777	0.0574	83.6
	60%	β_{rc}	-0.1565	0.1147	0.0673	55.0

TABLE 3: Simulation results for model (62).

Size	Rate	BIAS	β		
			SD	SE	CP
$n = 100$	10%	0.0101	0.0872	0.1182	96.8
	30%	0.0147	0.1287	0.1463	96.4
	60%	0.0268	0.1396	0.1538	97.2
$n = 400$	10%	0.0032	0.0719	0.0682	94.4
	30%	0.0094	0.0801	0.0834	96.2
	60%	0.0138	0.1246	0.1369	95.4

nominees. We treat the data set as right-censored data. Because 327 died before the end of 2001, hence the censoring ratio is about 57.3%. Han et al. [15] elaborated the description of the Oscar data. Whether winning an Oscar Award increases life expectancy is an interesting topic.

The number of total films in career, number of times the person was nominated for an Oscar (Num.), and whether the person won an Oscar (Id.) are chosen as covariates. We have intuitive thinking that there is a positive correlation between life expectancy and the number of total films in a career. To identify the parameter, we assume the coefficient of the number of total films in a career is 1. Random variable U follows a standard exponential distribution with 500 replications, and we choose $\sigma_n = o(n^{-1/2})$. Because the survival time of nominees belongs to censoring data, we exploit the parameter estimator $\tilde{\beta}$ for mean residual life transformation model (10) and β_{rc} for proportional mean residual life model (3) with censoring data.

From Table 4, the coefficients of whether the person won an Oscar for $\tilde{\beta}$ and β_{rc} are positive, and the 95% confidence intervals of whether the person won an Oscar contain zero. So, we can deduce that a performer winning Oscar may not have a longer life span than those without winning.

5. Discussion

The mean residual life $m(t)$ is often of interest and obviously important in many fields of application, such as industrial reliability study, actuarial research, and survival analysis. This paper presents a general class of mean residual life

TABLE 4: Results for Oscar data.

Method	Covariate	Coef.	\widehat{SE}	p value	95% CI
$\tilde{\beta}$	Num.	1.0017	0.2126	2.4628e-006	(0.5849, 1.4184)
β	Id.	0.3102	0.2603	0.2334	(-0.2000, 0.8204)
β_{rc}	Num.	1.3791	0.1837	6.0366e-14	(1.0190, 1.7391)
β_{rc}	Id.	0.5928	0.3791	0.1178	(-0.1502, 1.3358)

regression models, which contain the proportional mean residual life models and partially extends a class of mean residual life regression models studied by Sun and Zhao (2010). The link function is assumed to be unknown and increasing in its seconding argument, but it is permitted to be not differentiable. Traditional methods like the counting process approach cannot be used to estimate the unknown regression parameter β_0 . Under maximum rank correlation estimation, we propose the estimation procedures for model (10) with complete data and right censoring data.

In addition to right censoring data, length-biased data often appear in survival analysis. For example, in unemployment data surveys, short-term unemployment does not register with the unemployment center. In the modeling process, if the length-biased data is not corrected, the duration of unemployment will be overestimated. Therefore, one interesting aspect is the parameter estimation of model (10) with length-biased data. Second, it is possible to construct an estimator for the link function. We leave these possible extensions for future research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (11601083 and U1805263) and the Program for Probability and Statistics: Theory and Application (IRTL1704), and Innovative Research Team in Science and Technology in Fujian Province University (IRTSTFJ).

References

[1] H. J. He, D. Pan, X. Y. Song, and L. Sun, "Additive mean residual life model with latent variables under right censoring," *Statistica Sinica*, vol. 29, no. 1, pp. 47-66, 2019.

[2] D. Xu and Y. Zhou, "Proportional mean residual life model with varying coefficients for length-biased and right-censored data," *Acta Mathematica Sinica, English Series*, vol. 36, no. 5, pp. 578-596, 2020.

[3] H. Pourjafar and V. Zardasht, "Estimation of the mean residual life function in the presence of measurement errors," *Communications in Statistics - Simulation and Computation*, vol. 49, no. 2, pp. 532-555, 2020.

[4] D. Oakes and T. Dasu, "A note on residual life," *Biometrika*, vol. 77, no. 2, pp. 409-410, 1990.

[5] G. Maguluri and C.-H. Zhang, "Estimation in the mean residual life regression model," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 56, no. 3, pp. 477-489, 1994.

[6] Y. Q. Chen and S. Cheng, "Semiparametric regression analysis of mean residual life with censored survival data," *Biometrika*, vol. 92, no. 1, pp. 19-29, 2005.

[7] L. Sun and Q. Zhao, "A class of mean residual life regression models with censored survival data," *Journal of Statistical Planning and Inference*, vol. 140, no. 11, pp. 3425-3441, 2010.

[8] Y. Q. Chen, "Additive expectancy regression," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 153-166, 2007.

[9] R. P. Sherman, "The limiting distribution of the maximum rank correlation estimator," *Econometrica*, vol. 61, no. 1, pp. 123-138, 1993.

[10] T. Cai, L. Tian, and L. J. Wei, "Semiparametric Box-Cox power transformation models for censored survival observations," *Biometrika*, vol. 92, no. 3, pp. 619-632, 2005.

[11] S. Khan and E. Tamer, "Partial rank estimation of duration models with general forms of censoring," *Journal of Econometrics*, vol. 136, no. 1, pp. 251-280, 2007.

[12] X. Song, S. Ma, J. Huang, and X.-H. Zhou, "A semiparametric approach for the nonparametric transformation survival model with multiple covariates," *Biostatistics*, vol. 8, no. 2, pp. 197-211, 2007.

[13] D. Nolan and D. Pollard, "\$US\$-Processes: rates of convergence," *The Annals of Statistics*, vol. 15, no. 2, pp. 780-799, 1987.

[14] D. A. Redelmeier and S. M. Singh, "Survival in academy award-winning actors and actresses," *Annals of Internal Medicine*, vol. 134, no. 10, pp. 955-962, 2001.

[15] X. Han, D. S. Small, D. P. Foster, and V. Patel, "The effect of winning an Oscar award on survival: correcting for healthy performer application to a prevalent cohort study of dementia," *The Annals of Applied Statistics*, vol. 5, no. 2A, pp. 946-957, 2011.