



# Article Generalized Accelerated Failure Time Models for Recurrent Events

Xiaoyi Wen<sup>1</sup> and Jinfeng Xu<sup>2,\*</sup>

- <sup>1</sup> The Institute of Statistics and Big Data, Renmin University of China, Beijing 100872, China; wenxy123@ruc.edu.cn
- <sup>2</sup> School of Mathematics, Yunnan Normal University, Kunming 650092, China
- \* Correspondence: 204203@ynnu.edu.cn

**Abstract:** For analyzing recurrent event data, we consider a generalization of the classical accelerated failure time model. In the proposed approach, the general function is no longer assumed to be a singleton but allowed to be time-varying. This is in the same spirit as in quantile regression and the counting process techniques can be utilized. Theoretical properties such as consistency and asymptotic normality are obtained. The illustration of the methodology using simulation studies and then the application to the bladder cancer data is also given.

**Keywords:** accelerated failure time model; censored quantile regression; counting processes; recurrent events; time-varying general function

MSC: 46N30; 65C60



Citation: Wen, X.; Xu, J. Generalized Accelerated Failure Time Models for Recurrent Events. *Mathematics* 2022, 10, 2662. https://doi.org/10.3390/ math10152662

Academic Editors: Min Wang, Haijun Gong, Liucang Wu and Songfeng Zheng

Received: 29 June 2022 Accepted: 25 July 2022 Published: 28 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

# 1. Introduction

Recurrence event data refer to the situation in which events of interest occur repeatedly over time, and these are wildly used in science and technology, especially in medical research, e.g., epileptic seizures, tumor recurrences, or asthma attack. When the investigators studied the recurrent events, they are often interested in the estimation of covariates on recurrent event times, which can help them perform further predictions. The establishment of models can be approached in many ways; meanwhile, the concept of intensity functions and the counting process are also useful. The nonparametric method used to generalize the intensity estimator for the censored failure time data, independently started by [1,2], also denoted the Nelson–Aalen estimator, is one of the methods widely used by investigators. Refs. [3,4], among others, also studied multiplicative models for the rate and mean functions, an approach which can be used with more general models, including regression. Ref. [5] proposed a generalization of the accelerated failure time model, the so-called accelerated recurrence time model, which resembles quantile regression as well as allows for the evolution of covariate effects.

In this paper, we also extend the quantile regression approach and utilize the counting process techniques. For a time-to-event response *T*, Ref. [6] proposed that a form may assume  $Q_T(\tau|Z) = \exp\{X^T\beta_0\}$  for  $\tau \in (0,1)$ , where  $Q_T(\tau|Z) = \inf\{t : Pr(T \le t|Z) \ge \tau\}$ , so  $Q_T(\tau|Z)$  can also be seen as the  $\tau$ -th quantile of T given that the covariates *Z* and  $X = (1, Z^T)^T$ . Here,  $T, \tau$  and t denote the time to event, the quantile level, and a non-negative real number, respectively. Since its introduction, quantile regression has been widely used and researched, mainly in survival analysis. Ref. [7] extended the LAD estimation method to more general quantiles, and can also improve efficiency when the error terms are identically distributed. Ref. [5] developed a new quantile regression approach to the counting process, and Ref. [8] adopted a constant general function in quantile regression which it applied to recurrence event data. In the application and algorithm estimation of quantile regression, efforts have also been made by many people.

Ref. [9] proposed a locally weighted censored quantile regression approach that can solve covariate-dependent censoring. Ref. [10] proposed a semi-parametric approach using empirical likelihood to a random effects quantile regression model.

Overall, this work considers a situation wherein the general function of the accelerated failure time model is affected by time, sharing the same spirit of the parametric estimation method, while we will use the nonparametric method to estimate the effect. Then, we developed a new counting process model extended from the quantile regression models of this general function. This method provides some ideas for the development of more diverse estimation methods for the counting process.

## 2. Model

## 2.1. Accelerated Recurrent Events Time Model

Defining the  $R(\cdot) = \sum_{j=1}^{\infty} I(T^{(j)} \leq \cdot)$ , where  $I(\cdot)$  is the indicator function. The observed counting process  $N(\cdot) = R(\cdot \wedge L)$  shows that the observation is limited to followup time L and  $\wedge$  is the minimization operator, meanwhile, the at-risk process  $Y(\cdot)$  and  $Y(t) = I(L \geq t)$ . We also have  $\mu_Z = E\{R(t)|Z\}$ , and Z is a p-vector as the covariates of interest. The original accelerated recurrence time models considered covariates' effects as time scale changes, which also share the same spirit as quantile regression, and the inverse function is

$$\tau_Z(u) = \inf\{t : \mu_Z(t) > G(u)\},\$$

in a general setting, and G(u) = u denotes the time to expected frequency u. However, using a constant to estimate the expected frequency has some limitations. For example, the effect of the intervention of interest factors affecting the occurrence of an event may change over time. In this case, the model which combines a constant and variable coefficient may be more effective. As such, we proposed a new accelerated recurrence time model, and the right side of the inequality in the inverse function is  $G(u) = u \cdot m(t)$  instead of G(u) = u, where m(t) is the function of time *t*. Regardless of the form of inverse function, it can be written in the following form:

$$\log \tau_Z(u) = X^T \beta_0(u) \tag{1}$$

As can be seen from the above formula,  $\log \tau_Z(u)$  should be linear. Compared to the original accelerated recurrence model and the improved model, when the recurrent event T follows a homogeneous Poisson distribution, as shown in Figure 1, the estimated  $\log \tau_Z(u)$  by both models is approximately linear when the recurrent event is a non-homogeneous Poisson process, which has an intensity function  $\lambda_0(t)$  related to time t. In Figure 2, we compare three different situations, namely when  $\lambda_0(t) = b \times t$  and b = 1,5,10. As b increases, the estimation of the original accelerated recurrence model tends to be more and more nonlinear, while the proposed improved model still shows a strong linear relationship. In that way, the proposed recurrent event time model can be used in a more general situation.

In this paper, we consider the situation that G(u) = ut, then model 2 can be rewritten as

$$E(X\left[R\left\{\exp(X^{T}\beta)\right\}-\exp(X^{T}\beta)u\right])=0$$
(2)

if and only if  $\beta = \beta_0(u)$ , and the left-hand side should be monotone, then the above equation can be transformed into

$$\beta_0(u) = \operatorname{argmin}_{\beta} E\left\{\sum_{j=1}^{\infty} (X^T \beta - \log T^{(j)})^+ - \exp(X^T \beta)u\right\}.$$

where  $a^+ = a \lor 0$  and  $\lor$  denotes the maximization operator. Then, we can come up with the following theorem:

**Theorem 1.** For  $G(u) = \exp(X^T \beta(u))u$ , if (1) hold and  $R\{\tau_Z(u)\} \perp L \mid Z$ , define  $v^{\otimes 2} = vv^T$  for vector v,  $\mu_Z(\cdot)$  is the derivative of  $\mu_Z$ , let  $F_Z(x) = E\{Y(x) \mid Z\}$ , and assume (C1)  $E[X^{\otimes 2}I\{L > \tau_Z(u)\}]$  is non-singular;

(C2) 
$$\dot{\mu}_Z(e^{X^T\beta_0(u)})F_Z(e^{X^T\beta_0(u)}) > U$$
 for all  $u \in (0, U]$ .

Then,

$$\beta_0(u) = \operatorname*{argmin}_{\beta} \psi(\beta; u), \tag{3}$$

where

$$\psi(\beta; u) = E\left\{\sum_{j=1}^{M} (X^{T}\beta \wedge \log L - \log T^{(j)})^{+} - \exp(X^{T}\beta \wedge \log L)u\right\}.$$



Figure 1. Estimation results of homogeneous Poisson process.



Figure 2. Estimation results of non-homogeneous Poisson process.

The proof of this result, as well as the proofs for Theorems 2 and 3 below, are given in Appendix A. Powell extended these results for censored quantile regression when the censoring time is observed [7]. We will then combine the counting process model and apply this method to the recurrence events data.

#### 2.2. The Recurrent Events Model

For the counting process model, we begin with a review of Gill and Andersen's study [11], which extended the Cox proportional hazards model [12] to a multivariate counting process. The intimate connection of the study on the multivariate counting process and the use of martingale methods can allow us to derive more properties of the statistical estimation and testing procedures. As each component  $N_i(\cdot)$  of a multivariate counting process has a cumulative hazard function  $\Lambda_T(\cdot)$ , there exist local martingales  $M_i$  defined by  $M_i(t) = N_i(t) - \Lambda_T(t)$ . Since the cumulative hazard function  $\Lambda_T(\cdot)$  is the integral of the hazard function, then according to [11], the counting process formation of the Cox model can be given by

$$E\{N(t) \mid Z\} = E\left\{\int_{0}^{u} Y(s)\lambda_{0}(s)e^{X^{T}b}ds \mid Z\right\}, t > 0,$$
(4)

where  $\lambda_0(s)$  denotes the baseline hazard function, and this model can accommodate recurrent events data well. Peng and Huang conducted a re-examination work [5], which combined the quantile regression method for survival data with the above counting process model.

$$E\left\{N(e^{X^{T}\beta_{0}(u)}) \mid Z\right\} = E\left\{\int_{0}^{u} Y(e^{X^{T}\beta_{0}(s)})\frac{1}{1-s}ds \mid Z\right\}, u \in (0,1),$$
(5)

which has a nice monotonicity property. In model 6, the estimated cumulative hazard function  $\hat{\Lambda}_T(x) = -\log(1-x)$ . Huang also showed that a singleton model, where an estimated cumulative hazard function is a constant *u*, can also be applied in the counting process model [13]. When the singleton model is combined with quantile regression, this extended the following model to a more general situation of recurrent events [8], and the estimation formula is given by

$$E\left\{N(e^{X^{T}\beta_{0}(u)}) \mid Z\right\} = E\left\{\int_{0}^{u} Y(e^{X^{T}\beta_{0}(s)})g(s)ds \mid Z\right\}, u \in (0,1),$$
(6)

where  $g(s) \equiv 1$ . In fact, in model 7, the estimated cumulative hazard function G(u) = u, which is a constant. According to Theorem 1 and the above models, we can propose a new counting process model that takes the form

$$E\left\{N(e^{X^{T}\beta_{0}(u)}) \mid Z\right\} = E\left\{\int_{0}^{u} Y(e^{X^{T}\beta_{0}(s)})dG(s) \mid Z\right\}, u \in (0, U],$$
(7)

where in our model, we can also show that

$$\mu_Z(e^{X^T\beta_0(u)})=G(u)=ut=\int_0^u g(s)ds.$$

Therefore, we also have an alternative representation of the model that

$$\tau_Z(G(u)) = \exp\left\{X^T \beta_0(u)\right\}, u \in (0, U],\tag{8}$$

The proof of this recurrence events setting is also shown in the Appendix A. In the simulation part, we will show that our estimation method performs better when the recurrent event time follows the non-homogeneous Poisson distribution.

# 2.3. The Proposed Estimation Procedure

From Theorem 1, we can define the objective function

$$\Psi(\beta; u) = n^{-1} \sum_{i=1}^{n} \left\{ \sum_{j=1}^{M_i} (X_i^T \beta \wedge \log L_i - \log T_i^{(j)})^+ - \exp(X_i^T \beta \wedge \log L_i) u) \right\}.$$
(9)

Theorem 1 leads to the estimation of  $\beta_0(u)$ , that is  $\hat{\beta}(u) = \arg \min_{\beta} \Psi(\beta; u)$ .

It can be seen that the objective function is not convex. An algorithm was developed to find the local minimizer that is asymptotically equivalent to the global minimizer [13]. Since we proposed the counting process-based model (model 7), then we can propose an equation to estimate  $\beta_0(u)$ :

$$n^{1/2}S_n(\beta, u),\tag{10}$$

where

$$S_n(\beta, u) = n^{-1} \sum_{i=1}^n X_i \bigg\{ N_i(\exp\left\{X_i^T \beta(u)\right\}) - \int_0^u I(L_i \ge \exp\left\{X_i^T \beta(s)\right\}) dG(s) \bigg\}.$$

As the results proposed in [11], although  $M_h$  is a local square integrable martingale, it has the same asymptotic property as a global martingale, and if we let  $S(\beta, u) = E\{S_n(\beta, u)\}$ , then follows the martingale property of  $M(\cdot)$  given that  $S(\beta_0, u) = 0$ .

Equation (8) boils down to the estimation of censored quantile regression in [5]. A common method used to predict  $\beta_0(u)$  is the grid-based estimation procedure by denoting  $\beta_0(u)$  as a right-continuous piecewise-constant function that jumps on a grid. More specifically, we define  $S_{L(n)} = \left\{ 0 = u_0 < u_1 < \cdots < u_{L(n)} = U \right\}$ , and for our recurrence setting, the U in the grid can be greater than 1. The size of  $S_{L(n)}$ , denoted by  $\left\|S_{L(n)}\right\|$ , is the maximum value of  $u_j - u_{j-1}$  where  $j = 1, \cdots, L(n)$ . It is also noteworthy that  $\exp\{X^T\hat{\beta}(0)\} = 0$ . Thus, based on model 9, we can estimate  $\hat{\beta}(u_j)$  by sequentially solving the estimating equation

$$n^{-1/2} \sum_{i=1}^{n} X_i \left\{ R_i (\exp\left\{ X_i^T \beta(u_j) \right\}) - \sum_{k=0}^{j-1} I(L_i \ge \exp\left\{ X_i^T \hat{\beta}(u_k) \right\}) \int_{u_k}^{u_{k+1}} g(s) ds \right\} = 0, \quad (11)$$

Since Equation (10) is not continuous, the exact root may not exist. Fygenson and Ritov proposed a generalization solution of  $\hat{\beta}(u_j)$  for monotone estimating equations [14]. To find the generalization solution of Equation (10), we need to perform some simple algebraic manipulations, and then the solution-finding problem is equivalent to locating the minimizer of the following  $L_1$ -type convex function:

$$l_{j}(h) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} I(T_{i}^{(j)} \leq L_{i}) \left| \log T_{i}^{(j)} - X_{i}^{T}h \right| + \left| R^{*} - \left\{ \sum_{i=1}^{n} \sum_{j=1}^{\infty} I(T_{i}^{(j)} \leq L_{i})(-X_{i})^{T}h \right\} \right| + \left| R^{*} - \left\{ \sum_{i=1}^{n} 2X_{i}^{T}h \sum_{k=0}^{j-1} I(L_{i} \geq \exp\left\{X_{i}^{T}\hat{\beta}(u_{k})\right\}) \exp\left\{X_{i}^{T}\hat{\beta}(u_{k})\right\}(u_{k+1} - u_{k}) \right\} \right|.$$

where  $R^*$  is a very large number and  $j = 1, \dots, L(n)$ . One can also solve the  $l_j(h)$  by using statistical software, such as the rq() function in R package *quantreg* or the l1fit() function in S-PLUS.

# 2.4. Asymptotic Properties

In this section, we establish the uniform consistency and weak convergence of the proposed estimator  $\hat{\beta}(u)$ . Firstly, some regularity conditions should be stated.

Define  $\mu(b) = E\{XN(\exp(X^Tb))\}, \tilde{\mu}(b) = E\{XI(\exp(X^Tb) \le L)\}, \tilde{\mu}_Z(x) = E\{N(x)|Z\}, g_Z(x) = d\tilde{\mu}_Z(x)/dx, F(t|Z) = \Pr(x \le t|Z), \tilde{f}(x|Z) = -f(x|Z) = -dF(x|Z)/dx.$ By using simple algebra, we also define  $B(b) = d\mu(b)/db^T = E\{X^{\otimes 2}e^{X^Tb}g_Z(e^{X^Tb})\}, J(b) = d\tilde{\mu}(b)/db^T = E\{X^{\otimes 2}e^{X^Tb}f(e^{X^Tb}|Z)\}$ . Suppose the assumptions in Theorem 1 hold and  $\hat{\beta}(u)$  is strongly consistent for  $\beta_0(u)$ . Assume the following conditions:

(C1') Z and N(L) are bounded;

- (C2') (a)  $\mu_Z(\cdot)$  is bounded and continuous at  $t = \tau_Z(u)$ , as well as uniform in Z, and (b)  $\mu(\beta_0(u))$  is a Lipschitz function of u;
- (C3') (a) each component of  $J(b)B(b)^{-1}$  is uniformly bound on  $b \in \mathcal{B}(d)$ , where  $\mathcal{B}(d)$  is a neighborhood containing  $\beta_0(u)$ , (b)  $g_Z(x) > 0$  and  $E(Z^{\otimes 2})$  is positive definite;
- (C4')  $\inf_{u \in [v,U]} \operatorname{eigmin} B(\beta_0(u)) > 0$  for any  $v \in (0, U]$ , where  $\operatorname{eigmin}(\cdot)$  denotes the minimum eigenvalue of a matrix.

Condition (C1') implies the boundedness of covariates and the number of observed events, and (C2') gives the smoothness of  $\beta_0(u)$ , condition (C3') sets additional mild assumptions such as the positive definite. Noted that (C4') is a key condition that ensures the consistency of the proposed estimator. Then, we have the following theorems.

**Theorem 2.** Under conditions C1'-C4', when  $\lim_{n\to\infty} ||S_{L(n)}|| = 0$ , then

$$\sup_{u\in(v,U]} \left\|\hat{\beta}(u) - \beta_0(u)\right\| \xrightarrow{P} 0,$$

*where* 0 < v < U*.* 

**Theorem 3.** Under condition C1'-C4', when  $\lim_{n\to\infty} n^{1/2} \|S_{L(n)}\| = 0$ , then

$$n^{1/2}\{\hat{\beta}(u) - \beta_0(u)\}$$

converge weakly to a Gaussian process for  $u \in [v, U]$ , where  $v \in (0, U)$ .

The proofs are also in the Appendix A.

# 3. Simulation Studies

In this part, we conducted the Monte Carlo simulations to test our model. Combining the methods of Huang and Peng [13], recurrent events were generated from both homogeneous Poisson and non-homogeneous Poisson processes. We also generated two covariates,  $X_1$  and  $X_2$ , following the distribution *Bernoulli*(0.5) and *Unif*(-0.5, 0.5) separately. The recurrent event time was generated by:

$$T^{(j)} = \exp\left\{\min(1, \frac{T^{*(j)}}{1.5\gamma})X_1 + X_2\right\} T^{*(j)} / \gamma, \quad j = 1, 2, ...,$$

where in one case,  $\{T^{*(j)}, j = 1, 2, ...\}$  was a recurrent event time from the standard homogeneous Poisson process, in other words, the gap times of  $T^{*(j)} - T^{*(j-1)}$  are independent and identically *exponetial*(1) variables; in other cases,  $\{T^{*(j)}, j = 1, 2, ...\}$  is a recurrent event time from the non-homogeneous Poisson process with the intensity function  $\lambda_0(t) = t$ . Furthermore, the frailty  $\gamma$  followed the Gamma distribution for the homogeneous Poisson process, we considered two situations, that is the variance of  $\gamma$  was chosen to be 0 and 0.5, for the non-homogeneous Poisson process, and we only consider the variance of  $\gamma$  to be 0. Under our simulation setup, we have

$$\tau_Z(G(u)) = \exp\{\log(u) + \min(1, u/1.5)X_1 + X_2\}.$$

For censoring time *L*, we generated it from Unif(0, 12). For each selection of the variance of  $\gamma$ , we generated 500 datasets of sample size n = 100. Since we adopted the grid-based method to estimate the  $\beta$ , an equally spaced grid on  $u \in (0,3]$  with step size 0.02 was conducted in our simulation.

Figures 3 and 4 are the simulation results for the homogeneous Poisson process from the set-up with a Gamma frailty of variance 0 and 0.5. In the first row, we plot the empirical bias of the proposed estimator  $\hat{\beta}(u)$  (solid lines) and the empirical bias of the Sun's estimator (dashed lines) [8]. Sun considered the double censored situation while we only consider the right-censored event time to suit a more general situation. In the second row of the plots, we depicted the empirical mean squared error (MSE) versus the expected frequency u. The third row presents the coverage probabilities of 95% confidence intervals obtained from the proposed estimator (solid lines) and Sun's estimator (dashed lines). We can see that both methods have a slight bias which converges to 0, and the empirical MSE also tends to be stable as u increases. In the homogeneous Poisson process, when the variance of gamma frailty equals 0, the empirical MSE shows less fluctuation compared to the variance that is equal to 0.5.

Figure 5 depicts the same parameters as Figures 3 and 4, while the difference is that the results are from the non-homogeneous Poisson process. We can see that, in the first and second rows of Figure 5, both methods performed quite similarly in terms of the bias and SD. However, from the third row, in terms of the coverage probability, the proposed estimator (solid lines) performed slightly better than the Sun's estimator (dashed lines), than for the simulation of the non-homogeneous Poisson process.



Figure 3. Simulation results with sample size n = 100 and the set-up with Gamma frailty of variance 0.



**Figure 4.** Simulation results with sample size n = 100 and the set-up with Gamma frailty of variance 0.5.



**Figure 5.** Simulation results of non-homogeneous Poisson process with sample size n = 100 and the set-up with Gamma frailty of variance 0.

# 4. Application to the Bladder Tumor Studies

We applied our proposed estimated method to a well-known bladder tumor study [15]. This dataset was conducted to analyze the effect of two treatments, pyridoxine and thiotepa, based on the recurrence of bladder tumors. A total of 118 subjects were recorded, 48 were treated with placebo, 32 were pyridoxine, and 38 with thiotepa. The covariates also contained the initial number of tumors, the size of the largest initial tumor and others. The maximum observed number of recurrences is 9.

We selected four covariates and the intercept: the two treatment methods and the initial tumor size and number. In Figure 6, we displayed our estimation result of the proposed method (solid lines) and Sun's method (gray line) [8], surrounded by point-wise Wald 95% confidence intervals (dashed lines). The grid-based estimation method estimated the regression coefficients over [0, 1.6]. The intercept coefficient estimates represent the log time to the expected frequency of the bladder tumor recurrence, consisting of patients who had no pyridoxine and thiotepa treatment. The intercept term indicates that as time increases, the bladder tumor recurrence also increases, which is in line with expectations. The negative non-intercept coefficient estimates pyridoxine and thiotepa show that these two treatments can inhibit tumor recurrence. In contrast, the initial tumor size and number of covariates are positive, suggesting a negative effect on tumor recurrence. From Figure 6, we can see that our method and Sun's method do not have much difference, and the overall trend is the same.



Figure 6. Bladder data example: coefficient estimates and 95% pointwise confidence.

# 5. Discussion

In this article, we introduced the accelerated recurrence time model for recurrence events, and then considered a new situation when the expected frequency of the accelerated recurrence time model was affected by the time-varying and combined this model with the counting process model. This new counting process model with a non-singleton general function is similar to Cox's regression model but can be transformed into quantile regression.

This method can also estimate more general situations, such as the non-homogeneous Poisson process.

We can generalize our estimation procedure from double-censored to right-censored situations. In the recurrence events setting, the most popular choice of  $G(\cdot)$ , which may be G(u) = u, as well as  $G(u) = -\ln(1 - u)$ , we introduce our new estimation method with G(u) = ut. The assumption is that a constant hazard is rarely tenable in practical problems. Therefore, in the parametric estimation process, a more general distribution of the hazard function is the Weibull distribution, which is  $g(t) = \lambda \gamma t^{\gamma-1}$  for  $0 < t < \infty$ , so we have  $G(u) = ut^{\gamma}$ . When  $\gamma = 0$ , the Weibull distribution function becomes a constant, but resembles the accelerated recurrence time model. When  $\gamma = 1$ , the model became our proposed method. Thus, after combining the parametric and nonparametric methods, we can make the estimation method more diverse and adapt to different types of datasets.

**Author Contributions:** Conceptualization, J.X.; Data curation, X.W.; Formal analysis, X.W.; Investigation, X.W. and J.X.; Methodology, J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

# Appendix A

**Proof of Theorem 1.** If we want to estimate the  $\beta_0$ , the left-hand side of the following equation should be monotone,

$$E(X\left[R\left\{\exp(X^{T}\beta)\right\}-\exp(X^{T}\beta)u\right])=0$$
(A1)

taking derivative with respect to  $\beta$  on both sides of equation that

$$\dot{\mu}_Z(e^{X^T\beta(u)})F_Z(e^{X^T\beta_0(u)})e^{X^T\beta(u)}X^Td\beta-ue^{X^T\beta(u)}X^Td\beta,$$

then, under condition (2), the equation A1 can be monotone. Meanwhile, straightforward algebra gives

$$\begin{split} \phi(\beta; Z, L) &= E\left\{\sum_{j=1}^{M} (X^{T}\beta \wedge \log L - \log T^{(j)})^{+} - \exp(X^{T}\beta \wedge \log L)u \mid Z, L\right\} \\ &= \int_{-\infty}^{X^{T}\beta \wedge \log L} \mu_{Z} \{\exp(s)\} ds - \exp(X^{T}\beta \wedge \log L)u \end{split}$$

and for any  $\beta \neq \beta_0(u)$ ,

$$E[\phi(\beta_0(u); Z, L)] = \psi(\beta_0(u); u) < \psi(\beta; u) = E[\phi(\beta(u); Z, L)],$$

in order to eliminate the minimization operator, we consider the following situations:

• when  $\log L \leq X^T \beta$  and  $\log L \leq X^T \beta_0$ ,

$$\phi(\beta_0(u); Z, L) = \phi(\beta; Z, L).$$

• when  $\log L > X^T \beta$  and  $\log L \le X^T \beta_0$ ,

$$\begin{split} \phi(\beta_0(u); Z, L) - \phi(\beta; Z, L) &= \int_{X^T \beta}^{\log L} \mu_Z \{ \exp(s) \} ds - u \Big[ \exp(\log L) - \exp(X^T \beta) \Big] \\ &= \int_{X^T \beta}^{\log L} [\mu_Z \{ \exp(s) \} - u \{ \exp(s) \}] ds \le 0. \end{split}$$

• when  $\log L > X^T \beta_0$ , since  $\exp(\beta_0(u))$  is the unique value that minimizes

$$\int_{-\infty}^{x} \mu_Z \{ \exp(s) \} ds - e^x u,$$

considering the monotonicity of  $\mu_Z(\cdot)$ , therefore  $\phi(\beta_0(u); Z, L) \leq \phi(\beta; Z, L)$ . By the non-singularity of  $E[X^{\otimes 2}I\{L > \tau_Z(u)\}]$ ,

$$E\Big[\phi(\beta_0(u);Z,L)I\Big\{\log L > X^T\beta_0(u)\Big\}\Big] \le E\Big[\phi(\beta;Z,L)I\Big\{\log L > X^T\beta_0(u)\Big\}\Big].$$

Thus, the proof is completed.  $\Box$ 

**Proof of Theorem 2.** Define  $\alpha_0(u) = \mu(\beta_0(u)), \hat{\alpha}(u) = \mu(\hat{\beta}(u))$ , and

$$\mathcal{B}(d) = \left\{ b \in \mathbb{R}^{p+1} : \inf_{u \in (0,U]} \|\mu(b) - \mu(\beta_0(u))\| \le d \right\}, \mathcal{A}(d) = \{\mu(b) : b \in \mathcal{B}(d)\}.$$

Noted that for any  $b_1, b_2 \in \mathcal{B}(d)$ ,

$$(b_1 - b_2)\{\mu(b_1) - \mu(b_2)\} = E\Big[(X^T b_1 - X^T b_2)\Big\{N(exp(X^T b_1)) - N(exp(X^T b_2))\Big\}\Big] = 0,$$

which only occurs when  $b_1 = b_2$  by condition (C3'). Then, there exists a one-to-one map from  $\mathcal{B}(d)$  to  $\mathcal{A}(d)$  that we denoted by  $\kappa(\cdot)$  and  $\kappa\{\mu(b)\} = b$  for any  $b \in \mathcal{B}(d)$ .

Furthermore, by the properties of the martingale stochastic process

$$M(u) = N(\exp(X^{T}\beta_{0}(u)) - \int_{0}^{u} Y(e^{X^{T}\beta_{0}(s)})g(s)ds,$$

we have

$$n^{-1}\sum_{i=1}^{n} X_{i}N_{i}(\exp\left\{X_{i}^{T}\hat{\beta}(u)\right\}) = n^{-1}\sum_{i=1}^{n}\int_{0}^{u} X_{i}I\left[\exp\left\{X_{i}^{T}\hat{\beta}(s)\right\} \le L_{i}\right]g(s)ds + \xi_{n,k}$$

where  $\xi_{n,k} = O_p(n^{-1})$ .

Then, the following notations are defined:

$$v_n(b) = n^{-1} \sum_{i=1}^n X_i N_i(\exp\left\{X_i^T b\right\}) - \mu(b),$$
  
$$\tilde{v}_n(b) = n^{-1} \sum_{i=1}^n X_i I(\exp\left\{X_i^T b\right\} \le L_i) - \tilde{\mu}(b).$$

By the Glivenko-Cantelli theorem, we have

$$\sup_{b} \|v_{n}(b)\| \xrightarrow{a.s.} 0,$$
  
$$\sup_{b} \|\tilde{v}_{n}(b)\| \xrightarrow{a.s.} 0.$$

Since  $E(XN(\exp\{X^T\beta_0(u)\})) = E(X\int_0^u I(L \ge \exp\{X^T\beta_0(s)\})dG(u))$ . Then,  $\hat{\alpha}(u) - \alpha_0(u)$  is equivalent to

$$-v_{n}(\hat{\beta}(u)) + \int_{0}^{u} \tilde{v}_{n}\{\hat{\beta}(s)\} dG(s) + \sum_{k=1}^{j} \int_{u_{k-1}}^{u_{k}} \left[\tilde{\mu}(\hat{\beta}(s)) - \tilde{\mu}(\beta_{0}(s))\right] dG(s) + \xi_{n,k}$$

By the arguments as [5], the sup<sub> $u \in [u_{k-1}, u_k)$ </sub>  $\|\hat{\alpha}(u) - \alpha_0(u)\|$  can be bounded almost surely by  $\varepsilon_k = (1 + C_4 b_n)^{k-1} (C_1 + \varepsilon_0 C_4 b_n + C_2 n^{-1} + C_3 a_n)$ , where  $\varepsilon_0 = C_3 a_n$ . Given  $\lim_{n\to\infty} a_n = 0$  and  $L(n) = U/a_n$ , we have

$$\lim_{n \to \infty} (1 + C_4 b_n)^{L(n) - 1} = \exp\{C_4 U / (1 - U)\},\$$

where  $C_1$ ,  $C_2$ ,  $C_3$ , and  $C_4$  are some positive constant. Since  $C_i$  are arbitrarily and  $\lim_{n\to\infty} a_n = 0$ , then we can know that  $\sup_{u\in(0,U]} \|\hat{\alpha}(u) - \alpha_0(u)\| \xrightarrow{P} 0$ . By the application of the Taylor expansion of  $\kappa\{\hat{\alpha}(u)\}$  around  $\alpha_0(u)$ , the conclusion  $\sup_{u\in(v,U]} \|\hat{\beta}(u) - \beta_0(u)\| \xrightarrow{P} 0$  can be reached.  $\Box$ 

**Proof of Theorem 3.** From the arguments made by [13], we have the following lemma:

$$\sup_{u \in (0,U]} \left\| n^{-1/2} \sum_{i=1}^{n} X_i \left\{ N_i(e^{X_i^T \hat{\beta}(u)}) - N_i(e^{X_i^T \beta_0(u)}) \right\} - n^{-1/2} \left\{ \mu(\hat{\beta}(u)) - \mu(\beta_0(u)) \right\} \right\| \xrightarrow{P} 0,$$

$$\sup_{u \in (0,U]} \left\| n^{-1/2} \sum_{i=1}^{n} X_i \left\{ I(L_i \ge e^{X_i^T \hat{\beta}(u)}) - I(L_i \ge e^{X_i^T \beta_0(u)}) \right\} - n^{-1/2} \left\{ \tilde{\mu}(\hat{\beta}(u)) - \tilde{\mu}(\beta_0(u)) \right\} \right\| \xrightarrow{P} 0.$$
(A2)

and the above lemma implies that

$$-n^{1/2}S_n(\beta_0, u) = n^{1/2} \left[ \mu(\hat{\beta}(u)) - \mu(\beta_0(u)) \right] - \int_0^u n^{1/2} \left[ \tilde{\mu}(\hat{\beta}(s)) - \tilde{\mu}(\beta_0(s)) \right] dG(s) + o_{(0,U]}(1)$$
  
=  $n^{1/2} \left[ \mu(\hat{\beta}(u)) - \mu(\beta_0(u)) \right]$   
 $- \int_0^u \left[ J(\beta_0(u)) B(\beta_0(u))^{-1} + o_{(0,U]}(1) \right] n^{1/2} \left[ \mu(\hat{\beta}(u)) - \mu(\beta_0(u)) \right] dG(s) + o_{(0,U]}(1)$ 

since  $\lim_{n\to\infty} n^{1/2} \|S_{L(n)}\| = 0$ , then  $n^{1/2}S_n(\hat{\beta}, u) = o_{(0,U]}(1)$ . Then, we can use the product integration theory to obtain the following equation:

$$n^{1/2} \big[ \mu(\hat{\beta}(u)) - \mu(\beta_0(u)) \big] = \varphi \Big\{ -n^{1/2} S_n(\beta_0, u) \Big\} + o_{(0, U]}(1),$$

where  $\varphi(g)(u) = \int_0^u \mathcal{I}(s, u) dg(s)$ , which is a one-to-one map from  $\mathcal{F}$  to  $\mathcal{F}$ . By our definition,  $\mathcal{F} = \{g : [0, U] \to \mathbb{R}^{p+1}\}$  and g is left-continuous with a right limit, g(0) = 0. Meanwhile,

$$\mathcal{I}(s,t) = \mathcal{T}_{u \in (s,t]} \Big[ I_{p+1} + J(\beta_0(u)) B(\beta_0(u))^{-1} dG(u) \Big]$$

Since the definition given by [16] is followed,  $\{X_i N_i(\exp\{X_i\beta_0(s)\}), s \in (0, U]\}$  is a VCclass. For  $\int_0^u I(L_i \ge \exp\{X_i^T\beta_0(s)\}) dG(s)$ , let  $h(x) = \int_0^x I(L_i \ge \exp\{X_i^T\beta_0(s)\}) dG(s)$ , suppose that  $x_1 \le x_2$ , then

$$|h(u_1) - h(u_2)| = \left| \int_{u_2}^{u_1} I(L_i \ge \exp\{X_i^T \beta_0(s)\}) dG(s) \right| \le K|u_1 - u_2|,$$

where  $K = \max(\exp\{X_i^T \beta_0(s)\}, s \in [u_1, u_2])$ , this implies that h(x) is a Lipschitz in x. The above arguments show that

$$\left\{X_i N_i(\exp\{X_i\beta_0(s)\}) - \int_0^u I\left(L_i \ge \exp\{X_i^T\beta_0(s)\}\right) dG(s), u \in (0, U]\right\}$$

is a Donsker class. By the Donsker theorem, we can know that  $-n^{1/2}S_n$  converge weakly to a Gaussian process, then  $\varphi\left\{-n^{1/2}S_n(\beta_0, u)\right\}$  is also a Gaussian process for  $\varphi(\cdot)$  is a linear operator. By the continuous mapping theory and Taylor expansion of  $\kappa\left[\mu\left\{\hat{\beta}(u)\right\}\right]$ around  $\kappa\left[\mu\left\{\beta_0(u)\right\}\right]$ , we can know that  $n^{1/2}\left\{\hat{\beta}(u) - \beta_0(u)\right\}$  converge weakly to a Gaussian process with a mean of zero and a covariance matrix  $\Sigma(s, t) = E[\zeta_1(s)\zeta_1(t)]$ , where  $\zeta_i(u) = B(\beta_0(u))^{-1}\varphi(s_i)$ , and  $s_i(u) = N_i(\exp\left\{X_i^T\beta(u)\right\}) - \int_0^u I(L_i \ge \exp\left\{X_i^T\beta(s)\right\})dG(s)$ .  $\Box$ 

# Recurrent Event Setting

Suppose model 6 holds, then according to the following definition:

$$E\left\{N(e^{X^T\beta_0(u)})|Z\right\} = \mu_Z(e^{X^T\beta_0(u)} \wedge L),$$

and the inequality  $e^{X^T \beta_0(u)} \leq L$  can also be written as  $G(u) \leq \mu_Z(L)$  following the definition of  $\tau_Z(u)$ . Then

$$E\left\{\int_0^u Y(e^{X^T\beta_0(s)})g(s)ds|Z\right\} = E\left\{\int_0^\infty I(G(s) \le \mu_Z(L))dG(s)|L,Z\right\}.$$

In model 6,  $G(u) = \mu_Z \{ X^T \beta_0(u) \}$ , so it follows from the above equation that for  $u \in (0, U]$ ,

$$E\left\{\int_0^u Y(e^{X^T\beta_0(s)})g(s)ds \mid Z\right\} = \mu_Z(L \wedge e^{X^T\beta_0(u)})$$

Therefore, the model 5 is satisfied.

Suppose that model 5 holds, then taking the derivative with respect to u on both sides of equation in model 5, we have

$$\dot{\mu}_{Z}(e^{X^{T}\beta(u)})F_{Z}(e^{X^{T}\beta_{0}(u)})e^{X^{T}\beta(u)}X^{T}d\beta = F_{Z}(e^{X^{T}\beta_{0}(u)})g(u)du,$$

that is,  $d\mu_Z(e^{X^T\beta(u)}) = g(u)du$ , then  $\mu_Z(e^{X^T\beta(u)}) = G(u)$  and model 6 are satisfied.

## References

- 1. Nelson, W. Hazard plotting for incomplete failure data. J. Qual. Technol. 1969, 1, 27–52. [CrossRef]
- 2. Altshuler, B. Theory for the measurement of competing risks in animal experiments. Math. Biosci. 1970, 6, 1–11. [CrossRef]
- 3. Lawless, J.F. Regression methods for Poisson process data. J. Am. Stat. Assoc. 1987, 82, 808–815. [CrossRef]
- 4. Lin, D.Y.; Wei, L.J.; Yang, I.; Ying, Z. Semiparametric regression for the mean and rate functions of recurrent events. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 2000, 62, 711–730. [CrossRef]
- 5. Peng, L.; Huang, Y. Survival analysis with quantile regression models. J. Am. Stat. Assoc. 2008, 103, 637–649. [CrossRef]
- 6. Koenker, R.; Bassett, G., Jr. Regression quantiles. Econom. J. Econom. Soc. 1978, 46, 33–50. [CrossRef]
- 7. Powell, J.L. Censored regression quantiles. J. Econom. 1986, 32, 143–155. [CrossRef]
- 8. Sun, X.; Peng, L.; Huang, Y.; Lai, H.J. Generalizing quantile regression for counting processes with applications to recurrent events. *J. Am. Stat. Assoc.* **2016**, *111*, 145–156. [CrossRef]
- 9. Wang, H.J.; Wang, L. Locally weighted censored quantile regression. J. Am. Stat. Assoc. 2009, 104, 1117–1128. [CrossRef]
- 10. Kim, M.O.; Yang, Y. Semiparametric approach to a random effects quantile regression model. *J. Am. Stat. Assoc.* 2011, 106, 1405–1417. [CrossRef] [PubMed]
- 11. Andersen, P.K.; Gill, R.D. Cox's regression model for counting processes: A large sample study. *Ann. Stat.* **1982**, *10*, 1100–1120. [CrossRef]
- 12. Cox, D.R. Regression models and life-tables. J. R. Stat. Soc. Ser. B (Methodol.) 1972, 34, 187–202. [CrossRef]
- 13. Huang, Y.; Peng, L. Accelerated recurrence time models. Scand. J. Stat. 2009, 36, 636–648. [CrossRef] [PubMed]
- 14. Fygenson, M.; Ritov, Y. Monotone estimating equations for censored data. Ann. Stat. 1994, 22, 732–746. [CrossRef]

- Byar, D. The veterans administration study of chemoprophylaxis for recurrent stage i bladder tumours: Comparisons of placebo, pyridoxine and topical thiotepa. In *Bladder Tumors and Other Topics in Urological Oncology*; Springer: Berlin/Heidelberg, Germany, 1980; pp. 363–370.
- 16. Vaart, A.v.d.; Wellner, J.A. Weak convergence and empirical processes with applications to statistics. *J. R. Stat. Soc.-Ser. A Stat. Soc.* **1997**, *160*, 596–608.