KYBERNETIKA — VOLUME 57 (2021), NUMBER 5, PAGES 801–818

EMPIRICAL ANALYSIS OF CURRENT STATUS DATA FOR ADDITIVE HAZARDS MODEL WITH AUXILIARY COVARIATES

Jianling Zhang, Mei Yang and Xiuqing Zhou

In practice, it often occurs that some covariates of interest are not measured because of various reasons, but there may exist some auxiliary information available. In this case, an issue of interest is how to make use of the available auxiliary information for statistical analysis. This paper discusses statistical inference problems in the context of current status data arising from an additive hazards model with auxiliary covariates. An empirical log-likelihood ratio statistic for the regression parameter vector is defined and its limiting distribution is shown to be a standard chi-squared distribution. A profile empirical log-likelihood ratio statistic for a sub-vector of the parameters and its asymptotic distribution are also studied. To assess the finite sample performance of the proposed methods, simulation studies are implemented and simulation results show that the methods work well.

Keywords: current status data, auxiliary covariates, additive hazards model, empirical likelihood

Classification: 62E20, 62N01

1. INTRODUCTION

Current status data, also called as case I interval-censored data, means that the failure time T is unobservable, but it can be determined to lie below or above a random monitoring time C. Such data plays an important role in many fields including clinical medicine, econometrics, reliability studies and so on (Jewell and van der Laan [5], Huang [4], Rossini and Tsiatis [18]). In practice, both failure times and monitoring times can depend on some covariates. For such current status data, how to model it and make accurate inference for the unknown parameters in the model is particularly of interest to methodology researchers. Many procedures have been developed for current status failure time data under various models when covariate information is completely known. An extensive review of models for current status data can be found in Chapter 5 of Sun[19]. A popular choice is the Cox proportional hazards model, for which Huang [4] gave a profile likelihood approach for parameter estimation, and Wang et al. [20] proposed an EM algorithm under some conditions. Some authors considered analyzing the current status data by using the additive hazards model.(e. g. McKeague and Sasieni

DOI: 10.14736/kyb-2021-5-0801

[12], Lin et al. [4], Zhang et al. [21], Chen and Sun [2], Lu and Song [10], and so on), and their estimation method is based on the efficient score, which is used as an estimating equation in deriving parameter estimates.

However, due to budget constraints or other technique reasons, covariate information may only be collected in a randomly selected subset from the whole study cohort in biomedical studies. This subset is referred to as the validation set, and only subjects in it could provide true covariates. Meanwhile, to compensate the missing covariate information, some convenient information related to the true covariates is obtained by applying a cheaper substitute technique to all the study subjects. Such auxiliary information is referred to as auxiliary covariates.

In this paper, we consider the additive hazards model for current status data in situations when covariates are fully available only on some of the study subjects but are missing and compensated with auxiliary information on the others. To make statistical inference for unknown parameters, methods based on empirical likelihood will be applied. Empirical likelihood was introduced by Owen ([14, 15]) for a mean vector for i.i.d. observations, and has been extended to a wide range of applications. Zhao and Hsu [24] studied the additive risk models with right censoring; Zhou [26] considered empirical likelihood analysis for the accelerated failure time model; Zhang and Zhao [23] developed two empirical likelihood inference approaches for linear transformation models with interval-censored data; Liu *et al.* [9] gave empirical likelihood for the additive hazards model with current status data. However, it seems that there does not exist method for empirical analysis of current status data arising from the additive hazards model in the presence of auxiliary covariates. Therefore, this paper applies empirical likelihood to study the additive hazards model of current status data with auxiliary covariates.

The remainder of this paper is organized as follows. We begin in Section 2 with introducing some notations and models that will be used throughout the paper. In Section 3, we introduce the empirical likelihood method to the additive hazards model with current status data. We define an empirical log-likelihood ratio statistic for the unknown regression parameter vector, and show that its limiting distribution is a chi-squared distribution. A profile empirical log-likelihood ratio statistic for a sub-vector of the parameters and its asymptotic distribution are also proposed in Section 3. In Section 4, we conduct simulation studies to evaluate the performance of the proposed empirical likelihood methods. Some discussions are given in Section 5, and proof details are presented in Section 6.

2. MODELS AND NOTATIONS

Consider a failure time study that consists of n independent subjects. Let T_i denote the failure time of the i-th member and suppose that there exists a vector of covariates $Z_i(t)$, which may depend on time t. For the relationship between T_i and $Z_i(t)$, in the following, we assume that given the history of covariates up to time t, the hazard function of T_i has form

$$\lambda_T(t|Z_i(s), s \leq t) = \lambda_0(t) + \beta_0' Z_i(t), \tag{1}$$

where $\lambda_0(t)$ denotes an unknown marginal baseline hazard function, and β_0 is a p-dimensional vector of unknown regression parameters. That is, T_i follows the additive hazards model (Lin and Ying [7]). In current status data, T_i is not observed directly. Instead, each subject will be observed only once at time C_i, and only C_i and $\delta_i = I(T_i \geq C_i)$ are obtained. We assume that C_i is independent of Z_i and T_i.

In the following, we assume that covariate vector $Z_i(t)$ may be missing, but a vector of auxiliary covariates, which is denoted by $X_i(t)$, is observed for each subject. The relationship between $X_i(t)$ and $Z_i(t)$ is unspecified, but we assume that conditional on $Z_i(t)$, $X_i(t)$ provides no additional information to the regression model, that is,

$$\lambda_T(t|Z_i(t), X_i(t)) = \lambda_T(t|Z_i(t)).$$

Let V denote a simple random subset of $\{1, 2, \ldots, n\}$ which stands for the subjects whose true covariates $Z_i(t)$ are known, and $\bar{V} = \{1, 2, \ldots, n\} \setminus V$ is the complement of V. The set V is usually referred to as validation set. Let n_v and $n_{\bar{v}}$ denote the sizes of V and $\bar{V}$ respectively. For simplicity, we also assume that for each subject, all components of its covariate vector $Z_i(t)$ are either known or missing together. That means for each subject $i \in V$, we have $(C_i, \delta_i, Z_i(t), X_i(t))$ observed, and for each subject $i \in \bar{V}$, we have $(C_i, \delta_i, X_i(t))$ observed only. Some comments will be given in Discussion for the situation where missing happens to only some components of $Z_i(t)$.

Define $N_i(t) = I(C_i \leq \min(t, T_i))$ and $Y_i(t) = I(C_i \geq t)$. Then $N_i(t)$ is a counting process with intensity process

$$\begin{aligned} \lambda_i(t|Z_i(s), s \leq t) &= e^{-\Lambda_0(t)} e^{-\beta_0' Z_i^*(t)} \lambda^c(t) \\ &\triangleq \lambda_0^c(t) e^{-\beta_0' Z_i^*(t)} \end{aligned} \tag{2}$$

(Lin et al. [6]), where $\lambda^c(t)$ denotes hazard function of event $\{C_i = t\}$, $\Lambda_0(t) = \int_0^t \lambda_0(s)\,\mathrm{d}s$, $\lambda_0^c(t) = \lambda^c(t)e^{-\Lambda_0(t)}$ and $Z_i^*(t) = \int_0^t Z_i(s)\,\mathrm{d}s$. Equation (2) says that $\lambda_i(t|Z_i(s), s \leq t)$ satisfies Cox proportional hazards model with relative risk function $e^{-\beta_0' Z_i^*(t)}$. But it is apparent that $e^{-\beta_0' Z_i^*(t)}$ is not available for our situation because of the missing of $Z_i(t)$. Then following Prentice [16], we consider the intensity process conditional on $X_i(t)$ rather than on $Z_i(t)$, and the induced intensity process is written as

$$\tilde{\lambda}_i(t) = \lambda_0^c(t) E\left\{e^{-\beta_0' Z_i^*(t)} | Y_i(t) = 1, X_i(t)\right\}.$$

It is easy to see that this is still a proportional hazard model with relative risk function

$$\phi_i(t, \beta_0) \triangleq E\left\{e^{-\beta_0' Z_i^*(t)} | Y_i(t) = 1, X_i(t)\right\}.$$

Define

$$\Phi_i(t, \beta) = \phi_i(t, \beta) I(i \in \bar{V}) + \varphi_i(t, \beta) I(i \in V)$$

with $\varphi_i(t, \beta) = e^{-\beta' Z_i^*(t)}$. For $j = 0$, 1 and 2, let $a^{\otimes j}$ denote 1, a and aa' respectively. To estimate β_0, motivated by the efficient estimation procedure of Martinussen and Scheike [11], Chen et al. [1] proposed to use empirical efficient score function

$$U(\beta, \Lambda_0) = \sum_{i=1}^n \int_0^\tau \left(\frac{\Phi_i^{(1)}(t, \beta)}{\Phi_i(t, \beta)} - \frac{S_1(t, \beta)}{S_0(t, \beta)} \right) \left(\frac{p_i(t, \beta)}{1 - p_i(t, \beta)} \,\mathrm{d}N_i^*(t) - \mathrm{d}N_i(t) \right), \tag{3}$$

where $f^{(l)}(t,\beta)$ stands for the l-order partial derivative of function $f(t,\beta)$ with respect to β, and

$$\begin{aligned} N_i^*(t) &= (1-\delta_i)I(C_i \le t), \\ p_i(t,\beta) &= \exp(-\Lambda_0(t))\Phi_i(t,\beta), \\ S_j(t,\beta) &= \sum_{i=1}^n Y_i(t)\frac{p_i(t,\beta)}{1-p_i(t,\beta)}\left(\frac{\Phi_i^{(1)}(t,\beta)}{\Phi_i(t,\beta)}\right)^{\otimes j}, \\ &\qquad j=0,1,2. \end{aligned}$$

In next section, to address this issue, a empirical likelihood-based confidence region will be developed.

3. EMPIRICAL LIKELIHOOD PROCEDURE

In this section, we present an empirical likelihood procedure for obtaining confidence region and making inference about β in model (1). We first define compensated counting processes $M_i(t)$ and $M_i^*(t)$ as

$$\begin{aligned} M_i(t) &= N_i(t) - \int_0^t Y_i(s)p_i(s,\beta_0)\,\mathrm{d}\Lambda^c(s), \\ M_i^*(t) &= N_i^*(t) - \int_0^t Y_i(s)(1-p_i(s,\beta_0))\,\mathrm{d}\Lambda^c(s). \end{aligned}$$

Obviously, both $M_i(t)$ and $M_i^*(t)$ are martingales. Let

$$W_i(\beta) = \int_0^\tau \left(\frac{\Phi_i^{(1)}(t,\beta)}{\Phi_i(t,\beta)} - \frac{S_1(t,\beta)}{S_0(t,\beta)}\right)\left(\frac{p_i(t,\beta)}{1-p_i(t,\beta)}\,\mathrm{d}M_i^*(t) - \mathrm{d}M_i(t)\right),$$

then $E(W_i(\beta))=0$ if and only if $\beta=\beta_0$ and $U(\beta,\Lambda_0)$ defined in (3) can be rewritten as

$$U(\beta,\Lambda_0) = \sum_{i=1}^n W_i(\beta). \tag{4}$$

Define

$$W_{ni}(\beta) = \int_0^\tau \left(\frac{\hat{\Phi}_i^{(1)}(t,\beta)}{\hat{\Phi}_i(t,\beta)} - \frac{\hat{S}_1(t,\beta)}{\hat{S}_0(t,\beta)}\right)\left(\frac{\hat{p}_i(t,\beta)}{1-\hat{p}_i(t,\beta)}\,\mathrm{d}M_i^*(t) - \mathrm{d}M_i(t)\right)$$

where $\hat{\Lambda}_0$ be an estimate of Λ_0, $\hat{\Lambda}_0$ can refer to Zhu et al. [25], Φ_i replaced by

$$\hat{\Phi}_i(t,\beta) = \hat{\phi}_i(t,\beta)I(i\in\bar{V}) + \varphi_i(t,\beta)I(i\in V),$$

and

$$\hat{\phi}_i(t,\beta) = \frac{\sum_{j\in V} Y_j(t)Q\left(B^{-1}(X_j-X_i)\right)\varphi_j(t,\beta)}{\sum_{j\in V} Y_j(t)Q\left(B^{-1}(X_j-X_i)\right)},$$

where B is a positive-definite matrix with elements possibly depending on n, and Q is a kernel function with bandwidth matrix B. It can be seen that $\hat{U}(\beta, \hat{\Lambda}_0) = \sum_{i=1}^n W_{ni}(\beta)$ holds true because

$$\begin{aligned}
\hat{U}(\beta_0, \hat{\Lambda}_0) &= \sum_{i=1}^n \int_0^\tau \left(\frac{\hat{\Phi}_i^{(1)}(t,\beta_0)}{\hat{\Phi}_i(t,\beta_0)} - \frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)} \right) \left(\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} \mathrm{d}N_i^*(t) - \mathrm{d}N_i(t) \right) \\
&= \sum_{i=1}^n \int_0^\tau \left(\frac{\hat{\Phi}_i^{(1)}(t,\beta_0)}{\hat{\Phi}_i(t,\beta_0)} - \frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)} \right) \left[\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} \mathrm{d}M_i^*(t) - \mathrm{d}M_i(t) \right. \\
&\quad \left. + \left(\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} \left(\frac{e^{-\hat{\Lambda}_0} - e^{-\Lambda_0}}{e^{-\hat{\Lambda}_0}} \right) Y_i(t) d\Lambda^c(t) \right) \right] \\
&= \sum_{i=1}^n \int_0^\tau \left(\frac{\hat{\Phi}_i^{(1)}(t,\beta_0)}{\hat{\Phi}_i(t,\beta_0)} - \frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)} \right) \left(\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} \mathrm{d}M_i^*(t) - \mathrm{d}M_i(t) \right)
\end{aligned}$$

where $\hat{\Lambda}_0$ be an estimate of Λ_0.

Consider the empirical likelihood function:

$$L(\beta_0) = \sup \left\{ \prod_{i=1}^n q_i : \sum_{i=1}^n q_i = 1, \sum_{i=1}^n q_i W_i(\beta_0) = 0, q_i \geq 0 \right\}$$

and thus the estimated empirical likelihood function:

$$L_n(\beta_0) = \sup \left\{ \prod_{i=1}^n q_i : \sum_{i=1}^n q_i = 1, \sum_{i=1}^n q_i W_{ni}(\beta_0) = 0, q_i \geq 0 \right\} \tag{5}$$

for β_0, where $\mathbf{q}= (q_1, q_2, \ldots, q_n)'$ denotes the probability vector.

For without any restriction, the natural estimator of $\mathbf{q}$ is given by $(1/n, 1/n, \ldots, 1/n)'$. By section 2.3 of Owen [13], we obtain the following empirical likelihood ratio

$$R(\beta_0) = \sup \left\{ \prod_{i=1}^n nq_i : \sum_{i=1}^n q_i = 1, \sum_{i=1}^n q_i W_{ni}(\beta_0) = 0, q_i \geq 0 \right\}$$

for β_0. Using the lagrange multiplier method, we first define

$$G = \sum_{i=1}^n \log(nq_i) - n\lambda' \sum_{i=1}^n q_i W_{ni}(\beta_0) + \gamma \Big(\sum_{i=1}^n q_i - 1 \Big),$$

where λ, γ are Lagrange multipliers. Setting to zero the partial derivative of G with respect to q_i gives

$$\frac{\partial G}{\partial q_i} = \frac{1}{q_i} - n\lambda' W_{ni}(\beta_0) + \gamma = 0.$$

Therefore,

$$\sum_{i=1}^n q_i \frac{\partial G}{\partial q_i} = n - n \sum_{i=1}^n \lambda' q_i W_{ni}(\beta_0) + \gamma \sum_{i=1}^n q_i = 0.$$

We get the result

$$q_i = \frac{1}{n}\frac{1}{1+\lambda' W_{ni}(\beta_0)},$$

where the p-dimensional multiplier vector λ satisfies

$$\frac{1}{n}\sum_{i=1}^{n}\frac{W_{ni}(\beta_0)}{1+\lambda' W_{ni}(\beta_0)} = 0.$$

Then, the empirical log-likelihood ratio statistic is

$$\hat{l}(\beta_0) = -2\ln R(\beta_0) = 2\sum_{i=1}^{n}\ln\{1+\lambda' W_{ni}(\beta_0)\}.$$

In practice, one can calculate $\hat{l}(\beta_0)$ directly through $R(\beta_0)$ using the restricted maximization instead of calculating λ first.

It is apparent that one can easily develop a confidence region for β_0 based on $\hat{l}(\beta_0)$ if its distribution is known. Let

$$\begin{aligned}\Pi(\beta_0) &= E\left(W_i(\beta_0)W_i'(\beta_0)\right),\\ \Pi_n(\beta_0) &= \frac{1}{n}\sum_{i=1}^{n}W_i(\beta_0)W_i'(\beta_0),\\ \hat{\Pi}_n(\beta_0) &= \frac{1}{n}\sum_{i=1}^{n}W_{ni}(\beta_0)W_{ni}'(\beta_0).\end{aligned}$$

To derive asymptotic distribution of $\hat{l}(\beta_0)$, we need the following regularity conditions.

Condition 1. The longest follow-up time τ is finite, and the covariate vectors $Z_i(t)$'s are uniformly bounded such that $\| Z_i(t) \| \le C_1$ for some positive constant C_1, where $\| \cdot \|$ denotes the Euclidean norm.

Condition 2. $\int_0^t \lambda^c(s)\,\mathrm{d}s < \infty$ holds for any $0 < t < \infty$.

Condition 3. $\hat{\Lambda}_0(t) - \Lambda_0(t) = o_p(n^{-1/4})$ holds for any $0 < t < \infty$.

Condition 4. The matrix $\Pi(\beta_0)$ is positive definite.

The following theorem states the result for the limiting distribution of $\hat{l}(\beta_0)$.

Theorem 1. Under Conditions $1-4$, empirical log-likelihood ratio statistic $\hat{l}(\beta_0)$ converges in distribution to $\chi^2(p)$, where $\chi^2(p)$ is the standard chi-square distribution with p degree of freedom.

Using Theorem 1, an asymptotic $100(1-\alpha)\%$ confidence region for β_0 is given by

$$\mathcal{R}_1 = \{\beta : \hat{l}(\beta) \le \chi^2_\alpha(p)\}.$$

It is known that in general, one is often more interested in constructing confidence regions for a subvector of β. Rewrite β as $\beta = (\beta^{(1)'}, \beta^{(2)'})'$, where $\beta^{(1)}$ is the q-dimensional $(q < p)$ parameter vector we are interested in, and $\beta^{(2)}$ is nuisance parameter vector. For linear transformation models with interval-censored data and nuisance parameters, Zhang and Zhao [23] proposed a profile empirical likelihood method. Similar to their method, we propose a profile empirical likelihood for β_1 as

$$\tilde{l}(\beta^{(1)}) = \inf_{\beta^{(2)} \in R^{p-q}} \hat{l}(\beta).$$

Let $\beta_0^{(1)}$ denote the true value of $\beta^{(1)}$. The following theorem gives the asymptotic distribution of $\tilde{l}(\beta_0^{(1)})$.

Theorem 2. Assume that Conditions $1-4$ hold true, then as $n \to \infty$, the limiting distribution of $\tilde{l}(\beta_0^{(1)})$ is the standard chi-square distribution with q degree of freedom.

Theorem 2 suggests that an asymptotic $100(1-\alpha)\%$ confidence region for $\beta_0^{(1)}$ can be obtained as

$$\mathcal{R}_2 = \{\, \beta^{(1)} : \tilde{l}(\beta^{(1)}) \le \chi^2_\alpha(q) \,\}.$$

4. NUMERICAL STUDIES

In this section, a series of simulation studies are conducted to evaluate the performance of the empirical likelihood procedure proposed in this paper. For comparison, the confidence region based on the estimation equation method proposed in Feng et al. [3] is also computed.

In our simulations, covariate vectors $Z_i's$ are generated from Bernouli distribution $B(1, 0.5)$, and failure times $T_i's$ are drawn independently from model (1) with $\lambda_0(t) = 2$. To generate current status data, we assume observation times $C_i's$ satisfy proportional hazard model with hazard function given by

$$\lambda_C(t) = \lambda_{0C}(t)\exp(\gamma_0' Z_i),$$

and $\lambda_{0C}(t)$ is set to be $1, 2$ and 4 to make right-censored ratio (CR for short) be $33.3\%, 50\%$, and 66.7% respectively when $\beta_0 = \gamma_0 = 0$. Auxiliary covariate X_i is defined as $X_i = Z_i + e_i$, where $e_i's$ are generated from normal distribution $N(0, \sigma^2)$, and $\sigma = 0.1$ is chosen here. For each i, let ϱ stands for the probability that Z_i is observed, that is, $i \in V$ holds with probability ϱ and $i \in \bar{V}$ holds with probability $1-\varrho$. We set ϱ to be 0.3, 0.5, and 1 in different simulations. Gaussian function $Q(u) = (2\pi)^{-\frac{1}{2}}\exp(-u^2/2)$ is used as kernel and the bandwidth is set as $b = 2n_v^{-1/3}\hat{\sigma}_Z$, where $\hat{\sigma}_Z$ denotes the sample standard deviation of Z_i in validation set with sample size n_v.We consider $\beta_0 = -0.5, 0$ or 0.5, $\gamma_0 = -0.1, 0, 0.1$ to show the performance of our method under different models. For each set of parameters, we generate $n = 200$ independent subjects and each study is repeated 1000 times.

Simulation results are listed in tables 1−6, including the 95% empirical coverage probabilities(CP), Bias, empirical standard errors(SE) and mean errors(ME) of the estimation equation method(EE) and that of our empirical likelihood method(EL) under various cases. In order to show the simulation results more clearly, according to tables 1−3, we draw a line chart as shown in Fig.1 and the order of horizontal axis is of nonsense. In Fig.1, blue line and orange line respectively represent 95% empirical CP of EE and EL. And for convenient comparison, a horizontal line of 95% is drew as red line.

		CR=33.3%		CR=50%		CR=66.7%	
γ	β	EE	EL	EE	EL	EE	EL
-0.1	-0.5	91.0%	95.4%	90.0%	94.8%	89.9%	95.3%
	0	94.1%	93.8%	94.4%	95.4%	95.4%	95.5%
	0.5	95.4%	94.4%	95.2%	93.9%	97.0%	95.4%
0	-0.5	92.1%	94.3%	91.4%	93.0%	89.8%	92.4%
	0	96.8%	95.7%	94.1%	92.8%	93.1%	92.6%
	0.5	95.8%	92.8%	96.2%	93.8%	95.4%	93.8%
0.1	-0.5	92.6%	93.6%	92.4%	92.5%	91.5%	93.4%
	0	94.5%	92.5%	94.8%	94.6%	96.1%	95.1%
	0.5	96.3%	94.0%	96.3%	94.5%	95.7%	93.6%

Tab. 1. Simulation results of empirical CP with $\varrho = 30\%$.

		CR=33.3%		CR=50%		CR=66.7%	
γ	β	EE	EL	EE	EL	EE	EL
-0.1	-0.5	89.6%	94.9%	90.8%	94.2%	88.3%	94.0%
	0	93.4%	94.9%	93.9%	95.8%	93.7%	93.9%
	0.5	96.0%	93.9%	96.7%	95.8%	95.7%	94.4%
0	-0.5	90.5%	93.0%	91.6%	94.9%	89.8%	94.0%
	0	95.1%	94.8%	94.1%	93.6%	95.2%	95.8%
	0.5	95.6%	93.9%	96.6%	94.5%	95.0%	93.6%
0.1	-0.5	93.0%	93.1%	91.9%	93.4%	89.8%	93.4%
	0	95.4%	93.8%	95.6%	93.5%	95.7%	92.7%
	0.5	96.5%	94.7%	95.4%	93.6%	94.9%	93.9%

Tab. 2. Simulation results of empirical CP with $\varrho = 50\%$.

		CR=33.3%		CR=50%		CR=66.7%	
γ	β	EE	EL	EE	EL	EE	EL
-0.1	-0.5	90.1%	94.5%	90.7%	94.9%	89.7%	95.2%
	0	94.9%	95.8%	94.1%	94.6%	93.5%	94.3%
	0.5	95.4%	94.4%	94.6%	94.4%	94.8%	94.7%
0	-0.5	92.4%	94.7%	90.5%	93.8%	89.9%	93.1%
	0	93.9%	93.0%	94.6%	94.0%	95.5%	94.2%
	0.5	95.9%	95.3%	95.3%	94.2%	95.6%	94.1%
0.1	-0.5	92.3%	94.0%	93.1%	92.7%	91.5%	92.1%
	0	94.7%	93.3%	96.0%	94.3%	93.8%	92.9%
	0.5	96.6%	93.5%	95.2%	93.5%	96.5%	95.0%

Tab. 3. Simulation results of empirical CP with $\varrho = 100\%$.

		CR=33.3%						CR=50%						CR=66.7%					
		Bias		SE		ME		Bias		SE		ME		Bias		SE		ME	
γ	β	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL
-0.1	-0.5	-0.226	0.0214	0.3114	0.1738	0.2960	0.1695	-0.0594	-0.0404	0.6237	0.3320	0.5818	0.3248	-0.0239	-0.0252	1.032	0.9296	0.9520	0.8888
	0	-0.0025	0.0083	0.3489	0.1873	0.3423	0.1936	-0.0360	0.0156	0.7009	0.3778	0.6700	0.3659	0.0124	-0.0111	1.1020	0.9980	0.9886	0.9576
	0.5	0.0170	0.0170	0.4319	0.2358	0.4145	0.2358	0.0584	0.0179	0.8682	0.4319	0.7852	0.4260	0.0921	0.1603	1.3508	1.1800	1.1260	1.0743
0	-0.5	-0.0188	-0.0143	0.3045	0.1686	0.2966	0.1692	-0.0515	0.0461	0.6265	0.4390	0.5834	0.4270	-0.1245	-0.1322	1.0120	1.0023	0.9703	0.8903
	0	-0.0107	0.0049	0.3489	0.1973	0.3405	0.1936	-0.0122	0.0150	0.7138	0.3650	0.6643	0.3655	0.0231	-0.0111	1.0421	0.9980	0.9832	0.9576
	0.5	0.0241	0.0098	0.4473	0.2286	0.4166	0.2348	0.0888	0.0040	0.9419	0.4364	0.7989	0.4310	-0.06318	0.0547	1.1320	1.2398	1.1140	1.1065
0.1	-0.5	-0.0087	-0.0140	0.3056	0.1693	0.2968	0.1684	-0.0826	-0.0356	0.6397	0.3339	0.5875	0.3222	0.07371	-0.0616	0.9487	0.9390	0.9032	0.8745
	0	-0.0018	-0.0079	0.3367	0.1940	0.3429	0.1939	-0.0109	0.0135	0.7124	0.3750	0.6678	0.3662	-0.0531	0.0412	1.1047	1.0629	1.0045	0.9897
	0.5	0.0462	0.0217	0.4116	0.2390	0.4199	0.2377	0.0848	0.0555	0.8530	0.4662	0.8035	0.4336	0.1145	0.1019	1.3863	1.1606	1.1080	1.0700

Tab. 4. Simulation results of Bias, SE and ME with $\varrho = 30\%$.

		CR=33.3%						CR=50%						CR=66.7%					
		Bias		SE		ME		Bias		SE		ME		Bias		SE		ME	
γ	β	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL
-0.1	-0.5	-0.0112	-0.0109	0.2321	0.1623	0.2223	0.1624	-0.0020	-0.0358	0.4260	0.3059	0.4266	0.3081	-0.0697	-0.0796	1.4168	0.8716	1.2321	0.8384
	0	0.0036	0.0042	0.2570	0.1916	0.2564	0.1870	0.0104	0.0139	0.4859	0.3626	0.4803	0.3530	0.0118	-0.0162	1.5615	0.9495	1.3811	0.9151
	0.5	0.0057	0.0165	0.3161	0.2319	0.3092	0.2274	0.0166	0.0013	0.5846	0.4154	0.5698	0.4109	0.1248	0.1429	1.7929	1.1154	1.5371	1.0284
0	-0.5	0.0019	-0.0087	0.2199	0.1629	0.2220	0.1623	-0.0428	-0.0116	0.4544	0.3144	0.4345	0.3085	-0.2224	-0.1068	1.4909	0.9220	1.2520	0.8495
	0	0.0052	0.0037	0.2608	0.1845	0.2574	0.1865	-0.0058	0.0136	0.5066	0.3503	0.4864	0.3524	0.0128	-0.0082	1.6074	0.9837	1.3577	0.9129
	0.5	0.0119	0.0314	0.2582	0.2268	0.2642	0.2283	0.0620	0.0317	0.6062	0.4224	0.5754	0.4154	0.0789	0.0943	1.8946	1.1010	1.5624	1.0402
0.1	-0.5	-0.0196	0.0011	0.2363	0.1690	0.2233	0.1623	-0.0458	-0.0173	0.4341	0.3186	0.4250	0.3108	-0.0674	-0.0672	1.3499	0.9024	1.2187	0.8430
	0	-0.0028	0.0069	0.2404	0.1883	0.2567	0.1870	0.0061	-0.0135	0.5380	0.3554	0.4878	0.3521	0.0200	-0.0087	1.6104	1.0141	1.3722	0.9256
	0.5	0.1349	-0.0038	0.3144	0.2189	0.3127	0.2257	0.0547	0.0230	0.6180	0.4245	0.5839	0.4137	0.2092	0.0638	1.8202	1.0567	1.5556	1.0148

Tab. 5. Simulation results of Bias, SE and ME with $\varrho = 50\%$.

		CR=33.3%						CR=50%						CR=66.7%					
		Bias		SE		ME		Bias		SE		ME		Bias		SE		ME	
γ	β	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL	EE	EL
-0.1	-0.5	-0.0028	0.0050	0.1578	0.1583	0.1535	0.1522	-0.0238	-0.0404	0.3077	0.3320	0.2905	0.3248	-0.0740	-0.0534	0.8114	0.8454	0.7964	0.7876
	0	0.0042	-0.0018	0.1749	0.1794	0.1766	0.1758	-0.0018	-0.0096	0.3272	0.2993	0.3301	0.2907	-0.0249	0.0075	0.8899	0.9024	0.8763	0.8595
	0.5	0.0103	0.0081	0.2171	0.2167	0.2144	0.2129	0.0314	-0.0035	0.4043	0.3480	0.3902	0.3313	0.0859	0.0416	1.0562	1.0089	0.9734	0.9477
0	-0.5	-0.0074	0.0082	0.1572	0.1501	0.1535	0.1510	-0.0267	0.0088	0.3007	0.3751	0.2931	0.3842	-0.0718	0.0026	0.8414	0.7964	0.7925	0.7860
	0	-0.0033	-0.0077	0.1731	0.1820	0.1767	0.1757	0.0125	-0.0184	0.3398	0.3085	0.3322	0.2910	0.0221	0.0065	0.9444	0.9337	0.8768	0.8630
	0.5	0.0024	0.0142	0.2156	0.2229	0.2145	0.2137	0.0254	0.0093	0.3842	0.3334	0.3881	0.3304	0.0570	0.0754	1.0482	1.0180	0.9708	0.9491
0.1	-0.5	-0.0023	-0.0099	0.1562	0.1562	0.1534	0.1521	-0.0054	0.0486	0.3076	0.3862	0.2916	0.3848	-0.0432	-0.0830	0.8029	0.8917	0.7957	0.7988
	0	-0.0044	-0.0053	0.1832	0.1787	0.1773	0.1765	-0.0082	-0.0239	0.3493	0.2933	0.3302	0.2906	0.0244	-0.0096	0.9401	0.9298	0.8762	0.8594
	0.5	0.0155	0.0022	0.2214	0.2124	0.2146	0.2135	0.0260	0.0165	0.4080	0.3898	0.3825	0.3861	0.1074	0.0364	1.0825	0.9850	0.9843	0.9406

Tab. 6. Simulation results of Bias, SE and ME with $\varrho = 100\%$.

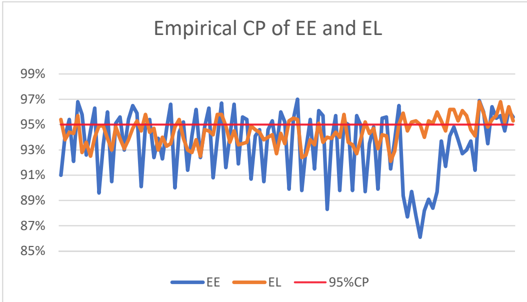

Fig. 1. line chart of 95% empirical CP of EE and EL.

The results show that EL method performs well under all situations we consider here. Details are as follows:

1. From the line chart plot, the 95% empirical CP line of EL fluctuates around 95%, but the line of the EE method fluctuates much more intense, and many results of EE method are much lower than 95%.

2. From concrete results in above tables 1-3, the minimal and maximal empirical CP of EL is 92.1% and 96.8%, whereas that of EE is 86.1% and 97%, which are both more far away from 95% than EL. As pointed out by Feng *et al.*[3] that the EE could underestimate the variance of the estimated parameters. So it could lead to lower coverage probabilities. Especially when both γ and β are less than 0, most of 95% empirical CP of EE are under 90%, which is obviously an awful result. However, under such situations, empirical CP of EL keeps around 95%.

3. From concrete results in above tables 4-6, the majority of results of EE are more far away from 0 than EL, whether it is |bias|, SE or ME. Especially, the results of ME are obviously better than those of EE.

In summary, it is indicates that the EL method performs more stable in different situations and is more practical than EE.

5. DISCUSSION

This paper discussed regression analysis of case I interval-censored failure time data arising from the additive hazards model when there exist missing covariates. Moreover, it has been assumed that some auxiliary covariates have been observed for all subjects. Corresponding to the estimating equation approach proposed in Chen et al. [1], this paper developed an empirical likelihood method for deriving the confidence regions for regression parameters.

The method has the advantages that it does not require any assumption on the distribution of the latent variables as well as the estimation of the unknown baseline hazard functions and can be easily implemented. In fact, one could easily obtain $\hat{\beta}$ by using some Matlab functions. For example, to obtain, one way is to directly minimize $U(\beta, \Lambda_0)$, this approaches, one could employ the Matlab functions 'fminsearch'. Furthermore, the numerical results indicated that the proposed approach works well for practical situations.

There exist several directions for future research related to the problem discuss here. One is that, throughout the paper, we are concerned with the current status data(type I interval censored data) , but in practice, one could face other types of interval-censored data, such as type II interval censored data, for which it does not seem to exist estimation procedures similar to the ones proposed above. Thus it is natural to develop some similar estimation procedures for them. The second one is that, we have assumed that the censored data is independent, and this may not be true in practice. In fact it is also straightforward to generalize the approach proposed above to the case where the censored data is dependent. In this case, we can rewrite the model (2) as

$$\begin{aligned}\lambda_i(t|Z_i(s), s \le t) &= e^{-\Lambda_0(t)} e^{-\beta_0' Z_i^*(t) + \gamma_0' Z_i(t)} \lambda^c(t) \\ &\triangleq \lambda_0^c(t) e^{-\beta_0' Z_i^*(t) + \gamma_0' Z_i(t)}\end{aligned}$$

where for the C_i, we assume that given Z_i, the hazard function of C_i has the form $\lambda_C(t|Z_i(s), s \le t) = \lambda^c(t) e^{\gamma_0' Z_i(t)}$. Similarly as the preceding sections, one can develop an inference procedure and establish some limit theories.

Finally, we have assumed that the current status data follow the additive hazards model. In practice, sometimes this may not be true and it would be useful to develop similar methods for other models such as the multiplicative hazards model. In the previous sections, we have assumed that all components of a covariate vector are either missing or observed together. In fact, the above inference method can be developed to the case where some components are observed and some components are missing but only auxiliary information obtained.

6. APPENDIX: PROOFS

To prove Theorem 1, we need the following lemmas.

Lemma 1. Under Conditions 1–4, as $n \to \infty$, we have

$$(i)\ \hat{\Pi}_n(\beta_0) \xrightarrow{P} \Pi(\beta_0)\ ,\ (ii)\ \tilde{\Pi}(\beta_0, \hat{\Lambda}_0) \xrightarrow{P} \Pi(\beta_0)\ ,$$

where "$\xrightarrow{P}$" means convergence in probability.

Proof. In order to prove (i), it is sufficient to show that $\hat{\Pi}_n = \Pi_n + o_p(1)$. In fact, for any $a \in R^p$, we have

$$\begin{aligned}a'(\hat{\Pi}_n - \Pi_n)a &= \frac{1}{n}\sum_{i=1}^{n}(a'(W_{ni}(\beta_0) - W_i(\beta_0)))^2 \\ &+ \frac{2}{n}\sum_{i=1}^{n}(a'W_i(\beta_0))(a'(W_{ni}(\beta_0) - W_i(\beta_0))).\end{aligned}$$

Denote $(W_{ni}(\beta_0) - W_i(\beta_0))$ as $\gamma_i(\beta_0)$, and note that

$$\begin{aligned} a'\gamma_i(\beta_0) &= a'\int_0^\tau \frac{\Phi_i^{(1)}(t,\beta_0)}{\Phi_i(t,\beta_0)}\left(\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} - \frac{p_i(t,\beta_0)}{1-p_i(t,\beta_0)}\right)\mathrm{d}M_i^*(t) \\ &+a'\int_0^\tau \left(\frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)} - \frac{S_1(t,\beta_0)}{S_0(t,\beta_0)}\right)\mathrm{d}M_i(t) \\ &-a'\int_0^\tau \left(\frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)}\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)} - \frac{S_1(t,\beta_0)}{S_0(t,\beta_0)}\frac{p_i(t,\beta_0)}{1-p_i(t,\beta_0)}\right)\mathrm{d}M_i^*(t). \end{aligned}$$

Under Conditions 1−3, it is easy to show that $a'\gamma_i(\beta_0) = o_p(1)$. On the other hand, $n^{-1}\sum_{i=1}^n a'W_i(\beta_0)$ converges to $E(a'W_i(\beta_0)) = 0$ in probability. Thus we have $a'(\hat{\Pi}_n - \Pi_n)a$ converges to 0 in probability, so (i) holds. By using the same proof as above, we can show that result (ii) holds true. This completes the proof of Lemma 1. □

Lemma 2. Under Condition 1−4, as $n \to \infty$, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n W_{ni}(\beta_0) \xrightarrow{D} N(0, \Pi(\beta_0)),$$

where "$\xrightarrow{D}$" means convergence in distribution.

Proof. Note the equation

$$\hat{U}(\beta_0, \hat{\Lambda}_0) = \sum_{i=1}^n \int_0^\tau \left(\frac{\hat{\Phi}_i^{(1)}(t,\beta_0)}{\hat{\Phi}_i(t,\beta_0)} - \frac{\hat{S}_1(t,\beta_0)}{\hat{S}_0(t,\beta_0)}\right)\left(\frac{\hat{p}_i(t,\beta_0)}{1-\hat{p}_i(t,\beta_0)}\mathrm{d}M_i^*(t) - \mathrm{d}M_i(t)\right),$$

and then under Condition 4, we can show that

$$\hat{U}(\beta_0, \hat{\Lambda}_0) = U(\beta_0, \Lambda_0) + o_p(1) \tag{A.1}$$

holds true by applying Lemma A.1 of Lin and Ying [8]. From the definition of $U(\beta_0, \Lambda_0)$, it can be easily seen that $U(\beta_0, \Lambda_0)/\sqrt{n}$ converges in distribution to a normal variable with mean zero and covariance matrix $\Pi(\beta_0)$. Combining this result with equation (A.1), Lemma 2 follows. □

Based on Lemma 1 and Lemma 2, the proof of Theorem 1 is given below.

Proof of Theorem 1: Using Conditions 1−2, it is easy to see that $E \parallel W_{ni}(\beta_0) \parallel^2 < \infty$. Applying Lemma 3 in Owen [14], we have that both

$$\max_{1\le i\le n} \|W_{ni}(\beta_0)\| = o(n^{1/2}) \tag{A.2}$$

and

$$\frac{1}{n}\sum_{i=1}^n \|W_{ni}(\beta_0)\|^3 = o(n^{1/2}) \tag{A.3}$$

hold true with probability 1. Write $\lambda = \rho\theta$, where $\rho \geq 0$ is a real number and $\| \theta \| = 1$. By equations (A.2), (A.3) and an argument similar to that used in Owen [14], we have

$$\| \lambda \| = \rho = O_p(n^{-1/2}). \tag{A.4}$$

Combining equations (A.2), (A.3) and (A.4), we have

$$\max_{1 \leq i \leq n} |\lambda' W_{ni}(\beta_0)| = o_p(1).$$

Next, we apply the Taylor's expansions to $\hat{l}(\beta_0) = 2\sum_{i=1}^n \ln(1 + \lambda' W_{ni}(\beta_0))$ and $\sum_{i=1}^n W_{ni}(\beta)/(1 + \lambda' W_{ni}(\beta))$, and obtain

$$\begin{aligned}
\hat{l}(\beta_0) &= 2\sum_{i=1}^n \left\{ \lambda' W_{ni}(\beta_0) - (\lambda' W_{ni}(\beta_0))^2 + O_p\left((\lambda' W_{ni}(\beta_0))^3\right) \right\} \\
&= \left(n^{-1/2} \sum_{i=1}^n W_{ni}(\beta_0) \right)' \hat{\Pi}_n^{-1}(\beta_0) \left(n^{-1/2} \sum_{i=1}^n W_{ni}(\beta_0) \right) + o_p(1).
\end{aligned}$$

Then Theorem 1 follows from Lammas 1 and 2 immediately. □

The following is the proof of Theorem 2.

Proof of Theorem 2: Rewrite β_0 as $\beta_0 = (\beta_0^{(1)'}, \beta_0^{(2)'})'$ corresponding to $\beta = (\beta^{(1)'}, \beta^{(2)'})'$. Assume that $\tilde{\beta}^{(2)}$ satisfies

$$\hat{l}((\beta_0^{(1)'}, \tilde{\beta}^{(2)})') = \inf_{\beta^{(2)} \in R^{p-q}} \hat{l}((\beta_0^{(1)'}, \beta^{(2)'})').$$

To simplify the notation we now write $(\beta_0^{(1)'}, \tilde{\beta}^{(2)'})'$ as $\tilde{\beta}_0$. Similar to (5), $\tilde{\beta}^{(2)}$ satisfies the estimated empirical likelihood function

$$L_n(\tilde{\beta}_0) = \sup\{\prod_{i=1}^n h_i : \sum_{i=1}^n h_i = 1, \sum_{i=1}^n h_i W_{ni}(\tilde{\beta}_0) = 0, h_i \geq 0\}.$$

Using the lagrange multiplier method, we get the result

$$h_i = \frac{1}{n} \frac{1}{1 + \tilde{\lambda}^{(2)'} W_{ni}(\tilde{\beta}_0)},$$

where the $p-q$-dimensional multiplier vector $\tilde{\lambda}^{(2)}$ satisfies

$$\frac{1}{n} \sum_{i=1}^n \frac{W_{ni}(\tilde{\beta}_0)}{1 + \tilde{\lambda}^{(2)'} W_{ni}(\tilde{\beta}_0)} = 0.$$

Using exactly the same arguments as in Lemma1 and Theorem1 of Qin and Lawless [17] (the conditions can be easily verified), we have

$$\sqrt{n}(\tilde{\beta}^{(2)} - \beta_0^{(2)}) = -(\Lambda_2' \Pi^{-1}(\beta_0) \Lambda_2)^{-1} \Lambda_2' \Pi^{-1}(\beta_0) n^{-1/2} \sum_{i=1}^n W_{ni}(\beta_0) + o_p(1)$$

and

$$\begin{aligned}\sqrt{n}\tilde{\lambda}^{(2)} &= \{\mathrm{I} - \Pi^{-1}(\beta_0)\Lambda_2(\Lambda_2'\Pi^{-1}(\beta_0)\Lambda_2)^{-1}\Lambda_2'\}\Pi^{-1}(\beta_0)n^{-1/2}\sum_{i=1}^{n} W_{ni}(\beta_0) + o_p(1) \\ &= Sn^{-1/2}\sum_{i=1}^{n} W_{ni}(\beta_0) + o_p(1),\end{aligned}$$

where $S = \left(\mathrm{I} - \Pi^{-1}(\beta_0)\Lambda_2(\Lambda_2'\Pi^{-1}(\beta_0)\Lambda_2)^{-1}\Lambda_2'\right)\Pi^{-1}(\beta_0)$, $\Lambda_2 = \sum_{i=1}^{n} \partial W_i(\beta_0)/\partial\beta^{(2)}$, and I is an identity matrix of order p. Following these, the profile empirical log-likelihood ratio statistic is

$$\tilde{l}(\beta_0^{(1)}) = 2\sum_{i=1}^{n} \ln\left\{ 1 + \tilde{\lambda}^{(2)'} W_{ni}(\tilde{\beta}_0) \right\}.$$

Next, using a similar Taylor expansion as that used in the proof of Theorem 1, we obtain

$$\begin{aligned}\tilde{l}(\beta_0^{(1)}) &= 2\sum_{i=1}^{n} \ln\{1 + \tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0)\} \\ &= 2\sum_{i=1}^{n}\{\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0) - \frac{1}{2}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0)W_{ni}(\tilde{\beta}_0)'\tilde{\lambda}^{(2)} + o_p((\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0))^2)\} \\ &= 2\sum_{i=1}^{n}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0) - \sum_{i=1}^{n}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0)W_{ni}(\tilde{\beta}_0)'\tilde{\lambda}^{(2)} + o_p(1) \\ &= \sum_{i=1}^{n}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0) + o_p(1).\end{aligned}$$

Moreover, $\sum_{i=1}^{n}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0)$ can be expanded as

$$\begin{aligned}\sum_{i=1}^{n}\tilde{\lambda}^{(2)'}W_{ni}(\tilde{\beta}_0) &= \sum_{i=1}^{n}\tilde{\lambda}^{(2)'}\{W_{ni}(\beta_0) + \frac{\partial W_i(\beta^*)}{\partial\beta^{(2)}}(\tilde{\beta}_2 - \beta_0^{(2)})\} \\ &= \tilde{\lambda}^{(2)'}\sum_{i=1}^{n}W_{ni}(\beta_0) + \tilde{\lambda}^{(2)'}\sum_{i=1}^{n}\frac{\partial W_i(\beta^*)}{\partial\beta^{(2)}}(\tilde{\beta}_2 - \beta_0^{(2)})\} \\ &= \Big[\sum_{i=1}^{n}W_{ni}(\beta_0)'/\sqrt{n}\Big]S\Big[\sum_{i=1}^{n}W_{ni}(\beta_0)/\sqrt{n}\Big] + o_p(1)\end{aligned}$$

because of the fact $S\Lambda_2 = 0$, where β^* is on the line segment between $\tilde{\beta}_0$ and β_0.

Therefore, we have that

$$\tilde{l}(\beta_0^{(1)}) = \Big[\sum_{i=1}^{n}W_{ni}(\beta_0)/\sqrt{n}\Big]'S\Big[\sum_{i=1}^{n}W_{ni}(\beta_0)/\sqrt{n}\Big] + o_p(1).$$

Using again the approximation adopted in Appendix C of Zhang et al. [22], i. e. Lemma 2 holds and the fact that $I - \Pi^{-\frac{1}{2}}(\beta_0)\Lambda_2(\Lambda_2'\Pi^{-1}(\beta_0)\Lambda_2)^{-1}\Lambda_2'\Pi^{-\frac{1}{2}}(\beta_0)$ is a symmetric and idempotent matrix with trace q, we have the desired result. □

7. ACKNOWLEDGEMENTS

This work was supported in part by Funding Project of National Natural Science (31872847), Social science planning project of Shandong Province (20CTJJ03) and Doctoral Funding of Weifang (2016BS11), and the work of the third author is partially supported by the Project of National Social Science Foundation(18BTJ040).

(Received July 14, 2020)

REFERENCES

[1] Y. Chen, Y. Feng, and J. Sun: Regression analysis of multivariate current status data with auxiliary covariates under the additive hazards model. Comput. Statist. Data Anal. *87* (2015), 34–45. DOI:10.1016/j.csda.2015.01.005

[2] L. Chen and J. Sun: A multiple imputation approach to the analysis of current status data with the additive hazards model. Commun. Stat. Theor. Meth. *38* (2009), 1009–1018.

[3] Y. Feng, L. Ma, and J. Sun: Additive hazards regression with auxiliary covariate for case I interval censored data. Scand. J. Stat. *42* (2015), 118–136. DOI:10.1111/sjos.12098

[4] J. Huang: Efficient estimation for the proportional hazards model with interval censoring. Ann. Statist. *24* (1996), 540–568.

[5] N. Jewell and M. van der Laan: Generalizations of current status data with applications. Lifetime Data Anal. *1* (1995), 101–110. DOI:10.1007/BF00985261

[6] D. Lin, D. Oakes, and Z. Ying: Additive hazards regression with current status data. Biometrika *85* (1998), 289–298. DOI:10.1093/biomet/85.2.289

[7] D. Lin and Z. Ying: Semiparametric analysis of the additive risk model. Biometrika *81* (1994), 61–71. DOI:10.1093/biomet/81.1.61

[8] D. Lin and Z. Ying: Semiparametric and nonparametric regression analysis of longitudinal data. J. Amer. Statist. Assoc. *96* (2001), 453, 103–126. DOI:10.1198/016214501750333018

[9] W. Liu, X. Lu, and C. Xie: Empirical likelihood for the additive hazards model with current status data. Comm. Statist. Simulation Comput. *45* (2016), 8, 2720–2732.

[10] X. Lu and Peter X. K. Song: On efficient estimation in additive hazards regression with current status data. Comput. Statist. Data Anal. *56* (2012), 6, 2051–2058. DOI:10.1016/j.csda.2011.12.011

[11] T. Martinussen and T. H. Scheike: Efficient estimation in additive hazards regression with current status data. Biometrika *89* (2002), 649–658. DOI:10.1093/biomet/89.3.649

[12] I. McKeague and P. Sasieni: A partly parametric additive risk model. Biometrika *81* (1994), 501–514. DOI:10.1093/biomet/81.3.501

[13] A. B. Owen: Empirical Likelihood. Chapman and Hall, 2001.

[14] A. B. Owen: Empirical likelihood and confidence regions. Ann. Statist. *18* (1990), 90–120. DOI: 10.1214/aos/1176347494

[15] A. B. Owen: Empirical likelihood ratio confidence intervals for a single functional. Biometrika *75* (1988), 237–249. DOI:10.1093/biomet/75.2.237

[16] R. Prentice: Covariate measurement errors and parameter estimation in a failure time regression model. Biometrika *69* (1982), 331–342. DOI:10.1093/biomet/69.2.331

[17] J. Qin and J. Lawless: Empirical likelihood and general estimating equations. Ann. Statist. *22* (1994), 300–325. DOI:10.1214/aos/1176325370

[18] J. Rossini and A. A. Tsiatis: A semiparametric proportional odds regression model for the analysis of current status data. J. Am. Stat. Assoc. *91* (1996), 713–721.

[19] J. Sun: The Statistical Analysis of Interval-censored Failure Time Data. Springer, New York 2006.

[20] N. Wang, L. Wang, and C. McMahan: Regression analysis of bivariate current status data under the gamma-frailty proportional hazards model using the EM algorithm. Comput. Stat. Data Anal. *83* (2015), 140–150. DOI:10.1016/j.csda.2014.10.013

[21] Z. Zhang, J. Sun, and L. Sun: Statistical analysis of current status data with informative observation times. Stat. Med. *24* (2005), 1399–1407. DOI:10.1002/sim.2001

[22] Z. Zhang, L. Sun, X. Zhao, and J. Sun: Regression analysis of interval-censored failure time data with linear transformation models. Canad. J. Statist. *33* (2005), 61–70. DOI:10.1002/cjs.5540330105

[23] Z. Zhang and Y. Zhao: Empirical likelihood for linear transformation models with interval-censored failure time data. J. Multivariate Anal. *116* (2013), 398–409. DOI:10.1016/j.jmva.2013.01.003

[24] Y. Zhao and Y. S. Hsu: Semiparametric analysis for additive risk model via empirical likelihood. Commun. Stat.-Simul. Comput. *34* (2005), 135–143. DOI:10.1081/sac-200047114

[25] L. Zhu, X. Tong, and J. Sun: A transformation approach for the analysis of interval-censored failure time data. Lifetime Data Anal. *14* (2008), 167–178. DOI:10.1007/s10985-007-9075-8

[26] M. Zhou: Empirical likelihood analysis of the rank estimator for the censored accelerated failure time model. Biometrika *92* (2005), 492–498. DOI:10.1093/biomet/92.2.492

Jianling Zhang, School of Mathematics and Information Sciences, Weifang University. P. R. China.
e-mail: zhangjianling@wfu.edu.cn

Mei Yang, School of Mathematics and Statistics, Wuhan University. P. R. China.
e-mail: 1414972663@qq.com

Xiuqing Zhou, School of Mathematical Sciences and Institute of Finance and Statistics, Nanjing Normal University. P. R. China.
e-mail: zhouxiuqing@njnu.edu.cn