Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 215–231
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Empirical likelihood inference for censored median regression model via nonparametric kernel estimation

Yichuan Zhao[a,*], Feiming Chen[b]

[a]*Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA*
[b]*Spectra Marketing Systems, Chicago, IL 60661, USA*

Received 1 March 2006
Available online 16 May 2007

Abstract

An alternative to the accelerated failure time model is to regress the median of the failure time on the covariates. In the recent years, censored median regression models have been shown to be useful for analyzing a variety of censored survival data with the robustness property. Based on missing information principle, a semiparametric inference procedure for regression parameter has been developed when censoring variable depends on continuous covariate. In order to improve the low coverage accuracy of such procedure, we apply an empirical likelihood ratio method (EL) to the model and derive the limiting distributions of the estimated and adjusted empirical likelihood ratios for the vector of regression parameter. Two kinds of EL confidence regions for the unknown vector of regression parameters are obtained accordingly. We conduct an extensive simulation study to compare the performance of the proposed methods with that normal approximation based method. The simulation results suggest that the EL methods outperform the normal approximation based method in terms of coverage probability. Finally, we make some discussions about our methods.

AMS 2000 subject classification: 62N02; 62G20

Keywords: Confidence regions; Conditional Nelson–Aalen estimator; Coverage probability; Least absolute deviations; Right censoring

1. Introduction

It is well known that Cox's regression model is popular in the analysis of survival data. An alternative to the Cox [5] proportional hazards model is the accelerate failure time (AFT) model. The AFT model is attractive due to its easy and simple interpretation. Some related early work

* Corresponding author. Fax: +1 404 651 2246.
E-mail addresses: yzhao@mathstat.gsu.edu (Y. Zhao), feiming_chen@spectramarketing.com (F. Chen).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2007.05.002

includes Buckley and James [2], Koul et al. [16], Lai and Ying [17], Ritov [32], Tsiatis [38], Wei et al. [42], among others.

Bassett and Koenker [1] and Koenker and Bassett [15] introduced the least absolute deviations (LAD) method for the robust estimation of median regression models with uncensored data. In recent years, the median regression models have shown many useful applications in economics, biology, ecology, finance and statistics, etc. In biomedical research, median life length provides accurate and meaningful information on the survival when the distribution is heavily skewed. Once the regression parameters of the model are estimated, it is straightforward to obtain the information about median life length. Thus, the median regression model offers an attractive alternative to both the Cox regression model and AFT model with the robustness property.

We consider the problem of fitting a median regression model from right censored data. Let T_i $(i = 1, \dots, n)$ be the response of interest. Let $Z_i = (Z_{0i}, Z_{1i}, \dots, Z_{pi})'$, where $Z_{0i} = 1$ be the corresponding $p+1$ dimensional covariate vector. The censoring variable C_i is assumed to be conditionally independent of T_i given the covariate Z_i for $1 \leqslant i \leqslant n$. We observe (X_i, δ_i), where $X_i = \min(T_i, C_i)$ and $\delta_i = I(X_i = T_i)$.

Suppose

$$T_i = \beta' Z_i + \epsilon_i, \tag{1.1}$$

where β is a $(p+1) \times 1$ vector of unknown regression parameter. The joint distribution of the observation error ϵ_i and the covariate Z_i is unknown, but the conditional median of ϵ_i is zero.

Ying et al. [44] investigated censored median regression model in which the censoring is independent of the covariate, or that Z takes only discrete values. More recently, Qin and Tsao [30] proposed an alternative semiparametric procedure based on the estimating equation proposed by Ying et al. [44]. The theoretical result holds only when censoring is independent of the covariate, or when the covariate is discrete. It is known that the independence of C_i from Z_i is rare in clinical trials. Thus, the independence assumption where the censoring variable is independent of the covariate is restrictive in the applications. Hence, the proposed method is not applicable when censoring depends on continuous covariate and it limits the application of the proposed method in practice.

Recently, McKeague et al. [21] applied missing information principle (MIP) to the median regression model (1.1) and proposed a new estimating equation. The MIP originated from Orchard and Woodbury [25] (cf. [18]). The idea of MIP is to replace a full-data estimating equation by its estimated conditional expectation given the observed data. It gives a general way to construct estimating equation in incomplete data problem. Moreover, the simulation study by McKeague et al. [21] demonstrates that the new estimator has improved moderate-sample performance of the estimator of Ying et al. [44]. Applying Dabrowska's [6] kernel conditional Kaplan–Meier estimator for the conditional survival function to estimate $F(t, z)$ for a one-dimensional continuous covariate, Subramanian [36] proposed an estimating equation with MIP. Subramanian [36] removes the independence assumption of censoring variable from covariates in Ying et al. [44], and allows censoring variable to depend on a one-dimensional covariate which is more suitable in clinical trials. In addition, he derived the large sample properties of the MIP estimator. The simulation study shows that the estimator of Ying et al. [44] performs significantly worse than the MIP estimator. However, Subramanian [36] had some overcoverage problems, sometimes severely for small sample and heavy censoring.

Empirical likelihood (EL) method is a powerful nonparametric method for constructing confidence regions of parameters. It is well known that it holds some unique and desirable properties, such as range respecting, transformation-preserving, asymmetric confidence interval, Bartlett correctability, and better coverage probability for small sample, see for example DiCiccio et al. [7], DiCiccio and Romano [8,9], Hall [13]. The EL approach does not require to estimate the limiting variance matrix. Moreover, the confidence region is adapted to the data set and not necessarily symmetric. Thus, it reflects the nature of the underlying data and hence give a more representative way to make inferences about the parameter of interest.

The use of empirical likelihood in survival analysis goes back to Thomas and Grunkemeier [37] who derived pointwise confidence intervals for survival function with right censored data. In the seminar papers, Owen [26,27] introduced novel empirical likelihood confidence regions for the mean of a random vector based on i.i.d. complete data. Since then, the empirical likelihood has been widely applied to do statistical inferences in various statistical settings due to its excellent properties and well-recognized advantage compared to other methods such as normal-approximation and bootstrap methods. Recent work of empirical likelihood includes construction of simultaneous confidence band for right-censored data under a variety of setting [14,11,19,22–24], linear model [28,3,4], linear model for missing data [40,41], partial linear regression model [39,33], regression analysis of long-term survival rate [45], the semiparametric additive risk model [46], missing response problem and the application in observational studies [30], nonlinear errors-in-covariables models with validation data [35], weighted empirical likelihood [12], among others.

In order to overcome the limitation of existing methods, we adopt EL approach. We consider censored median regression model (1.1) with continuous covariate, make full use of the estimating equation of Subramanian [36], and propose two kinds of empirical likelihood-based confidence regions for unknown vector of regression parameter. The profiled EL is used to obtain the estimator of regression parameter. The MIP and EL estimators are identical (cf. Section 2.2). The EL and MIP approaches differ from which the confidence regions are derived. Thus, our objective is to build proper confidence region for the unknown regression parameter and compare their confidence regions other than for the purpose of efficiency estimates. We propose an estimated EL confidence region, which is based on the weighted chi-square distribution. The corresponding constrained maximization of empirical likelihood can be done reliably by Newton–Raphson method. Moreover, in order to avoid the simulation for critical value of weighted chi-square distribution, we also develop an adjusted empirical likelihood confidence region for the vector of regression parameter. The simulation results demonstrate the proposed EL methods are more accurate than existing method in terms of coverage probability.

The rest of the paper is organized as follows. The proposed estimated EL and adjusted EL confidence regions and main asymptotic results are presented in Section 2. In Section 3, we conduct extensive simulation study to compare the performance of EL based methods with that of normal approximation based method. In Section 4, we make some discussions about the proposed method. Proofs are put in the Appendix.

2. Main results

2.1. Preliminaries

We consider the median regression model (1.1). For uncensored data, the LAD estimator of β is obtained by minimizing $\sum_{i=1}^{n} |T_i - \beta' Z_i|$, or by solving the LAD estimating equation

$$U_1(\beta) = \sum_{i=1}^{n} \left(I(T_i \geqslant \beta' Z_i) - \frac{1}{2} \right) Z_i \approx 0. \tag{2.1}$$

A "root" of this estimating equation is a minimizer of the Euclidean norm of the estimating function.

Suppose the censoring variable depends on the covariate $Z = (1, W)'$ and the W is a one-dimensional continuous variable (Note that 1st component 1 corresponds to the intercept and $p = 1$). Subramanian [36] applied MIP to the uncensored LAD estimating equation (2.1). The idea is to replace the unobservable $I(T_i \geqslant \beta' Z_i)$ by an estimate of its conditional expectation given the data. It can be shown that (cf. [10, p. 840, Equation (7.4)]) the condition expectation $E_i(\beta)$ is given by

$$
\begin{aligned}
E_i(\beta) &= E(I(T_i \geqslant \beta' Z_i)|(X_i, \delta_i, Z_i)) \\
&= I(X_i \geqslant \beta' Z_i) + I(X_i < \beta' Z_i, \delta_i = 0)\frac{F(\beta' Z_i, Z_i)}{F(X_i, Z_i)},
\end{aligned}
$$

where $F(t, z) = P(T > t|Z = z)$ is the conditional survival function of T given Z. Let

$$
\hat{E}_i(\beta) = I(X_i \geqslant \beta' Z_i) + I(X_i < \beta' Z_i, \delta_i = 0)\frac{\hat{F}(\beta' Z_i, Z_i)}{\hat{F}(X_i, Z_i)},
$$

where $\hat{F}(t, z)$ is an appropriate estimator of $F(t, z)$ which is defined using the product integral by

$$
\hat{F}(t, z) = \prod_{s \leqslant t}(1 - d\hat{\Lambda}(s, z)), \tag{2.2}
$$

where $\hat{\Lambda}(t, z)$ is the kernel conditional Nelson–Aalen estimator of the conditional cumulative hazard function of T given $Z = z$, denoted by $\Lambda(t, z)$, and is given by

$$
\hat{\Lambda}(t, z) = \int_{-\infty}^{t} \frac{d\hat{H}_1(s, z)}{\hat{H}(s-, z)}, \tag{2.3}
$$

and

$$
\begin{aligned}
\hat{H}_1(t, z) &= \hat{g}(z)^{-1}\left((na_n)^{-1}\sum_{i=1}^{n} I(X_i \leqslant t, \delta_i = 1)K(a_n^{-1}(z - Z_i))\right), \\
\hat{H}(t, z) &= \hat{g}(z)^{-1}\left((na_n)^{-1}\sum_{i=1}^{n} I(X_i > t)K(a_n^{-1}(z - Z_i))\right), \\
\hat{g}(z) &= (na_n)^{-1}\sum_{i=1}^{n} K(a_n^{-1}(z - Z_i)),
\end{aligned}
$$

where $\hat{H}_1(t, z)$ is a kernel estimator of the conditional subdistribution function of X given $Z = z$, namely $H_1(t, z) = P(X \leqslant t, \delta = 1|Z = z)$, $\hat{H}(t, z)$ is a kernel estimator of the conditional survival function of X given $Z = z$, namely $H(t, z) = P(X > t|Z = z)$, K is an appropriate kernel function, and a_n is a bandwidth sequence.

The MIP estimating equation takes

$$
U(\beta) = \sum_{i=1}^{n}\left(\hat{E}_i(\beta) - \frac{1}{2}\right) Z_i \approx 0, \tag{2.4}
$$

and define $\hat{\beta}$ to be a minimizer of the Euclidean norm of $U(\beta)$. Under mild conditions the estimating equation (4.1) has a unique solution $\hat{\beta}$.

Throughout the paper, we only concern about the case where the censoring variable is dependent on a one-dimensional continuous covariate W. We assume the censoring variable C_i is assumed to be conditionally independent of T_i given the covariate Z_i for $1 \leqslant i \leqslant n$. Let D be a bounded convex region. Suppose that the true value β_0 of β is in its interior. In order to derive the asymptotic normality of $\hat{\beta}$, we need the following conditions (cf. [36]). Assume the following conditions hold:

1. The sequence, $\{a_n\}$, of bandwidths satisfies (i) $na_n^2/\log(a_n^{-1}) \to \infty$ as $n \to \infty$ and (ii) $na_n^4 \to 0$.
2. The kernel K (i) is a probability density function with support on $[-M, M]$ for some $M > 0$, (ii) is bounded and continuous, and of bounded variation and (iii) satisfies $\int_{-M}^{M} sK(s)\,\mathrm{d}s = 0$.
3. The covariate Z (i) is bounded, i.e., $\|Z\| \leqslant L$ for some positive constant L, where $\|\cdot\|$ is the Euclidean norm, (ii) has a bounded density g which is bounded away from 0, and with support on $[-L, L]$; the density is uniformly continuous on $[-L, L]$. Furthermore, (iii) the density g is twice differentiable in $[-L, L]$ and the first and second derivatives are bounded, continuous, and bounded away from 0.
4. For $\beta \in D$, there exists a constant t_1, such that $P(X \geqslant t_1|Z) > 0$ and $\beta' Z \leqslant t_1$ with probability 1. Moreover, (i) the functions $H_1(x, z)$ and $H(x, z)$ are bounded away from 0 over the range $(-\infty, t_1] \times [-L, L]$, (ii) the first, second and third partial derivatives of $H_1(x, z)$ with respect to z and with respect to x are, respectively, continuous and uniformly bounded in $(x, z) \in (-\infty, t_1] \times [-L, L]$, and (iii) the same conditions are satisfied by the first and second partial derivatives of $H(x, z)$ with respect to z or x.
5. The matrix A is positive definite, where $A = E[ZZ' f(0|Z)]$ and $f(t|z) = f(t + \beta_0' z, z)$ is the conditional density of $\epsilon = T - \beta_0' Z$ given $Z = z$.

Remark. The above regularity conditions are commonly used in survival analysis, see Subramanian [36] for discussion. These conditions are essential for the consistency and asymptotic normality of the estimator $\hat{\beta}$. Condition 3 is somewhat strong and the bounded assumption in Condition 3 is standard for continuous covariate used in survival analysis. Condition 4 is not difficult to check. These assumptions are satisfied by the most common distributions in survival analysis. Condition 5 guarantees that the estimators are well defined asymptotically.

Let

$$\Gamma(\beta) = \frac{1}{4} E\left(Z_1 Z_1' \int_{-\infty}^{\beta^T Z_1} \frac{d\Lambda(s, Z_1)}{H(s, Z_1)} \right). \tag{2.5}$$

Under conditions 1–5, it is shown in Subramanian [36],

$$n^{1/2}(\hat{\beta} - \beta_0) \xrightarrow{\mathcal{D}} N(0, A^{-1}\Gamma A^{-1}). \tag{2.6}$$

Assume that

$$\hat{\Gamma} = \frac{1}{4n} \sum_{i=1}^{n} Z_i Z_i' \int_{-\infty}^{\hat{\beta}^T Z_i} \frac{d\hat{\Lambda}(s, Z_i)}{\hat{H}(s, Z_i)}.$$

From Lemma A.1, Γ is consistently estimated by $\hat{\Gamma}$. Thus, an asymptotic $100(1-\alpha)\%$ confidence region for β based on the above normal approximation is given by

$$\mathcal{R}_1 = \{\beta : nU'(\beta)\hat{\Gamma}^{-1}U(\beta) \leqslant \chi^2_{p+1}(\alpha)\}, \tag{2.7}$$

where $\chi^2_{p+1}(\alpha)$ is the upper α-quantile of the chi-squared distribution with degrees of freedom $p+1$.

2.2. EL confidence region

Now consider empirical likelihood approach. For $1 \leqslant i \leqslant n$, we define

$$W_i = \left(E_i(\beta_0) - \frac{1}{2}\right) Z_i,$$
$$W_{ni} = \left(\hat{E}_i(\beta_0) - \frac{1}{2}\right) Z_i.$$

It is easy to check that $EW_i = 0$ by the definition of W_i. Then, the empirical likelihood is given by

$$L(\beta_0) = \sup\left\{\prod_{i=1}^n p_i : \sum p_i = 1, \sum_{i=1}^n p_i W_i = 0, p_i \geqslant 0, i = 1, \ldots, n\right\}.$$

However, the W_i's depend on $F(t, z)$ which is unknown, we replace them by the W_{ni}'s. Therefore, using the notation L_n, an estimated empirical likelihood at the true value β_0 is given by

$$L_n(\beta_0) = \sup\left\{\prod_{i=1}^n p_i : \sum p_i = 1, \sum_{i=1}^n p_i W_{ni} = 0, p_i \geqslant 0, i = 1, \ldots, n\right\}.$$

Let $p = (p_1, \ldots, p_n)$ be a probability vector, i.e., $\sum_{i=1}^n p_i = 1$ and $p_i \geqslant 0$ for $1 \leqslant i \leqslant n$. Note that $\prod_{i=1}^n p_i$ attains its maximum at $p_i = 1/n$. Thus, the empirical likelihood ratio at the true value β_0 is defined by

$$R(\beta_0) = \sup\left\{\prod_{i=1}^n np_i : \sum p_i = 1, \sum_{i=1}^n p_i W_{ni} = 0, p_i \geqslant 0, i = 1, \ldots, n\right\}.$$

We profile the estimated empirical likelihood ratio and obtain the profile estimator of β, i.e., $\hat{\beta}_P = \operatorname{argmax}_\beta R(\beta)$. We note that $\prod_{i=1}^n np_i$ attains its maximum 1 at $p_i = 1/n$, and then $\hat{\beta}_P$ satisfies estimating equation (2.6). Thus $\hat{\beta}_P = \hat{\beta}$.

By using Lagrange multipliers, we know that $R(\beta_0)$ is maximized when

$$p_i = \frac{1}{n}\{1 + \lambda' W_{ni}\}^{-1}, \quad i = 1, \ldots, n,$$

where $\lambda = (\lambda_1, \ldots, \lambda_{p+1})'$ satisfies the equation

$$\frac{1}{n}\sum_{i=1}^n \frac{W_{ni}}{1 + \lambda' W_{ni}} = 0. \tag{2.8}$$

The value of λ may be found by Newton–Raphson algorithm. Thus, from the above equalities we have

$$\hat{l}(\beta_0) = -2\log R(\beta_0) = -2\log \prod_{i=1}^{n}(np_i) = 2\sum_{i=1}^{n}\log\{1+\lambda' W_{ni}\}, \tag{2.9}$$

where λ satisfies Eq. (2.8).

Let

$$\Gamma_1 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\left(E_i(\beta_0)-\frac{1}{2}\right)^2 Z_i Z_i'.$$

Now we state our main result and explain how it can be used to construct confidence region for β.

Theorem 2.1. *Under the above conditions*, $-2\log R(\beta_0)$ *converges in distribution to* $r_1\chi^2_{1,1} + \cdots + r_{p+1}\chi^2_{1,p+1}$, *where* $\chi^2_{1,1},\ldots,\chi^2_{1,p+1}$ *are independent chi-square random variables with* 1 *degree of freedom and* $r_1,\ldots,r_{p+1}$ *are the eigenvalues of* $\Gamma_1^{-1}\Gamma$.

From Theorem 2.1, we note that the limiting distribution of the EL ratio is a weighted sum of i.i.d. χ^2_1's instead of the standard χ^2_{p+1} distribution, where the weights can be consistently estimated. Because the W_{ni}'s are dependent, $-2\log R(\beta_0)$ is no longer a sum of standard independent random variables. The similar phenomenon has occurred in various contexts such as right censoring and missing data settings (cf. [40,20,45]).

From Lemma A.1(i), Γ_1 is consistently estimated by

$$\hat{\Gamma}_1 = n^{-1}\sum_{i=1}^{n}\left(\hat{E}_i(\hat{\beta})-\frac{1}{2}\right)^2 Z_i Z_i'.$$

Hence, the r_i's can be estimated by the $\hat{r}_i$'s which are the eigenvalues of $\hat{\Gamma}_1^{-1}\hat{\Gamma}$.

An asymptotic $100(1-\alpha)\%$ estimated empirical likelihood (EEL) confidence region for β is given by

$$\mathcal{R}_2 = \{\beta : -2\log R(\beta) \leqslant c(\alpha)\}, \tag{2.10}$$

where $c(\alpha)$ is the upper α-quantile of the distribution of $\hat{r}_1\chi^2_{1,1} + \cdots + \hat{r}_{p+1}\chi^2_{1,p+1}$ and can be obtained by simulation method.

However, the accuracy of the EEL depends on the values of the $\hat{r}_i$'s. Alternatively, the above EEL approach can be adjusted to avoid the weighted sum expression. Let $\rho(\beta)=(p+1)/\mathrm{tr}\{\Gamma_1^{-1}(\beta)\Gamma(\beta)\}$ with $\mathrm{tr}(\cdot)$ denoting the trace vector. Then, following Rao and Scott [31], the distribution of $\rho(\beta_0)(r_1\chi^2_{1,1}+\cdots+r_{p+1}\chi^2_{1,p+1})$ may be approximated by χ^2_{p+1}. This implies that the asymptotic distribution of the Rao–Scott adjusted empirical likelihood ratio, $\tilde{l}_{ad}(\beta_0) = \hat{\rho}(\beta_0)\hat{l}(\beta_0)$, may be approximated by χ^2_{p+1}, where the adjustment factor $\hat{\rho}(\beta)$ is $\rho(\beta)$ with $\Gamma_1(\beta)$ and $\Gamma(\beta)$ replaced by $\hat{\Gamma}_1(\beta)$ and $\hat{\Gamma}(\beta)$, respectively.

Wang and Rao [40,41] and Li and Wang [20], among others proposed the adjusted empirical likelihood method. Now, we define an adjusted empirical likelihood ratio, by modifying $\rho(\beta_0)$

in $\tilde{l}_{ad}(\beta_0)$, whose asymptotic distribution is exactly a χ^2_{p+1}. Note that

$$\hat{\rho}(\beta) = \frac{\mathrm{tr}\{\hat{\Gamma}^{-1}(\beta)\hat{\Gamma}(\beta)\}}{\mathrm{tr}\{\hat{\Gamma}_1^{-1}(\beta)\hat{\Gamma}(\beta)\}}.$$

We define $\hat{r}(\beta)$ to be $\hat{\rho}(\beta)$ with $\hat{\Gamma}(\beta)$ replaced by $\hat{S}(\beta) = \{\sum_{i=1}^n W_{ni}(\beta)/n\} \times \{\sum_{i=1}^n W_{ni}(\beta)/n\}'$. That is,

$$\hat{r}(\beta) = \frac{\mathrm{tr}\{\hat{\Gamma}^{-1}(\beta)\hat{S}(\beta)\}}{\mathrm{tr}\{\hat{\Gamma}_1^{-1}(\beta)\hat{S}(\beta)\}}.$$

We define an adjusted empirical likelihood ratio by

$$\hat{l}_{ad}(\beta) = \hat{r}(\beta)\hat{l}(\beta).$$

Theorem 2.2. *Under the above conditions, the EL statistic $\hat{l}_{ad}(\beta_0)$ converges in distribution to χ^2_{p+1}.*

Based on Theorem 2.2, an asymptotic $100(1-\alpha)\%$ adjusted empirical likelihood (AEL) confidence region for β is given by

$$\mathcal{R}_3 = \{\beta : \hat{l}_{ad}(\beta) \leqslant \chi^2_{p+1}(\alpha)\},$$

where $\chi^2_{p+1}(\alpha)$ is as before.

3. Simulation study

In this section, we compare the performance of the proposed empirical likelihood (EEL and AEL) confidence regions with the normal approximation (NA) based confidence region. Our goal is to compare coverage probabilities. We consider the following simulation models:

In each simulation example, we use $Z = (1, W)'$, with W taking Uniform(0, 1). The true value of the intercept is 0, and the true value of the slope is 1, i.e., $\beta_0 = (0, 1)'$. A grid search over the rectangle $(-2, 2) \times (-1, 3)$ is used to find the solution of the estimating equation (4.1). The kernel function $K(u)$ on $[-1, 1]$ is chosen to be a biweight kernel

$$K(u) = \tfrac{15}{16}(1-u^2)_+^2 = \tfrac{15}{16}(1-u^2)^2 I_{[-1,1]}(u).$$

The sequence of bandwidths is chosen to be $a_n = n^{-1/2+0.15}$ which satisfies condition 1.

We consider model M_1: $T|Z \sim$ Exponential$(\log(2)/w)$. It is a median regression model (the conditional median of T given Z is $\beta_0' z$), and also a Cox proportional hazard model with covariate $\log w$, regression parameter -1, and baseline hazard $\log 2$. The conditional hazard function of C given Z is $\lambda_c(t|w) = C_0 w$, where C_0 is used to control the censoring rate.

We consider model M_2: $T|Z \sim \mathrm{N}(w, 0.5)$, which departs from the Cox proportional hazard model. A Cox proportional hazard model is used for the censoring time: $\lambda_c(t|w) = C_0 w$, i.e., a Cox regression model with covariate $\log w$, regression parameter 1, and baseline hazard 0.25, where C_0 is used to calibrate the censoring rate.

We consider model M_3: The error distribution is taken to be normal with mean 0, and constant variance 1, i.e., $T|Z \sim \mathrm{N}(w, 1)$. The censoring is $C|Z \sim \log \mathrm{Uniform}(0, C_0 w)$, where C_0 is chosen to give the censoring rate.

Table 1
Coverage probabilities for the regression parameter β under simulation model M_1: $T|Z \sim$ Exponential$(\log(2)/w)$ and $\lambda_c(t|w) = C_0 w$, where C_0 was used to calibrate the censoring rate

Model: M_1		$1-\alpha = 0.90$			$1-\alpha = 0.95$		
CR	n	EEL	AEL	NA	EEL	AEL	NA
10%	30	0.9207	0.8644	0.9571	0.9607	0.9013	0.9751
	70	0.9196	0.8907	0.9640	0.9617	0.9290	0.9810
	120	0.9143	0.9009	0.9610	0.9600	0.9400	0.9823
	200	0.9146	0.9026	0.9566	0.9598	0.9464	0.9824
20%	30	0.9408	0.8628	0.9606	0.9694	0.8995	0.9794
	70	0.9383	0.8955	0.9685	0.9744	0.9299	0.9850
	120	0.9263	0.9001	0.9661	0.9664	0.9358	0.9857
	200	0.9188	0.9038	0.9602	0.9638	0.9415	0.9832
30%	30	0.9331	0.8486	0.9639	0.9600	0.8846	0.9804
	70	0.9461	0.8891	0.9691	0.9756	0.9259	0.9835
	120	0.9406	0.9059	0.9687	0.9698	0.9389	0.9865
	200	0.9363	0.9043	0.9630	0.9710	0.9432	0.9837

Table 2
Coverage probabilities for the regression parameter β under simulation model M_2: $T|Z \sim \mathrm{N}(w, 0.5)$ and $\lambda_c(t|w) = C_0 w$, where C_0 was used to calibrate the censoring rate

Model: M_2		$1-\alpha = 0.90$			$1-\alpha = 0.95$		
CR	n	EEL	AEL	NA	EEL	AEL	NA
10%	30	0.9231	0.8491	0.9803	0.9618	0.8893	0.9969
	70	0.9233	0.8914	0.9482	0.9630	0.9273	0.9835
	120	0.9149	0.8984	0.9316	0.9593	0.9387	0.9738
	200	0.9110	0.9006	0.9200	0.9552	0.9418	0.9653
20%	30	0.9278	0.8535	0.9793	0.9606	0.8891	0.9962
	70	0.9327	0.8836	0.9539	0.9694	0.9196	0.9873
	120	0.9249	0.8948	0.9394	0.9639	0.9303	0.9769
	200	0.9218	0.9056	0.9287	0.9629	0.9457	0.9696
30%	30	0.9075	0.8275	0.9672	0.9424	0.8625	0.9877
	70	0.9349	0.8739	0.9481	0.9642	0.9076	0.9829
	120	0.9393	0.8953	0.9389	0.9670	0.9307	0.9756
	200	0.9312	0.9012	0.9316	0.9631	0.9390	0.9696

We take 0.90, 0.95 as the nominal confidence level $1-\alpha$, respectively. We use 10%, 20%, and 30% censoring rates (CR), respectively, for Tables 1–3, which represent light censoring, middle censoring, and heavy censoring, respectively. The sample size n is chosen to be 30, 70, 120, and 200, respectively, which represent small sample, moderate sample, large sample and very large sample, respectively. The simulated coverage probabilities of the normal approximation based

Table 3
Coverage probabilities for the regression parameter β under simulation model M_3: $T|Z \sim \mathrm{N}(w, 1)$ and $C|Z \sim \log \mathrm{Uniform}(0, C_0 w)$, where C_0 was used to calibrate the censoring rate

Model: M_3		$1-\alpha=0.90$			$1-\alpha=0.95$		
CR	n	EEL	AEL	NA	EEL	AEL	NA
10%	30	0.9219	0.8450	0.9852	0.9617	0.8815	0.9979
	70	0.9078	0.8758	0.9439	0.9534	0.9144	0.9829
	120	0.9092	0.8945	0.9263	0.9518	0.9335	0.9723
	200	0.9070	0.8950	0.9195	0.9547	0.9367	0.9623
20%	30	0.9292	0.8388	0.9867	0.9645	0.8750	0.9982
	70	0.9235	0.8726	0.9588	0.9624	0.9118	0.9887
	120	0.9221	0.8910	0.9328	0.9627	0.9278	0.9747
	200	0.9190	0.9021	0.9236	0.9610	0.9411	0.9669
30%	30	0.9289	0.8332	0.9907	0.9620	0.8703	0.9983
	70	0.9349	0.8685	0.9592	0.9706	0.9028	0.9892
	120	0.9334	0.8794	0.9323	0.9690	0.9179	0.9734
	200	0.9319	0.8907	0.9175	0.9669	0.9299	0.9609

method and the empirical likelihood methods are estimated from 10,000 simulated data sets. The simulation results for the three models are reported in Tables 1–3, respectively. The models are also specified in the titles of the tables.

From these tables, we make the following observations:

1. At each nominal level, the coverage accuracies for empirical likelihood and normal approximation methods in the three models tend to decrease as censoring rates increase, and tend to increase when sample size increases.
2. The coverage probabilities for the normal approximation method and estimated empirical likelihood method are consistently larger than the nominal level. While the coverage probability for adjusted empirical likelihood undercovers the nominal level for small sample size. However, NA method performs poorly in all cases.
3. The empirical likelihood outperforms the normal approximation method in these models. In particular, under very large sample ($n = 200$), the adjusted empirical likelihood confidence region has more accurate coverage probabilities than the normal approximation based and estimated empirical likelihood confidence region.
4. The coverage probability of NA confidence region is more conservative than that of EEL confidence region. The normal approximation based confidence regions overcover the true regression parameter and the coverage probabilities are far above the nominal levels.

From Tables 1–3, we find that the normal approximation based method does not always work well. One reason is that the NA based confidence region needs to estimate Γ (cf. (2.7)). The variance estimates are not very stable and may contain values outside their ranges. Another possible reason is that we may not use an optimal bandwidth to obtain the coverage probability. However, the EL method is relatively robust and not sensitive to the choice of bandwidth.

In summary, our simulation study indicates that the proposed EL methods give very competitive coverage probabilities and suggests that the EL based confidence regions outperform the NA based confidence region.

4. Discussion

To choose an optimal bandwidth for the estimator, a number of methods have been proposed in the kernel density estimation, see the discussion in Silverman [34]. However, the bandwidth selection for obtaining the accurate coverage probability has not been studied by researchers. A worthwhile direction for future research would be to investigate bandwidth selection in censored median regression model.

We have investigated the continuous covariate case based on MIP in the article. When the censoring variable depends on the covariate vector Z and Z takes only discrete values, the EL approach also works. In the discrete case, denote the possible values of Z by z_k, $k = 1, \dots, K$, and assume each occurs with positive probability. Let n_k denote the number of Z_j, $j = 1, \dots, n$, taking value z_k. Rewrite the sample (X_i, δ_i, Z_i), $i = 1, \dots, n$ as $(X_{k,m}, \delta_{k,m})$, $m = 1, \dots, n_k$, for $k = 1, \dots, K$, where $(X_{k,m}, \delta_{k,m})$ corresponds to (X_i, δ_i) with covariate Z_i having the value z_k. Let $\hat{F}(t, z_k)$ be the local Kaplan–Meier estimator based on the pairs $(X_{k,m}, \delta_{k,m})$, $m = 1, \dots, n_k$.

The MIP estimating equation takes

$$U(\beta) = \sum_{i=1}^{n} \left(\hat{E}_i(\beta) - \frac{1}{2} \right) Z_i \approx 0, \tag{4.1}$$

and define $\hat{\beta}$ to be a minimizer of the Euclidean norm of $U(\beta)$. Under mild conditions the estimating Eq. (4.1) has a unique solution $\hat{\beta}$.

Following the same argument as before we can define estimated empirical likelihood ratio and adjusted empirical likelihood ratio, derive the limiting distributions of the two empirical likelihood ratios for the regression parameter. The estimated EL and adjusted EL confidence regions for the vector of regression parameter can be obtained. We also compare the proposed methods with that normal approximation based method through extensive simulation study and the results are omitted here to save the space. The simulation results suggest that our proposed methods outperform the normal approximation based method in terms of coverage probability which is the same as that for a one-dimensional continuous covariate case.

In this article, we consider the median regression model when censoring variable depends on a one-dimensional continuous covariate and EL works well in this case. However, the drawback of our method is the curse of dimensionality problem when the dimension of covariate is higher. Because it is difficult to estimate F nonparametrically in high dimensions—the higher the dimension the more spread apart are the data points, and the larger the data set required for a sensible analysis. One way to avoid this problem is to make an independence assumption in Ying et al. [44]. However, these independence assumptions limit the application of the proposed method in clinical trials. Thus, a more flexible approach is to fit Cox regression model or semiparametric additive risk model. It may provide a relatively good approximation for the unknown conditional distribution in the estimating equation. In addition, Yang [43] proposed alternative estimators. They are based on some weighted empirical survival and hazard functions. In the future, we will explore this challenging issue using EL.

Acknowledgments

The authors would like to thank Professor Yanqing Sun for making available some computer programs. The authors are grateful to the editor, the two referees for their helpful and valuable comments which improved the presentation of the paper greatly. The research of Yichuan Zhao is partially supported by National Security Agency Grant and Faculty Mentoring Grant, Georgia State University.

Appendix A. Proofs of Theorems

We need the following lemma in order to prove Theorem 2.1.

Lemma A.1. *Under the conditions of Theorem* 2.1, *we have*

(i) $\sum_{i=1}^n W_{ni} W_{ni}'/n \xrightarrow{\mathcal{P}} \Gamma_1$, (ii) $\hat{\Gamma}_1 \xrightarrow{\mathcal{P}} \Gamma_1$, (iii) $\hat{\Gamma} \xrightarrow{\mathcal{P}} \Gamma$.

Proof. Let

$$\hat{\Gamma}_{1n} = \frac{1}{n}\sum_{i=1}^n W_{ni} W_{ni}', \quad \Gamma_{1n} = \frac{1}{n}\sum_{i=1}^n W_i W_i'.$$

In order to prove (i), we only need to show $\hat{\Gamma}_{1n} = \Gamma_{1n} + o_P(1)$.

For any $a \in R^{p+1}$, the following decomposition holds:

$$\begin{aligned} a'(\hat{\Gamma}_{1n} - \Gamma_{1n})a &= \frac{1}{n}\sum_{i=1}^n (a'(W_{ni} - W_i))^2 + \frac{2}{n}\sum_{i=1}^n (a'W_i)(a'(W_{ni} - W_i)) \\ &= I_1 + 2I_2. \end{aligned} \tag{A.1}$$

First note that

$$\begin{aligned} & a'(W_{ni} - W_i) \\ &= a'(\hat{E}_i(\beta_0) - E_i(\beta_0))Z_i \\ &= \frac{I(X_i < \beta_0' Z_i, \delta_i = 0)(\hat{F}(\beta_0' Z_i, Z_i)F(X_i, Z_i) - F(\beta_0' Z_i, Z_i)\hat{F}(X_i, Z_i))}{\hat{F}(X_i, Z_i)F(X_i, Z_i)}(a'Z_i). \end{aligned}$$

The conditions 1–4 imply that the kernel conditional Kaplan–Meier estimator $\hat{F}(t, z)$ is strongly uniformly consistent (cf. [6]). It follows from conditions 1–4 that

$$\begin{aligned} |I_2| &\leqslant \frac{1}{n}\sum_{i=1}^n \frac{I(X_i < \beta_0' Z_i)|\hat{F}(\beta_0' Z_i, Z_i)F(X_i, Z_i) - F(\beta_0' Z_i, Z_i)\hat{F}(X_i, Z_i)|}{\hat{F}(X_i, Z_i)F(X_i, Z_i)}(aZ_i)^2 \\ &= \max_{X_i < \beta_0' Z_i, 1 \leqslant i \leqslant n} \left(\frac{|\hat{F}(\beta_0' Z_i, Z_i) - F(\beta_0' Z_i, Z_i)|}{\hat{F}(X_i, Z_i)} + \frac{|\hat{F}(X_i, Z_i) - F(X_i, Z_i)|}{\hat{F}(X_i, Z_i)} \right)(\|a\|L)^2 \\ &= o_P(1). \end{aligned} \tag{A.2}$$

Similarly, we can show that $I_1 = o_P(1)$. Thus by (A.1), (A.2), we prove Lemma A.1(i).

In order to prove Lemma A.1(ii), we only need to show that $\hat{\Gamma}_1 = \hat{\Gamma}_{1n} + o_P(1)$. Let

$$\begin{aligned} J_i &= I(X_i \geqslant \hat{\beta}' Z_i) - I(X_i \geqslant \beta_0' Z_i) \\ &\quad + \frac{I(X_i < \hat{\beta}' Z_i, \delta_i = 0)\hat{F}(\hat{\beta}' Z_i, Z_i)}{\hat{F}(X_i, Z_i)} - \frac{I(X_i < \beta_0' Z_i, \delta_i = 0)\hat{F}(\beta_0' Z_i, Z_i)}{\hat{F}(X_i, Z_i)}. \end{aligned}$$

We have

$$
\begin{aligned}
|J_i| \leqslant{} & |(I(X_i \geqslant \hat{\beta}' Z_i) - I(X_i \geqslant \beta_0' Z_i)| \left(1 + \frac{\hat{F}(\hat{\beta}' Z_i, Z_i)}{\hat{F}(X_i, Z_i)}\right) \\
& + |I(X_i < \beta_0' Z_i, \delta_i = 0)| \frac{|\hat{F}(\hat{\beta}' Z_i, Z_i) - F(\hat{\beta}' Z_i, Z_i)|}{\hat{F}(X_i, Z_i)} \\
& + |I(X_i < \beta_0' Z_i, \delta_i = 0)| \frac{|F(\hat{\beta}' Z_i, Z_i) - F(\beta_0' Z_i, Z_i)|}{\hat{F}(X_i, Z_i)} \\
& + |I(X_i < \beta_0' Z_i, \delta_i = 0)| \frac{|\hat{F}(\beta_0' Z_i, Z_i) - F(\beta_0' Z_i, Z_i)|}{\hat{F}(X_i, Z_i)}.
\end{aligned}
$$

By (2.6), we have the following equality as (A.7) in Ying et al. [44]

$$
\sum_{i=1}^{n} |I(Y_i \geqslant \hat{\beta}' Z_i) - I(Y_i \geqslant \beta_0' Z_i)| = O_P(n^{2/3}).
$$

It follows by combining conditions 1–4 and (2.6)

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} J_i^2 + \frac{1}{n}\sum_{i=1}^{n} |J_i| &\leqslant \frac{4}{n}\sum_{i=1}^{n} |J_i| + \frac{1}{n}\sum_{i=1}^{n} |J_i| \\
&\leqslant \frac{10}{n}\sum_{i=1}^{n} |I(Y_i \geqslant \hat{\beta}' Z_i) - I(Y_i \geqslant \beta_0' Z_i)| + o_P(1) \\
&= o_P(1). \qquad \text{(A.3)}
\end{aligned}
$$

Then for any $a \in R^{p+1}$, by $\|Z\| \leqslant L$ and (A.3) we have

$$
\begin{aligned}
|a'(\hat{\Gamma}_1 - \hat{\Gamma}_{1n})a| &\leqslant \frac{1}{n}\sum_{i=1}^{n}(a'Z_i)^2 J_i^2 + \frac{2}{n}\sum_{i=1}^{n}(a'Z_i)^2 |J_i| \left|\hat{E}_i(\beta_0) - \frac{1}{2}\right| \\
&\leqslant \sup_{1 \leqslant i \leqslant n} (a'Z_i)^2 \frac{1}{n}\sum_{i=1}^{n} J_i^2 + \sup_{1 \leqslant i \leqslant n} (a'Z_i)^2 \frac{1}{n}\sum_{i=1}^{n} |J_i| \\
&\leqslant (\|a\| L)^2 \left(\frac{1}{n}\sum_{i=1}^{n} J_i^2 + \frac{1}{n}\sum_{i=1}^{n} |J_i|\right) \\
&= o_P(1).
\end{aligned}
$$

Therefore, we have $a'(\hat{\Gamma}_1 - \hat{\Gamma}_{1n})a = o_P(1)$. Lemma A.1(ii) follows. Following the same line as above, we can show Lemma A.1(iii). □

Proof of Theorem 2.1. By conditions 3, 4, and the uniform consistency of $\hat{F}(t, z)$

$$
\begin{aligned}
\max_{1 \leqslant i \leqslant n} \|W_{ni}\| &\leqslant \max_{1 \leqslant i \leqslant n} \left|\hat{E}_i(\beta_0) + \frac{1}{2}\right| \max_{1 \leqslant i \leqslant n} \|Z_i\| \\
&= O_P(1). \qquad \text{(A.4)}
\end{aligned}
$$

Let $\lambda = \rho\theta$, where $\rho \geqslant 0$ and $\|\theta\| = 1$. Recall that $\hat{\Gamma}_{1n} = \Gamma_1 + o_P(1)$ (cf. Lemma A.1(i)). Let $\sigma_1 > 0$ be the smallest eigenvalue of Γ_1. Then,

$$\theta' \hat{\Gamma}_{1n} \theta \geqslant \sigma_1 + o_P(1). \tag{A.5}$$

From the Appendix of Subramanian [36], we have

$$\left\| \frac{1}{n} \sum_{i=1}^{n} W_{ni} \right\| = O_P(n^{-1/2}). \tag{A.6}$$

Then, it follows from (2.8), (A.5), (A.6), and the argument used in Owen [27] that

$$\|\lambda\| = O_P(n^{-1/2}). \tag{A.7}$$

Applying Taylor's expansion to (2.9), we have

$$-2 \log R(\beta_0) = 2 \sum_{i=1}^{n} \left(\lambda' W_{ni} - \frac{1}{2} (\lambda' W_{ni})^2 \right) + r_n, \tag{A.8}$$

where

$$|r_n| \leqslant C \sum_{i=1}^{n} |\lambda' W_{ni}|^3 \text{ in probability.}$$

Hence, by (A.5), (A.8)

$$|r_n| \leqslant C n \|\lambda\|^3 \left(\max_{1 \leqslant i \leqslant n} \|W_{ni}\| \right)^3 = O_P(n^{-1/2}). \tag{A.9}$$

Note that

$$\begin{aligned} 0 = \frac{1}{n} \sum_{i=1}^{n} \frac{W_{ni}}{1 + \lambda' W_{ni}} &= \frac{1}{n} \sum_{i=1}^{n} W_{ni} \left(1 - \lambda' W_{ni} + \frac{(\lambda' W_{ni})^2}{1 + \lambda' W_{ni}} \right) \\ &= \frac{1}{n} \sum_{i=1}^{n} W_{ni} - \left(\frac{1}{n} \sum_{i=1}^{n} W_{ni} W_{ni}' \right) \lambda + \frac{1}{n} \sum_{i=1}^{n} \frac{W_{ni} (\lambda' W_{ni})^2}{1 + \lambda' W_{ni}}. \end{aligned} \tag{A.10}$$

By (A.4), (A.7), (A.10), and Lemma A.1(i), it follows that

$$\lambda = \left(\sum_{i=1}^{n} W_{ni} W_{ni}' \right)^{-1} \sum_{i=1}^{n} W_{ni} + o_P(n^{-1/2}). \tag{A.11}$$

By (A.10), we have

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \frac{\lambda' W_{ni}}{1 + \lambda' W_{ni}} \\ &= \sum_{i=1}^{n} (\lambda' W_{ni}) - \sum_{i=1}^{n} (\lambda' W_{ni})^2 + \sum_{i=1}^{n} \frac{(\lambda' W_{ni})^3}{1 + \lambda' W_{ni}}. \end{aligned} \tag{A.12}$$

Similarly as before, we have

$$\sum_{i=1}^{n} \frac{(\lambda' W_{ni})^3}{1+\lambda' W_{ni}} = o_P(n^{-1/2}). \tag{A.13}$$

Combining (A.12) and (A.13) we have

$$\sum_{i=1}^{n} (\lambda' W_{ni})^2 = \sum_{i=1}^{n} \lambda' W_{ni} + o_P(1). \tag{A.14}$$

By (A.8), (A.9), (A.11), (A.14) and Lemma A.1(i), we have

$$\begin{aligned} -2\log R(\beta_0) &= \sum_{i=1}^{n} \lambda' W_{ni} + o_P(1) \\ &= \left(n^{-1/2}\sum_{i=1}^{n} W_{ni}\right)' \left(n^{-1}\sum_{i=1}^{n} W_{ni} W_{ni}'\right)^{-1} \left(n^{-1/2}\sum_{i=1}^{n} W_{ni}\right) + o_P(1) \\ &= \left(\Gamma^{-1/2} n^{-1/2}\sum_{i=1}^{n} W_{ni}\right)' (\Gamma^{1/2}\Gamma_1^{-1}\Gamma^{1/2}) \left(\Gamma^{-1/2} n^{-1/2}\sum_{i=1}^{n} W_{ni}\right) + o_P(1). \end{aligned}$$

By the proof of Theorem 1 in Subramanian [36], we have $\Gamma^{-1/2}(n^{-1/2}\sum_{i=1}^{n} W_{ni}) \xrightarrow{\mathcal{D}} N(0, I_{p+1})$. Because $\Gamma^{1/2}\Gamma_1^{-1}\Gamma^{1/2}$ and $\Gamma_1^{-1}\Gamma$ have the same eigenvalues. Using [29, Lemma A.3] to re-express the limiting distribution of (2.9) as a weighted sum of independent χ_1^2 distribution, we complete the proof of Theorem 2.1. □

Proof of Theorem 2.2.. Recall the definition of $\hat{l}_{ad}(\beta)$. It follows that, by (A.8),

$$\hat{l}_{ad}(\beta_0) = \left(n^{-1/2}\sum_{i=1}^{n} W_{ni}\right)' \hat{\Gamma}^{-1} \left(n^{-1/2}\sum_{i=1}^{n} W_{ni}\right) + o_P(1).$$

We can show that $\hat{\Gamma} \xrightarrow{\mathcal{P}} \Gamma$. Using $\Gamma^{-1/2}(n^{-1/2}\sum_{i=1}^{n} W_{ni}) \xrightarrow{\mathcal{D}} N(0, I_{p+1})$ again, we complete the proof of Theorem 2.2. □

References

[1] G. Bassett, R. Koenker, Asymptotic theory of least absolute regression, J. Amer. Statist. Assoc. 73 (1978) 618–622.

[2] J. Buckley, I. James, Linear regression with censored data, Biometrika 66 (1979) 429–436.

[3] S.X. Chen, On the accuracy of empirical likelihood confidence regions for linear regression model, Ann. Inst. Statist. Math. 45 (1993) 621–637.

[4] S.X. Chen, Empirical likelihood confidence intervals for linear regression coefficients, J. Multivariate Anal. 49 (1994) 24–40.

[5] D.R. Cox, Regression models and life tables (with discussion), J. R. Statist. Soc. B 34 (1972) 187–220.

[6] D.M. Dabrowska, Uniform consistency of the kernel conditional Kaplan–Meier estimate, Ann. Statist. 17 (1989) 1157–1167.

[7] T.J. DiCiccio, P. Hall, J.P. Romano, Empirical likelihood is Bartlett-correctable, Ann. Statist. 19 (1991) 1053–1061.

[8] T.J. DiCiccio, J.P. Romano, On adjustments based on the signed root of the empirical likelihood ratio statistic, Biometrika 76 (1989) 447–456.

[9] T.J. DiCiccio, J.P. Romano, Nonparametric confidence-limits by resampling methods and least favorable families, Int. Statist. Rev. 58 (1990) 59–76.

[10] B. Efron, Two sample problem with censored data, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. IV, Prentice-Hall, Engewood Cliffs, NJ, 1967, pp. 831–853.
[11] J.H. Einmahl, I.W. McKeague, Confidence tubes for multiple quantile plots via empirical likelihood, Ann. Statist. 27 (1999) 1348–1367.
[12] N.L. Glenn, Y. Zhao, Weighted empirical likelihood estimates and their robustness properties, Comput. Statist. Data Anal. 51 (2007) 5130–5141.
[13] P. Hall, Pseudo-likelihood theory for empirical likelihood, Ann. Statist. 18 (1990) 121–140.
[14] M. Hollander, I.W. McKeague, J. Yang, Likelihood ratio-based confidence bands for survival functions, J. Amer. Statist. Assoc. 92 (1997) 215–226.
[15] R. Koenker, G. Bassett, Regression quantiles, Econometrica 46 (1978) 33–50.
[16] H. Koul, V. Susarla, J. Van Ryzin, Regression analysis with randomly right-censored data, Ann. Statist. 9 (1981) 1276–1288.
[17] T.L. Lai, Z. Ying, Linear rank statistics in regression analysis with censored or truncated data, J. Multivariate Anal. 40 (1992) 13–45.
[18] N.M. Laird, Missing information principle, Encycl. Statist. Sci. 5 (1985) 548–552.
[19] G. Li, I. Van Keilegom, Likelihood ratio confidence bands in nonparametric regression with censored data, Scand. J. Statist. 29 (2002) 547–562.
[20] G. Li, Q.H. Wang, Empirical likelihood regression analysis for right censored data, Statist. Sinica 13 (2003) 51–68.
[21] I.W. McKeague, S. Subramanian, Y. Sun, Median regression and the missing information principle, J. Nonparametric Statist. 13 (2001) 709–727.
[22] I.W. McKeague, Y. Zhao, Simultaneous confidence bands for ratios of survival functions via empirical likelihood, Statist. Probab. Lett. 60 (2002) 405–415.
[23] I.W. McKeague, Y. Zhao, Comparing distribution functions via empirical likelihood, Int. J. Biostatist., article 5 (2005) 1–20.
[24] I.W. McKeague, Y. Zhao, Width-scaled confidence bands for survival functions, Statist. Probab. Lett. 76 (2006) 327–339.
[25] T. Orchard, M.A. Woodbury, A missing information principle: theory and applications, in: Proceedings of the Fifth Berkeley Symposium Mathematical Statistics and Probability, vol. I, Prentice-Hall, Engewood Cliffs, NJ, 1972, pp. 697–715.
[26] A. Owen, Empirical likelihood ratio confidence intervals for a single functional, Biometrika 75 (1988) 237–249.
[27] A. Owen, Empirical likelihood and confidence regions, Ann. Statist. 18 (1990) 90–120.
[28] A. Owen, Empirical likelihood for linear models, Ann. Statist. 19 (1991) 1725–1747.
[29] G. Qin, M. Tsao, Empirical likelihood inference for median regression models for censored survival data, J. Multivariate Anal. 85 (2003) 416–430.
[30] J. Qin, B. Zhang, Empirical-likelihood-based inference in missing response problems and its application in observational studies, J. R. Statist. Soc. B 69 (2007) 101–122.
[31] J.N.K. Rao, A.J. Scott, The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables, J. Amer. Statist. Assoc. 76 (1981) 221–230.
[32] Y. Ritov, Estimating a linear regression model with censored data, Ann. Statist. 18 (1990) 303–328.
[33] J. Shi, T.S. Lau, Empirical likelihood for partially linear models, J. Multivariate Anal. 72 (2000) 132–148.
[34] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman & Hall, London, 1986.
[35] W. Stute, L. Xue, L. Zhu, Empirical likelihood inference in nonlinear errors-in-covariables models with validation data, J. Amer. Statist. Assoc. 102 (2007) 332–346.
[36] S. Subramanian, Median regression using nonparametric kernel estimation, J. Nonparametric Statist. 14 (2002) 583–605.
[37] D.R. Thomas, G.L. Grunkemeier, Confidence interval estimation for survival probabilities for censored data, J. Amer. Statist. Assoc. 70 (1975) 865–871.
[38] A.A. Tsiatis, Estimating regression parameters using linear rank tests for censored data, Ann. Statist. 18 (1990) 354–372.
[39] Q.H. Wang, B.Y. Jing, Empirical likelihood for partial linear models with fixed designs, Statist. Probab. Lett. 41 (1999) 425–433.
[40] Q.H. Wang, J.N.K. Rao, Empirical likelihood for linear regression models under imputation for missing responses, Canad. J. Statist. 29 (2001) 597–608.
[41] Q.H. Wang, J.N.K. Rao, Empirical likelihood-based inference in linear models with missing data, Scand. J. Statist. 29 (2002) 563–576.
[42] L.J. Wei, Z. Ying, D.Y. Lin, Linear regression analysis of censored survival data based on rank tests, Biometrika 19 (1990) 845–851.

[43] S. Yang, Censored median regression using weighted empirical survival and hazard functions, J. Amer. Statist. Assoc. 94 (1999) 137–145.
[44] Z. Ying, S. Jung, L.J. Wei, Survival analysis with median regression models, J. Amer. Statist. Assoc. 90 (1995) 178–184.
[45] Y. Zhao, Regression analysis for long-term survival rate via empirical likelihood, J. Nonparametric Statist. 17 (2005) 995–1007.
[46] Y. Zhao, Y.S. Hsu, Semiparametric analysis for additive risk model via empirical likelihood, Comm. Statist. Simulation Comput. 34 (2005) 135–143.