Journal of Multivariate Analysis **79**, 295–315 (2001)
doi:10.1006/jmva.2000.1974, available online at http://www.idealibrary.com on IDEAL®

Inference for the Mean Difference in the Two-Sample Random Censorship Model

Qihua Wang

Chinese Academy of Science, Beijing, People's Republic of China; Heilongjiang University, Harbin, People's Republic of China; and Peking University, Beijing, People's Republic of China

and

Jane-Ling Wang

University of California, Davis

Received November 23, 1999; published online June 19, 2001

Inference for the mean difference in the two-sample random censorship model is an important problem in comparative survival and reliability test studies. This paper develops an adjusted empirical likelihood inference and a martingale-based bootstrap inference for the mean difference. A nonparametric version of Wilks' theorem for the adjusted empirical likelihood is derived, and the corresponding empirical likelihood confidence interval of the mean difference is constructed. Also, it is shown that the martingale-based bootstrap gives a correct first order asymptotic approximation of the corresponding estimator of the mean difference, which ensures that the martingale-based bootstrap confidence interval has asymptotically correct coverage probability. A simulation study is conducted to compare the adjusted empirical likelihood, the martingale-based bootstrap, and Efron's bootstrap in terms of coverage accuracies and average lengths of the confidence intervals. The simulation indicates that the proposed adjusted empirical likelihood and the martingale-based bootstrap confidence procedures are comparable, and both seem to outperform Efron's bootstrap procedure.

AMS 1991 subject classifications: 62G15; 62E20.

Key words and phrases: empirical likelihood; martingale-based bootstrap; confidence interval.

1. INTRODUCTION

In the analysis of survival data in comparative survival and reliability studies, it is important to make statistical inference for the mean difference of two populations based on the life data of two treatment groups or a treatment group and a control group. Let $X_1^0, ..., X_n^0$ and $Y_1^0, ..., Y_m^0$ be random samples of survival times from two different populations with

0047-259X/01 $35.00

distribution function F and G, respectively. Let μ_1 and μ_2 be the means of the two populations X^0 and Y^0, respectively, and $\theta = \mu_2 - \mu_1$. In practice, we need to compare the two population means, i.e., test $\theta = 0$, or provide a confidence interval for θ. With complete observations, Qin (1997) used an empirical likelihood ratio statistic to test hypothesis and construct confidence interval for θ. This paper extends the method to the two sample random censorship model. In addition, a martingale-based bootstrap inference for the mean difference is also established.

In the two-sample random censorship model, the variables $\{X_1^0, ..., X_n^0\}$ and $\{Y_1^0, ..., Y_m^0\}$ are randomly censored by two sequences of random variables $\{U_1, ..., U_n\}$ and $\{V_1, ..., V_n\}$, with distribution function K and Q, respectively. So instead of observing the X_i^0's and Y_j^0's directly, one only observes (X_i, δ_i) for $i = 1, 2, ..., n$ and $(Y_j, \triangle_j)$ for $j = 1, 2, ..., m$, where

$$X_i = \min(X_i^0, U_i), \qquad \delta_i = I(X_i^0 \leqslant U_i),$$

$$Y_j = \min(Y_j^0, V_j), \qquad \triangle_j = I(Y_j^0 \leqslant V_j).$$

Here $I(\cdot)$ denotes the indicator function. We shall assume that X_i^0, U_i, Y_j^0, V_j for $i = 1, ..., n$ and $j = 1, ..., m$ are mutually independent.

As Owen (1988) pointed out, empirical likelihood methods were first used by Thomas and Grunkemeier (1975) to construct confidence intervals for survival probabilities. In their method, Thomas and Grunkemeier used product type constraints by decomposing the survival probability to a product of some conditional probabilities. However, this limits the applicability of this method to other cases. For example, it is difficult to extend this method to inference for the mean difference in the two sample random censoring model considered here. The reason is that proper product type constraints are difficult to find. It is noted that Owen's empirical likelihood is based on linear constraints and hence has very general applicability in the absence of censoring (see, e.g., Owen, 1988, 1990, 1991; Hall and Scala, 1990; DiCiccio *et al.*, 1991; Chen, 1993, 1994; Qin and Lawless, 1994; Qin 1996; Chen and Qin, 1993; Kolaczyk, 1994, and Wang and Jing, 1999). We show in this paper that for the two-sample random censorship model, the empirical likelihood idea is also useful in order to develop an adjusted empirical likelihood inference for the mean difference. Under the assumption that the censoring distributions are known, one could extend Owen's idea to define an empirical log-likelihood function (ELLF). However, in practice, the censoring distributions are usually unknown. Naturally, we replace the unknown censoring distribution functions in ELLF with their Kaplan–Meier product-limit estimators (Kaplan and Meier, 1958) to define an estimated ELLF. The estimated ELLF involves the estimation of the unknown censoring distributions and hence is not asymptotically standard chi-square distributed. This motivates us to adjust the estimated ELLF in such a way that the adjusted ELLF retains an asymptotic standard chi-square distribution. Hence, such an adjustment also achieves the construction of confidence intervals. The adjusted empirical likelihood has the same advantages as the standard empirical likelihood except that an unknown adjusting factor must be estimated. The adjusting factor is a quantity which reflects the loss of information due to censoring. A different approach to extend empirical likelihood for censored data is given in Pan and Zhou (2000).

Another accomplishment of this paper is the construction of a martingale-based bootstrap confidence interval of θ. This procedure is as follows: First, we define an estimator of θ, say $\hat{\theta}_n$, and represent it as a stochastic integral with respect to a martingale. Second, we replace the martingale process by the products of the corresponding point processes and standard normal random variables. Finally, we use the conditional distribution of the resulting statistic to approximate that of $\hat{\theta}_n$ and apply the approximate distribution to construct a confidence interval of θ. Such a procedure was first applied by Lin *et al.* (1993) for checking the Cox model. Later, Lin and Spiekerman (1996) also applied it for model checking for a parametric regression. Recently, Wang (1998) applied this method to inference for a class of functionals of survival distribution and termed it "martingale-based bootstrap." An obvious advantage of this method is that it doesn't use the variance estimators. Another advantage is that its computation is simple since it involves only resampling from a standard normal population.

The rest of this paper is organized as follows. In Section 2, the adjusted empirical log-likelihood ratio is described with Wilks' Theorem (Theorem 2.1) established. From there, an adjusted empirical likelihood confidence interval of θ is derived. We introduce in Section 3 the martingale-based bootstrap method and provides the approximation theorems. A simulation study is conducted in Section 4 to compare the coverage accuracies of the confidence intervals constructed from the adjusted empirical likelihood, the martingale-based bootstrap and the Efron's bootstrap method. Proofs are given in the Appendix.

2. AN ADJUSTED EMPIRICAL LIKELIHOOD INFERENCE

2.1. *Description of Methods*

We first give some motivation for defining the adjusted empirical likelihood. Let $\theta(F, G) = EY^0 - EX^0$. Then we have

$$E\frac{\triangle Y}{1 - Q(Y-)} - E\frac{\delta X}{1 - K(X-)} = \theta(F, G) \tag{2.1}$$

by the facts that $E\,\frac{\delta X}{1-K(X-)}=EX^0$ and $E\,\frac{\triangle Y}{1-Q(Y-)}=EY^0$. In the location shift model, θ is just $\theta(F,G)$. Let $\mathbf{P}_1=(p_{11},...,p_{1n})$ and $\mathbf{P}_2=(p_{21},...,p_{2m})$ be probability vectors, i.e., $\sum_{i=1}^n p_{1i}=1$, $\sum_{j=1}^m p_{2j}=1$, $p_{1i}>0$, $p_{2j}>0$ for $i=1,2,...,n$ and $j=1,2,...,m$. Let F_{p_1} and G_{p_2} be the distribution functions which assign probabilities p_{1i} and p_{2j} at the points $X_i\delta_i/(1-K(X_i-))$ and $Y_j\triangle_j/(1-Q(Y_j-))$, respectively, for $i=1,2,...,n$ and $j=1,2,...,m$. Then, we have

$$\theta(F_{p_1},G_{p_2})=\sum_{j=1}^m p_{2j}\frac{\triangle_j Y_j}{1-Q(Y_j-)}-\sum_{i=1}^n p_{1i}\frac{\delta_i X_i}{1-K(X_i-)}. \tag{2.2}$$

The empirical log-likelihood ratio can be defined as

$$l_0(\theta)=-2\max_{\substack{\theta(F_{p1},G_{p2})=\theta\\ \sum_{i=1}^n p_{1i}=1\\ \sum_{j=1}^n p_{2j}=1}}\left(\sum_{i=1}^n\log(np_{1i})+\sum_{j=1}^m\log(mp_{2j})\right). \tag{2.3}$$

Notice that K and Q in the definition of $\theta(F_{p_1},G_{p_2})$ are usually assumed unknown. Hence, a natural way is to replace K and Q in $l_0(\theta)$ by their Kaplan-Meier estimators, say $\hat{K}_n$ and $\hat{Q}_m$, which are defined by

$$1-\hat{K}_n(t)=\prod_{i=1}^n\left[\frac{n-i}{n-i+1}\right]^{I[X_{(i)}\leqslant t,\,\delta_{(i)}=0]}$$

and

$$1-\hat{Q}_m(t)=\prod_{j=1}^m\left[\frac{m-j}{m-j+1}\right]^{I[Y_{(j)}\leqslant t,\,\triangle_{(j)}=0]},$$

where $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$ and $Y_{(1)}\leqslant Y_{(2)}\leqslant\cdots\leqslant Y_{(n)}$ are the order statistics of the X-sample and Y-sample, and $\delta_{(i)}$ and $\triangle_{(j)}$ are the δ and $\triangle$ associated with $X_{(i)}$ and $Y_{(j)}$ respectively. That is, we can define an estimated empirical likelihood, evaluated at θ, by

$$\hat{l}(\theta)=-2\max\left[\sum_{i=1}^n\log(np_{1i})+\sum_{j=1}^m\log(mp_{2j})\right] \tag{2.4}$$

subject to the restrictions

$$\begin{cases}\displaystyle\sum_{j=1}^m p_{2j}\left(\frac{Y_j\triangle_j}{1-\hat{Q}_m(Y_j-)}\right)-\sum_{i=1}^n p_{1i}\left(\frac{X_i\delta_i}{1-\hat{K}_n(X_i-)}\right)=\theta,\\ \displaystyle\sum_{i=1}^n p_{1i}=1 \text{ and } \sum_{j=1}^m p_{2j}=1.\end{cases} \tag{2.5}$$

Let $V_{ni} = X_i \delta_i / (1 - \hat{K}_n(X_i -))$, $U_{mj} = Y_j \triangle_j / (1 - \hat{Q}_m(Y_j -))$, $S_{n,mi} = \frac{1}{m} \sum_{j=1}^{m} U_{mj} - V_{ni}$, and $T_{n,mj} = U_{mj} - \frac{1}{n} \sum_{i=1}^{n} V_{ni}$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

By using the Lagrange multiplier method, $\hat{l}(\theta)$ can be proved to be

$$\hat{l}(\theta) = 2 \left\{ \sum_{i=1}^{n} \log \left[1 + \lambda \left(1 + \frac{m}{n} \right) (S_{n,mi} - \theta) \right] \right. \tag{2.6}$$

$$\left. + \sum_{j=1}^{m} \log \left[1 + \lambda \left(1 + \frac{n}{m} \right) (T_{n,mj} - \theta) \right] \right\}, \tag{2.7}$$

where λ satisfies

$$\left(1 + \frac{m}{n}\right) \sum_{i=1}^{n} \frac{S_{n,mi} - \theta}{1 + \lambda \left(1 + \frac{m}{n}\right)(S_{n,mi} - \theta)} + \left(1 + \frac{n}{m}\right) \times \sum_{j=1}^{m} \frac{T_{n,mj} - \theta}{1 + \lambda \left(1 + \frac{n}{m}\right)(T_{n,mj} - \theta)} = 0. \tag{2.8}$$

Let $\hat{F}_n(t)$ and $\hat{G}_m(t)$ be the Kaplan–Meier estimators of F and G, respectively. It is easy to check the jumps of $\hat{F}_n(t)$ and $\hat{G}_m(t)$ at X_i and Y_j are $\delta_i / n(1 - \hat{K}_n(X_i -))$ and $\triangle_j / m(1 - \hat{Q}_m(Y_j -))$, respectively. This implies that $\frac{1}{n} \sum_{i=1}^{n} V_{ni} = \int_0^\infty t \, d\hat{F}_n(t)$ and $\frac{1}{m} \sum_{i=1}^{m} U_{mj} = \int_0^\infty t \, d\hat{G}_m(t)$. Hence, the modified jackknife variance estimators for $\int_0^\infty t \, d\hat{F}_n(t)$ and $\int_0^\infty t \, d\hat{G}_m(t)$ due to Stute (1996) can be used to estimate the asymptotic variance of $\frac{1}{n} \sum_{i=1}^{n} V_{ni}$ and of $\frac{1}{m} \sum_{j=1}^{m} U_{mj}$ consistently. Let us denote by $\hat{\sigma}^2_{x,Jk}$ and $\hat{\sigma}^2_{y,Jk}$ the modified jackknife estimators of the asymptotic variances, respectively. Further, let

$$\hat{\sigma}^2_{n,m} = 4 \left(1 + \frac{m}{n}\right) \hat{\sigma}^2_{x,JK} + 4 \left(1 + \frac{n}{m}\right) \hat{\sigma}^2_{y,JK}, \tag{2.9}$$

$$\hat{D}^2_{n,m} = \left(\frac{n+m}{n}\right) \frac{1}{n} \sum_{i=1}^{n} (S_{n,mi} - \theta)^2 + \left(\frac{n+m}{m}\right) \frac{1}{m} \sum_{j=1}^{m} (T_{n,mj} - \theta)^2 \tag{2.10}$$

and

$$\eta_{n,m} = \frac{\hat{D}^2_{n,m}}{\hat{\sigma}^2_{n,m}}. \tag{2.11}$$

Then, the adjusted empirical log-likelihood is defined as

$$\hat{l}_{ad}(\theta) = \eta_{n,m} \hat{l}(\theta), \tag{2.12}$$

and $\hat{l}_{ad}(\theta)$ can be proved to be asymptotically standard chi-square distributed with 1 degree of freedom because of the use of the estimated adjusting factor $\eta_{n,m}$. This is an estimator of a quantity indicating the information loss due to censoring.

Let $\bar{H}(s)=P(X_1>s)$, $\bar{L}(s)=P(Y_1>s)$, $\tilde{H}_0(s)=P(X_1>s, \delta_1=0)$, $\tilde{L}_0(s)=P(Y_1>s, \triangle_1=0)$, $\tilde{H}_1(s)=P(X_1>s, \delta_1=1)$, $\tilde{L}_1(s)=P(Y_1>s, \triangle_1=1)$, $\gamma_{0,H}(x)=\exp\{\int_0^{x-} \frac{d\tilde{H}_0(s)}{\bar{H}(s))}\}$, $C_H(x)=\int_0^{x-} \frac{dK(s)}{(1-H(s))(1-K(s))}$, $\tau_H=\inf\{t: H(t)=1\}$. Similar definitions also apply to $\gamma_{0,L}$, $C_L(x)$ and τ_L.

The following assumptions are needed for our results:

(A1)(i) $\int_0^{\tau_H} x\gamma_{0,H}^2(x)\, d\tilde{H}_1(x)<\infty$,

(ii) $\int_0^{\tau_L} y\gamma_{0,L}^2(y)\, d\tilde{L}_1(y)<\infty$,

(A2)(i) $\int_0^{\tau_H} xC_H^{1/2}(x)\, dF(x)<\infty$,

(ii) $\int_0^{\tau_L} yC_L^{12}(y)\, dG(y)<\infty$

(A3)(i) $\int_0^{\tau_H} (x^2\, dF(x)/(1-K(x-)))<\infty$,

(ii) $\int_0^{\tau_L} (y^2\, dG(y)/(1-Q(Y-)))<\infty$,

(A4)(i) $\tau_F=\tau_H$ and $F(\tau_F)=F(\tau_F-)$

(ii) $\tau_G=\tau_L$ and $G(\tau_G)=G(\tau_G-)$,

(A5) $\frac{m}{n}\to\rho>0$.

Remark 2.1. Conditions (A1) and (A2) are used in Stute (1996, 1995). Condition (A3) is to ensure that the second moment of $\Delta Y/(1-Q(Y-))$ and $\delta X/(1-K(X-))$ exists. Condition (A4) is used in Stute and Wang (1993) to ensure estimability of F and G in the right tails, which in turn ensure estimability of the means of F and G.

THEOREM 2.1. *Under assumptions* (A1)–(A5), $\hat{l}_{ad}(\theta)$ *has an asymptotic standard chi-square distribution with* 1 *degree of freedom, that is,*

$$\hat{l}_{ad}(\theta)\xrightarrow{\mathscr{L}}\chi_1^2.$$

Theorem 2.1 can be used to construct an α-level confidence interval, i.e.

$$I_\alpha=\{\tilde{\theta}: \hat{l}(\tilde{\theta})\leqslant c_\alpha\},$$

with $P(\chi_1^2\leqslant c_\alpha)=1-\alpha$.

THEOREM 2.2. *Under the conditions of Theorem* 2.1, I_α *has asymptotically the correct coverage probability* $1-\alpha$, *i.e.*,

$$P(\theta\in I_\alpha)=1-\alpha+o(1).$$

Theorem 2.1 can be used to test the hypothesis $H_0\colon \theta = \theta_0$. According to the duality between confidence intervals and hypothesis tests, we can define an α-level empirical likelihood test for the null hypothesis H_0 by

$$\phi = \begin{cases} 1, & \text{if} \quad \hat{l}_{ad}(\theta_0) > c_\alpha \\ 0, & \text{otherwise.} \end{cases}$$

By Theorem 2.1, we can get

$$P(\phi = 1 \mid H_0) = \alpha + o(1),$$

which means the asymptotic significant level of ϕ is α. That is, we reject H_0 at asymptotic confidence level α if $\hat{l}_{ad}(\theta_0) > c_\alpha$, where c_α is as defined before.

3. MARTINGALE-BASED BOOTSTRAP INFERENCE

Note that $\theta = EY^0 - EX^0 = \int_0^\infty (1 - G(t))\, dt - \int_0^\infty (1 - F(t))\, dt$. Hence, a natural estimator of θ is

$$\hat{\theta}_{n,m} = \int_0^{Y_{(m)}} (1 - \hat{G}_m(t))\, dt - \int_0^{X_{(n)}} (1 - \hat{F}_n(t))\, dt, \tag{3.1}$$

where $\hat{F}_n(t)$ and $\hat{G}_m(t)$ are the Kaplan–Meier estimators.

Let $H_n(t) = \frac{1}{n} \sum_{i=1}^n I[X_i \leqslant t]$, $L_m(t) = \frac{1}{m} \sum_{j=1}^m I[Y_j \leqslant t]$, $N_{1i}(t) = I[X_i \leqslant t, \delta_i = 1]$, $N_{2j}(t) = I[Y_j \leqslant t, \triangle_j = 1]$, $\Lambda^F(t) = \int_0^t \frac{dF(s)}{1 - F(s)}$, and $\Lambda^G(t) = \int_0^t \frac{dG(s)}{1 - G(s)}$. Let

$$M_i^F(t) = N_{1i}(t) - \int_0^t I[X_i \geqslant s]\, d\Lambda^F(s)$$

and

$$M_j^G(t) = N_{2j}(t) - \int_0^t I[Y_j \geqslant s]\, d\Lambda^G(s).$$

By Shorack and Wellner (1986), the $M_i^F(t)$'s and the $M_j^G(t)$'s are square integrable martingales on $[0, +\infty)$.

Let $J_{mY}(y) = I[0 \leqslant y \leqslant Y_{(m)}]$ and $J_{nX}(x) = I[0 \leqslant x \leqslant X_{(n)}]$.

THEOREM 3.1. *Assume* $\sqrt{n}\int_{X_{(n)}}^{\tau_F}(1-F(t))\,dt \xrightarrow{p} 0$ *and* $\sqrt{m}\int_{Y_{(m)}}^{\tau_G}(1-G(t))\,dt \xrightarrow{p} 0$. *We have*

$$\hat{\theta}_{n,m}-\theta=\alpha_m-\beta_n+o_p(n^{-1/2})+o_p(m^{-1/2}),$$

where

$$\alpha_m=-\frac{1}{m}\sum_{j=1}^{m}\int_0^{Y_{(m)}}\left[\int_y^{Y_{(m)}}(1-G(x))\,dx\right]\times\frac{1-\hat{G}_m(y-)}{1-G(y)}\,\frac{J_{mY}(y)}{1-L_m(y-)}\,dM_j^G(y)$$

and

$$\beta_n=-\frac{1}{n}\sum_{i=1}^{n}\int_0^{X_{(n)}}\left[\int_x^{X_{(n)}}(1-F(x))\,dx\right]\times\frac{1-\hat{F}_n(x-)}{1-F(x)}\,\frac{J_{nX}(x)}{1-H_n(x-)}\,dM_i^F(x).$$

Theorem 3.1 gives a martingale representation for $\hat{\theta}_{n,m}-\theta$. From this theorem, the asymptotic distribution of $\sqrt{n+m}(\hat{\theta}_{n,m}-\theta)$ is the same as that of $\sqrt{n+m}(\alpha_m-\beta_n)$ as $\frac{m}{n}\to\rho$. Following the idea of Lin *et al.* (1993), the limiting distribution of $\sqrt{n+m}(\alpha_m-\beta_n)$ can be approximated through a Monte Carlo simulation. Let $\{\xi_{1i}, 1\leqslant i\leqslant n\}$ and $\{\xi_{2j}, 1\leqslant j\leqslant m\}$ be independent standard normal random variables which are independent of each other, $\{(X_i,\delta_i), 1\leqslant i\leqslant n\}$, and $\{(Y_j,\triangle_j) 1\leqslant j\leqslant m\}$, respectively. We replace $\{M_i^F(t)\}$ and $\{M_j^G(t)\}$ in α_m and β_n by $\{N_{1i}(t)\,\xi_{1i}\}$ and $\{N_{2j}(t)\,\xi_{2j}\}$, and F and G by $\hat{F}_n(t)$ and $\hat{G}_m(t)$, respectively. The resulting statistic is then

$$W_{n,m}^*=\alpha_m^*-\beta_n^*,$$

where

$$\alpha_m^*=-\frac{1}{m}\sum_{j=1}^{m}\int_0^{Y_{(m)}}\left(\int_y^{(Y_{(m)}}(1-\hat{G}_m(s))\,ds\times\frac{1-\hat{G}_m(y-)}{1-\hat{G}_m(y)}\,\frac{J_{mY}(y)}{1-L_m(y-)}\right)\xi_{2j}\,dN_{2j}(y)$$

and

$$\beta_n^* = -\frac{1}{n}\sum_{j=1}^{n}\int_0^{X_{(n)}}\left(\int_x^{(X_{(n)}}(1-\hat{F}_n(t))\,dt\right.$$
$$\left.\times\frac{1-\hat{F}_n(x-)}{1-\hat{F}_n(x)}\,\frac{J_{nX}(x)}{1-H_n(x-)}\right)\xi_{1j}\,dN_{1j}(x).$$

The following theorem shows that the distribution of $\sqrt{n+m}(\hat{\theta}_{n,m}-\theta)$ can be approximated by $K^*(x)=P^*(\sqrt{n+m}\ W_{n,m}^*\leqslant x)$, where P^* denotes the conditional probability given $\{X_i,\delta_i\}_{i=1}^{n}$ and $\{Y_j,\triangle_j\}_{j=1}^{m}$.

THEOREM 3.2. *Under the following conditions*:

(C1) $\int_0^{\tau_H}\frac{1}{1-K(s-)}\,dF(s)$ *and* $\int_0^{\tau_L}\frac{1}{1-Q(s-)}\,dG(s)<\infty$,

(C2) $\sup_t|\int_t^{\tau_F}(1-F(s))\,ds/(1-F(t))|<\infty$ *and* $\sup_t|\int_t^{\tau_G}(1-G(s))\,ds/(1-G(t))|<\infty$,

(C3) $\sqrt{n}\int_{(X_{(n)}}^{\tau_F}(1-F(t))\,dt\xrightarrow{p}0$ *and* $\sqrt{m}\int_{Y_{(m)}}^{\tau_G}(1-G(t))\,dt\xrightarrow{p}0$,

if F and K, and G and Q have no common jumps and $F(\{\tau_H\})$, *and* $G(\{\tau_L\})=0$, *we have with probability* 1,

$$\sup_x|P(\sqrt{n+m}\,(\hat{S}_{n,m}-S)\leqslant x)-K^*(x)|\xrightarrow{a.s.}0.$$

Remark 3.1. There are many examples where the conditions of Theorem 3.2 are satisfied. For instance, the first parts of Conditions (C1) and (C2) are clearly satisfied when $F(t)=1-e^{-2rt}$ and $K(t)=1-e^{-rt}$ for $t\geqslant 0$ and some constant $r>0$. Now let us check the first part of Condition (C3) for this example. Notice that $P(X_{(n)}>\log n^{5/16r})=1-(1-e^{-\log n^{15/16}})^n=1-(1-n^{-15/16})^n$. This proves $P(X_{(n)}>\log n^{5/16r})\to 1$. Hence, we have in probability $\sqrt{n}\int_{X_{(n)}}^{\tau_F}(1-F(t)\,dt\leqslant\sqrt{n}e^{-\log n^{5/8}}\to 0$.

Clearly, $K^*(x)$ can be calculated using Monte carlo simulation by repeatedly generating $\{\xi_{1i}\}_{i=1}^{n}$ and $\{\xi_{2j}\}_{j=1}^{m}$, respectively, from the standard normal distribution while keeping $\{X_i,\delta_i\}_{i=1}^{n}$ and $\{Y_j,\triangle_j\}_{j=1}^{m}$ fixed. This method is introduced by Lin *et al.* (1993) and is termed the martingale-based bootstrap in Wang and Jing (1998).

Theorem 3.2 can also be applied for the construction of confidence intervals of θ. From Theorem 3.2, the confidence interval for θ at level α can be written as

$$I_{MB,\alpha}=(\hat{\theta}_{n,m}-\hat{q}_{1-\alpha/2}^*(n+m)^{-1/2},\hat{\theta}_{n,m}-\hat{q}_{\alpha/2}^*(n+m)^{-1/2}),$$

where $\hat{q}_\gamma$ satisfies $K^*(\hat{q}_\gamma^*) = \gamma$ for $0 < \gamma < 1$. The following theorem shows that the confidence interval has the correct coverage probability.

THEOREM 3.3. *Under conditions of Theorem 3.2, we have*

$$P(\theta \in \hat{I}_{MB,\alpha}) = 1 - \alpha + o(1).$$

4. SIMULATION RESULTS

In the introduction, we presented some advantages of the adjusted empirical likelihood (AEL) and the martingale-based bootstrap method (MBB). We now compare the performances of the AEL, the MBB, and Efron's bootstrap (EB) method in terms of the coverage probabilities and the average lengths of their confidence intervals via simulation studies. The coverage probabilities and average lengths of the EB confidence sets are calculated based on the bootstrap estimator of $\hat{\theta}_{n,m}$ in (3.1).

We consider the two-sample random censorship model with $F(t) = 1 - e^{-2t}$, $t \geqslant 0$, $G(t) = 1 - e^{-2(t-1)}$, $t \geqslant 1$, $K(t) = 1 - e^{-c_1 t}$, $t \geqslant 0$ and $Q(t) = 1 - e^{-c_2(t-1)}$, $t \geqslant 1$ with c_1 and c_2 chosen to accommodate certain preselected censoring percentage. That is, in the model, the life data of two groups were generated from $F(t)$ and $G(t)$, and the corresponding censoring times were generated from $K(t)$ and $Q(t)$, respectively. The simulations were run with sample sizes of $(n, m) = (10, 20)$, $(15, 10)$, $(25, 30)$, $(30, 25)$ and $(60, 60)$, respectively. The coverage probabilities and average lengths of the confidence intervals are calculated for the AEL, MB, and EB method from 1000 simulated data sets of each sample size (n, m). The nominal level is taken to be 0.90. Table I gives the simulation results.

From Table I, we observe the following:

(1) The AEL and the MBB method do perform competitively in comparison to Efron's bootstrap method, as their confidence intervals have relatively high coverage accuracies and short average lengths. Actually, the standard bootstrap confidence intervals are too conservative in terms of the coverage probabilities, and this suggests lack of consistency.

(2) The AEL works uniformly well in terms of the average lengths of the confidence intervals. In terms of coverage accuracies, it seems that the AEL also performs better than the MBB for small and moderate sample sizes (e.g., $(n, m) = (10, 15)$ and $(n, m) = (25, 30)$) in the cases where the CP are 0.10 and 0.25, respectively. For large sample sizes (e.g., $(n, m) = (60, 60)$), the MBB seems to perform slightly better than the AEL. Also, the MBB is more preferable than AEL in the worst case where the CP is

TABLE I

Coverage Probabilities (COPR) and Average Lengths (AVLE) for the Confidence Intervals of θ under Different Censoring Percentages (CP) When the Nominal Level is 0.90

		AEL		MB		EB	
CP	(n, m)	COPR	AVLE	COPR	AVLE	COPR	AVLE
	(10,15)	0.8760	0.4010	0.9520	0.6060	0.9800	1.5784
	(15,10)	0.8680	0.3720	0.9480	0.6295	0.9720	1.4296
0.10	(25,30)	0.9180	0.3480	0.8770	0.4534	0.9760	0.7323
	(30,25)	0.8860	0.3430	0.9390	0.4515	0.9580	0.5803
	(60,60)	0.9080	0.2860	0.9020	0.3358	0.9500	0.3504
	(10,15)	0.8420	0.4190	0.9560	0.6372	0.9820	1.4815
	(15,10)	0.8340	0.4240	0.9510	0.6466	0.9760	1.3137
0.25	(25,30)	0.9160	0.3830	0.9320	0.4753	0.9820	1.1353
	(30,25)	0.9220	0.3970	0.8630	0.4698	0.9680	0.9302
	(60,60)	0.9170	0.3120	0.9140	0.3677	0.9700	0.8711
	(10,15)	0.7840	0.4450	0.8220	0.6693	0.9580	1.2075
	(15,10)	0.8120	0.4220	0.8420	0.5587	0.9840	1.2293
0.40	(25,30)	0.8540	0.4010	0.8540	0.4745	0.9820	0.8259
	(30,25)	0.8580	0.4200	0.8600	0.4692	0.9660	0.7698
	(60,60)	0.8810	0.3810	0.8830	0.3823	0.9900	0.5647

0.40 in terms of coverage accuracies. However, in terms of the average lengths, we have the opposite conclusion in this case.

(3) The performances of both AEL and MBB depend on the censoring percentages and sample size. For every fixed sample size (n, m), the coverage accuracies for both methods generally decrease as the censoring percentage increases. The coverage accuracies increase for every fixed censoring percentage as sample size increases. It seems that the CP and sample size affect the coverage accuracies of the bootstrap much less. The reason may be that the bootstrap is not consistent.

Based on (B.2) and the "plug method" to estimate the asymptotic variance, we also calculated the coverage probabilities and average lengths of normal approximation (NA) confidence intervals of θ. The simulation results show that the NA confidence intervals have uniformly lower coverage accuracies than the AEL and MBB and similar average lengthes to MBB.

APPENDIX A

Proofs of Theorems 2.1 *and* 2.2

Let

$$\gamma_{1,H}(x) = \frac{1}{\bar{H}(x)} \int I[x < s]\, s\gamma_{0,H}(s)\, d\tilde{H}_1(s)$$

and

$$\gamma_{2,H}(x) = \iint \frac{I[s < x, s < t]\, t\gamma_{0,H}(t)}{\bar{H}^2(s)}\, d\tilde{H}_0(s)\, d\tilde{H}_1(t),$$

where $\gamma_{0,H}$ is defined in Section 2. Similarly, we can define $\gamma_{1,L}(x)$ and $\gamma_{2,L}(x)$.

To prove Theorem 2.1, Lemma is needed.

LEMMA 4.1. *Under Assumptions* (A1), (A3), *and* (A4), *we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (V_{ni} - EX^0) \xrightarrow{\mathscr{L}} N(0, \sigma_1^2),$$

and

$$\frac{1}{\sqrt{m}} \sum_{j=1}^{m} (U_{mj} - EY^0) \xrightarrow{\mathscr{L}} N(0, \sigma_2^2),$$

where

$$\sigma_1^2 = Var[X_1 \gamma_{0,H}(X_1)\, \delta_1 + \gamma_{1,H}(X_1)(1-\delta_1) - \gamma_{2,H}(X_1)],$$

and

$$\sigma_2^2 = Var[Y_1 \gamma_{0,L}(Y_1)\, \triangle_1 + \gamma_{2,L}(Y_1)(1-\triangle_1) - \gamma_{2,L}(Y_1)].$$

Proof of Lemma. Notice that

$$\frac{1}{n} \sum_{i=1}^{n} (V_{ni} - EX^0) = \int_0^\infty x\, d(\hat{F}_n - F)$$

and the similar expressions apply to $\frac{1}{m} \sum_{j=1}^{m} (U_{mj} - EY^0)$. Hence, Corollary 1.2 of Stute (1995) proves Lemma A.1.

Proof of Theorem 2.1. To prove Theorem 2.1, we need to prove

(a) $\max_{1 \leqslant i \leqslant n} |S_{n,mi}| = o_p((n+m)^{1/2})$ and $\max_{1 \leqslant j \leqslant m} |T_{n,mj}| = o_p((n+m)^{1/2})$.

(b) $\lambda = O_p((n+m)^{-1/2})$, where λ is that satisfying (2.8).

Let us first prove (a). Notice that

$$\max_{1 \leqslant i \leqslant n} |S_{n,mi}| = \max_{1 \leqslant i \leqslant n} \left| V_{ni} + \frac{1}{m} \sum_{j=1}^{m} \frac{\triangle_j Y_j}{1 - \hat{Q}_m(Y_j-)} - \theta \right| \leqslant \max_{1 \leqslant i \leqslant n} |V_{ni} - EX^0| + \max_{1 \leqslant j \leqslant m} |U_{mj} - EY^0|. \tag{A.1}$$

Let $V_i = \delta_i X_i / (1 - K(X_i-))$. It is clear that $\{V_i\}_{i=1}^n$ are iid non-negative random variables and $EV_i^2 = \int (x^2/(1-K(x-)))\, dF(x) < \infty$. Hence, by Lemma 3 of Owen (1990), we have

$$\max_{1 \leqslant i \leqslant n} V_i = o_p(n^{1/2}). \tag{A.2}$$

This together with the fact

$$\sup_{0 \leqslant x \leqslant X_{(n)}} \left| \frac{\hat{K}_n(x-) - K(x-)}{1 - \hat{K}_n(x-)} \right| = O_p(1) \tag{A.3}$$

(see, e.g., Zhou, 1991) shows

$$\max_{1 \leqslant i \leqslant n} |V_{ni}| \leqslant \max_{1 \leqslant i \leqslant n} |V_i| + \max_{1 \leqslant i \leqslant n} \left| \frac{\delta_i X_i (\hat{K}_n(X_i-) - K(X_i-))}{(1 - K(X_i-))(1 - \hat{K}_n(X_i-))} \right| \leqslant o_p(n^{1/2}) + \sup_{0 \leqslant s \leqslant X_{(n)}} \left| \frac{\hat{K}_n(s) - K(s)}{1 - \hat{K}_n(s)} \right| \max_{1 \leqslant i \leqslant n} |V_i| = o_p(n^{1/2}). \tag{A.4}$$

Similarly, we can demonstrate

$$\max_{1 \leqslant j \leqslant m} |U_{mj}| = o_p(m^{1/2}). \tag{A.5}$$

Relations (A.4) and (A.5) together prove

$$\max_{1 \leqslant j \leqslant m} |S_{n,mi}| = o_p((n+m)^{1/2}), \tag{A.6}$$

since $\frac{m}{n} \to \rho > 0$. The same arguments can be used to obtain

$$\max_{1 \leqslant j \leqslant m} |T_{n,mj}| = o_p((n+m)^{1/2}). \tag{A.7}$$

Relations (A.6) and (A.7) together yield part (a).

Next, we show part (b). Notice that

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i\delta_i}{1-\hat{K}_n(X_i-)}\right)^2 \\
&\quad = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i\delta_i}{1-\hat{K}_n(X_i-)}\right)^2 I[1-\hat{K}_n(X_i-) \leqslant 2(1-K(X_i-))] \\
&\quad\quad + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i\delta_i}{1-\hat{K}_n(X_i-)}\right) I[1-\hat{K}_n(X_i-) > 2(1-K(X_i-))] \\
&\quad := \zeta_{n1} + \zeta_{n2}.
\end{aligned} \tag{A.8}$$

Under (A3), we have $(1/n)\sum_{i=1}^{n}(X_i\delta_i/(1-K(X_i-)))^2 \xrightarrow{a.s.} \int (x^2\, dF(x)/(1-K(x))) < \infty$. Hence, it follows that with probability 1

$$\zeta_{n1} \geqslant \frac{1}{4}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i\delta_i}{1-K(X_i-)}\right)^2 \geqslant \frac{1}{4}\int \frac{x^2\, dF(x)}{1-K(x-)} > 0. \tag{A.9}$$

On the other hand, for any $\varepsilon > 0$ we have

$$\begin{aligned}
P(|\zeta_{n1}| > \varepsilon) &\leqslant P\left(\bigcup_{i=1}^{n}\left\{1-\hat{K}_n(X_i-) > 2(1-K(X_i-))\right\}\right) \\
&\leqslant P\left(\bigcup_{i=1}^{n}\left\{|\hat{K}_n(X_i-)-K(X_i-)| > 1-K(X_i-)\right\}\right) \\
&\leqslant P\left(\sup_{1 \leqslant x \leqslant X_{(n)}}\left|\frac{\hat{K}_n(x)-K(x)}{1-K(x)}\right| > 1\right) \to 0.
\end{aligned} \tag{A.10}$$

From (A.8), (A.9), and (A.10), we get

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i\delta_i}{1-\hat{K}_n(X_i-)}\right)^2 \geqslant \frac{1}{4}\int \frac{x^2\, dF(x)}{1-K(x-)} + o_p(1). \tag{A.11}$$

Similarly, we have

$$\frac{1}{m}\sum_{j=1}^{m}\left(\frac{Y_j\triangle_j}{1-\hat{Q}_m(Y_j-)}\right)^2 \geqslant \frac{1}{4}\int \frac{y^2\, dG(y)}{1-Q(Y-)} + o_p(1). \tag{A.12}$$

By (A.11) and (A.12), it follows that

$$\frac{1}{n}\sum_{i=1}^{n} S_{n,\,mi}^2 \geqslant \frac{1}{4}\left(\int \frac{x^2\,dF(x)}{1-K(x-)} + \int \frac{y^2\,dG(y)}{1-Q(Y-)}\right) + o_p(1), \quad \text{(A.13)}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} T_{n,\,mj}^2 \geqslant \frac{1}{4}\left(\int \frac{x^2\,dF(x)}{1-K(x-)} + \int \frac{y^2\,dG(y)}{1-Q(Y-)}\right) + o_p(1). \quad \text{(A.14)}$$

Lemma A.1 implies that

$$\frac{1}{n}\sum_{i=1}^{n} S_{n,\,mi} - \theta = O_p((n+m)^{-1/2}), \quad \text{(A.15)}$$

and

$$\frac{1}{m}\sum_{j=1}^{m} T_{n,\,mj} - \theta = O_p((n+m)^{-1/2}), \quad \text{(A.16)}$$

as $\frac{m}{n} \to \rho > 0$.

By (A.13)–(A.16) and the same arguments as in the proof of (2.14) in Owen (1990), we can prove (b).

From (2.8), we have

$$\begin{aligned}
&\left(1+\frac{m}{n}\right)\sum_{i=1}^{n}(S_{n,\,mi}-\theta) \\
&\quad\times\left[1-\lambda\left(1+\frac{m}{n}\right)(S_{n,\,mi}-\theta)+\frac{\lambda^2(1+m/n)^2\,(S_{n,\,mi}-\theta)^2}{1+\lambda(S_{n,\,mi}-\theta)}\right] \\
&\quad+\left(1+\frac{n}{m}\right)\sum_{i=1}^{m}(T_{n,\,mi}-\theta) \\
&\quad+\left[1-\lambda\left(1+\frac{n}{m}\right)(T_{n,\,mi}-\theta)+\frac{\lambda^2(1+n/m)^2\,(T_{n,\,mi}-\theta)^2}{1+\lambda(T_{n,\,mi}-\theta)}\right]=0.
\end{aligned} \quad \text{(A.17)}$$

Solving the equation, we get

$$\lambda = \frac{\left(1+\frac{m}{n}\right)\sum_{i=1}^{n}(S_{n,\,mi}-\theta)+\left(1+\frac{n}{m}\right)\sum_{j=1}^{m}(T_{n,\,mj}-\theta)}{\left(1+\frac{m}{n}\right)^2\sum_{i=1}^{m}(S_{n,\,mi}-\theta)^2+\left(1+\frac{n}{m}\right)^2\sum_{j=1}^{m}(T_{n,\,mj}-\theta)^2}+\gamma_n \quad \text{(A.18)}$$

with

$$\gamma_n = \frac{\lambda^2((m+n)/n)^3 \sum_{i=1}^n ((S_{n,mi}-\theta)^3/(1+\lambda(1+m/n)(S_{n,mi}-\theta))) + \lambda^2((m+n)/m)^3 \sum_{j=1}^m ((T_{n,mj}-\theta)^3/(1+\lambda(1+n/m)(T_{n,mj}-\theta)))}{((n+m)/n)^2 \sum_{i=1}^n (S_{n,mi}-\theta)^2 + ((n+m)/m)^2 \sum_{j=1}^m (T_{n,mj}-\theta)^2}.$$

Under Assumption (A3), we have $(1/n)\sum_{i=1}^n (\delta_i X_i/(1-K(X_i-)))^k = O_p(1)$ for $k=1,2$. This together with Eq. (A.3) proves $(1/n)\sum_{i=1}^n (\delta_i X_i/(1-\hat{K}_n(X_i-)))^k = O_p(1)$, $k=1,2$. Similarly, we can prove that $(1/m)\sum_{j=1}^m (\triangle_j Y_j/(1-\hat{Q}_m(Y_j-)))^k = O_p(1)$, $k=1,2$. Hence, we have

$$\frac{1}{n}\sum_{i=1}^n (S_{n,mi}-\theta)^2 = O_p(1), \tag{A.19}$$

and

$$\frac{1}{m}\sum_{j=1}^m (T_{n,mj}-\theta)^2 = O_p(1). \tag{A.20}$$

By result (b), Eqs. (A.6), (A.7), (A.19), and (A.20), we get

$$\begin{aligned} |\gamma_n| &\leqslant O_p((n+m)^{-1})(\max_{1\leqslant i\leqslant n}|S_{n,mi}-\theta| + \max_{1\leqslant j\leqslant m}|T_{n,mj}-\theta|) \\ &= o_p((n+m)^{-12}). \end{aligned} \tag{A.21}$$

Using Taylor's expansion in (2.12), we get

$$\begin{aligned} l_{ad} = 2\sum_{i=1}^n &\left\{\lambda\left(1+\frac{m}{n}\right)(S_{n,mi}-\theta) - \frac{1}{2}\left[\lambda\left(1+\frac{m}{n}\right)(S_{n,mi}-\theta)\right]^2\right\} \\ &+ 2\sum_{j=1}^m \left\{\lambda\left(1+\frac{n}{m}\right)(T_{n,mj}-\theta)\right. \\ &\left. - \frac{1}{2}\left[\lambda\left(1+\frac{n}{m}\right)(T_{n,mj}-\theta)\right]^2\right\} + \zeta_{n,m}, \end{aligned} \tag{A.22}$$

where

$$\zeta_{n,m} \leqslant \lambda^3 \sum_{i=1}^n \left[\left(1+\frac{m}{n}\right)(S_{n,mi}-\theta)\right]^3 + \lambda^3 \sum_{j=1}^m \left[\left(1+\frac{n}{m}\right)(T_{m,j}-\theta)\right]^2.$$

Again using result (b) and Eqs. (A.6), (A.7), (A.19), and (A.20), it follows that

$$\zeta_{n,m} = o_p(1). \tag{A.23}$$

Denote by $g(\lambda)$ the right hand side of Eq. (2.8). Similar to (A.22) and (A.23), it follows that

$$0=\lambda g(\lambda)=\frac{n+m}{n}\left\{\sum_{i=1}^{n}\lambda(S_{n,mi}-\theta)-\left(1+\frac{m}{n}\right)\sum_{i=1}^{n}[\lambda(S_{n,mi}-\theta)]^2\right\}$$
$$+\frac{n+m}{m}\left\{\sum_{j=1}^{m}\lambda(T_{n,mj}-\theta)-\left(1+\frac{n}{m}\right)\sum_{j=1}^{m}[\lambda(T_{n,mj}-\theta)]^2\right\}+o_p(1). \tag{A.24}$$

That is,

$$\frac{n+m}{n}\sum_{i=1}^{n}(S_{n,mi}-\theta)\,\lambda+\frac{n+m}{m}\sum_{j=1}^{m}(T_{n,mj}-\theta)\,\lambda$$
$$=\left(\frac{n+m}{n}\right)^2\sum_{i=1}^{n}[\lambda(S_{n,mi}-\theta)]^2+\left(\frac{n+m}{m}\right)^2\sum_{j=1}^{m}[\lambda(T_{n,mj}-\theta)]^2+o_p(1). \tag{A.25}$$

Equations (A.22), (A.23), and (A.25) together yield

$$l_{ad}(\theta)=\lambda^2\left(1+\frac{m}{n}\right)^2\sum_{i=1}^{n}(S_{n,mi}-\theta)^2+\lambda^2\left(1+\frac{n}{m}\right)^2\sum_{j=1}^{m}(T_{n,mj}-\theta)^2+o_p(1). \tag{A.26}$$

From (A.18), (A.21), (A.26), (A.15), (A.16), (A.19), and (A.20), it follows that

$$l_{ad}(\theta)=\frac{\Gamma_{n,m}^2}{D_{n,m}^2}+o_p(1), \tag{A.27}$$

where $D_{n,m}^2$ is defined as in (2.10) and

$$\Gamma_{n,m}=\frac{\sqrt{n+m}}{n}\sum_{i=1}^{n}(S_{n,mi}-\theta)+\frac{\sqrt{n+m}}{m}\sum_{j=1}^{m}(T_{n,mj}-\theta).$$

A simple calculation yields

$$\Gamma_{n,m}=2\sqrt{\frac{n+m}{m}}\left[\frac{1}{\sqrt{m}}\sum_{j=1}^{m}(U_{n,mj}-EY^0)\right]$$
$$-2\sqrt{\frac{n+m}{n}}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(V_{n,mi}-EX^0)\right].$$

By Lemma A.1, we get

$$\Gamma_{n,m} \xrightarrow{\mathscr{L}} N(0, \sigma^2), \tag{A.28}$$

where

$$\sigma^2 = 4(1+\rho)\,\sigma_1^2 + 4\left(1+\frac{1}{\rho}\right)\sigma_2^2.$$

Recalling the definition of $\hat{l}_{ad}(\theta)$, we have

$$\hat{l}_{ad}(\theta) = \frac{\Gamma_{n,m}^2}{\hat{\sigma}_{n,m}^2} + o_p(1).$$

By Stute (1996),we have $\hat{\sigma}_{i,JK}^2 \xrightarrow{a.s.} \sigma_i^2$ for $i=1, 2$. This proves $\hat{\sigma}_{n,m}^2 \xrightarrow{p} \sigma^2$ and hence

$$\hat{l}_{ad}(\theta) \xrightarrow{\mathscr{L}} \chi_1^2 \tag{A.29}$$

by (A.28).

Proof of Theorem 2.2. Theorem 2.2 is a direct result of Theorem 2.1.

APPENDIX B

Proofs of Theorems 3.1, 3.2, *and* 3.3

Proof of Theorem 3.1. Theorem 3.1 is a direct result of (3.5) in Wang and Jing (2000) under assumptions $\sqrt{n}\int_{X_{(n)}}^{\tau_F}(1-F(t))\,dt \xrightarrow{p} 0$ and $\sqrt{m}\int_{Y_{(m)}}^{\tau_G}(1-G(t))\,dt \xrightarrow{p} 0$.

Proof of Theorem 3.2. Note that

$$\begin{aligned}
&\sqrt{n+m}\,(\hat{\theta}_{n,m}-\theta)\\
&\quad= \sqrt{1+\frac{n}{m}}\left[\sqrt{m}\left(\int_0^{Y_{(m)}}(1-\hat{G}_m(t))\,dt - \int_0^{\infty}(1-G(t))\,dt\right)\right]\\
&\qquad - \sqrt{1+\frac{m}{n}}\left[\sqrt{n}\left(\int_0^{X_{(n)}}(1-\hat{F}_n(t))\,dt - \int_0^{\infty}(1-F(t))\,dt\right)\right].
\end{aligned} \tag{B.1}$$

By Theorem 2.1 of Wang and Jing (2000) and the fact that $(X_1, \delta_1), ..., (X_n, \delta_n)$ are independent of $(Y_1, \triangle_1), ..., (Y_m, \triangle_m)$, it follows that

$$\sqrt{n+m}\,(\hat{\theta}_{n,m} - \theta) \xrightarrow{\mathscr{L}} N\left(0, \left(1+\frac{1}{\rho}\right)\tilde{\sigma}_2^2 + (1+\rho)\,\tilde{\sigma}_1^2\right) \tag{B.2}$$

as $\frac{m}{p} \to \rho > 0$, where

$$\tilde{\sigma}_1^2 = \int_0^{\tau_H} \left(\int_s^{\tau_H} (1 - F(x)\, dx\right)^2 \frac{1-F(s-)}{1-F(s)} \frac{1}{1-H(s-)}\, d\Lambda^F(s),$$

and

$$\tilde{\sigma}_2^2 = \int_0^{\tau_L} \left(\int_s^{\tau_L} (1 - G(x))\, dx\right)^2 \frac{1-G(s-)}{1-G(s)} \frac{1}{1-L(s-)}\, d\Lambda^G(s).$$

Next, we prove with probability 1

$$\sqrt{n+m}\; W_n^* \xrightarrow{\mathscr{L}^*} N\left(0, (1+\rho)\,\tilde{\sigma}_1^2 + \left(1+\frac{1}{\rho}\right)\tilde{\sigma}_2^2\right). \tag{B.2}$$

It is easy to see that $\sqrt{n+m}\; W_n^*$ is a sequence of normal variables with zero mean and variance

$$\begin{aligned}\hat{\sigma}_{n,m}^2 = &\left(1+\frac{n}{m}\right)\frac{1}{m}\sum_{j=1}^{m}\left(\int_0^{Y_{(m)}}\left[\int_s^{Y_{(m)}} (1 - G_m(x))\, dx\right]\right.\\ &\left.\times \frac{1-\hat{G}_m(s-)}{1-\hat{G}_m(s)} \frac{1}{1-L_m(s-)}\, dN_{2j}(s)\right)^2\\ &+\left(1+\frac{m}{n}\right)\frac{1}{n}\sum_{i=1}^{n}\left(\int_0^{X_{(n)}}\left[\int_s^{X_{(n)}} (1 - \hat{F}_n(x))\, dx\right]\right.\\ &\left.\times \frac{1-\hat{F}_n(s-)}{1-\hat{F}_n(s)} \frac{1}{1-H_n(s-)}\, dN_{1i}(s)\right)^2.\end{aligned}$$

To prove (B.2), it is sufficient to prove

$$\sigma_{n,m}^2 \xrightarrow{p} (1+\rho)\,\tilde{\sigma}_1^2 + \left(1+\frac{1}{\rho}\right)\tilde{\sigma}_2^2. \tag{B.3}$$

Observe that

$$\hat{\sigma}^2_{n,m} = \left(1+\frac{n}{m}\right)\frac{1}{m}\sum_{j=1}^{m}\int_0^{Y_{(m)}}\left[\left(\int_s^{Y_{(m)}}(1-G_m(x))\,dx\right)\right.$$
$$\left.\times\frac{1-\hat{G}_m(s-)}{1-\hat{G}_m(s)}\,\frac{1}{1-L_m(s-)}\right]^2 dN_{2j}(s)$$
$$+\left(1+\frac{m}{n}\right)\frac{1}{n}\sum_{i=1}^{n}\int_0^{X_{(n)}}\left[\left(\int_s^{X_{(n)}}(1-\hat{F}_n(x))\,dx\right)\right.$$
$$\left.\times\frac{1-\hat{F}_n(s-)}{1-\hat{F}_n(s)}\,\frac{1}{1-H_n(s-)}\right]^2 dN_{1i}(s).$$

By Stute and Wang (1993), $\hat{F}_n$ and $\hat{G}_m$ are strong uniform consistent on $[0, \tau_H]$ since F and K, and G and Q have no common jumps and $F(\{\tau_H\})=0$ and $G(\{\tau_L\})=0$. Hence,

$$\hat{\sigma}^2_{n,m} \xrightarrow{a.s.} \left(1+\frac{1}{\rho}\right)\int_0^{\tau_G}\left[\left(\int_s^{\tau_G}(1-G(x))\,dx\right)\frac{1-G(s-)}{1-G(s)}\,\frac{1}{1-L(s-)}\right]^2$$
$$\times(1-Q(s-))\,dG(s)$$
$$+(1+\rho)\int_0^{\tau_F}\left[\left(\int_s^{\tau_F}(1-F(x))\,dx\right)\frac{1-F(s-)}{1-F(s)}\,\frac{1}{1-H(s-)}\right]^2$$
$$\times(1-K(s-))\,dF(s)$$
$$=(1+\rho)\,\tilde{\sigma}_1^2+\left(1+\frac{1}{\rho}\right)\tilde{\sigma}_2^2.$$

This proves (B.3), and (B.2) and Theorem 2.2.

Proof of Theorem 3.3. Theorem 3.3 is a direct result of Theorem 3.2.

ACKNOWLEDGMENTS

The authors thank two referees for their valuable comments and suggestions. Our thanks are also due to Professors W. Stute and G. Li for their careful reading and beneficial suggestions which improved the presentation of this work. The work is supported partially by an NSF Grant DMS-98-03627, the NNSF of China, and a grant to the first author for his excellent Ph.D. dissertation work in China.

REFERENCES

1. S. X. Chen, On the accuracy of empirical likelihood confidence regions for linear regression model, *Ann. Inst. Statist. Math.* **45** (1993), 621–637.
2. J. H. Chen and J. Qin, Empirical likelihood estimation for finite populations and the effective usage of auxiliary information, *Biometrika* **80** (1993), 107–116.
3. S. X. Chen, Empirical likelihood confidence intervals for linear regression coefficients, *J. Multivariate Anal.* **49** (1994), 24–40.
4. T. J. DiCiccio, P. Hall, and J. Romano, "Bartlett Adjustment for Empirical Likelihood," Technical Report 298, Department of Statistics, Stanford University, 1988.
5. T. J. DiCiccio, P. Hall, and J. P. Romano, Bartlett adjustment for empirical likelihood, *Ann. Statist.* **19** (1991), 1053–1061.
6. P. Hall and B. La Scala, Methodology and algorithms of empirical likelihood, *Internat. Statist. Rev.* **58** (1990), 109–127.
7. E. Kaplan and R. Meier, Nonparametric estimation from incomplete observations, *J. Am. Statist. Assoc.* **53** (1958), 457–481.
8. Y. Kitamura, Empirical likelihood methods with weakly dependent processes, *Ann. Statist.* **25** (1997), 2084–2102.
9. E. D. Kolaczyk, Empirical likelihood for generalized linear models, *Statist. Sinica* **4** (1994), 199–218.
10. Y. Lin and C. F. Spiekerman, Model checking techniques for parametric regression with censored data, *Scand. J Statist.* **23** (1996), 157–177.
11. D. Y. Lin, L. J. Wei, and Z. Ying, Checking the Cox model with cumulative sums of martingale-based residuals, *Biometrika* **80** (1993), 557–572.
12. A. Owen, Empirical likelihood ratio confidence intervals for single functional, *Biometrika* **75** (1988), 237–249.
13. A. Owen, Empirical likelihood ratio confidence regions, *Ann. Statist.* **18** (1990), 90–120.
14. A. Owen, Empirical likelihood for linear models, *Ann. Statist.* **19** (1991), 1725–1747.
15. X.-R. Pan and M. Zhou, Using 1-parameter sub-family of distributions in empirical likelihood ratio with censored data, manuscript, 2000.
16. J. Qin, Semi-empirical likelihood ratio confidence intervals for the difference of two sample means, *Ann. Inst. Statist. Math.* **46** (1994), 117–126.
17. J. Qin and J. F. Lawless, Empirical likelihood and general estimating equations, *Ann. Statist.* **22** (1994), 300–325.
18. J. Qin, Empirical likelihood in a semiparametric Model, *Scand. J. Statist.* **23** (1996), 209–219.
19. G. R. Shorack and J. A. Wellner, "Empirical Processes with Applications to Statistics," Wiley, New York, 1986.
20. W. Stute, The central limit theorem under random censorship, *Ann. Statist.* **23** (1995), 422–439.
21. W. Stute, The jackknife estimate of variance of a Kaplan–Meier integral, *Ann. Statist.* **24** (1996), 2679–2704.
22. W. Stute and J.-L. Wang, The strong law under random censorship, *Ann. Statist.* **21** (1993), 1591–1607.
23. D. R. Thomas and G. L. Grunkemeier, Confidence interval estimation of survival probabilities for censored data, *J. Amer. Statist. Assoc.* **70** (1975), 865–871.
24. Q. H. Wang and B. Y. Jing, Empirical likelihood for partial linear model with fixed design, *Statist. Probab. Lett.* **41** (1999), 425–433.
25. Q. H. Wang and B. Y. Jing, A martingale-based bootstrap inference with censored data, *Comm. Statist.* **29** (2000), 401–416.
26. M. Zhou, Some properties of the Kaplan–Meier estimator, for independent, non-identically distributed random variables, *Ann. Statist.* **19** (1991), 2266–2276.