Published in final edited form as:

Scand Stat Theory Appl. 2008 June 1; 35(2): 212–227. doi:10.1111/j.1467-9469.2007.00580.x.

Semiparametric Inference for ROC Curves with Censoring

Hua Liang and

PhD, Associate Professor, Department of Biostatistics and Computational Biology, University of Rochester Medical Center, 601 Elmwood Avenue, Box 630, Rochester, New York 14642, USA
hliang@bst.rochester.edu

Yong Zhou

Shanghai University of Finance and Economics and Chinese Academy of Sciences

Abstract

Comparison of two samples can sometimes be conducted on the basis of analysis of ROC curves. A variety of methods of point estimation and confidence intervals for ROC curves have been proposed and studied well. We develop smoothed empirical likelihood based confidence intervals for ROC curves when the samples are censored and generated from semiparametric models. The resulting empirical log-likelihood function is shown to be asymptotically chi-squared. Simulation studies illustrate that the proposed empirical likelihood confidence interval is advantageous over the normal approximation based confidence interval. A real data set is analyzed using the proposed method.

Keywords

Estimating equation; confidence interval; coverage; Kaplan-Meier estimation; empirical likelihood ratio; empirical likelihood function

1 Introduction

Receiver operating characteristic (ROC) curve, a plot of false positive rate versus $1 -$ the false negative rate when the response of a test is continuous, is perhaps the most common measure to assess the accuracy of a test in diagnostic medicine. A diagnostic test works well if it has small false positive and false negative rates across a reasonable range of cut off values.

Conversely, a diagnostic test does not work well if the cut offs that make the false positive rate low have a high false negative rate and vice versa. A diagnostic test may not be flexible if its ROC curve follows a diagonal path from the lower left hand corner to the upper right hand corner. This means that every improvement in false positive rate is matched by a corresponding decline in the false negative rate. ROC curves can also be used to compare the diagnostic performance of two or more laboratory or diagnostic tests (Griner *et al.*, 1981).

Linnet (1987) developed parametric and non-parametric methods to construct confidence intervals for the sensitivity of a test at a fixed value of specificity. Platt *et al.* (2000) pointed out several shortcomings in Linnet's methods and suggested the use of Efron's bias-corrected acceleration (BCa) bootstrap interval. Zhou & Qin (2005) further showed that the BCa bootstrap interval can still has poor coverage accuracy in many circumstances, and proposed two new intervals for ROC curves.

Most approaches of study of ROC curves are limited to complete data regardless of parametric or nonparametric setup. In this paper, we are concerned with a situation that the observed data are censored, and one model is parametric and another one is nonparametric because additional information from either treatment or the best currently available treatment (i.e., control treatment) is available. For example, the goal is to compare a new treatment with a control treatment, for which we already have some auxiliary information. However, we may lack information about the new treatment. We are interested in deriving confidence intervals of ROC curves in this special framework. A more general but different situation was studied by Zhou & Liang (2005), in which the authors proposed a semiparametric inference for the parameters which can reflect the difference of two samples of censored data, established asymptotic distribution of the resulting statistics, and derived empirical likelihood-based confidence intervals of the parameters of interest under mild assumptions. A crucial assumption of Zhou & Liang (2005) is that the objective function is continuous on the nuisance parameter. This assumption excludes ROC curves. To remedy the deficiency of the discontinuity of the ROC curves, we propose a smoothed version of the ROC curves. The smoothing strategy has been widely used in literature. For example, Chen & Hall (1993) introduced smoothed empirical likelihood based confidence intervals for quantiles on one population. Zou *et al.* (1997), Lloyd (1998), and Ren *et al.* (2003) proposed various smoothed estimators for ROC curves. Zhou & Jing (2003) developed smoothed empirical likelihood confidence intervals for the difference of quantiles. Claeskens *et al.* (2003) proposed a smoothed empirical likelihood method for confidence intervals of ROC curves. The principle of our smoothed empirical likelihood is similar to that of Claeskens *et al.* (2003) in spirit.

Let X and Y be the responses of two samples, for example, the observations of diseased and non-diseased subjects, and let $F(\cdot)$ and $G(\cdot)$ be the distribution functions of X and Y , respectively. The ROC curve, a plot of $1 - F(t)$ versus $1 - G(t)$ for $t \in \mathbb{R}$, can be represented by $\Delta_p = 1 - F\{G^{-1}(1 - p)\}$ for $0 < p < 1$, where p is the false positive rate in $(0, 1)$ as the corresponding cutoff point ranges over $\mathbb{R}$, $G^{-1}(\cdot)$ is the inverse function of G , i.e., $G^{-1}(p) = \inf\{t : G(t) \geq p\}$. We are interested in the confidence intervals of the ROC curve, Δ_p , when F is nonparametric while G is parametric, and both X and Y are censored. We propose combining smoothing technique and the approaches developed by Zhou & Liang (2005) to derive an appropriate estimating equation because the naive estimating function is nondifferentiable and the corresponding estimating equation is inconsistent. See Zhou & Liang (2005) for a detailed discussion on this violation. This study is an extension of Zhou & Liang (2005). However, this extension is by no means trivial because (i) study of inference for ROC curves under semiparametric set-up is of independent interest, and (ii) the proofs of the theoretical results are difficult in contrast to the case discussed in Zhou & Liang (2005).

The paper is organized as follows. Section 2 presents a normal approximation based confidence interval and establishes the asymptotic result. Section 3 develops an empirical likelihood based confidence interval, and states the main results. Section 4 includes a simulation experiment that illustrates the numerical performance of the two methods. Section 5 analyzes real data by using the two methods. Section 6 contains a discussion. All technical details are presented in the Appendix.

2 Normal approximation based confidence interval

Assume that the variables $\{X_1^0, X_2^0, \dots, X_n^0\}$ and $\{Y_1^0, Y_2^0, \dots, Y_m^0\}$, two samples following X and Y , are randomly censored by two sequences of random variables $\{U_i, i = 1, \dots, n\}$ and $\{V_j, j = 1, \dots, m\}$, respectively. The elements of each sequence of random variables $\{X_i^0\}$, $\{Y_i^0\}$, $\{U_i\}$ and $\{V_i\}$ are independent and identically distributed with survival functions $S(x) = 1 - F(x)$, $D(y)$, $K(u)$, and $Q(v)$, respectively. $Q(v)$, $S(x)$ and $K(u)$ are unknown functions. The distribution function $G(y) = 1 - D(y)$ is parametric and depends on an unknown d -dimensional parameter

θ , i.e., $G(y) = G_\theta(y)$. Assume that $G_\theta(y)$ has the density function $g_\theta(y)$. We can only observe $\{(X_i, \delta_i), i = 1, 2, \dots, n\}$ and $\{(Y_j, \eta_j), j = 1, \dots, m\}$, where

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

, $i = 1, 2, \dots, n$, and $Y_j = \min(Y_j^0, V_j), \eta_j = I(Y_j^0 \leq V_j), j = 1, 2, \dots, m$, and $I(A)$ denotes the indicator function of the set A .

$$\psi(X, \theta, \Delta_p)$$

Write $\phi(x, \theta, \Delta_p) = 1 - I\{X \leq G_\theta^{-1}(1 - p)\} - \Delta_p$. Define $\phi(x, \theta, \Delta_p) = \int \psi(x, \theta - bw, \Delta_p) dL(w)$, where $b = b_n$ is a sequence of bandwidths with $b_n \rightarrow 0$ and $nb_n \rightarrow \infty$ as $n \rightarrow \infty$, and $L(\cdot)$ is a r th-order kernel function with finite support and satisfies:

$$\int u^j L(u) du = \begin{cases} 1, & \text{if } j=0, \\ 0, & \text{if } 1 \leq j \leq r-1, \\ c_r \neq 0, & \text{if } j=r, \end{cases}$$

for some integer $r \geq 2$.

Note that

$$\begin{aligned} E_r \phi(X, \theta_0, \Delta_p) &= E_r \left\{ \int \Psi(X, \theta_0 - bw, \Delta_p) dL(w) \right\} \\ &= \int \int \Psi(x, \theta_0 - bw, \Delta_p) dF(x) dL(w) \\ &= \int \int \Psi(x, u, \Delta_p) dF(x) dL\left(\frac{\theta_0 - u}{b}\right) \\ &\rightarrow E_r \Psi(X, \theta_0, \Delta_p) = 0, \end{aligned}$$

provided that $D(\theta) = E_F \Psi(X, \theta, \Delta_p)$ is continuous in the neighborhood of θ_0 . This means that $E_F \phi(X, \theta_0, \Delta_p)$ is asymptotically unbiased to $E_F \Psi(X, \theta_0, \Delta_p)$. Note that $\phi(x, \theta, \Delta_p)$ is continuous on θ , therefore the results obtained by Zhou & Liang (2005) are applicable to the smoothing version of ROC curves.

Let $\hat{\theta}_{MLE}$ be the maximum likelihood estimator based on the sample (Y_j, η_j) , $\hat{K}$ be the Kaplan-Meier estimator of K , i.e.,

$$\hat{K}(t) = \prod_{u \leq t} \left\{ 1 - \frac{dN^c(u)}{Y(u)} \right\},$$

$$N^c(u) = \sum_{i=1}^n I(X_i \leq u, \delta_i = 0) \text{ and } Y(u)$$

where $= \sum_{i=1}^n I(X_i \geq u)$. Our estimator of Δ_p is defined as the solution of the following estimating equation:

$$\sum_{i=1}^n \frac{\delta_i}{\hat{K}(X_i)} \phi(X_i, \hat{\theta}_{MLE}, \Delta_p) = 0. \quad (1)$$

A direct derivation yields

$$\widehat{\Delta}_p = \frac{\sum_{i=1}^n \frac{\delta_i}{\widehat{K}(X_i)} [1 - I\{X_i \leq G_{\theta_{MLE}}^{-1}(1-p)\}]}{\sum_{i=1}^n \frac{\delta_i}{\widehat{K}(X_i)}}.$$

Assume that $m/n \rightarrow \zeta > 0$ as $n, m \rightarrow \infty$. We need the following assumptions to establish the asymptotic normality for the estimator $\widehat{\Delta}_p$.

Assumptions

1. $\tau_S \leq \tau_K$, where $\tau_S = \sup\{x : S(x) > 0\}$ and $\tau_K = \sup\{x : K(x) > 0\}$, and

$$\int_0^{\tau_1} \frac{\Psi^2(u, \theta, \Delta)}{K(u)} dF(u) < \infty,$$

where τ_1 is the maximum of uncensored observations.

2. The distribution $F(x)$ is r times continuously differentiable.
3. The density function $g_\theta(y)$ is $r-1$ times differentiable with respect to θ on $A = \{y : g_\theta(y) > 0\}$, which is assumed to be independent of θ .
4. The information matrix $I(\theta)$ with the entries

$$I_{ij} = - \int_0^{\tau_2} \frac{\partial^2 \log g_\theta(y)}{\partial \theta_i \partial \theta_j} \{1 - Q(y)\} g_\theta(y) dy - \int_0^{\tau_2} \frac{\partial^2 \log \{1 - G_\theta(y)\}}{\partial \theta_i \partial \theta_j} \{1 - G_\theta(y)\} dQ(y)$$

($i, j = 1, \dots, d$), which are continuous, is positive definite. Here $\tau_2 = \sup\{t : D(t)Q(t) > 0\}$.

5. $nb \rightarrow \infty$, $nb^r \rightarrow 0$ for some r as $n \rightarrow \infty$. $L(\cdot)$ is continuously differentiable and $L'(x)$ is bounded.
6. $\tau_S < \tau_K$, and the $(r-1)$ th times derivation $g_\theta^{(r-1)}(t)$ of $g_\theta(t)$ with respect to θ is uniformly bounded for $t \in [1 - G(\tau_S), 1]$.

Theorem 1—Suppose that Assumptions 1–5 hold. Then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi(X_i, \widehat{\theta}_{MLE}, \Delta_p) \xrightarrow{L} N(0, \sigma_\psi^2),$$

$$W_{ni} = \delta_i / \widehat{K}(X_i), \sigma_\psi^2 = \sigma^2(\theta, \Delta_p) + \beta_0^T(\theta, \Delta_p) \sum^{-1} \beta_0(0, \Delta_p)$$

where $\sum = \zeta^{-1} I(\theta)$ with $I(\theta)$ given in Assumption 4,

$$\sigma^2(\theta, \Delta_p) = \Delta_p(1 - \Delta_p)$$

$$+ \Delta_p \int_0^{\tau_1} \{1 - \Delta_p/S(u)\} \lambda^c(u)/K(u) du, \beta_0(\theta, \Delta_p) = L'(0)(1 - \Delta_p) - F\{G_\theta^{-1}(1-p) \wedge \tau_1\},$$

and $\lambda^c(t)$ is the hazard function of U . As a consequence, (i) $\sqrt{n}(\widehat{\Delta}_p - \Delta_p) \xrightarrow{L} N(0, \sigma_\psi^2)$. (ii)

Furthermore, if Assumption 6 holds, then $\sqrt{n}(\widehat{\Delta}_p - \Delta_p) \Rightarrow W(p)$, a Gaussian process with mean zero and variance function $\sigma_\psi^2(p)$. Here “ $\Rightarrow$ ” denotes weak convergence on the functional space $\mathcal{D}$, which contains all functions that are right continuous and have left limits on $[1 - G(\tau_1), 1]$.

In a similar way to Zhou & Liang (2005), we can construct the normal approximation based confidence interval of Δ_p by substitution of $\widehat{\Delta}_p$, $\widehat{\theta}_{MLE}$, and $\widehat{K}$ in $\widehat{\sigma}_\psi^2$ and associated components, that is,

$$\widehat{\sigma}_\psi^2(\widehat{\theta})$$

$$, \widehat{\Delta}_p) = \widehat{\Delta}_p(1 - \widehat{\Delta}_p) + \widehat{\Delta}_p \int_0^{\tau_1} \frac{(1 - \widehat{\Delta}_p)}{\widehat{S}(u)\widehat{K}(u)} d\widehat{\Lambda}^c(u)$$

$$\widehat{\beta}_0(\widehat{\theta}, \widehat{\Delta}_p)$$

$$= L'(0)(1 - \widehat{\Delta}_p)$$

and $-\widehat{F}\{G_\theta^{-1}(1 - p) \wedge \tau_1\}$, where $\widehat{S} = 1 - \widehat{F}$, the Kaplan-Meier estimator of S ,

$$d\widehat{\Lambda}^c(u) = \frac{\sum_{i=1}^n I(X_i \leq u, \delta_i = 0) du}{\sum_{i=1}^n I(X_i \geq u)}.$$

As a result, the confidence interval is of form $(\widehat{\Delta}_p - z_{1-\alpha/2} \sqrt{\widehat{\sigma}_\psi^2/n}, \widehat{\Delta}_p + z_{1-\alpha/2} \sqrt{\widehat{\sigma}_\psi^2/n})$, where $z_{1-\alpha/2}$ denotes the $1 - \alpha/2$ quantile of the standard normal distribution.

3 Smoothed empirical likelihood

In the previous section, we establish the asymptotically normal theory of the estimators, the solutions of the estimating equation, in large samples. The inference based on the result of Theorem 1, in small or moderate sample sizes, may not be applicable or available because the covariance of the proposed estimator is difficult to get or complicated to implement. As a result the numerical performance may be poor. To avoid these deficiencies, we propose an empirical-likelihood-based method for inference of Δ_p . See Owen (2001) for a comprehensive survey of the empirical likelihood method, and Qin (1994, 1999) and Qin & Lawless (1994) for related topics. In this section, we develop an empirical likelihood based confidence interval for Δ_p by combining the smoothing technique and the method proposed by Zhou & Liang (2005).

Note that the likelihood function based on the samples $\{(X_i, \delta_i), i = 1, \dots, n\}$ and $\{(Y_j, \eta_j), j = 1, \dots, m\}$, denoted by $L(F, \theta)$, can be rewritten as

$$\prod_{i=1}^n \{F(X_i) - F(X_i-)\}^{\delta_i} \{1 - F(X_i)\}^{1-\delta_i} \prod_{j=1}^m g_\theta^{\eta_j}(Y_j) \{1 - G_\theta(Y_j)\}^{1-\eta_j}.$$

In a similar way to Zhou & Liang (2005), we define the log-likelihood ratio function as

$$\ell(\theta, \Delta_p) = - \sum_{i=1}^n \log\{1 + \lambda(\theta) W_{ni} \phi(X_i, \theta, \Delta_p)\} + \sum_{j=1}^m \log\{g_\theta^{\eta_j}(Y_j) \{1 - G_\theta(Y_j)\}^{1-\eta_j}\}, \quad (2)$$

where W_{ni} is given in Theorem 1, $\lambda(\theta)$ is subject to the following constraint

$$\frac{1}{n} \sum_{i=1}^n \frac{\phi(X_i, \theta, \Delta_p) W_{ni}}{1 + \lambda(\theta) W_{ni} \phi(X_i, \theta, \Delta_p)} = 0. \quad (3)$$

The following theorem indicates that the maximizer of (2) is a consistent estimator of the true value, and the estimated log-likelihood ratio is asymptotically weighted chi-squared.

Theorem 2

Suppose that Assumptions 1–5 in Section 2 hold. Then (i) the maximum estimator of $\ell(\theta, \Delta_p)$,

$$2\pi(\theta, \Delta_p) \ell(\hat{\theta}_{EL}, \Delta_p) \xrightarrow{\mathcal{L}} \chi_{(1)}^2$$

$\hat{\theta}_{EL}$, is strongly consistent, and (ii) $\chi_{(1)}^2$ is the standard chi-square distribution function with one degree of freedom, and

$$\pi(\theta, \Delta_p) = \frac{\sigma^2(\theta, \Delta_p) + \beta_0^T(\theta, \Delta_p) \Sigma^{-1} \beta_0(\theta, \Delta_p)}{\sigma_0^2(\theta, \Delta_p) + \beta_0^T(\theta, \Delta_p) \Sigma^{-1} \beta_0(\theta, \Delta_p)},$$

in which $\sigma_0^2(\theta, \Delta_p) = \int_0^{\tau_1} \psi^2(u, \theta, \Delta_p) / K(u) dF(u)$ and $\Sigma = \zeta^{-1} I(\theta)$ with $I(\theta)$ and $\sigma(\theta, \Delta_p)$ given in Theorem 1. Furthermore, if Assumption 6 holds, then

$$\sup_{1-G(\tau_1) < p < 1} 2\pi(\theta, \Delta_p) \ell(\hat{\theta}_{EL}, \Delta_p) \xrightarrow{\mathcal{L}} \chi_{(1)}^2.$$

After estimating $\pi(\theta, \Delta_p)$ by substituting θ and Δ_p by their estimators, we can derive an empirical-likelihood-based confidence interval on a basis of $\pi(\hat{\theta}_{EL}, \hat{\Delta}_p) \ell(\hat{\theta}_{EL}, \hat{\Delta}_p)$.

4 Simulation Study

In this section, we present simulation results to investigate the small and moderate sample size behavior of the CIs given in Sections 2 and 3. We mainly compare the lengths and the coverage probabilities of two CIs based on the empirical likelihood and the estimating equation methods.

Two iid sequences $\{X_i^0\}$ and $\{Y_i^0\}$ are generated from the exponential random variables $1/6 \exp(-x/6)$ and $1/10 \exp(-x/10)$, respectively. The censored random variables U and V follow the uniform distributions $U(0, c_x)$ and $U(0, c_y)$, respectively, in which c_x and c_y reflect the rates of censored data; the larger the c_x and c_y , the lighter the censorship. We consider the CIs of the ROC curve for a combination of $c_x = 12.5, 15, 20$, and 25 , and $c_y = 15$ and 20 at the points $p = 0.1, 0.2, \dots, 0.9$. For each of the 8 censoring configurations, sample sizes $(n, m) = (50, 50), (100, 100)$, and $(200, 200)$ are considered. At each censoring configuration and sample size, 1000 independent sets of data are generated. The CIs are based on the 95% level. We use the biweight kernel, $15/16(1-t^2)^2 I_{(|t| \leq 1)}$, and bandwidth $b = n^{-3/20}$ in the simulation study and real data analysis below. It is easy to see that Assumptions 1–5 are satisfied. Table 1 presents the results of this simulation study for $p = 0.4$ for an illustration. The coverage probabilities of the empirical-likelihood-based CIs are closer to the nominal level, especially when the sample size

is bigger, while the coverage probabilities of the normal-approximation-based CI are all larger than those of the empirical-likelihood method and the nominal level. The lengths of the CIs are inversely proportional to the sample size and censoring degree of sample Y , and the CI based on the empirical likelihood method is consistently narrower than its normal-approximation competitor. It is worthy pointing out that the values of lower confidence bound based on normal-approximation methods are sometimes negative, and this may result in a misleading conclusion.

5 Real Example

We apply the proposed method to analyze a dataset from a study of bone marrow transplantation (BMT) for leukemia. This study was designed in 1984 as a single institution (Ohio State University Hospitals, OSU) study and was modified in 1987 to include five institutions. One hundred thirty-seven patients (99 AML, 38 ALL) with acute myelocytic leukemia (AML) and acute lymphoblastic leukemia (ALL) were treated from March 1, 1984 through June 30, 1989 at the five separate centers with allogeneic BMT. More details of the study are referred to Copelan *et al.* (1991).

For an illustration, we only consider two groups of the ALL and AML low-risk patients. There are 38 (ALL) and 54 (AML) observations in these two groups, respectively. Let X and Y be the survival times of the ALL and AML low-risk patients. An individual was said to be disease free survival time at a given time after transplant if the individual was alive without the recurrence of leukemia. The event indicator for disease-free survival is if the patient was dead or relapsed. The time on study for the patient is the smaller number of relapse or death time. We check the exponential assumption of Y by calculating the KM estimate of the survival distribution of Y . A QQ-plot shows that the exponential distribution assumption for Y is fairly appropriate. The estimated values of the ROC curves for $p \in (0.2, 0.9)$, the 95% confidence regions of the ROC curves based on the normal approximation and empirical likelihood methods in Figure 1, in which the diagonal line describes $R(p) = p$, the dotted and dashed lines are the lines of the confidence bands based on the normal approximation method and the empirical likelihood method. The confidence band based on the normal approximation method almost contains the whole diagonal line, while the confidence band based on the empirical likelihood method rarely contains part of the diagonal line, and is enveloped by the normal approximation based confidence band. For a comparison, we also conduct a log-rank test on the data set and obtain a p -value to be 0.03, which indicates that the difference between these two groups are statistically significant, and the conclusion deduced from the empirical likelihood method is preferable.

6 Discussion

We developed the smoothed semiparametric empirical likelihood based confidence intervals for ROC curves when the observations are censored. The finite-sample behavior of the inference outperforms that of the common estimating equation, and shows promising. The method can easily be extended to the case where the objective function $\psi(x, \theta, \Delta_p)$ is a general discontinuous function of the nuisance parameter θ . Consequently, we can develop similar empirical likelihood inference for other situations such as quantile comparison when the observations are censored.

Acknowledgments

Liang's research was supported by NIH/NIAID grants AI62247 and AI59773. Zhou's research was partially supported by Funds of National Natural Science (No.10471140 and No. 10571169) of China. The authors thank the Editor, an associate editor and two referees for their constructive comments that substantially improved an earlier version of this paper.

References

- Chen SX, Hall P. Smoothed empirical likelihood confidence intervals for quantiles. *Ann. Statist* 1993;21:1166–1181.
- Claeskens G, Jing BY, Peng L, Zhou W. An empirical likelihood confidence interval for an ROC curve. *Can. J. Statist* 2003;31:173–190.
- Copelan EA, Biggs JC, Thompson JM, Crilley P, Szer J, Klein JP, Kapoor N, Avalos BR, Cunningham I, Atkinson K. Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with BuCy2. *Blood* 1991;78:838–843. [PubMed: 1859895]
- Gill, RD. Mathematical Centre Tracts. Censoring and Stochastic Integrals. Vol. 124. Amsterdam: Mathematisch Centrum; 1980.
- Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. *Ann. Internal Med* 1981;94:555–600.
- Linnet K. Comparison of quantitative diagnostic tests: type I error, power, and sample size. *Statist. Med* 1987;6:147–158.
- Owen, AB. Empirical likelihood. London: Chapman and Hall/CRC; 2001.
- Platt RW, Hanley JA, Yang H. Bootstrap confidence intervals for the sensitivity of a quantitative diagnostic test. *Statist. Med* 2000;19:313–322.
- Qin J. Semi-empirical likelihood ratio confidence intervals for the difference of two sample means. *Ann. Statist* 1994;46:117–126.
- Qin J. Empirical likelihood ratio based confidence intervals for mixture proportions. *Ann. Statist* 1999;27:1368–1384.
- Qin J, Lawless J. Empirical likelihood and general estimating equations. *Ann. Statist* 1994;22:300–325.
- Ren HB, Zhou XH, Liang H. A flexible method for estimating ROC curve. *J. Appl. Statist* 2004;31:773–784.
- Robins, J.; Rotnitzky, A. Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell, N.; Dietz, K.; Farewell, V., editors. *AIDS Epidemiology – Methodological Issues*. Boston: Birkhäuser; 1992. p. 297–331.
- Stute W. The central limit theorem under random censorship. *Ann. Statist* 1995;23:422–439.
- van der Varrrt, AW.; Wellner, JA. Weak convergence and empirical processes with application to statistics. New York: Springer-Verlag; 1996.
- Zhou W, Jing BY. Smoothed empirical likelihood confidence intervals for the difference of quantiles. *Statist. Sinica* 2003;13:83–95.
- Zhou XH, Qin GS. Improved confidence intervals for the sensitivity at a fixed level of specificity of a continuous-scale diagnostic test. *Statist. Med* 2005;24:465–477.
- Zhou Y, Liang H. Empirical-likelihood-based semiparametric inference for the treatment effect in the two-sample problem with censoring. *Biometrika* 2005;92:271–282. [PubMed: 20179776]
- Zou KH, Hall WJ, Shapiro DE. Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. *Statist. Med* 1997;16:2143–2156.

Appendix

The proofs of Theorems 1 and 2 can be finished following a technique similar to that used by Zhou & Liang (2005) for their main results. But most of the preliminary lemmas utilized by Zhou & Liang (2005) need to be proven in a different way. In this appendix, we first prove these preliminary lemmas. Lemmas A.1–A.3 will be used for the proof of Theorem 1, while the proof of Theorem 2 depends on Lemmas A.3 and A.5. The proof of Lemma A.5 requires Lemmas A.2–A.4.

Denote $\rho(\theta, \Delta) = E_F \{ \psi(X, \theta, \Delta) \}$, $\rho_1(\theta, \Delta) = E_F \{ \psi(X, \theta, \Delta)/K(X) \}$, and $\zeta(\theta_1, \theta_2) = E_F \{ \psi(X, \theta_1, \Delta)\psi(X, \theta_2, \Delta)/K(X) \}$.

Lemma A.1

Suppose that Assumptions 1–3 and 5 are satisfied. Then if $nb^{2r} \rightarrow 0$, as $n \rightarrow \infty$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi(X_i, \theta_0, \Delta) \xrightarrow{\mathcal{L}} N\{0, \sigma^2(\theta_0, \Delta)\},$$

where $\sigma^2(\theta_0, \Delta)$ is defined in Theorem 2, and this condition is also necessary if $\rho_{\theta}^{(r)}(\theta_0, \Delta) \neq 0$ for any Δ . Furthermore, if Assumption 6 is satisfied, then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi(X_i, \theta_0, \Delta) \Rightarrow W(p).$$

Proof

It follows from Corollary 1.2 of Stute (1995) that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} [\phi(X_i, \theta_0, \Delta) \\ - E\{\phi(X_i, \theta_0, \Delta)\}] \xrightarrow{\mathcal{L}} N\{0, \sigma^2(\theta_0, \Delta)\}. \end{aligned} \quad (4)$$

Let $\mathcal{F}(u)$ be the set of σ -algebras generated by $\{I(C_i \leq t), t \leq u; I(X_i \leq x), 0 \leq x < \infty, i = 1, 2, \dots, n\}$, which is the increasing information derived by all the survival times and cost histories throughout all time and the censoring times that are observed up to time u . It is convenient that we consider the counting process of the number of individuals censored over time instead of the number of deaths. The corresponding martingale of such a counting process for $\mathcal{F}(u)$ is

$$\mu_i^c(u) = N_i^c(u) - \int_0^u \lambda^c(t) Y_i(t) dt$$

defined by $\begin{aligned} & \text{, where } N_i^c(u) = I(X_i \leq u, \delta_i = 0) \\ & \text{, } Y_i(u) = I(X_i \geq u) \end{aligned}$. Note that $E\mu_i^c(u) = 0$.

Let $\mu^c(u) = \sum \mu_i^c(u)$. It follows from Robins & Rotnitzky's (1992, P.313) result that

$$\frac{\delta_i}{K(X_i)} = 1 - \int_0^\infty \frac{d\mu_i^c(u)}{K(u)}.$$

By the convergence of martingale, we have

$$\begin{aligned} n^{-1} \sum_{i=1}^n W_{ni} \\ = 1 - \frac{1}{n} \int_0^\infty \frac{d\mu^c(u)}{K(u)} = 1 + o_p(1). \end{aligned}$$

It follows from the representation of Kaplan-Meier estimator (Gill; 1980, P.37) that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi_c(X_i, \theta_0, \Delta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \phi_c(X_i, \theta_0, \Delta) \\ &- \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^{\tau_1} \frac{d\mu_i^c(u)}{K(u)} \{ \phi_c(X_i, \theta_0, \Delta) - G(u, \theta_0, \Delta) \} + o_p(1), \end{aligned} \quad (5)$$

where $\phi_c(X_i, \theta_0, \Delta) = \phi(X_i, \theta_0, \Delta) - E\{\phi(X_i, \theta_0, \Delta)\}$, a centered functional sequence, and $G(u, \theta_0, \Delta) = E\{\phi_c(X_i, \theta_0, \Delta)I(T \geq u)\}/S(u)$.

Write $\mathcal{F}_n = \{\phi_c(\cdot, \theta, \Delta_p), p \in [1 - G(\tau_1), 1]\}$. The class $\mathcal{F}_n$ has an envelope function $2V(L)$, where $V(L)$ is the total variation boundary of L , and then $\mathcal{F}_n$ is a VC class. Using Theorem 2.11.2 of van der Vaart and Wellner (1996), we can prove that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} [\phi(X_i, \theta_0, \Delta) \\ - E\{\phi(X_i, \theta_0, \Delta)\}] \Rightarrow W(p). \end{aligned} \quad (6)$$

By the assumption on $L(\cdot)$ and a direct calculation, we know

$$\begin{aligned} E\{\phi(X, \theta_0, \Delta)\} \\ = b^r \rho^{(r)}(\theta_0, \Delta) \int_{-\infty}^{+\infty} u^r l(u) du \{1 + o(1)\}. \end{aligned} \quad (7)$$

Therefore

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} E\{\phi(X_i, \theta_0, \Delta)\} \\ = n^{1/2} b^r \rho^{(r)}(\theta_0, \Delta) \int_{-\infty}^{+\infty} u^r l(u) du \{1 + o(1)\}. \end{aligned} \quad (8)$$

Combining (4) and (8), we complete the proof of Lemma A.1.

Lemma A.2

Suppose that Assumptions 1–3 and 5 are satisfied. Then

$$\begin{aligned} E \left\{ \frac{1}{K(X)} \phi^2(X_i, \theta, \Delta) \right\} \\ = \sigma_0^2(\theta, \Delta) + O(b^r), \end{aligned}$$

where $\sigma_0^2(\theta, \Delta) = E\{\psi^2(X, \theta, \Delta)/K(X)\}$.

Proof

Note that $E\{\phi^2(X_i, \theta, \Delta)/K(X)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \zeta(\theta - bw, \theta - bu, \Delta) dL(w) dL(u)$. By the definition of $\zeta(\theta_1, \theta_2, \Delta)$, we have

$$\begin{aligned} & \int \int \zeta(\theta - bu, \theta - bv, \Delta) dL(u) dL(v) \\ &= \int \int \left[\int \frac{1}{K(x)} I\{x > G_{\theta - bu}^{-1}(1 - p)\} I\{x > G_{\theta - bv}^{-1}(1 - p)\} dF(x) \right] dL(u) dL(v) \\ & \quad - 2 \int \int \left[\int \frac{\Delta}{K(x)} I\{x > G_{\theta - bu}^{-1}(1 - p)\} dF(x) \right] dL(u) dL(v) \\ & \quad - \int \int \left\{ \int \frac{\Delta}{K(x)} dF(x) \right\} dL(u) dL(v) \stackrel{\text{def}}{=} I_1 - 2I_2 + I_3. \end{aligned} \quad (9)$$

Write

$$\begin{aligned} \xi(\theta) &= E \left[\frac{1}{K(X)} I\{X > G_{\theta}^{-1}(1 - p)\} \right] \\ &= \int_{G_{\theta}^{-1}(1-p)}^{\tau_F} \frac{1}{K(x)} dF(x). \end{aligned}$$

Note that $\xi(\theta)$ is r times differentiable with respect to θ by Assumptions 1–3. It follows that

$$I_2 = \int \int \xi(\theta - bu) dL(u) dL(v) \Delta = \xi(\theta) \Delta + O(b^r). \quad (10)$$

A direct calculation yields

$$I_1 = \int \int \xi(\theta - bu) \vee \xi(\theta - bv) dL(u) dL(v) = \xi(\theta) + O(b^r), \quad (11)$$

where and below $a \vee b = \max(a, b)$. Combining (9), (10) with (11), we have

$$\begin{aligned} & E \left\{ \frac{1}{K(X-)} \phi^2(X_i, \theta, \Delta) \right\} \\ &= E \{ \psi^2(X \\ & \quad , \theta, \Delta) / K(X) \} + O(b^r). \end{aligned}$$

This completes the proof of Lemma A.2.

Lemma A.3

Suppose that Assumptions 1, 2, and 5 hold. Then

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_i-)} \right\}^2 = \sigma_0^2(\theta, \Delta) + o_p(1), \\ & \frac{1}{n} \sum_{i=1}^n \frac{\dot{\phi}_{\theta}(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_i-)} = \beta_0(\theta, \Delta) + o_p(1), \text{ and} \\ & \frac{1}{n} \sum_{i=1}^n \frac{\dot{\phi}_{\Delta}(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_i-)} = \gamma(\theta, \Delta) + o_p(1), \end{aligned}$$

where $\sigma_0^2(\theta, \Delta)$ was given in Lemma A.2, $\beta_0(\theta, \Delta) = \dot{\beta}(\theta, \Delta) = L'(0)(1-\Delta_p) - F\{G^{-1}(1-p) \vee \tau_1\}$ and $\gamma(\theta, \Delta) = \int_0^{\tau_1} \dot{\psi}_\Delta(x, \theta, \Delta) dx$.

Proof

We only prove the first part, because similar arguments can be used to derive the second and third formulae. At first, we show that

$$\frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_{i-})} \right\}^2 = O_p(1). \quad (12)$$

Note that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_{i-})} \right\}^2 &\leq \frac{2}{n} \sum_{i=1}^n \left[\frac{\{\phi(X_i, \theta, \Delta) - E\phi(X_i, \theta, \Delta)\} \delta_i}{K(X_{i-})} \right]^2 \\ &\quad + \frac{4}{n} \sum_{i=1}^n \frac{\{\phi(X_i, \theta, \Delta) - E\phi(X_i, \theta, \Delta)\} \{E\phi(X_i, \theta, \Delta)\} \delta_i}{K(X_{i-})} \\ &\quad + \frac{2}{n} \sum_{i=1}^n \left[\frac{E\phi(X_i, \theta, \Delta) \delta_i}{K(X_{i-})} \right]^2 \\ &\quad + \frac{2}{n} \sum_{i=1}^n \left\{ \frac{K(X_{i-}) - \widehat{K}(X_{i-})}{\widehat{K}(X_{i-}) K(X_{i-})} \right\}^2 \phi^2(X_i, \theta, \Delta) \delta_i^2 \\ &\stackrel{\text{def}}{=} J_1 + J_2 + J_3 + J_4. \end{aligned}$$

By the law of large numbers, we have

$$J_1 = \left\{ \frac{\phi^2(X_i, \theta, \Delta) - 2\phi(X_i, \theta, \Delta)E\phi(X_i, \theta, \Delta) + E^2\phi(X_i, \theta, \Delta)}{K(X_{i-})} \right\} + o_p(1).$$

It follows from Lemma A.2 and (7) that

$$J_1 \rightarrow 2\sigma_0^2(\theta, \Delta) < \infty. \quad (13)$$

Using the law of large numbers again, we have

$$\begin{aligned} J_2 &= 4E[\phi(X, \theta, \Delta) - E\phi(X, \theta, \Delta)]E\phi(X, \theta, \Delta) + o_p(1) = o_p(1). \\ J_3 &= E\left\{\frac{1}{K(X)}\right\}\{E\psi(X, \theta, \Delta)\}^2 + o_p(1) + O(nb^r) = o_p(1) + O(nb^r). \end{aligned}$$

On the other hand,

$$J_4 \leq \sup_{0 \leq x \leq X_{(n)}} \left| \frac{K(x) - \widehat{K}(x)}{K(x)} \right|^2 \frac{2}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_{i-})} \right\}^2.$$

The first term is $o_p(1)$ by Gill (1980, P.37). The argument similar to that for (13) and a direct calculation yield $J_4 = o_p(1)$. We complete the proof of (12).

By using an argument similar to that of (12),

$$\begin{aligned}
 \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{\widehat{K}(X_{i-})} \right\}^2 &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi(X_i, \theta, \Delta) \delta_i}{K(X_{i-})} \right\}^2 \\
 &\quad + \frac{1}{n} \sum_{i=1}^n \left\{ \frac{K(X_{i-}) - \widehat{K}(X_{i-})}{\widehat{K}(X_{i-}) K(X_{i-})} \right\}^2 \phi^2(X_i, \theta, \Delta) \delta_i^2 \\
 &\quad + \frac{2}{n} \sum_{i=1}^n \frac{K(X_{i-}) - \widehat{K}(X_{i-})}{K^2(X_{i-}) - \widehat{K}^2(X_{i-})} \phi^2(X_i, \theta, \Delta) \delta_i^2 \\
 &\stackrel{\text{def}}{=} M_1 + M_2 + M_3.
 \end{aligned} \tag{14}$$

It follows from the same argument as for (13) that $M_1 = \sigma_0^2(\theta, \Delta) + o_p(1)$. By the similar proof of J_4 and (12), we have $M_2 = o_p(1)$ and $M_3 = o_p(1)$. Combining these statements, we complete the proof of the first result of Lemma A.3.

Lemma A.4

Suppose that Assumptions 1–5 hold. Then we have

$$\begin{aligned}
 \frac{1}{n} \sum_{i=1}^n W_{ni} \phi(X_i, \theta_0, \Delta) \\
 &= \frac{1}{n} \sum_{i=1}^n W_{ni} \phi(X_i, \theta_0, \Delta) \\
 &\quad + O_p(n^{-\rho})
 \end{aligned}$$

for any $\theta \in \{\theta : \|\theta - \theta_0\| \leq cn^{-\rho}\}$, where $0 < \rho \leq 1/2$.

Proof

Consider the following quantity,

$$J = \frac{1}{n} \sum_{i=1}^n \{W_{ni} \phi(X_i, \theta_0, \Delta) - W_{ni} \phi(X_i, \theta, \Delta)\} \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n Z_i.$$

Note that

$$EZ_i = \int \{\rho(\theta - bw, \Delta) - \rho(\theta_0 - bw, \Delta)\} dL(w) = O(|\theta - \theta_0|) = O(n^{-\rho}). \tag{15}$$

By similar arguments to the proof of Lemma A.2, we obtain

$$\begin{aligned}
 EZ_i^2 &= E \left[\frac{1}{K(X_i)} \left[\int \{\psi(X_i, \theta - bw, \Delta) - \psi(X_i, \theta_0 - bw, \Delta)\} dL(w) \right]^2 \right] \\
 &\leq E \left[\frac{1}{K(X_i)} \int \{\psi(X_i, \theta - bw, \Delta) - \psi(X_i, \theta_0 - bw, \Delta)\}^2 dL(w) \right] \\
 &= E\{\psi^2(X, \theta, \Delta)/K(X)\} + E\{\psi^2(X, \theta_0, \Delta)/K(X)\} \\
 &\quad - 2E\{\psi(X, \theta, \Delta)\psi(X, \theta_0, \Delta)/K(X)\} + O(b^r).
 \end{aligned} \tag{16}$$

Note that

$$\begin{aligned}
& E \left[\frac{\psi(X, \theta, \Delta) \{\psi(X, \theta, \Delta) - \phi(X, \theta, \Delta)\}}{K(X)} \right] \\
&= E \left[\frac{I\{X \geq G_{\theta}^{-1}(1-p)\}}{K(X)} I\{G_{\theta}^{-1}(1-p) \leq X \leq G_{\theta_0}^{-1}(1-p)\} \right] \\
&\quad - \Delta E \left[\frac{1}{K(X)} I\{G_{\theta}^{-1}(1-p) \leq X \leq G_{\theta_0}^{-1}(1-p)\} \right] \\
&= (1-\Delta) \int_{G_{\theta}^{-1}(1-p)}^{G_{\theta_0}^{-1}(1-p)} \frac{1}{K(x)} dF(x) \\
&= O(|G_{\theta}^{-1}(1-p) - G_{\theta_0}^{-1}(1-p)|) = O(n^{-\rho}).
\end{aligned} \tag{17}$$

It follows from (15) and (17) that $\text{var}(Z_i) = O(n^{-\rho} + b^r)$.

$$\begin{aligned}
& \text{Note that } \frac{1}{\sqrt{n}} \sum_{i=1}^n (Z_i - EZ_i) = O_p(\{\text{var}(Z_i)\}^{1/2}). \text{ Hence,} \\
& J = \frac{1}{n} \sum_{i=1}^n Z_i = O_p(n^{-\rho} + n^{-1/2}(n^{-\rho/2} + b^{r/2})) \\
& \quad = O_p(n^{-\rho} + (n^{-1}b^r)^{1/2}).
\end{aligned}$$

Lemma A.5

Suppose that Assumptions 1–4 hold. Then we have $\lambda(\theta) = O_p(n^{-\varrho} + b^r)$ uniformly on $\{\theta : \|\theta - \theta_0\| \leq cn^{-\varrho}\}$, where $1/3 < \varrho < 1/2$, c is some constant, and

$$\lambda(\theta) = \frac{\sum_{i=1}^n \phi(X_i, \theta, \Delta) W_{ni}}{\sum_{i=1}^n \{\phi(X_i, \theta, \Delta) W_{ni}\}^2} + O_p(n^{-\varrho} + b^r)$$

uniformly on $\{\theta : \|\theta - \theta_0\| \leq cn^{-\varrho}\}$, and $\lambda(\theta)$ satisfies (2).

Proof

It follows from (2) and Taylor expansion that

$$\begin{aligned}
0 &= \frac{1}{n} \sum_{i=1}^n \frac{\phi(X_i, \theta, \Delta) W_{ni}}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}} \\
&= \frac{1}{n} \sum_{i=1}^n \left[\{1 - \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}\} \phi(X_i, \theta, \Delta) W_{ni} + \frac{\lambda^2(\theta) \phi^3(X_i, \theta, \Delta) W_{ni}^3}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}} \right].
\end{aligned}$$

$$\lambda(\theta) = \frac{\sum_{i=1}^n \phi(X_i, \theta, \Delta) W_{ni}}{\sum_{i=1}^n \phi^2(X_i, \theta, \Delta) W_{ni}^2} + R_n, \text{ where}$$

This means that

$$R_n = \lambda^2(\theta) \left\{ \sum_{i=1}^n \frac{\phi^3(X_i, \theta, \Delta) W_{ni}^3}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}} \right\} \left\{ \sum_{i=1}^n \phi^2(X_i, \theta, \Delta) W_{ni}^2 \right\}^{-1}.$$

We separately deal with the numerator and denominator of the first term in $\lambda(\theta)$. We will assume $\theta \in \{\|\theta - \theta_0\| \leq n^{-\varrho}\}$ below without specification. Note that

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \frac{\phi(X_i, \theta, \Delta) W_{ni}}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}} \\ &= \frac{1}{n} \sum_{i=1}^n \left\{ 1 - \frac{\lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}} \right\} \phi(X_i, \theta, \Delta) W_{ni}. \end{aligned}$$

It follows from Lemma A.4 that

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \phi(X_i, \theta_0, \Delta) W_{ni} + O_p(n^{-\varrho}) \\ &= \frac{\lambda(\theta)}{n} \sum_{i=1}^n \frac{\phi^2(X_i, \theta, \Delta) W_{ni}^2}{1 + \lambda(\theta) \phi(X_i, \theta, \Delta) W_{ni}}. \end{aligned} \quad (18)$$

By Assumption 2 and a similar proof to Lemma 3 of Owen (1990), we obtain

$$\max_{1 \leq i \leq n} \left| \frac{\phi(X_i, \theta, \Delta) \delta_i}{K(X_i -)} \right| = o_p(n^{1/3}). \quad (19)$$

An argument similar to that for I_3 in Lemma A.3 yields

$$\max_{1 \leq i \leq n} |\phi(X_i, \theta, \Delta) W_{ni}| = o_p(n^{1/3}). \quad (20)$$

It follows from (18) and (20) that

$$\frac{1}{n} \sum_{i=1}^n \phi(X_i, \theta, \Delta) W_{ni} + O_p(n^{-\varrho}) \geq \frac{\lambda(\theta)}{1 + |\lambda(\theta)| Z_{(n)}} \frac{1}{n} \sum_{i=1}^n \phi^2(X_i, \theta, \Delta) W_{ni}^2, \quad (21)$$

where $Z_{(n)} = \max_{1 \leq i \leq n} |\phi(X_i, \theta, \Delta) W_{ni}|$. By using arguments similar to Lemma A.3, we have

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \phi^2(X_i, \theta_0, \Delta) W_{ni}^2 \\ &= O_p(1), \text{ and } \frac{1}{n} \sum_{i=1}^n \phi(X_i, \theta_0, \Delta) W_{ni} \\ &= O_p(n^{-1/2} + b^r); \end{aligned}$$

therefore, $\lambda(\theta) = O_p(n^{-\varrho} + b^r)$. This statement along with Lemma A.3 and (19) yields $R_n \leq O_p(\{n^{-\varrho} + b^r\}^2) \max_{1 \leq i \leq n} |\phi(X_i, \theta, \Delta) W_{ni}| = o_p(n^{-\varrho} + b^r)$ uniformly on $\{\theta : \|\theta - \theta_0\| < cn^{-\varrho} + b^r\}$. This completes the proof of Lemma A.5.

We now finish the proofs of Theorems 1 and 2. Only the key steps are given here. Because the results of Lemmas A.1–A.5 uniformly hold under Assumptions 1–6, the uniform convergence in Theorems 1 and 2 can be proved in the same way as for the point convergence. We present only the proofs for point convergence, but omit the details for uniform convergence.

Proof of Theorem 1—By Assumption 5, we have

$$\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi(X_i, \widehat{\theta}_{\text{MLE}}, \Delta) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n W_{ni} \phi(X_i, \theta, \Delta) \\
&+ \beta_0(\theta, \Delta) \sqrt{n}(\widehat{\theta}_{\text{MLE}} - \theta) + o_p(1).
\end{aligned}$$

By Lemmas A.1 and A.3, we complete the proof of Theorem 1.

Proof of Theorem 2—The result in (i) is a direct consequence of Lemma A.5. We now prove result (ii). By using an argument similar to the proof of Lemma A.4 of Zhou & Liang (2005) and the Taylor expansion, we can show that

$$\begin{aligned}
-2\mathcal{L}(\widehat{\theta}_{\text{EL}}, \Delta) &= -2\widehat{\lambda} \sum_{i=1}^n \phi(X_i, \widehat{\theta}_{\text{EL}}, \Delta) W_{ni} + \widehat{\lambda}^2 \sum_{i=1}^n \{\phi(X_i, \widehat{\theta}_{\text{EL}}, \Delta) W_{ni}\}^2 \\
&- 2 \frac{\partial \ell_s^T(\widehat{\theta}_{\text{MLE}})}{\partial \theta} (\widehat{\theta}_{\text{EL}} - \widehat{\theta}_{\text{MLE}}) + (\widehat{\theta}_{\text{EL}} - \widehat{\theta}_{\text{MLE}})^T \frac{\partial^2 \ell_s(\widehat{\theta}_{\text{MLE}})}{\partial \theta \partial \theta^T} (\widehat{\theta}_{\text{EL}} - \widehat{\theta}_{\text{MLE}}) + o_p(1).
\end{aligned}$$

(3) and Lemma A.5 imply that

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n \phi(X_i, \widehat{\theta}_{\text{EL}}, \Delta) W_{ni} \\
&= \frac{\widehat{\lambda}}{n} \sum_{i=1}^n \{\phi(X_i, \widehat{\theta}_{\text{EL}}, \Delta) W_{ni}\}^2 \\
&+ o_p(n^{-\varrho}).
\end{aligned} \tag{22}$$

These statements together the arguments similar to the proof of Theorem 2 of Zhou & Liang (2005) yield

$$\begin{aligned}
& 2\mathcal{L}(\widehat{\theta}_{\text{EL}}, \Delta) = (\sqrt{n\widehat{\lambda}})^2 \left[\frac{1}{n} \sum_{i=1}^n \{\phi(X_i, \widehat{\theta}_{\text{EL}}, \Delta) W_{ni}\}^2 + \beta_0^T \sum^{-1} \beta_0 \right] + o_p(1).
\end{aligned}$$

Lemmas A.3 and Lemma A.6 of Zhou & Liang (2005) indicate that

$$2\mathcal{L}(\widehat{\theta}_{\text{EL}}, \Delta) \xrightarrow{\mathcal{L}} \frac{\sigma_0^2(\theta, \Delta) + \beta_0^T \Sigma^{-1} \beta_0}{\sigma^2(\theta, \Delta) + \beta_0^T \Sigma^{-1} \beta_0} \chi_{(1)}^2.$$

This completes the proof of Theorem 2.

Figure 1.

The confidence regions of the ROC curve (solid line) for the data from the BMT study. The dotted lines: the CI based on empirical likelihood method; the broken lines: the CI based on normal approximation method; and the dash-dotted line: $R(p) = p$.

Table 1

Confidence intervals (CI) of the ROC curves at $p = 0.4$, and the corresponding coverage probabilities (CP) based on the empirical likelihood (EL) and normal approximation (Norm) methods for the simulated data.

Sample size			c_x		c_y	EL		CI(95%)		CP	
n	m		c_x		c_y	EL		Norm		EL	Norm
50	50	12.5	12.5		20	(0.047, 0.413)		(-0.265, 0.711)		94.55	99.45
						(0.013, 0.375)		(-0.160, 0.533)		95.55	99.75
						(0.082, 0.409)		(-0.088, 0.584)		91.60	98.15
						(0.054, 0.383)		(-0.187, 0.629)		94.15	99.25
						(0.129, 0.425)		(-0.040, 0.612)		91.85	93.65
100	100	15	20		25	(0.096, 0.391)		(-0.075, 0.579)		93.95	97.15
						(0.159, 0.438)		(-0.026, 0.644)		94.40	98.95
						(0.125, 0.403)		(-0.070, 0.619)		92.05	94.45
						(0.086, 0.345)		(-0.060, 0.488)		95.80	99.85
						(0.057, 0.314)		(-0.080, 0.448)		95.25	99.65
200	200	12.5	15		20	(0.121, 0.356)		(-0.010, 0.493)		92.15	98.60
						(0.092, 0.325)		(-0.034, 0.457)		94.50	99.45
						(0.167, 0.378)		(0.040, 0.514)		93.35	93.25
						(0.137, 0.348)		(-0.002, 0.498)		93.10	96.95
						(0.195, 0.395)		(0.063, 0.539)		93.50	94.25
	200	25	20		25	(0.165, 0.365)		(0.037, 0.506)		92.75	96.10
						(0.118, 0.302)		(0.018, 0.403)		96.35	99.80
						(0.086, 0.269)		(-0.012, 0.369)		95.20	99.20
						(0.152, 0.319)		(0.070, 0.406)		94.62	97.90
						(0.124, 0.290)		(0.027, 0.391)		94.05	99.50
	200	15	20		25	(0.196, 0.347)		(0.099, 0.450)		94.21	97.50
						(0.166, 0.335)		(0.088, 0.401)		95.90	96.80
						(0.154, 0.366)		(0.140, 0.457)		94.19	97.70
						(0.191, 0.334)		(0.101, 0.43)		93.85	94.32