Journal of Multivariate Analysis 101 (2010) 1520–1531

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Smoothed jackknife empirical likelihood method for ROC curve

Yun Gong [a], Liang Peng [a,*], Yongcheng Qi [b]

[a] School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA

[b] Department of Mathematics and Statistics, University of Minnesota-Duluth, 1117 University Drive, Duluth, MN 55812, USA

ARTICLE INFO

Article history:
Received 25 September 2009
Available online 1 February 2010

AMS subject classifications:
62G15

Keywords:
Confidence interval
Empirical likelihood
Jackknife
ROC curve

ABSTRACT

In this paper we propose a smoothed jackknife empirical likelihood method to construct confidence intervals for the receiver operating characteristic (ROC) curve. By applying the standard empirical likelihood method for a mean to the jackknife sample, the empirical likelihood ratio statistic can be calculated by simply solving a single equation. Therefore, this procedure is easy to implement. Wilks' theorem for the empirical likelihood ratio statistic is proved and a simulation study is conducted to compare the performance of the proposed method with other methods.

1. Introduction

In diagnostic medicine, it is important to assess the accuracy of a diagnostic test in discriminating diseased patients from non-diseased ones. When the response of a test is continuous, its accuracy is measured by the receiver operating characteristic (ROC) curve; see, e.g., [1,2]. ROC curves can also be used to compare the diagnostic performance of two or more laboratory or diagnostic tests [3].

Let F and G be the distribution functions of the diseased and non-diseased populations, respectively. Then the ROC curve can be written as $R(t) = 1 - F(G^{-}(1-t))$ for $0 < t < 1$, where G^{-} denotes the inverse of G and is defined by $G^{-}(u) = \inf\{x : G(x) \geq u\}$ for $u \in (0, 1)$.

Throughout we assume that $X_1, \ldots, X_m$ are independent and identically distributed (i.i.d.) test responses of m patients from the diseased population with distribution F and $Y_1, \ldots, Y_n$ are i.i.d. test responses of n patients from the non-diseased population with distribution G. A simple estimator of $R(t)$ is defined as

$$R_{m,n}(t) = 1 - F_m(G_n^{-}(1-t)), \tag{1}$$

where F_m and G_n are the empirical distribution functions of F and G given by

$$F_m(x) = \frac{1}{m}\sum_{j=1}^{m} I(X_j \leq x), \quad G_n(y) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \leq y).$$

For the study of the estimator $R_{m,n}(t)$ and its smooth version, we refer to [4–9]. For some inference problems related to the ROC curve see, e.g., [10,11].

Using the fact that

$$\sqrt{m+n}\{R_{m,n}(t) - R(t)\} \xrightarrow{d} N(0, \sigma^2(t)), \tag{2}$$

* Corresponding author.
E-mail addresses: ygong@math.gatech.edu (Y. Gong), peng@math.gatech.edu (L. Peng), yqi@d.umn.edu (Y. Qi).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2010.01.012

where

$$\sigma^2(t) = \left(1 + \frac{1}{r}\right) R(t)(1 - R(t)) + (1 + r)t(1 - t)\left\{\frac{F'(G^-(1-t))}{G'(G^-(1-t))}\right\}^2, \tag{3}$$

and $r := \lim_{m,n\to\infty} m/n \in (0, \infty)$, one can construct a confidence interval for $R(t)$ via estimating the density functions of F and G or bootstrap methods. As an alternative way to construct confidence intervals without estimating the asymptotic variance explicitly, Claeskens et al. [12] proposed an empirical likelihood method based on the smoothing estimators of the functions F and G via some link variable. Molanes-Lopez, Van Keilegom and Veraverbeke [13] studied the empirical likelihood method based on empirical estimators. Qin and Zhou [14] employed the empirical likelihood method to construct confidence intervals for the area under the ROC curve.

The empirical likelihood, introduced in [15,16], is a well-known nonparametric method for constructing confidence regions. Like the bootstrap and the jackknife, the empirical likelihood method does not assume a parametric family of distributions for the data. One of the advantages of the empirical likelihood method is that it enables the shape of a region, such as the degree of asymmetry in a confidence interval, to be determined automatically by the sample. We refer to [17] for overviews. Some recent developments of empirical likelihood methods include inferences for: censored median regression model [18,19], two-sample problems [20–25], time series models [26–31], longitudinal data and single-index models [32–35] and Copula [36]. However, all these applications and extensions of empirical likelihood methods work under linear constraints. In case of nonlinear functionals such as variance, ROC curves and copulas, a common way is to transform nonlinear constraints to linear constraints by introducing some link variables as in [12,36]. Unfortunately, this method does not always work and the introduced link variables create more linear constraints, which increases the computational burden. Seeking a general method to deal with nonlinear functionals becomes important.

Recently, Jing, Yuan and Zhou [37] proposed a so-called jackknife empirical likelihood method for a U-statistic. The procedure is as follows. For a U-statistic, construct a jackknife sample (see, e.g., [38]) first, and then treat this jackknife pseudo-sample as a sample of i.i.d. observations and apply the standard empirical likelihood method for the mean of i.i.d. observations to obtain the empirical likelihood ratio statistic for the U statistic. Hence, the procedure is easy to implement.

In this paper, we study the possibility of extending the jackknife empirical likelihood method in [37] to construct confidence intervals for the ROC curve so as to avoid adding extra constraints due to the link variable in [12]. It turns out that we have to work with a smooth version of the empirical estimator of the ROC curve. We organize this paper as follows. Section 2 gives the detailed methodology and main results. A simulation study is presented in Section 3. All proofs are put in Section 4.

2. Methodology

Let w be a symmetric density function with support $[-1, 1]$ and put $K(x) = \int_{-\infty}^{x} w(y)dy$. Define the smooth version of $R_{m,n}(t)$ as

$$\hat{R}_{m,n}(t) = 1 - \frac{1}{m}\sum_{j=1}^{m} K\left(\frac{1 - t - G_n(X_j)}{h}\right),$$

where $h = h(n) > 0$ is a bandwidth. In fact, this smooth estimator of R is obtained via replacing F_m in (1) by its smoothed version and G_n is still the empirical distribution of G. Thus, this smoothed estimator of the ROC curve R is different from the one in [12]. The reason why we have to work with a smooth version is given in Remark 1 below. Define

$$\hat{R}_{m,n,i}(t) = 1 - \frac{1}{m-1}\sum_{1\le j\le m, j\ne i} K\left(\frac{1 - t - G_n(X_j)}{h}\right), \quad 1 \le i \le m,$$

$$\hat{R}_{m,n,i}(t) = 1 - \frac{1}{m}\sum_{j=1}^{m} K\left(\frac{1 - t - G_{n,i-m}(X_j)}{h}\right), \quad m < i \le m + n,$$

where

$$G_{n,k}(y) = \frac{1}{n-1}\sum_{1\le i\le n, i\ne k} I(Y_i \le y), \quad k = 1, \ldots, n.$$

The jackknife pseudo-sample is therefore defined as

$$\hat{V}_i(t) = (m + n)\hat{R}_{m,n}(t) - (m + n - 1)\hat{R}_{m,n,i}(t), \quad i = 1, \ldots, m + n.$$

Next, we form the empirical likelihood at $R(t) = \theta$ based on the jackknife pseudo-sample as

$$L_{m,n}(t, \theta) = \sup\left\{\prod_{i=1}^{m+n} p_i : p_1 > 0, \ldots, p_{m+n} > 0, \sum_{i=1}^{m+n} p_i = 1, \sum_{i=1}^{m+n} p_i\hat{V}_i(t) = \theta\right\}.$$

By the standard Lagrange multiplier argument, we obtain that the above maximization is achieved at

$$p_i = \frac{1}{(m+n)\{1+\lambda(\hat{V}_i(t)-\theta)\}}, \quad i = 1, \ldots, m+n,$$

where $\lambda = \lambda(t, \theta)$ satisfies

$$\frac{1}{m+n}\sum_{i=1}^{m+n}\frac{\hat{V}_i(t)-\theta}{1+\lambda(\hat{V}_i(t)-\theta)} = 0,$$

which gives the log empirical likelihood ratio as

$$l_{m,n}(t,\theta) = -2\log L_{m,n}(t,\theta) = 2\sum_{i=1}^{m+n}\log\{1+\lambda(\hat{V}_i(t)-\theta)\}.$$

In order to show that the above log empirical likelihood ratio converges in distribution to a χ^2 limit, one has to show that the jackknife variance estimator

$$v_{m,n}(t) = \frac{1}{m+n}\sum_{i=1}^{m+n}\left\{\hat{V}_i(t) - \frac{1}{m+n}\sum_{j=1}^{m+n}\hat{V}_j(t)\right\}^2$$

is a consistent estimator of $(m+n)Var(\hat{R}_{m,n}(t))$.

Theorem 1. *Assume that w is a symmetric density with support $[-1, 1]$ and the first derivative of w is bounded. Further assume that the second derivative of $R(t)$ is continuous at $t_0 \in (0, 1)$, and $\lim_{n\to\infty} m/n = r \in (0, \infty)$. If $h = h(n) \to 0$, $nh^2/\log n \to \infty$ and $nh^4 \to 0$ as $n \to \infty$, then*

$$v_{m,n}(t_0) \xrightarrow{p} \sigma^2(t_0) \quad \text{as } n \to \infty.$$

Remark 1. Although we cannot show that the above jackknife variance estimator based on $R_{m,n}(t)$ instead of $\hat{R}_{m,n}(t)$ is inconsistent, our simulation study does confirm this conjecture. This explains why we have to work with a smooth version of the empirical estimator of the ROC curve.

Theorem 2. *Under the conditions of Theorem 1, we have*

$$l_{m,n}(t_0, R(t_0)) \xrightarrow{d} \chi^2(1) \quad \text{as } n \to \infty.$$

Based on Theorem 2, a confidence interval with level γ for $R(t_0)$ can be constructed as

$$I_\gamma(t_0, m, n) = \{\theta : l_{m,n}(t_0, \theta) \le \chi^2_{1,\gamma}\},$$

where $\chi^2_{1,\gamma}$ is the γ quantile of $\chi^2(1)$.

3. Simulation study

In this section, we compare the coverage accuracy of the proposed jackknife empirical likelihood method with the normal approximation method and the empirical likelihood method in [12], where an extra constraint and smooth distribution estimation for both populations are required.

We consider three cases: (**A**) $F \sim N(0, 1)$, $G \sim N(1, 0.5)$, (**B**) $F \sim N(0, 1)$, $G \sim \text{Exp}(1)$ and (**C**)$F \sim \text{Exp}(1)$, $G \sim \text{Exp}(1)$, where Exp(1) denotes the standard exponential distribution function. We generate 10,000 random samples from the above cases with sample sizes $m = 50$, 100, 200 and $n = 50$, 100, 200. We use the kernel $w(x) = \frac{15}{16}(1-t^2)^2 I(|t| \le 1)$ for both methods, and we choose $h = m^{-1/3}$ for the jackknife empirical likelihood method and $h_1 = m^{-1/3}$ and $h_2 = n^{-1/3}$ for the empirical likelihood method in [12]. Note that Chen, Peng and Zhao [36] pointed out that the above choices of bandwidth for the method in [12] are valid. For the naive bootstrap method based on $R_{m,n}(t)$, we employ 1000 bootstrap samples. We compute the coverage probabilities for $t_0 = 0.05, 0.10, 0.25$ with confidence levels 0.9 and 0.95. From Tables 1–3, we observe that both the proposed jackknife empirical likelihood method and the empirical likelihood method in [12] perform much better than the naive bootstrap method. When $t = 0.05$ and 0.10, the proposed jackknife empirical likelihood method performs best in most cases. Both empirical likelihood methods are comparable in case of $t = 0.25$. However, the proposed jackknife empirical likelihood method is less computationally intensive since the empirical likelihood method in [12] has more constraints in the optimization procedure. Indeed, we employ the "emplik" R package for the proposed jackknife empirical likelihood method.

Table 1
Coverage probabilities for the ROC curve $R(0.05)$ are reported for the intervals based on the naive bootstrap method for $R_{m,n}(t)$ (NBM), the proposed jackknife empirical likelihood method (JELM) and the empirical likelihood method (ELM) in [12] for levels $\gamma = 0.9, 0.95$ and various sample sizes.

(*m*, *n*, *Case*)	NBM $\gamma = 0.9$	JELM $\gamma = 0.9$	ELM $\gamma = 0.9$	NBM $\gamma = 0.95$	JELM $\gamma = 0.95$	ELM $\gamma = 0.95$
(50, 50, A)	0.5383	0.8530	0.6816	0.5510	0.8867	0.7042
(50, 100, A)	0.5545	0.8356	0.6786	0.5692	0.8838	0.7119
(50, 200, A)	0.5324	0.8183	0.6442	0.5424	0.8708	0.6855
(100, 50, A)	0.7517	0.8950	0.8157	0.7667	0.9314	0.8488
(100, 100, A)	0.7858	0.8903	0.8329	0.8015	0.9311	0.8706
(100, 200, A)	0.7763	0.8719	0.8058	0.7880	0.9236	0.8509
(200, 50, A)	0.7331	0.9070	0.8998	0.7489	0.9473	0.9302
(200, 100, A)	0.8006	0.9147	0.9185	0.8133	0.9552	0.9495
(200, 200, A)	0.7992	0.9050	0.9144	0.8102	0.9496	0.9493
(50, 50, B)	0.1631	0.9138	0.9284	0.1645	0.9547	0.9758
(50, 100, B)	0.1431	0.8326	0.9404	0.1439	0.9351	0.9877
(50, 200, B)	0.1040	0.6433	0.9520	0.1044	0.8293	0.9897
(100, 50, B)	0.2456	0.9377	0.9544	0.2498	0.9636	0.9678
(100, 100, B)	0.2490	0.8952	0.9695	0.2522	0.9623	0.9786
(100, 200, B)	0.1962	0.7255	0.9800	0.1970	0.8845	0.9873
(200, 50, B)	0.3531	0.9448	0.9236	0.3611	0.9647	0.9288
(200, 100, B)	0.3699	0.9364	0.9415	0.3781	0.9759	0.9477
(200, 200, B)	0.3211	0.8203	0.9626	0.3248	0.9374	0.9669
(50, 50, C)	0.6505	0.9056	0.8363	0.6727	0.9550	0.8570
(50, 100, C)	0.7041	0.8686	0.8897	0.7262	0.9379	0.9149
(50, 200, C)	0.7010	0.8223	0.8944	0.7187	0.9052	0.9269
(100, 50, C)	0.7359	0.9151	0.8033	0.7572	0.9589	0.8330
(100, 100, C)	0.8208	0.9058	0.8797	0.8424	0.9532	0.9135
(100, 200, C)	0.8433	0.8656	0.9141	0.8601	0.9350	0.9507
(200, 50, C)	0.7518	0.9078	0.7349	0.7916	0.9473	0.8055
(200, 100, C)	0.8244	0.9135	0.8184	0.8681	0.9585	0.8845
(200, 200, C)	0.8562	0.8973	0.8950	0.8940	0.9508	0.9409

Table 2
Coverage probabilities for the ROC curve $R(0.1)$ are reported for intervals based on the naive bootstrap method for $R_{m,n}(t)$ (NBM), the proposed jackknife empirical likelihood method (JELM) and the empirical likelihood method (ELM) in [12] for levels $\gamma = 0.9, 0.95$ and various sample sizes.

(*m*, *n*, *Case*)	NBM $\gamma = 0.9$	JELM $\gamma = 0.9$	ELM $\gamma = 0.9$	NBM $\gamma = 0.95$	JELM $\gamma = 0.95$	ELM $\gamma = 0.95$
(50, 50, A)	0.7673	0.8685	0.8292	0.7797	0.9001	0.8664
(50, 100, A)	0.7659	0.8601	0.8101	0.7734	0.9013	0.8557
(50, 200, A)	0.7497	0.8469	0.7772	0.7561	0.8928	0.8237
(100, 50, A)	0.7478	0.8997	0.9066	0.7768	0.9364	0.9423
(100, 100, A)	0.7559	0.8991	0.9065	0.7773	0.9412	0.9411
(100, 200, A)	0.7526	0.8961	0.8955	0.7727	0.9396	0.9345
(200, 50, A)	0.8150	0.8910	0.8976	0.8739	0.937	0.9516
(200, 100, A)	0.8347	0.9040	0.9060	0.8936	0.9496	0.9594
(200, 200, A)	0.8369	0.9019	0.9019	0.9032	0.9478	0.9548
(50, 50, B)	0.4936	0.9015	0.5875	0.5121	0.9449	0.6052
(50, 100, B)	0.4539	0.8672	0.6000	0.4661	0.9348	0.6121
(50, 200, B)	0.4429	0.7871	0.6065	0.4508	0.8946	0.6206
(100, 50, B)	0.6660	0.9173	0.7102	0.6809	0.9511	0.7273
(100, 100, B)	0.6670	0.9122	0.7466	0.6758	0.9574	0.7637
(100, 200, B)	0.6615	0.8443	0.7616	0.6690	0.9302	0.7805
(200, 50, B)	0.6190	0.9140	0.7846	0.6401	0.9453	0.8116
(200, 100, B)	0.6191	0.9215	0.8356	0.6353	0.9596	0.8643
(200, 200, B)	0.6195	0.8947	0.8769	0.6319	0.9544	0.9039
(50, 50, C)	0.8103	0.9068	0.8784	0.8339	0.9524	0.9232
(50, 100, C)	0.8257	0.9078	0.9114	0.8540	0.9553	0.9502
(50, 200, C)	0.8472	0.9040	0.9094	0.8731	0.9573	0.9529
(100, 50, C)	0.7946	0.8851	0.8168	0.8521	0.9354	0.8856
(100, 100, C)	0.8360	0.9060	0.8841	0.8900	0.9530	0.9397
(100, 200, C)	0.8531	0.9068	0.9069	0.9033	0.9575	0.9516
(200, 50, C)	0.7717	0.8771	0.7655	0.8342	0.9185	0.8212
(200, 100, C)	0.8026	0.8926	0.8422	0.8668	0.9375	0.8949
(200, 200, C)	0.8274	0.9005	0.8886	0.8880	0.9512	0.9369

Next we examine the interval lengths of the proposed jackknife empirical likelihood method and the naive bootstrap method based on $R_{m,n}(t)$ since the computation for the other empirical likelihood interval is quite intensive. Note that $l_{m,n}(t, \theta) \geq 0$ is a convex function of θ and $l_{m,n}(t, \frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(t)) = 0$. So by increasing and decreasing θ from

Table 3
Coverage probabilities for the ROC curve $R(0.25)$ are reported for intervals based on the naive bootstrap method for $R_{m,n}(t)$ (NBM), the proposed jackknife empirical likelihood method (JELM) and the empirical likelihood method (ELM) in [12] for levels $\gamma = 0.9, 0.95$ and various sample sizes.

$(m, n, Case)$	NBM $\gamma = 0.9$	JELM $\gamma = 0.9$	ELM $\gamma = 0.9$	NBM $\gamma = 0.95$	JELM $\gamma = 0.95$	ELM $\gamma = 0.95$
(50, 50, A)	0.8320	0.9047	0.9172	0.8503	0.9417	0.9587
(50, 100, A)	0.8424	0.8984	0.9070	0.8588	0.9398	0.9479
(50, 200, A)	0.8464	0.9001	0.9069	0.8604	0.9402	0.9434
(100, 50, A)	0.8369	0.8878	0.9018	0.8662	0.9407	0.9518
(100, 100, A)	0.8657	0.9013	0.9039	0.8957	0.9481	0.9516
(100, 200, A)	0.8760	0.9041	0.9008	0.9028	0.9512	0.9501
(200, 50, A)	0.8305	0.8820	0.9032	0.8786	0.9348	0.9508
(200, 100, A)	0.8577	0.8963	0.9045	0.9045	0.9453	0.9517
(200, 200, A)	0.8628	0.8980	0.9003	0.9137	0.9505	0.9507
(50, 50, B)	0.6957	0.8895	0.9002	0.7142	0.9354	0.9568
(50, 100, B)	0.7424	0.9022	0.9089	0.7601	0.9473	0.9582
(50, 200, B)	0.7647	0.9087	0.9002	0.7804	0.9509	0.9545
(100, 50, B)	0.7505	0.8739	0.8924	0.7862	0.9285	0.9399
(100, 100, B)	0.8129	0.8982	0.9085	0.8578	0.9473	0.9558
(100, 200, B)	0.8269	0.9056	0.9067	0.8782	0.9512	0.9539
(200, 50, B)	0.7512	0.8526	0.8791	0.8014	0.9057	0.9265
(200, 100, B)	0.8115	0.8794	0.9018	0.8576	0.9287	0.9449
(200, 200, B)	0.8438	0.9007	0.9098	0.8927	0.9465	0.9537
(50, 50, C)	0.8040	0.8878	0.8970	0.8599	0.9368	0.9434
(50, 100, C)	0.8417	0.9006	0.9060	0.8907	0.9465	0.9515
(50, 200, C)	0.8576	0.9083	0.9035	0.9108	0.9553	0.9537
(100, 50, C)	0.8137	0.8705	0.8785	0.8651	0.9260	0.9239
(100, 100, C)	0.8549	0.8915	0.9049	0.9109	0.9462	0.9507
(100, 200, C)	0.8708	0.9019	0.9083	0.9286	0.9507	0.9531
(200, 50, C)	0.7992	0.8638	0.8729	0.8554	0.9154	0.9197
(200, 100, C)	0.8371	0.8837	0.8957	0.8949	0.9339	0.9413
(200, 200, C)	0.8540	0.8916	0.9029	0.9155	0.9414	0.9496

$\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(t)$ with a step 0.001 till $l_{m,n}(t,\theta) > \chi^2_{1,\gamma}$ we can easily obtain the upper and lower endpoints of the jackknife empirical likelihood interval $I_\gamma(t_0, m, n)$. In Table 4, we report the interval lengths for the jackknife empirical likelihood method and the naive bootstrap method. We observe that the jackknife empirical likelihood method results in a shorter interval than the naive bootstrap method for almost all of cases except case C with $\gamma = 0.95$.

4. Proofs

We need the following lemmas to prove Theorems 1 and 2.

Lemma 1. *Assume conditions in Theorem 1 hold. Then there exists an interval $(a, b) \subset (0, 1)$ such that $t_0 \in (a, b)$ and*

$$\sqrt{m+n}\{\hat{R}_{m,n}(t) - R(t)\} \xrightarrow{D} \sqrt{1+\frac{1}{r}}B_1(1-R(t)) + \sqrt{1+r}R'(t)B_2(t) \tag{4}$$

in $D((a, b))$, where $B_1(t)$ and $B_2(t)$ are two independent Brownian bridges.

Proof. Since R'' is continuous at $t_0 \in (0, 1)$, there exists a subset (a, b) containing t_0 such that R' and R'' are bounded in (a, b). It is known that

$$\sqrt{m}\{F_m(x) - F(x)\} \xrightarrow{D} W_1(x) \quad \text{and} \quad \sqrt{n}\{G_n(y) - G(y)\} \xrightarrow{D} W_2(y) \tag{5}$$

in $D((-\infty, \infty))$, where W_1 and W_2 are two independent Wiener processes with zero means and covariances

$$\begin{cases} EW_1(x_1)W_1(x_2) = F(x_1 \wedge x_2) - F(x_1)F(x_2) \\ EW_2(y_1)W_2(y_2) = G(y_1 \wedge y_2) - G(y_1)G(y_2). \end{cases}$$

Write

$$\begin{aligned} 1 - \frac{1}{m}\sum_{j=1}^{m} K\left(\frac{1-t-G(X_j)}{h}\right) - R(t) &= F(G^-(1-t)) - \int_{-\infty}^{\infty} K\left(\frac{1-t-G(x)}{h}\right) \mathrm{d}F_m(x) \\ &= F(G^-(1-t)) - \int_{-\infty}^{\infty} F_m(x)w\left(\frac{1-t-G(x)}{h}\right) h^{-1}\,\mathrm{d}G(x) \\ &= F(G^-(1-t)) - \int_{-1}^{1} F_m(G^-(1-t-xh))w(x)\,\mathrm{d}x \end{aligned}$$

Table 4
Interval lengths are reported for the ROC curve $R(t)$ based on the naive bootstrap method for $R_{m,n}(t)$ (NBM) and the proposed jackknife empirical likelihood method (JELM) for levels $\gamma = 0.9$, 0.95 and various sample sizes.

$(m, n, Case)$	NBM $\gamma = 0.9$ $t = 0.1$	JELM $\gamma = 0.9$ $t = 0.1$	NBM $\gamma = 0.95$ $t = 0.1$	JELM $\gamma = 0.95$ $t = 0.1$	NBM $\gamma = 0.9$ $t = 0.25$	JELM $\gamma = 0.9$ $t = 0.25$	NBM $\gamma = 0.95$ $t = 0.25$	JELM $\gamma = 0.95$ $t = 0.25$
(50, 50, A)	0.0879	0.0582	0.1031	0.0818	0.1346	0.1128	0.1590	0.1567
(50, 100, A)	0.0746	0.0573	0.0873	0.0779	0.1271	0.1098	0.1500	0.1532
(50, 200, A)	0.0700	0.0579	0.0814	0.0780	0.1205	0.1072	0.1420	0.1532
(100, 50, A)	0.0711	0.0448	0.0844	0.0653	0.1089	0.0923	0.1294	0.1334
(100, 100, A)	0.0623	0.0466	0.0736	0.0646	0.0975	0.0848	0.1158	0.1296
(100, 200, A)	0.0571	0.0447	0.0672	0.0634	0.0910	0.0804	0.1080	0.1285
(200, 50, A)	0.0599	0.0387	0.0710	0.0568	0.0883	0.0776	0.1051	0.1190
(200, 100, A)	0.0495	0.0374	0.0589	0.0539	0.0765	0.0672	0.0908	0.1135
(200, 200, A)	0.0441	0.0344	0.0524	0.0525	0.0681	0.0604	0.0811	0.1102
(50, 50, B)	0.0766	0.0415	0.0948	0.0674	0.1767	0.1284	0.2071	0.1791
(50, 100, B)	0.0540	0.0427	0.0662	0.0624	0.1572	0.1221	0.1851	0.1706
(50, 200, B)	0.0439	0.0449	0.0533	0.0610	0.1436	0.1142	0.1691	0.1682
(100, 50, B)	0.0702	0.0320	0.0855	0.0536	0.1519	0.1134	0.1793	0.1624
(100, 100, B)	0.0495	0.0317	0.0601	0.0482	0.1296	0.1038	0.1534	0.1517
(100, 200, B)	0.0395	0.0335	0.0471	0.0461	0.1124	0.0914	0.1329	0.1463
(200, 50, B)	0.0637	0.0268	0.0772	0.0453	0.1355	0.1035	0.1597	0.1527
(200, 100, B)	0.0444	0.0247	0.0535	0.0393	0.1111	0.0930	0.1316	0.1407
(200, 200, B)	0.0340	0.0255	0.0407	0.0359	0.0912	0.0764	0.1084	0.1320
(50, 50, C)	0.2139	0.1381	0.2519	0.1969	0.2873	0.2363	0.3399	0.3894
(50, 100, C)	0.1804	0.1259	0.2137	0.1903	0.2545	0.2065	0.3018	0.3789
(50, 200, C)	0.1583	0.1132	0.1870	0.1830	0.2290	0.1879	0.2711	0.3665
(100, 50, C)	0.1863	0.1245	0.2210	0.1805	0.2488	0.2137	0.2958	0.3739
(100, 100, C)	0.1480	0.1065	0.1755	0.1670	0.2059	0.1741	0.2448	0.3521
(100, 200, C)	0.1276	0.0916	0.1515	0.1622	0.1793	0.1515	0.2129	0.3439
(200, 50, C)	0.1686	0.1138	0.1982	0.1677	0.2245	0.2011	0.2660	0.3615
(200, 100, C)	0.1294	0.0977	0.1537	0.1569	0.1765	0.1550	0.2101	0.3403
(200, 200, C)	0.1026	0.0776	0.1221	0.1469	0.1453	0.1269	0.1728	0.3246

$$
\begin{aligned}
&= F(G^-(1-t)) - F_m(G^-(1-t)) - \int_{-1}^{1}\{F(G^-(1-t-xh)) - F(G^-(1-t))\}w(x)\,dx \\
&\quad - \int_{-1}^{1}\{F_m(G^-(1-t-xh)) - F(G^-(1-t-xh)) - F_m(G^-(1-t)) + F(G^-(1-t))\}w(x)\,dx
\end{aligned}
\tag{6}
$$

and

$$
\begin{aligned}
\int_{-1}^{1}\{F(G^-(1-t-xh)) - F(G^-(1-t))\}w(x)\,dx &= -\int_{-1}^{1} R'(t)xhw(x)\,dx - \frac{1}{2}\int_{-1}^{1} R''(t^*)(xh)^2 w(x)\,dx \\
&= -\frac{1}{2}h^2\int_{-1}^{1} R''(t^*)x^2 w(x)\,dx,
\end{aligned}
\tag{7}
$$

where t^* is between t and $t + xh$. If follows from conditions in Lemma 1 and (7) that

$$
\int_{-1}^{1}\{F(G^-(1-t-xh)) - F(G^-(1-t))\}w(x)\,dx = O(h^2)
\tag{8}
$$

uniformly in $t \in (a, b)$. Using the conditions on h, (5) and the continuity of W_1, we have

$$
\begin{aligned}
&\int_{-1}^{1}\{F_m(G^-(1-t-xh)) - F(G^-(1-t-xh)) - F_m(G^-(1-t)) + F(G^-(1-t))\}w(x)\,dx \\
&= \int_{-1}^{1}\{F_m(G^-(1-t-xh)) - F(G^-(1-t-xh)) - m^{-1/2}W_1(G^-(1-t-xh))\}w(x)\,dx \\
&\quad - \int_{-1}^{1}\{F_m(G^-(1-t)) - F(G^-(1-t)) - m^{-1/2}W_1(G^-(1-t))\}w(x)\,dx \\
&\quad + \int_{-1}^{1}\{m^{-1/2}W_1(G^-(1-t-xh)) - m^{-1/2}W_1(G^-(1-t))\}w(x)\,dx \\
&= o_p(m^{-1/2}).
\end{aligned}
$$

Hence

$$\sqrt{m}\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G(X_j)}{h}\right)-R(t)\right\}\xrightarrow{D}W_1(G^-(1-t)) \tag{9}$$

in $D((a,b))$.

Write

$$\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G(X_j)}{h}\right)=\frac{1}{m}\sum_{j=1}^{m}\frac{G(X_j)-G_n(X_j)}{h}w\left(\frac{1-t-G(X_j)}{h}\right)$$
$$+\frac{1}{2m}\sum_{j=1}^{m}\left(\frac{G(X_j)-G_n(X_j)}{h}\right)^2 w'\left(\frac{1-t-G(X_j)+\xi_{n,j}}{h}\right), \tag{10}$$

where $\xi_{n,j}$ is between 0 and $G(X_j)-G_n(X_j)$. It follows from Theorem A of Silverman [39] that

$$\sup_{t\in(a,b)}\left|\frac{1}{mh}\sum_{j=1}^{m}\left|w'\left(\frac{1-t-G(X_j)}{h}\right)\right|-R'(t)\int_{-1}^{1}|w'(x)|\,\mathrm{d}x\right|=o_p(1), \tag{11}$$

where $R'(1-x)$ is the density of $G(X_1)$. By (5), (10) and (11), we have

$$\sqrt{n}\left\{\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G(X_j)}{h}\right)\right\}$$
$$=-\int_{-\infty}^{\infty}W_2(x)h^{-1}w\left(\frac{1-t-G(x)}{h}\right)\mathrm{d}F(x)+O_p(n^{-1/2}h^{-1})$$
$$=\int_{-1}^{1}W_2(G^-(1-t-xh))h^{-1}w(x)\,\mathrm{d}F(G^-(1-t-hx))+O_p(n^{-1/2}h^{-1})$$
$$=-R'(t)W_2(G^-(1-t))+o_p(1) \tag{12}$$

uniformly in $t\in(a,b)$. Hence the lemma follows from (9) and (12) with $B_1(1-R(t))=W_1(G^-(1-t))$ and $B_2(t)=W_2(G^-(1-t))$. This completes the proof of the lemma. □

Lemma 2. *Under conditions of Theorem 1, we have*

$$\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(t)-R(t)\right\}\xrightarrow{d}N(0,\sigma^2(t))$$

as $n\to\infty$ *for* $t=t_0$.

Proof. Throughout we assume $t=t_0$. It follows from the definition of $\hat{V}_i(t)$ that

$$\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(t)=\frac{1}{m+n}\left\{m+n-\frac{m+n}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right.$$
$$\left.+\frac{m+n-1}{m}\sum_{k=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,k}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\right\}. \tag{13}$$

Write

$$\sum_{k=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,k}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}$$
$$=\sum_{k=1}^{n}\sum_{j=1}^{m}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}w\left(\frac{1-t-G_n(X_j)}{h}\right)+\sum_{k=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_n(X_j)-G_{n,k}(X_j)}{h}\right\}^2 w'\left(\frac{1-t-\xi_{n,k,j}}{h}\right)$$
$$=\sum_{j=1}^{m}\left\{\sum_{k=1}^{n}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}\right\}w\left(\frac{1-t-G_n(X_j)}{h}\right)+\sum_{k=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_n(X_j)-G_{n,k}(X_j)}{h}\right\}^2 w'\left(\frac{1-t-\xi_{n,k,j}}{h}\right)$$
$$=\sum_{k=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_n(X_j)-G_{n,k}(X_j)}{h}\right\}^2 w'\left(\frac{1-t-\xi_{n,k,j}}{h}\right), \tag{14}$$

where $\xi_{n,k,j}$ is a random variable between $G_{n,k}(X_j)$ and $G_n(X_j)$. Since

$$G_n(X_j) - G_{n,k}(X_j) = \frac{1}{n-1}\{G_n(X_j) - I(Y_k \le X_j)\} = O_p\left(\frac{1}{n-1}\right)$$

uniformly in $1 \le k \le n$ and $1 \le j \le m$, we can write

$$\xi_{n,k,j} = G_n(X_j) + O_p\left(\frac{1}{n-1}\right) = G(X_j) + O_p\left(n^{-\frac{1}{2}}\right). \tag{15}$$

It follows from (14), (15) and (11) that

$$\sum_{k=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,k}(X_j)}{h}\right) - K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\} = O_p\left\{\frac{mn}{(n-1)^2 h}\right\}. \tag{16}$$

By (13), (16) and Lemma 1, we have

$$\begin{aligned}
&\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(t) - R(t)\right\} \\
&= \sqrt{m+n}\left\{1 - \frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right) + O_p\left\{\frac{(m+n-1)n}{(m+n)(n-1)^2 h}\right\} - R(t)\right\} \\
&= \sqrt{m+n}\left\{\hat{R}_{m,n}(t) - R(t) + O_p\left\{\frac{(m+n-1)n}{(m+n)(n-1)^2 h}\right\}\right\} \\
&\xrightarrow{d} N(0, \sigma^2(t)),
\end{aligned}$$

i.e., Lemma 2 holds. □

Lemma 3. *Under conditions of Theorem 1, we have*

$$\frac{1}{m+n}\sum_{i=1}^{m+n}\{\hat{V}_i(t) - R(t)\}^2 \xrightarrow{p} \sigma^2(t)$$

as $n \to \infty$ for $t = t_0$.

Proof. Throughout we assume $t = t_0$. For $1 \le i \le m$, we can write that

$$\hat{V}_i(t) = 1 + \frac{n}{(m-1)m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right) - \frac{m+n-1}{m-1}K\left(\frac{1-t-G_n(X_i)}{h}\right)$$

and

$$\begin{aligned}
\hat{V}_i^2(t) &= \left\{1 - \frac{m+n-1}{m-1}K\left(\frac{1-t-G_n(X_i)}{h}\right)\right\}^2 + \left\{\frac{n}{(m-1)m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}^2 \\
&\quad + 2\left\{\frac{n}{(m-1)m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\left\{1 - \frac{m+n-1}{m-1}K\left(\frac{1-t-G_n(X_i)}{h}\right)\right\},
\end{aligned}$$

which imply that

$$\begin{aligned}
\sum_{i=1}^{m}\hat{V}_i^2(t) &= m - \frac{2(m+n-1)}{m-1}\sum_{i=1}^{m}K\left(\frac{1-t-G_n(X_i)}{h}\right) + \frac{(m+n-1)^2}{(m-1)^2}\sum_{i=1}^{m}K^2\left(\frac{1-t-G_n(X_i)}{h}\right) \\
&\quad + \frac{mn^2}{(m-1)^2 m^2}\left\{\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}^2 \\
&\quad + \frac{2n}{(m-1)m}\left\{\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\left\{m - \frac{m+n-1}{m-1}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_i)}{h}\right)\right\}.
\end{aligned} \tag{17}$$

Since K^2 is a distribution function, it follows from Lemma 1 that

$$\frac{1}{m}\sum_{i=1}^{m}K^2\left(\frac{1-t-G_n(X_i)}{h}\right)\xrightarrow{p}F(G^-(1-t)). \tag{18}$$

Hence, by (17), (18) and Lemma 1,

$$\begin{aligned}\frac{1}{m+n}\sum_{i=1}^{m}\hat{V}_i^2(t) &\xrightarrow{p} \frac{r}{1+r}-2F(G^-(1-t))+\left(1+\frac{1}{r}\right)F(G^-(1-t))\\ &\quad+\frac{1}{r(1+r)}F^2(G^-(1-t))+\frac{2}{1+r}F(G^-(1-t))-\frac{2}{r}F^2(G^-(1-t))\\ &= \frac{r}{1+r}+\frac{1+2r-r^2}{r(1+r)}F(G^-(1-t))-\frac{1+2r}{r(1+r)}F^2(G^-(1-t))\\ &= \frac{r+1}{r}R(t)-\frac{1+2r}{r(1+r)}R^2(t).\end{aligned} \tag{19}$$

Next, for $m<i\le m+n$, we can write that

$$\hat{V}_i(t)=1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)+\frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,i-m}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}$$

and

$$\begin{aligned}\hat{V}_i^2(t) &= \left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}^2\\ &\quad+\left\{\frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,i-m}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\right\}^2\\ &\quad+2\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,i-m}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}.\end{aligned} \tag{20}$$

It follows from (11) that

$$\begin{aligned}A_k &:= \left\{\sum_{j=1}^{m}\left\{K\left(\frac{1-t-G_{n,k}(X_j)}{h}\right)-K\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}\right\}^2\\ &= \left\{\sum_{j=1}^{m}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}w\left(\frac{1-t-G_n(X_j)}{h}\right)+\sum_{j=1}^{m}\frac{\{G_n(X_j)-G_{n,k}(X_j)\}^2}{2h^2}w'\left(\frac{1-t-\xi_{n,k,j}}{h}\right)\right\}^2\\ &= \left\{\sum_{j=1}^{m}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}w\left(\frac{1-t-G_n(X_j)}{h}\right)+O_p(mn^{-2}h^{-1})\right\}^2\\ &= \left\{\sum_{j=1}^{m}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}w\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}^2+O_p(n^{-1}h^{-1}),\end{aligned}$$

which implies that

$$\begin{aligned}\frac{1}{m+n}\sum_{k=1}^{n}A_k &= \frac{1}{m+n}\sum_{k=1}^{n}\left\{\sum_{l=1}^{m}\sum_{j=1}^{m}\frac{G_n(X_l)-G_{n,k}(X_l)}{h}\frac{G_n(X_j)-G_{n,k}(X_j)}{h}\right.\\ &\quad\left.\times\ w\left(\frac{1-t-G_n(X_l)}{h}\right)w\left(\frac{1-t-G_n(X_j)}{h}\right)\right\}+O_p(n^{-1}h^{-1})\\ &= \frac{1}{m+n}\frac{n}{(n-1)^2h^2}\sum_{l=1}^{m}\sum_{j=1}^{m}\{G_n(X_l\wedge X_j)-G_n(X_l)G_n(X_j)\}\end{aligned}$$

$$
\begin{aligned}
&\times w\left(\frac{1-t-G_n(X_l)}{h}\right) w\left(\frac{1-t-G_n(X_j)}{h}\right) + O_p(n^{-1}h^{-1}) \\
=\ & \frac{1}{m+n}\frac{n}{(n-1)^2h^2}\sum_{l=1}^{m}\sum_{j=1}^{m}\{G(X_l\wedge X_j)-G(X_l)G(X_j)\} \\
&\times w\left(\frac{1-t-G(X_l)}{h}\right) w\left(\frac{1-t-G(X_j)}{h}\right)\{1+o_p(1)\} + O_p(n^{-1}h^{-1}) \\
\xrightarrow{p}\ & \frac{r^2}{1+r}\{1-t-(1-t)^2\}\{R'(t)\}^2 \\
=\ & \frac{r^2}{1+r}t(1-t)\{R'(t)\}^2. \qquad (21)
\end{aligned}
$$

By (20), (21), (16) and Lemma 1, we have

$$
\frac{1}{m+n}\sum_{i=m+1}^{m+n}\hat{V}_i^2(t) \xrightarrow{p} \frac{1}{1+r}R^2(t) + (r+1)t(1-t)\{R'(t)\}^2. \qquad (22)
$$

Hence, it follows from (19), (22) and Lemma 2 that

$$
\begin{aligned}
\frac{1}{m+n}\sum_{i=1}^{m+n}\{\hat{V}_i(t)-R(t)\}^2 &= \frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i^2(t) + R^2(t) - \frac{2}{m+n}R(t)\sum_{i=1}^{m+n}\hat{V}_i(t) \\
&\xrightarrow{p} \sigma^2(t).
\end{aligned}
$$

This completes the proof of Lemma 3. □

Proof of Theorem 1. It follows immediately from Lemmas 2 and 3. □

Proof of Theorem 2. Throughout let $\theta = R(t_0)$. Define $g(\lambda) = \frac{1}{m+n}\sum_{i=1}^{m+n}\frac{\hat{V}_i(t_0)-\theta}{1+\lambda(\hat{V}_i(t_0)-\theta)}$. It is easy to check that

$$
\begin{aligned}
0 = |g(\lambda)| &= \frac{1}{m+n}\left|\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta) - \lambda\sum_{i=1}^{m+n}\frac{(\hat{V}_i(t_0)-\theta)^2}{1+\lambda(\hat{V}_i(t_0)-\theta)}\right| \\
&\geq \left|\frac{\lambda}{m+n}\sum_{i=1}^{m+n}\frac{(\hat{V}_i(t_0)-\theta)^2}{1+\lambda(\hat{V}_i(t_0)-\theta)}\right| - \left|\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)\right| \\
&\geq \frac{|\lambda|S_{m+n}}{1+|\lambda|Z_{m+n}} - \left|\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)\right|,
\end{aligned}
$$

where $S_{m+n} = \frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)^2$ and $Z_{m+n} = \max_{1\leq i\leq m+n}|\hat{V}_i(t_0)-\theta|$. Using similar arguments in proving Lemma 2, we can show that Z_{m+n} is bounded in probability. Hence, by Lemma 2, Lemma 3 and the fact that Z_{m+n} is bounded in probability, we have

$$
|\lambda| = O_p\left\{(m+n)^{-\frac{1}{2}}\right\}. \qquad (23)
$$

Put $\gamma_i = \lambda(\hat{V}_i(t_0)-\theta)$. Then, we have that

$$
\max_{1\leq i\leq m+n}|\gamma_i| = o_p(1). \qquad (24)
$$

Using (23), (24) and Taylor expansion, we have

$$
\begin{aligned}
0 = g(\lambda) &= \frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)\left(1-\gamma_i+\frac{\gamma_i^2}{1+\gamma_i}\right) \\
&= \frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta) - S_{m+n}\lambda + \frac{1}{m+n}\sum_{i=1}^{m+n}\frac{(\hat{V}_i(t_0)-\theta)\gamma_i^2}{1+\gamma_i} \\
&= \frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta) - S_{m+n}\lambda + O_p\left(\frac{1}{m+n}\right),
\end{aligned}
$$

which implies that

$$\lambda = S_{m+n}^{-1}\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)+\beta_n, \tag{25}$$

where $\beta_n = O_p(\frac{1}{m+n})$. Hence, it follows from (23), (25), Lemmas 1 and 2 that

$$\begin{aligned} l_{m,n}(t_0,\theta) &= 2\sum_{i=1}^{m+n}\gamma_i - \sum_{i=1}^{m+n}\gamma_i^2 + 2\sum_{i=1}^{m+n}\eta_i \\ &= 2(m+n)\lambda\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta) - (m+n)S_{m+n}\lambda^2 + 2\sum_{i=1}^{m+n}\eta_i \\ &= \frac{(m+n)\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)\right\}^2}{S_{m+n}} - (m+n)S_{m+n}\beta_n^2 + 2\sum_{i=1}^{m+n}\eta_i \\ &= \frac{(m+n)\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}(\hat{V}_i(t_0)-\theta)\right\}^2}{S_{m+n}} + o_p(1) \\ &\xrightarrow{d} \chi_1^2, \end{aligned}$$

i.e., Theorem 2 holds. □

Acknowledgments

We thank two reviewers for their helpful comments. Peng's research was supported by NSA grant H98230-10-1-0170 and Qi's research was supported by NSA grant H98230-10-1-0161.

References

[1] C.Z. Metz, Basic principles of ROC analysis, Seminars in Nuclear Medicine 8 (1978) 283–298.
[2] M.H. Zweig, G. Campbell, Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine, Clinical Chemistry 39 (1993) 561–577.
[3] P.F. Griner, R.J. Mayewski, A.I. Mushlin, P. Greenland, Selection and interpretation of diagnostic tests and procedures, Annals of Internal Medicine 94 (1981) 555–600.
[4] F. Hsieh, B.W. Turnbull, Nonparametric and semiparametric estimation of the receiver operating characteristic curve, Annals of Statistics 24 (1996) 25–40.
[5] K.H. Zhou, W.J. Hall, D.E. Shapiro, Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests, Statistics in Medicine 16 (1997) 2143–2156.
[6] C.J. Lloyd, The use of smoothed ROC curves to summarize and compare diagnostic systems, Journal of the American Statistical Association 93 (1998) 1356–1364.
[7] C.J. Lloyd, Z. Yong, Kernel estimators of the ROC curve are better than empirical, Statistics & Probability Letters 44 (1999) 221–228.
[8] P. Hall, R.J. Hyndman, Y. Fan, Nonparametric confidence intervals for receiver operating characteristic curves, Biometrika 91 (2004) 743–750.
[9] L. Peng, X. Zhou, Local linear smoothing of receiver operator characteristic (ROC) curves, Journal of Statistical Planning and Inference 118 (2004) 129–143.
[10] J. Gu, S. Ghosal, A. Roy, Bayesian bootstrap estimation of ROC curve, Statistics in Medicine 27 (2008) 5407–5420.
[11] W. Zhou, Statistical inference for $P(X < Y)$, Statistics in Medicine 27 (2008) 257–279.
[12] G. Claeskens, B. Jing, L. Peng, W. Zhou, Empirical likelihood confidence regions for comparison distributions and ROC curves, The Canadian Journal of Statistics 31 (2) (2003) 173–190.
[13] E. Molanes-Lopez, I. Van Keilegom, N. Veraverbeke, Empirical likelihood for non-smooth criterion functions, Scandinavian Journal of Statistics 36 (2009) 413–432.
[14] G.S. Qin, X.H. Zhou, Empirical likelihood inference for the area under the ROC curve, Biometrics 62 (2006) 613–622.
[15] A. Owen, Empirical likelihood ratio confidence intervals for a single functional, Biometrika 75 (1988) 237–249.
[16] A. Owen, Empirical likelihood ratio confidence regions, Annals of Statistics 18 (1990) 90–120.
[17] A. Owen, Empirical Likelihood, Chapman & Hall, CRC, 2001.
[18] Y. Zhao, F. Chen, Empirical likelihood inference for censored median regression model via nonparametric kernel estimation, Journal of Multivariate Analysis 99 (2008) 215–231.
[19] Y. Zhao, S. Yang, Empirical likelihood inference for censored median regression with weighted empirical hazard functions, Annals of the Institute of Statistical Mathematics 60 (2008) 441–457.
[20] Y. Liu, C. Zou, R. Zhang, Empirical likelihood for the two-sample mean problem, Statistics & Probability Letters 78 (2008) 548–556.
[21] J. Ren, Weighted empirical likelihood in some two-sample semiparametric models with various types of censored data, Annals of Statistics 36 (2008) 147–166.
[22] A. Keziou, S. Leoni-Aubin, On empirical likelihood for semiparametric two-sample density ratio models, Journal of Statistical Planning and Inference 138 (2008) 915–928.
[23] J. Shen, S. He, Empirical likelihood for the difference of quantiles under censorship, Statistical Papers 48 (2007) 437–457.
[24] R. Cao, I. Van Keilegom, Empirical likelihood tests for two-sample problems via nonparametric density estimation, Canadian Journal Statistics 34 (2006) 61–77.

[25] Y. Zhou, H. Liang, Empirical-likelihood-based semiparametric inference for the treatment effect in the two-sample problem with censoring, Biometrika 92 (2005) 271–282.
[26] P. Guggenberger, R.J. Smith, Generalized empirical likelihood tests in time series models with potential identification failure, Journal of Econometrics 142 (2008) 134–161.
[27] S. Chen, J. Gao, An adaptive empirical likelihood test for parametric time series regression models, Journal of Econometrics 141 (2007) 950–972.
[28] D.J. Nordman, P. Sibbertsen, S.N. Lahiri, Empirical likelihood confidence intervals for the mean of a long-range dependent process, Journal of Time Series Analysis 28 (2007) 576–599.
[29] D.J. Nordman, S.N. Lahiri, A frequency domain empirical likelihood for short- and long-range dependence, Annals of Statistics 34 (2006) 3019–3050.
[30] T. Otsu, Generalized empirical likelihood inference for nonlinear and time series models under weak identification, Econometric Theory 22 (2006) 513–527.
[31] N.H. Chan, S. Ling, Empirical likelihood for GARCH models, Econometric Theory 22 (2006) 403–428.
[32] Y. Zhao, W. Jian, Analysis of longitudinal data in the case-control studies via empirical likelihood, Communications in Statistics—Simulation and Computation 36 (2007) 565–578.
[33] L. Xue, L. Zhu, Empirical likelihood for single-index models, Journal of Multivariate Analysis 97 (2006) 1295–1312.
[34] L. Xue, L. Zhu, Empirical likelihood for a varying coefficient model with longitudinal data, Journal of American Statistical Association 102 (2007) 642–654.
[35] J. You, G. Chen, Y. Zhou, Block empirical likelihood for longitudinal partially linear regression models, Canadian Journal of Statistics 34 (2006) 79–96.
[36] J. Chen, L. Peng, Y. Zhao, Empirical likelihood based confidence intervals for copulas, Journal of Multivariate Analysis 100 (2009) 137–151.
[37] B. Jing, J. Yuan, W. Zhou, Jackknife empirical likelihood, Journal of American Statistical Association 104 (2009) 1224–1232.
[38] J. Shao, D. Tu, The Jackknife and Bootstrap, Springer, 1995.
[39] B.W. Silverman, Weak and strong uniform consistency of the kernel estimate of a density and its derivatives, The Annals of Statistics 6 (1978) 177–184.