Smoothed jackknife empirical likelihood method for ROC curve

Yun Gong, Liang Peng, Yongcheng Qi

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA
Department of Mathematics and Statistics, University of Minnesota-Duluth, 1117 University Drive, Duluth, MN 55812, USA

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 \( R(t) = 1 - F(G^{-1}(1-t)) \) for \( 0 < t < 1 \), where \( G^{-1} \) denotes the inverse of \( G \) and is defined by \( G^{-1}(u) = \inf\{x : G(x) \geq u\} \) 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

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

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

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

© 2010 Elsevier Inc. All rights reserved.

doi:10.1016/j.jmva.2010.01.012
where
\[
\sigma^2(t) = \left(1 + \frac{1}{r}\right) R(t)(1 - R(t)) + (1 + r)t(1 - t) \left\{ \frac{F'(G^-(1 - t))}{G'(G^-(1 - t))} \right\}^2,
\]
and \( r := \lim_{n \to \infty} m/n \in (0, \infty) \), 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
\[
\hat{R}_{m,n}(t) = 1 - \frac{1}{m} \sum_{i=1}^{m} K \left( \frac{1 - G_n(X_j)}{h} \right),
\]
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 \( K \) 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
\[
\hat{R}_{m,n,i}(t) = 1 - \frac{1}{m - 1} \sum_{1 \leq j \leq m, j \neq i} K \left( \frac{1 - G_n(X_j)}{h} \right), \quad 1 \leq i \leq m,
\]
\[
\hat{R}_{m,n,i}(t) = 1 - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - G_n,i-m(X_j)}{h} \right), \quad m < i \leq m + n,
\]
where
\[
G_{n,k}(y) = \frac{1}{n - 1} \sum_{1 \leq l \leq n, l \neq k} I(Y_l \leq y), \quad k = 1, \ldots, n.
\]
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
\[ p_i = \frac{1}{(m + n)(1 + \lambda(\hat{V}(t) - \hat{\theta}))}, \quad i = 1, \ldots, m + n, \]
where \( \lambda = \lambda(t, \hat{\theta}) \) satisfies
\[ \frac{1}{m + n} \sum_{i=1}^{m+n} \frac{\hat{V}(t) - \hat{\theta}}{1 + \lambda(\hat{V}(t) - \hat{\theta})} = 0, \]
which gives the log empirical likelihood ratio as
\[ l_{m,n}(t, \hat{\theta}) = -2 \log l_{m,n}(t, \hat{\theta}) = 2 \sum_{i=1}^{m+n} \log(1 + \lambda(\hat{V}(t) - \hat{\theta})). \]

In order to show that the above log empirical likelihood ratio converges in distribution to a \( \chi^2 \) limit, one has to show that the jackknife variance estimator
\[ v_{m,n}(t) = \frac{1}{m + n} \sum_{i=1}^{m+n} \left( \frac{\hat{V}(t) - \hat{\theta}}{m + n} \sum_{j=1}^{m+n} \hat{V}(t) \right)^2 \]
is a consistent estimator of \((m + n)\text{Var}(\hat{R}_{m,n}(t))\).

**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
\[ v_{m,n}(t_0) \to \sigma^2(t_0) \quad \text{as} \quad n \to \infty. \]

**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
\[ l_{m,n}(t_0, R(t_0)) \to \chi^2(1) \quad \text{as} \quad n \to \infty. \]

Based on **Theorem 2**, a confidence interval with level \( \gamma \) for \( R(t_0) \) can be constructed as
\[ I_{\gamma, m, n} = \{ \hat{\theta} : l_{m, n}(t_0, \hat{\theta}) \leq \chi^2_{1, \gamma} \}, \]
where \( \chi^2_{1, \gamma} \) is the \( \gamma \) 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 \( \text{Exp}(1) \) 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 \( w(x) = \frac{15}{16}(1 - t^2)^2(|t| \leq 1) \) 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 $R_{n,t}(t)$ (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.

<table>
<thead>
<tr>
<th>(m, n, Case)</th>
<th>NBM $\gamma = 0.9$</th>
<th>JELM $\gamma = 0.9$</th>
<th>ELM $\gamma = 0.9$</th>
<th>NBM $\gamma = 0.95$</th>
<th>JELM $\gamma = 0.95$</th>
<th>ELM $\gamma = 0.95$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(50, 50, A)</td>
<td>0.5383</td>
<td>0.8530</td>
<td>0.6816</td>
<td>0.5510</td>
<td>0.8867</td>
<td>0.7042</td>
</tr>
<tr>
<td>(50, 100, A)</td>
<td>0.5545</td>
<td>0.8356</td>
<td>0.6786</td>
<td>0.5602</td>
<td>0.8838</td>
<td>0.7119</td>
</tr>
<tr>
<td>(50, 200, A)</td>
<td>0.5324</td>
<td>0.8183</td>
<td>0.6442</td>
<td>0.5424</td>
<td>0.8708</td>
<td>0.6855</td>
</tr>
<tr>
<td>(100, 50, A)</td>
<td>0.7517</td>
<td>0.8950</td>
<td>0.8157</td>
<td>0.7667</td>
<td>0.9314</td>
<td>0.8488</td>
</tr>
<tr>
<td>(100, 100, A)</td>
<td>0.7858</td>
<td>0.8903</td>
<td>0.8329</td>
<td>0.8015</td>
<td>0.9311</td>
<td>0.8706</td>
</tr>
<tr>
<td>(100, 200, A)</td>
<td>0.7763</td>
<td>0.8719</td>
<td>0.8058</td>
<td>0.7880</td>
<td>0.9236</td>
<td>0.8509</td>
</tr>
<tr>
<td>(200, 50, A)</td>
<td>0.7331</td>
<td>0.9070</td>
<td>0.8998</td>
<td>0.7489</td>
<td>0.9473</td>
<td>0.9302</td>
</tr>
<tr>
<td>(200, 100, A)</td>
<td>0.8006</td>
<td>0.9147</td>
<td>0.9185</td>
<td>0.8133</td>
<td>0.9552</td>
<td>0.9495</td>
</tr>
<tr>
<td>(200, 200, A)</td>
<td>0.7902</td>
<td>0.9050</td>
<td>0.9144</td>
<td>0.8102</td>
<td>0.9496</td>
<td>0.9493</td>
</tr>
<tr>
<td>(50, 50, B)</td>
<td>0.1631</td>
<td>0.9138</td>
<td>0.9284</td>
<td>0.1645</td>
<td>0.9547</td>
<td>0.9758</td>
</tr>
<tr>
<td>(50, 100, B)</td>
<td>0.1431</td>
<td>0.8326</td>
<td>0.9404</td>
<td>0.1439</td>
<td>0.9351</td>
<td>0.9877</td>
</tr>
<tr>
<td>(50, 200, B)</td>
<td>0.1040</td>
<td>0.6433</td>
<td>0.9520</td>
<td>0.1044</td>
<td>0.8293</td>
<td>0.9897</td>
</tr>
<tr>
<td>(100, 50, B)</td>
<td>0.2456</td>
<td>0.9377</td>
<td>0.9544</td>
<td>0.2496</td>
<td>0.9636</td>
<td>0.9678</td>
</tr>
<tr>
<td>(100, 100, B)</td>
<td>0.2490</td>
<td>0.8952</td>
<td>0.9695</td>
<td>0.2522</td>
<td>0.9623</td>
<td>0.9766</td>
</tr>
<tr>
<td>(100, 200, B)</td>
<td>0.2352</td>
<td>0.7255</td>
<td>0.9600</td>
<td>0.2397</td>
<td>0.8845</td>
<td>0.9873</td>
</tr>
<tr>
<td>(200, 50, B)</td>
<td>0.3531</td>
<td>0.9448</td>
<td>0.9236</td>
<td>0.3611</td>
<td>0.9647</td>
<td>0.9288</td>
</tr>
<tr>
<td>(200, 100, B)</td>
<td>0.3699</td>
<td>0.9364</td>
<td>0.9415</td>
<td>0.3781</td>
<td>0.9759</td>
<td>0.9477</td>
</tr>
<tr>
<td>(200, 200, B)</td>
<td>0.3211</td>
<td>0.8203</td>
<td>0.9626</td>
<td>0.3248</td>
<td>0.9374</td>
<td>0.9669</td>
</tr>
<tr>
<td>(50, 50, C)</td>
<td>0.6505</td>
<td>0.9056</td>
<td>0.8363</td>
<td>0.6727</td>
<td>0.9550</td>
<td>0.8570</td>
</tr>
<tr>
<td>(50, 100, C)</td>
<td>0.7041</td>
<td>0.8686</td>
<td>0.8897</td>
<td>0.7262</td>
<td>0.9379</td>
<td>0.9149</td>
</tr>
<tr>
<td>(50, 200, C)</td>
<td>0.7010</td>
<td>0.8223</td>
<td>0.8944</td>
<td>0.7187</td>
<td>0.9052</td>
<td>0.9269</td>
</tr>
<tr>
<td>(100, 50, C)</td>
<td>0.7359</td>
<td>0.9151</td>
<td>0.9033</td>
<td>0.7572</td>
<td>0.9589</td>
<td>0.8330</td>
</tr>
<tr>
<td>(100, 100, C)</td>
<td>0.8208</td>
<td>0.9058</td>
<td>0.8797</td>
<td>0.8424</td>
<td>0.9532</td>
<td>0.9135</td>
</tr>
<tr>
<td>(100, 200, C)</td>
<td>0.8433</td>
<td>0.8656</td>
<td>0.9141</td>
<td>0.8601</td>
<td>0.9350</td>
<td>0.9507</td>
</tr>
<tr>
<td>(200, 50, C)</td>
<td>0.7518</td>
<td>0.9078</td>
<td>0.7349</td>
<td>0.7916</td>
<td>0.9473</td>
<td>0.8055</td>
</tr>
<tr>
<td>(200, 100, C)</td>
<td>0.8244</td>
<td>0.9135</td>
<td>0.8184</td>
<td>0.8681</td>
<td>0.9585</td>
<td>0.8845</td>
</tr>
<tr>
<td>(200, 200, C)</td>
<td>0.8562</td>
<td>0.8973</td>
<td>0.8950</td>
<td>0.8940</td>
<td>0.9508</td>
<td>0.9409</td>
</tr>
</tbody>
</table>

Next we examine the interval lengths of the proposed jackknife empirical likelihood method and the naive bootstrap method based on $K_{m,n}(t)$ 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 $\theta$ and $l_{m,n}(t, \frac{1}{m+n} \sum_{i=1}^{m+n} V(t)) = 0$. So by increasing and decreasing $\theta$ 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.

<table>
<thead>
<tr>
<th>(m, n, Case)</th>
<th>NBM ( \gamma = 0.9 )</th>
<th>JELM ( \gamma = 0.9 )</th>
<th>ELM ( \gamma = 0.9 )</th>
<th>NBM ( \gamma = 0.95 )</th>
<th>JELM ( \gamma = 0.95 )</th>
<th>ELM ( \gamma = 0.95 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(50, 50, A)</td>
<td>0.8320</td>
<td>0.9047</td>
<td>0.9172</td>
<td>0.8503</td>
<td>0.9417</td>
<td>0.9587</td>
</tr>
<tr>
<td>(50, 100, A)</td>
<td>0.8424</td>
<td>0.8984</td>
<td>0.9070</td>
<td>0.8588</td>
<td>0.9398</td>
<td>0.9479</td>
</tr>
<tr>
<td>(200, 100, A)</td>
<td>0.8369</td>
<td>0.8901</td>
<td>0.9069</td>
<td>0.8564</td>
<td>0.9402</td>
<td>0.9434</td>
</tr>
<tr>
<td>(100, 50, A)</td>
<td>0.8369</td>
<td>0.8878</td>
<td>0.9018</td>
<td>0.8662</td>
<td>0.9407</td>
<td>0.9518</td>
</tr>
<tr>
<td>(100, 100, A)</td>
<td>0.8657</td>
<td>0.9013</td>
<td>0.9039</td>
<td>0.8957</td>
<td>0.9481</td>
<td>0.9516</td>
</tr>
<tr>
<td>(200, 200, A)</td>
<td>0.8760</td>
<td>0.9041</td>
<td>0.9008</td>
<td>0.9028</td>
<td>0.9512</td>
<td>0.9501</td>
</tr>
<tr>
<td>(200, 50, A)</td>
<td>0.8305</td>
<td>0.8820</td>
<td>0.9032</td>
<td>0.8786</td>
<td>0.9348</td>
<td>0.9508</td>
</tr>
<tr>
<td>(200, 100, A)</td>
<td>0.8577</td>
<td>0.8963</td>
<td>0.9045</td>
<td>0.9045</td>
<td>0.9453</td>
<td>0.9517</td>
</tr>
<tr>
<td>(200, 200, A)</td>
<td>0.8826</td>
<td>0.8880</td>
<td>0.9003</td>
<td>0.9137</td>
<td>0.9505</td>
<td>0.9507</td>
</tr>
<tr>
<td>(50, 50, B)</td>
<td>0.6958</td>
<td>0.8895</td>
<td>0.9002</td>
<td>0.7142</td>
<td>0.9354</td>
<td>0.9568</td>
</tr>
<tr>
<td>(50, 100, B)</td>
<td>0.7424</td>
<td>0.9022</td>
<td>0.9089</td>
<td>0.7601</td>
<td>0.9473</td>
<td>0.9582</td>
</tr>
<tr>
<td>(50, 200, B)</td>
<td>0.7647</td>
<td>0.9087</td>
<td>0.9002</td>
<td>0.7804</td>
<td>0.9509</td>
<td>0.9545</td>
</tr>
<tr>
<td>(100, 100, B)</td>
<td>0.7505</td>
<td>0.8739</td>
<td>0.8924</td>
<td>0.7862</td>
<td>0.9285</td>
<td>0.9399</td>
</tr>
<tr>
<td>(100, 100, C)</td>
<td>0.8129</td>
<td>0.8982</td>
<td>0.9085</td>
<td>0.8578</td>
<td>0.9473</td>
<td>0.9558</td>
</tr>
<tr>
<td>(100, 100, C)</td>
<td>0.8269</td>
<td>0.9056</td>
<td>0.9067</td>
<td>0.8782</td>
<td>0.9512</td>
<td>0.9539</td>
</tr>
<tr>
<td>(100, 200, B)</td>
<td>0.7512</td>
<td>0.8526</td>
<td>0.8791</td>
<td>0.8014</td>
<td>0.9057</td>
<td>0.9265</td>
</tr>
<tr>
<td>(100, 200, B)</td>
<td>0.8115</td>
<td>0.8794</td>
<td>0.9018</td>
<td>0.8576</td>
<td>0.9287</td>
<td>0.9449</td>
</tr>
<tr>
<td>(100, 200, B)</td>
<td>0.8438</td>
<td>0.9007</td>
<td>0.9098</td>
<td>0.8927</td>
<td>0.9495</td>
<td>0.9537</td>
</tr>
<tr>
<td>(50, 50, C)</td>
<td>0.8040</td>
<td>0.8878</td>
<td>0.8970</td>
<td>0.8599</td>
<td>0.9368</td>
<td>0.9434</td>
</tr>
<tr>
<td>(50, 100, C)</td>
<td>0.8417</td>
<td>0.9006</td>
<td>0.9060</td>
<td>0.8907</td>
<td>0.9465</td>
<td>0.9515</td>
</tr>
<tr>
<td>(50, 200, C)</td>
<td>0.8576</td>
<td>0.9083</td>
<td>0.9035</td>
<td>0.9108</td>
<td>0.9553</td>
<td>0.9537</td>
</tr>
<tr>
<td>(100, 50, C)</td>
<td>0.8137</td>
<td>0.8705</td>
<td>0.8785</td>
<td>0.8651</td>
<td>0.9260</td>
<td>0.9239</td>
</tr>
<tr>
<td>(100, 100, C)</td>
<td>0.8549</td>
<td>0.8915</td>
<td>0.9049</td>
<td>0.9109</td>
<td>0.9462</td>
<td>0.9507</td>
</tr>
<tr>
<td>(100, 200, C)</td>
<td>0.8708</td>
<td>0.9019</td>
<td>0.9083</td>
<td>0.9286</td>
<td>0.9507</td>
<td>0.9531</td>
</tr>
<tr>
<td>(200, 50, C)</td>
<td>0.7992</td>
<td>0.8638</td>
<td>0.8729</td>
<td>0.8554</td>
<td>0.9154</td>
<td>0.9197</td>
</tr>
<tr>
<td>(200, 100, C)</td>
<td>0.8371</td>
<td>0.8837</td>
<td>0.8957</td>
<td>0.8949</td>
<td>0.9339</td>
<td>0.9413</td>
</tr>
<tr>
<td>(200, 200, C)</td>
<td>0.8540</td>
<td>0.8916</td>
<td>0.9029</td>
<td>0.9155</td>
<td>0.9414</td>
<td>0.9496</td>
</tr>
</tbody>
</table>

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

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

\[
\sqrt{m + n} \{\tilde{R}_{m,n}(t) - R(t)\} \xrightarrow{D} \sqrt{1 + \frac{1}{r}B_1(1 - R(t)) + \sqrt{1 + rR(t)}B_2(t)} \quad (4)
\]

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

\[
\sqrt{m} \{F_n(x) - F(x)\} \xrightarrow{D} W_1(x) \quad \text{and} \quad \sqrt{n} \{G_n(y) - G(y)\} \xrightarrow{D} W_2(y)
\]

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

\[
\begin{align*}
EW_1(x_1)W_2(x_2) &= F(x_1 \wedge x_2) - F(x_1)F(x_2) \\
EW_2(y_1)W_2(y_2) &= G(y_1 \wedge y_2) - G(y_1)G(y_2).
\end{align*}
\]

Write

\[
1 - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G(X_j)}{h} \right) - R(t) = F(G^- (1 - t)) - \int_{-\infty}^{\infty} \sum_{j=1}^{m} K \left( \frac{1 - t - G(X_j)}{h} \right) dF_m(x)
\]

\[
= F(G^- (1 - t)) - \int_{-\infty}^{\infty} F_m(x)w \left( \frac{1 - t - G(x)}{h} \right) h^{-1} dG(x)
\]

\[
= F(G^- (1 - t)) - \int_{-\infty}^{t} F_m(G^- (1 - t - xh))w(x) dx
\]
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.

<table>
<thead>
<tr>
<th>(m, n, Case)</th>
<th>NBM</th>
<th>JELM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma = 0.9$</td>
<td>$\gamma = 0.9$</td>
</tr>
<tr>
<td></td>
<td>$t = 0.1$</td>
<td>$t = 0.1$</td>
</tr>
<tr>
<td>(50, 50, A)</td>
<td>0.0879</td>
<td>0.0582</td>
</tr>
<tr>
<td>(50, 100, A)</td>
<td>0.0746</td>
<td>0.0573</td>
</tr>
<tr>
<td>(50, 200, A)</td>
<td>0.0700</td>
<td>0.0579</td>
</tr>
<tr>
<td>(100, 50, A)</td>
<td>0.0711</td>
<td>0.0448</td>
</tr>
<tr>
<td>(100, 100, A)</td>
<td>0.0623</td>
<td>0.0466</td>
</tr>
<tr>
<td>(100, 200, A)</td>
<td>0.0571</td>
<td>0.0447</td>
</tr>
<tr>
<td>(200, 50, A)</td>
<td>0.0599</td>
<td>0.0387</td>
</tr>
<tr>
<td>(200, 100, A)</td>
<td>0.0495</td>
<td>0.0374</td>
</tr>
<tr>
<td>(200, 200, A)</td>
<td>0.0441</td>
<td>0.0344</td>
</tr>
<tr>
<td>(50, 50, B)</td>
<td>0.0766</td>
<td>0.0415</td>
</tr>
<tr>
<td>(50, 100, B)</td>
<td>0.0540</td>
<td>0.0427</td>
</tr>
<tr>
<td>(50, 200, B)</td>
<td>0.0439</td>
<td>0.0449</td>
</tr>
<tr>
<td>(100, 50, B)</td>
<td>0.0702</td>
<td>0.0320</td>
</tr>
<tr>
<td>(100, 100, B)</td>
<td>0.0495</td>
<td>0.0317</td>
</tr>
<tr>
<td>(100, 200, B)</td>
<td>0.0395</td>
<td>0.0335</td>
</tr>
<tr>
<td>(200, 50, B)</td>
<td>0.0637</td>
<td>0.0268</td>
</tr>
<tr>
<td>(200, 100, B)</td>
<td>0.0444</td>
<td>0.0247</td>
</tr>
<tr>
<td>(200, 200, B)</td>
<td>0.0340</td>
<td>0.0255</td>
</tr>
<tr>
<td>(50, 50, C)</td>
<td>0.2139</td>
<td>0.1381</td>
</tr>
<tr>
<td>(50, 100, C)</td>
<td>0.1804</td>
<td>0.1259</td>
</tr>
<tr>
<td>(50, 200, C)</td>
<td>0.1583</td>
<td>0.1132</td>
</tr>
<tr>
<td>(100, 50, C)</td>
<td>0.1863</td>
<td>0.1245</td>
</tr>
<tr>
<td>(100, 100, C)</td>
<td>0.1480</td>
<td>0.1065</td>
</tr>
<tr>
<td>(100, 200, C)</td>
<td>0.1276</td>
<td>0.0916</td>
</tr>
<tr>
<td>(200, 50, C)</td>
<td>0.1686</td>
<td>0.1138</td>
</tr>
<tr>
<td>(200, 100, C)</td>
<td>0.1294</td>
<td>0.0977</td>
</tr>
<tr>
<td>(200, 200, C)</td>
<td>0.1026</td>
<td>0.0776</td>
</tr>
</tbody>
</table>

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

$$\int_{-1}^{1} \{F(G^-(1-t-xh)) - F(G^-(-1-t))\}w(x)\,dx = \frac{1}{2}h^2 \int_{-1}^{1} R''(t^*)x^2w(x)\,dx,$$

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

$$\int_{-1}^{1} \{F(G^-(1-t-xh)) - F(G^-(-1-t))\}w(x)\,dx = \frac{1}{2}h^2 \int_{-1}^{1} R''(t^*)x^2w(x)\,dx.$$

$$= F(G^-(1-t)) - F_m(G^-(1-t)) - \int_{-1}^{1} \{F(G^-(1-t-xh)) - F(G^-(1-t))\}w(x)\,dx$$

$$- \int_{-1}^{1} \{F_m(G^-(1-t-xh)) - F(G^-(1-t-xh)) - F_m(G^-(1-t)) + F(G^-(1-t))\}w(x)\,dx$$

(6)

$$\int_{-1}^{1} [F(G^-(1-t-xh)) - F(G^-(1-t))w(x)\,dx = -\int_{-1}^{1} R(t)\,dhw(x)\,dx - \frac{1}{2} \int_{-1}^{1} R''(t)(x)^2w(x)\,dx$$

$$= -\frac{1}{2}h^2 \int_{-1}^{1} R''(t^*)x^2w(x)\,dx,$$

(7)
Hence
\[
\sqrt{m} \left\{ 1 - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G(X_j)}{h} \right) - R(t) \right\} \overset{d}{\to} W_1(G^-(1-t)) \tag{9}
\]
in \(D((a, b))\).

Write
\[
\frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G(X_j)}{h} \right) = \frac{1}{m} \sum_{j=1}^{m} G(X_j) - G_n(X_j) w \left( \frac{1 - t - G(X_j)}{h} \right)
\]
\[
+ \frac{1}{2m} \sum_{j=1}^{m} \left( \frac{G(X_j) - G_n(X_j)}{h} \right)^2 w' \left( \frac{1 - t - G(X_j) + \xi_{n,j}}{h} \right),
\tag{10}
\]
where \(\xi_{n,j}\) is between 0 and \(G(X_j) - G_n(X_j)\). It follows from Theorem A of Silverman [39] that
\[
\sup_{t \in (a, b)} \left| \frac{1}{mh} \sum_{j=1}^{m} w' \left( \frac{1 - t - G(X_j)}{h} \right) - R'(t) \int_{-1}^{1} |w'(x)| \, dx \right| = o_p(1),
\tag{11}
\]
where \(R'(1-x)\) is the density of \(G(X_1)\). By (5), (10) and (11), we have
\[
\sqrt{n} \left\{ \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G(X_j)}{h} \right) \right\}
\]
\[
= - \int_{-\infty}^{\infty} W_2(x) h^{-1} w \left( \frac{1 - t - G(x)}{h} \right) \, dF(x) + O_p(n^{-1/2}h^{-1})
\]
\[
= \int_{-1}^{1} W_2(G^-(1 - tx)) h^{-1} w(x) \, dF(G^-(1 - tx)) + O_p(n^{-1/2}h^{-1})
\]
\[
= -R'(t)W_2(G^-(1-t)) + o_p(1)
\tag{12}
\]
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. \(\Box\)

**Lemma 2.** Under conditions of Theorem 1, we have
\[
\sqrt{m+n} \left\{ \frac{1}{m+n} \sum_{i=1}^{m+n} \tilde{V}_i(t) - R(t) \right\} \overset{d}{\to} N(0, \sigma^2(t))
\]
as \(n \to \infty\) for \(t = t_0\).

**Proof.** Throughout we assume \(t = t_0\). It follows from the definition of \(\tilde{V}_i(t)\) that
\[
\frac{1}{m+n} \sum_{i=1}^{m+n} \tilde{V}_i(t) = \frac{1}{m+n} \left\{ m+n - \frac{m+n}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right. \\
+ \frac{m+n}{m} \sum_{j=1}^{m} \sum_{k=1}^{m} \left\{ K \left( \frac{1 - t - G_{n,k}(X_j)}{h} \right) - K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\} \right\}.
\tag{13}
\]
Write
\[
\sum_{j=1}^{m} \sum_{k=1}^{m} \left\{ K \left( \frac{1 - t - G_{n,k}(X_j)}{h} \right) - K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\}
\]
\[
= \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{G_n(X_j) - G_{n,k}(X_j)}{h} w \left( \frac{1 - t - G_n(X_j)}{h} \right) + \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{1}{2} \left\{ \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \right\}^2 w' \left( \frac{1 - t - \xi_{n,k,j}}{h} \right)
\]
\[
= \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{1}{2} \left\{ \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \right\}^2 w' \left( \frac{1 - t - \xi_{n,k,j}}{h} \right),
\tag{14}
\]
where \( \xi_{n,k,j} \) is a random variable between \( G_{n,k}(X_j) \) and \( G_n(X_j) \). Since
\[
G_n(X_j) - G_{n,k}(X_j) = \frac{1}{n-1} \{ G_n(X_j) - I(Y_k \leq X_j) \} = O_p \left( \frac{1}{n-1} \right)
\]
uniformly in \( 1 \leq k \leq n \) and \( 1 \leq j \leq m \), we can write
\[
\xi_{n,k,j} = G_n(X_j) + O_p \left( \frac{1}{n-1} \right) = G(X_j) + O_p \left( n^{-\frac{1}{2}} \right).
\]
(15)

It follows from (14), (15) and (11) that
\[
\sum_{k=1}^{n} \sum_{j=1}^{m} \left\{ K \left( \frac{1 - t - G_{n,k}(X_j)}{h} \right) - K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\} = O_p \left\{ \frac{mn}{(n-1)^2h} \right\}.
\]
(16)

By (13), (16) and Lemma 1, we have
\[
\sqrt{m+n} \left\{ \frac{1}{m+n} \sum_{i=1}^{m+n} \hat{V}_i(t) - R(t) \right\} = \sqrt{m+n} \left\{ 1 - \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) + O_p \left\{ \frac{(m+n-1)n}{(m+n)(n-1)^2h} \right\} - R(t) \right\}
\]
\[
= \sqrt{m+n} \left\{ \hat{K}_{m,n}(t) - R(t) + O_p \left\{ \frac{(m+n-1)n}{(m+n)(n-1)^2h} \right\} \right\}
\]
\[
\xrightarrow{d} N(0, \sigma^2(t)),
\]
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 \( 1 \leq i \leq m \), 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
\[
\hat{V}_i^2(t) = \left\{ 1 - \frac{m+n-1}{m-1} K \left( \frac{1 - t - G_n(X_i)}{h} \right) \right\}^2 + \left\{ \frac{n}{(m-1)m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\}^2
\]
\[
+ 2 \left\{ \frac{n}{(m-1)m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\} \left\{ 1 - \frac{m+n-1}{m-1} K \left( \frac{1 - t - G_n(X_i)}{h} \right) \right\},
\]
which imply that
\[
\sum_{i=1}^{m} \hat{V}_i^2(t) = m - \frac{2(m+n-1)}{m-1} \sum_{i=1}^{m} K \left( \frac{1 - t - G_n(X_i)}{h} \right) + \frac{(m+n-1)^2}{(m-1)^2} \sum_{i=1}^{m} K^2 \left( \frac{1 - t - G_n(X_i)}{h} \right)
\]
\[
+ \frac{mn^2}{(m-1)^2m^2} \left\{ \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\}^2
\]
\[
+ \frac{2n}{(m-1)m} \left\{ \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\} \left\{ m - \frac{m+n-1}{m-1} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\}.
\]
(17)
Since $K^2$ is a distribution function, it follows from Lemma 1 that
\begin{equation}
\frac{1}{m} \sum_{i=1}^{m} K^2 \left( \frac{1 - t - G_n(X_i)}{h} \right) \overset{P}{\to} F(G^-(1 - t)).
\end{equation}

Hence, by (17), (18) and Lemma 1,
\begin{align}
\frac{1}{m + n} \sum_{i=1}^{m+n} \hat{V}_i^2(t) &\overset{P}{\to} \frac{r}{1 + r} - 2 F(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 + 2 r - r^2}{r(1 + r)} F(G^-(1 - t)) - \frac{1 + 2 r}{r(1 + r)} F^2(G^-(1 - t)) \\
&= \frac{r + 1}{r} R(t) - \frac{1 + 2 r}{r(1 + r)} R^2(t).
\end{align}

Next, for $m < i \leq m + n$, we can write that
\begin{equation}
\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} K \left( \frac{1 - t - G_{n,i-m}(X_j)}{h} \right) - K \left( \frac{1 - t - G_n(X_i)}{h} \right)
\end{equation}
and
\begin{align}
\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_i)}{h} \right) \right] \right\}^2 \\
&\quad + 2 \left\{ \frac{1}{m} \sum_{j=1}^{m} K \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\} \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_i)}{h} \right) \right] \right\}.
\end{align}

It follows from (11) that
\begin{align}
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 \\
&= \sum_{j=1}^{m} \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \left( \frac{1 - t - G_n(X_j)}{h} \right) + \sum_{j=1}^{m} \left\{ \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \right\}^2 \left( 1 - \frac{\xi_{n,k,j}}{h} \right) \\
&= \left\{ \sum_{j=1}^{m} \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \left( \frac{1 - t - G_n(X_j)}{h} \right) \right\}^2 + O_p(n^{-1} h^{-1})
\end{align}
which implies that
\begin{align}
\frac{1}{m + n} \sum_{k=1}^{n} A_k &= \frac{1}{m + n} \sum_{k=1}^{n} \left\{ \sum_{i=1}^{m} \frac{G_n(X_i) - G_{n,k}(X_i)}{h} \frac{G_n(X_j) - G_{n,k}(X_j)}{h} \right\} \\
&\quad \times \left( \frac{1 - t - G_n(X_i)}{h} \right) \left( \frac{1 - t - G_n(X_j)}{h} \right) + O_p(n^{-1} h^{-1}) \\
&= \frac{1}{m + n} \left( \frac{n}{(n - 1)^2 h^2} \sum_{j=1}^{m} \sum_{j=1}^{m} \left[ G_n(X_i \wedge X_j) - G_n(X_i)G_n(X_j) \right] \right) \\
&\quad + O_p(n^{-1} h^{-1}).
\end{align}
\[
\times w \left( \frac{1 - t - G_n(X_i)}{h} \right) \times w \left( \frac{1 - t - G_n(X_j)}{h} \right) + O_p(n^{-1}h^{-1})
\]
\[
= \frac{1}{m + n} \sum_{i=1}^{m+n} \left\{ G(X_i \wedge X_j) - G(X_i)G(X_j) \right\} \times w \left( \frac{1 - t - G_n(X_i)}{h} \right) \times w \left( \frac{1 - t - G_n(X_j)}{h} \right) \{1 + o_p(1)\} + O_p(n^{-1}h^{-1})
\]
\[
\xrightarrow{p} \frac{r^2}{1 + r} (1 - t)(1 - t)^2 \{R(t)^2\}.
\]
By (20), (21), (16) and Lemma 1, we have
\[
\frac{1}{m + n} \sum_{i=m+1}^{m+n} \hat{\nu}_i^2(t) \xrightarrow{p} \frac{1}{1 + r} R^2(t) + (r + 1)t(1 - t)\{R(t)^2\}.
\]  
Hence, it follows from (19), (22) and Lemma 2 that
\[
\frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t) - R(t))^2 = \frac{1}{m + n} \sum_{i=1}^{m+n} \hat{\nu}_i^2(t) + R^2(t) - \frac{2}{m + n} R(t) \sum_{i=1}^{m+n} \hat{\nu}_i(t) \xrightarrow{p} \sigma^2(t).
\]  
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 \( g(\lambda) = \frac{1}{m+n} \sum_{i=1}^{m+n} \hat{\nu}_i(t_0) - \theta \). It is easy to check that
\[
0 = |g(\lambda)| \geq \left| \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta) - \lambda \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta)^2 \right| \left| 1 + \lambda(\hat{\nu}_i(t_0) - \theta) \right|
\]
\[
\geq \left| \frac{\lambda}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta)^2 \right| - \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta)
\]
\[
\geq \frac{|\lambda| S_{m+n}}{1 + |\lambda| Z_{m+n}} - \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta),
\]
where \( S_{m+n} = \frac{1}{m+n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta)^2 \) and \( Z_{m+n} = \max_{1 \leq i \leq m+n} |\hat{\nu}_i(t_0) - \theta| \). 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
\[
|\lambda| = O_p \left\{ (m + n)^{-\frac{1}{2}} \right\}.
\]  
Put \( \gamma_1 = \lambda(\hat{\nu}_i(t_0) - \theta) \). Then, we have that
\[
\max_{1 \leq i \leq m+n} |\gamma_1| = o_p(1).
\]  
Using (23), (24) and Taylor expansion, we have
\[
0 = g(\lambda) = \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta) \left( 1 - \gamma_1 + \frac{\gamma_1^2}{1 + \gamma_1} \right)
\]
\[
= \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta) - S_{m+n} \lambda + \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta) \gamma_1^2
\]
\[
= \frac{1}{m + n} \sum_{i=1}^{m+n} (\hat{\nu}_i(t_0) - \theta) - S_{m+n} \lambda + O_p \left( \frac{1}{m + n} \right).
\]
which implies that
\[ \lambda = S_{m+n}^{-1} \frac{1}{m+n} \sum_{i=1}^{m+n} (\hat{V}(t) - \theta) + \beta_n, \]  
(25)
where \( \beta_n = O_p\left(\frac{1}{m+n}\right) \). Hence, it follows from (23), (25), Lemmas 1 and 2 that
\[ l_{m,n}(t_0, \theta) = 2 \sum_{i=1}^{m+n} \gamma_i - \sum_{i=1}^{m+n} \gamma_i^2 + 2 \sum_{i=1}^{m+n} \eta_i \]
\[ = 2(m+n) \lambda + \frac{1}{m+n} \sum_{i=1}^{m+n} \left(\hat{V}(t_0) - \theta\right) - (m+n) S_{m+n} \lambda^2 + 2 \sum_{i=1}^{m+n} \eta_i \]
\[ = \frac{(m+n) \left\{ \frac{1}{m+n} \sum_{i=1}^{m+n} (\hat{V}(t_0) - \theta) \right\}^2}{S_{m+n}} - (m+n) S_{m+n} \beta_n^2 + 2 \sum_{i=1}^{m+n} \eta_i \]
\[ = \frac{(m+n) \left\{ \frac{1}{m+n} \sum_{i=1}^{m+n} (\hat{V}(t_0) - \theta) \right\}^2}{S_{m+n}} + o_p(1) \]
\[ \overset{d}{\rightarrow} \chi_1^2. \]

i.e., Theorem 2 holds. \( \Box \)

**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**


