Smoothed jackknife empirical likelihood inference for ROC curves with missing data

Hanfang Yang and Yichuan Zhao*

Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303

email: yichuan@gsu.edu

December 3, 2011

Abstract

In this paper, we apply smoothed jackknife empirical likelihood (JEL) method to construct confidence intervals for the receiver operating characteristic (ROC) curve with missing data. After using hot deck imputation, we generate pseudo-jackknife sample to develop jackknife empirical likelihood. Comparing to traditional empirical likelihood method, the smoothed JEL has a great advantage in saving computational cost. Under mild conditions, the smoothed jackknife empirical likelihood ratio converges to a scaled chi-square distribution. Furthermore, simulation studies in terms of coverage probability and average length of confidence intervals demonstrate this proposed method has the good performance in small sample sizes.

KEY WORDS: Jackknife; Smoothed empirical likelihood; Missing data; ROC curves.

1 Introduction

The ROC curve has received considerable attention over past decades, and has been widely used in epidemiology, medical research, industrial quality control and signal detection, diagnostic medicine and material testing. In medical studies, the sensitivity or true positive rate (TPR) of the diagnostic test is the proportion of the diseased patients who have positive tests among diseased patients. The specificity or true negative rate (TNR) of the test is the proportion of the healthy people who have negative test among non-diseased people. A plot of sensitivity (TPR) against 1-specificity (FPR) defines the ROC curve, which is a graphical summary of the discriminatory accuracy of diagnostic tests. Furthermore, the ROC curve function can be represented by $ROC(p) = 1 - F(G^{-1}(1-p))$, where F and G are continuous cumulative distribution functions of positive population and negative population, respectively. Recent interesting literatures include Swets and Pickett (1982), Tosteson and Begg (1988), Hsieh and Turnbull (1996), Zou et al. (1997), Lloyd (1998), Pepe (1997), Metz et al. (1998) and Lloyd and Yong (1999), among others. Moreover, Pepe (2003) provided an excellent summary for recent research work and useful applications of ROC curves. Claeskens et al. (2003) developed smoothed empirical likelihood confidence intervals for the continuous-scale ROC curve in the absence of censoring.

Empirical likelihood (EL) is a nonparametric method for statistical inference, which employs the maximum likelihood method without having to assume a known distribution family of data. Owen (1988, 1990) introduced EL method to construct confidence regions for the mean vector. Some related literatures include the Bartlett-correctability (DiCiccio and Hall, 1991), general estimating equations (Qin and Lawless, 1994), the general plug-in EL (Hjort et al., 2009) and so on. For ROC curves, a common way of EL is to transform nonlinear constraints to linear constraints by introducing some link variables as in Claeskens et al. (2003) and Chen et al. (2009), etc. More recently, jackknife empirical likelihood method, based on jackknife pseudo-sample, becomes more attractive. Jing et al. (2009) proposed the jackknife empirical likelihood method for a U-statistic. Gong et al. (2010) demonstrated that the smoothed jackknife empirical likelihood method for the continuous-scale ROC curve can outperform EL methods with more accurate coverage probability in a smaller sample size.

The imputation-based procedure is one of the most common methods to deal with missing data problem, which highly relies on the mechanism of missing. In this paper, we assume that data are missing completely at random (MCAR), which indicates the causality of missing data is not associated with other values of observed or unobserved variables (Little and Rubin, 2002). Some imputation methods were introduced by previous researchers. For instance, using linear regression imputation, Wang and Rao (2002) addressed missing response questions based on empirical likelihood methods. By empirical likelihood method, missing data problem was also studied by Wang and Rao (2001), Qin and Zhang (2008) and Qin and Qian (2009). In this paper, we consider hot deck imputation, which is the procedure in which missing data randomly substituted by values from observed sample data. In addition, An (2010) derived smoothed empirical likelihood for the ROC curve with missing data. However, the selection of bandwidth is still disputable about kernel estimators, especially with regard to missing data.

To the best of our knowledge, no paper has addressed the problem on how to construct confidence intervals for the continuous-scale ROC curve with missing completely at random data by jackknife EL methods. In this paper, we apply smoothed jackknife EL to construct confidence intervals for the ROC curve with missing data to avoid adding extra constraints. The paper is organized as follows. Major procedures for jackknife empirical likelihood ratio are proposed in Section 2, including methods to develop smoothed empirical likelihood and asymptotic results of jackknife empirical likelihood ratio. In Section 3, we conduct simulation studies to evaluate smoothed jackknife empirical likelihood confidence intervals for continuous-scale ROC curves in small and moderate samples in terms of coverage probability and average length of confidence intervals. We make a brief discussion in Section 4. All proofs are given in the Appendix.

2 Inference Procedure

2.1 Missing data and hot deck imputation

Consider the random samples of $x_i,\ i = 1, ..., m$ in distribution F and independent missing indicators $\delta_{xi},\ i = 1, ..., m$ in Bernoulli distribution with response rate P_1, which means $P_1 = P(\delta_{xi} = 1|x_i)$. Similarly, the random samples are denoted by $y_i,\ i = 1, ..., n$ in distribution G and missing indicators $\delta_{yi},\ i = 1, ..., n$ in Bernoulli distribution with response rate P_2. Thus, we have $P_2 = P(\delta_{yi} = 1|y_i)$. Combining x_i with δ_{xi}, we can define $x_{i,m} = x_i * \delta_{xi},\ i = 1, ..., m$ as completely random missing data. Also, we have $y_{i,m} = y_i * \delta_{yi},\ i = 1, ..., n$. Denote the observed set as $X_{obs} = \{x_i : \delta_{xi} = 1\ , i = 1, ..., m\}$ and $Y_{obs} = \{y_i : \delta_{yi} = 1\ , i = 1, ..., n\}$. Then, we adopt the procedure of the hot deck imputation, replacing the missing value with values from observed set X_{obs} and Y_{obs}. Denote $r_1 = \sum_{i=1}^{m} \delta_{x_i}$, $r_2 = \sum_{j=1}^{n} \delta_{y_j}$, $m_1 = m - r_1$ and $m_2 = n - r_2$. Let $S_{rx} = \{i : \delta_{x_i} = 1\}$, $S_{mx} = \{i : \delta_{x_i} = 0\}$, $S_{ry} = \{j : \delta_{y_j} = 1\}$ and $S_{my} = \{j : \delta_{y_j} = 0\}$. x_i^* are generated by the discrete uniform distribution from observed data set X_{obs}, and y_i^* are generated by the discrete uniform distribution from observed data set Y_{obs}. Finally, we obtain the data after hot deck imputation $x_{I,i} = x_{i,m} + x_i^* * (1 - \delta_{xi}),\ i = 1, ..., m$ and $y_{I,i} = y_{i,m} + y_i^* * (1 - \delta_{yi}),\ i = 1, ..., n$.

2.2 Smoothed empirical likelihood ratio

Let F and G be the distribution functions of the diseased and non-diseased populations, respectively. The ROC curve can be written as $ROC(p) = 1 - F(G^{-1}(1 - p))$, where $0 < p < 1$ and G^{-1} denotes the quantile function of G. Denote $F_m(x) = 1/m \sum_{i=1}^{m} I(x_{I,i} \le x)$ and $G_n(y) = 1/n \sum_{j=1}^{n} I(y_{I,j} \le y)$. Let $K(p)$ be the smooth kernel function which satisfies

$$K(p) = \int_{u \le p} w(u) du,$$

where w is a symmetric density function with support $[-1, 1]$. Define the smooth estimator of ROC(p) as

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

where $h = h(n) > 0$ is a bandwidth. Define

$$\begin{aligned}
\hat{R}_{m,n,i}(p) &= 1 - \frac{1}{m-1}\sum_{1\le j\le m, j\ne i} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right), 1 \le i \le m, \\
\hat{R}_{m,n,i}(p) &= 1 - \frac{1}{m-1}\sum_{j=1}^{m} K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right), m+1 \le i \le m+n,
\end{aligned}$$

where

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

The jackknife pseudo-sample is defined is

$$\hat{V}_i(p) \quad = \quad (m+n)\hat{R}_{m,n}(p) - (m+n-1)\hat{R}_{m,n,i}(p), i = 1, \ldots, m+n. \tag{2.1}$$

The empirical likelihood ratio at $\tilde{R}(p)$, based on the $\hat{V}_i(p)$ as

$$L(\tilde{R}, p) \quad = \quad \frac{\sup\{\prod_{i=1}^{m+n}\{p_i\} : \sum_{i=1}^{m+n} p_i = 1, \sum_{i=1}^{m+n} p_i \hat{V}_i(p) = \tilde{R}, p_i > 0, i = 1, \ldots, m+n\}}{\sup\{\prod_{i=1}^{m+n}\{p_i\}\}}.$$

By using the Lagrange multiplier method, we have

$$l(\tilde{R}, p) = -2\log L(\tilde{R}, p) \quad = \quad 2\sum_{i=1}^{n} \log\{1 + \lambda(\hat{V}_i(p) - \tilde{R})\}, \tag{2.2}$$

where λ satisfies the equation

$$\sum_{i=1}^{m+n} \frac{\hat{V}_i(p) - \tilde{R}}{1 + \lambda\{\hat{V}_i(p) - \tilde{R}\}} \quad = \quad 0.$$

Define

$$v_{m,n}(p) = \frac{1}{m+n}\sum_{i=1}^{m+n}\{\hat{V}_i(p) - \frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(p)\}^2. \tag{2.3}$$

We will develop the asymptotic properties of empirical variance $v_{m,n}(p)$ and the empirical likelihood ratio statistic for the true value $R(p)$ of the ROC curve at point p based on $x_{I,i},\ i = 1, ..., m,\ y_{I,j},\ j = 1, ..., n$. These results are used to construct an asymptotic confidence interval for $R(p)$. Assume that $m/n \to \gamma$, as $m + n \to \infty$. We give the following regularity conditions. A.1. $p \in (a, b)$ for any subset $(a, b) \subset (0, 1)$;

A.2. $F(x)$ and $G(y)$ are continuous functions;

A.3. $\sup |f(x)| < \infty$ and $\sup |g(y)| < \infty$, where $f(x) = dF(x)/dx$ and $g(y) = dG(y)/dy$;

A.4. $1 > P_1 > 0$ and $1 > P_2 > 0$;

A.5. $w^{'}(u)$ is bounded by $M < \infty$ for $u \in (-1, 1)$;

A.6. The distribution function $F(x) \in \mathcal{F}$, where $\mathcal{F}$ and $\mathcal{G}$ are Donsker classes, i.e., $\mathcal{F} \in CLT(P_F)$ and $\mathcal{G} \in CLT(P_G)$, where $\mathcal{F} \in CLT(P_F)$ means $\sqrt{n}(P_{F_n} - P_F)$ converges weakly to P_F-Brownian bridge B_p which has bounded uniformly continuous sample paths almost surely.

Theorem 2.1. *Under assumptions A.1-A.6, assume conditions $h = h(n) \to 0$, $nh^2/\log n \to \infty$ and $nh^4 \to 0$ as $n \to \infty$. Then, for $p \in (a, b)$,*

$$v_{m,n}(p) \xrightarrow{\mathcal{P}} \sigma_1^2(p),$$

where

$$\sigma_1^2(p) = (1 - P_1 + P_1^{-1})\left(1 + \frac{1}{r}\right)R(p)\{1 - R(p)\} + (1 - P_2 + P_2^{-1})(1 + r)R^{'2}(p)p(1 - p).$$

Theorem 2.2. *Under the conditions of Theorem 2.1, for $p \in (a, b)$, we have*

$$l(R(p), p) \xrightarrow{\mathcal{D}} c(p)\chi_1^2,$$

where $R(p)$ is true ROC curve at p,

$$c(p) = \frac{\sigma_1^2(p)}{\sigma_2^2(p)},$$

$$\sigma_2^2(p) = \left(1 + \frac{1}{r}\right) R(p)(1 - R(p)) + (1 + r)R^{'2}(p)p(1 - p).$$

We may use the consistent estimator $\hat{c}$ of $c(p)$ to construct our confidence intervals of $R(p)$ given p. Thus, the asymptotic $100(1-\alpha)\%$ smoothed jackknife EL confidence interval for $R(p)$ is given by

$$I(p) = \left\{\tilde{R} : l(\tilde{R}, p) \leq \hat{c}\chi_1^2(\alpha)\right\},$$

where $\chi_1^2(\alpha)$ is the upper α-quantile of χ_1^2.

3 Numerical Studies

In this section, we conduct simulation studies to compare the performance of jackknife empirical likelihood (JEL) method and smoothed empirical likelihood (SEL) proposed by An (2010) for the ROC curve in terms of coverage accuracy and average length of confidence intervals with various distributions, response rates and sample sizes. In the simulation studies, distributions of the diseased population (X) and the non-diseased population (Y) are represented by $F(x)$ and $G(y)$. We consider three scenarios, which are $(\mathbf{A})F \sim N(0.2, 0.5), G \sim N(0, 0.5)$, $(\mathbf{B})F \sim Exp(1), G \sim N(0, 0.5)$ and $(\mathbf{C})F \sim Exp(1), G \sim Exp(1)$. Random samples x and y are independently drawn from populations X and Y. The response rates for data x and y are chosen as, $(P_1, P_2) = (0.7, 0.6)$ or $(0.9, 0.8)$. The sample sizes for x and y are $(m, n) = (50, 50)$, $(100, 100)$ and $(200, 150)$. For certain response rate and sample size, we generate 1000 independent random samples of missing data. Without the loss of generality, we use both methods to construct confidence intervals for ROC curves at p$= 0.2, 0.3$. The significant level of the confidence intervals is $1 - \alpha = 0.95$. Then, the Epanechnikov kernel

$$K(u) = \begin{cases} \frac{3}{4}(1 - u^2) & \text{if } |u| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

is used for both JEL method and SEL method (see An, 2010) and the smoothing parameter is chosen to be $h = n^{-1/3}$ for JEL method and $h_1 = m^{-1/3}$ and $h_2 = n^{-1/3}$ for SEL method. The simulation results of coverage probability are illustrated in Table 1. From Table 1, we find out that JEL method has much better performance than SEL (smoothed empirical likelihood) method in the most simulation settings.

Next, we investigate the performance of average length of ROC curves using JEL method and SEL method. We arrange the same simulation settings as ones which are specified for coverage probability. To obtain the average length, we applied the bisection method to search solutions. It does not involve high computation costs because jackknife method can dramatically simplify the complexity of equations. This is one of main advantages of the smoothed jackknife EL method. The results are given in Table 2 which shows comparable results between JEL method and SEL method based on average length. Generally, the JEL method has better coverage probability and similar average length in small samples, compared with the traditional SEL method.

[Table 1 about here.]

[Table 2 about here.]

In addition, we study empirical likelihood confidence intervals for ROC curves at different specificities generated by simulated data. Data A is employed in this case. We choose two sample sizes $(50, 50)$ and $(200, 200)$ with different response rates $(1, 1)$ and $(0.8, 0.9)$. We select 100 points on the ROC curve evenly to test JEL confidence intervals respectively. As Figure 1 shows, it is clear that the confidence intervals of ROC curves are located above the diagonal line, which indicates two distributions can be distinguished clearly by ROC curves. In addition, the ROC curve has the shorter confidence intervals when the sample sizes are larger.

[Figure 1 about here.]

4 Discussion

In this paper, we apply jackknife empirical likelihood method to construct confidence intervals for the continuous-scale ROC curve with missing data. The theoretical results provide the asymptotic properties, including asymptotic variance and limiting distribution of empirical likelihood ratio statistics. The simulation results demonstrate that coverage probability of EL confidence interval can be close to nominal level at various response rates, distributions of data and different locations of the ROC curve. Comparing with traditional SEL methods, JEL methods can be recognized as less computational cost and a more precise coverage probability and similar average length.

There are other topics, which should be investigated in the future. For instance, combining jackknife empirical likelihood method, imputation methods could be applied to solve other missing data problems. Moreover, we may consider to develop smoothed jackknife empirical likelihood method for ROC curves with other kinds of incomplete data, such as right censoring data and current status data. Furthermore, if we generally assume that data are missing at random rather than missing completely at random, it is challenging to prove the corresponding Donsker's theorem and Glivenko-Cantelli Theorem.

References

An, Yueheng, Smoothed empirical likelihood inference for ROC curves with missing data, *Open Journal of Statistics*, 2011.

Bickel, P. J. and Freedman, D. A., Some asymptotic theory for the bootstrap, *The Annals of Statistics*, Vol. 9, 1196-1217, 1981.

Claeskens, G., Jing, B.Y., Peng, L. and Zhou, W., Empirical likelihood confidence regions for comparison distributions and ROC curves, *Canadian J. Statist.*, Vol. 31, 173-190, 2003.

Chen, J. and Rao, J.N.K., Asymptotic normality under two-phase sampling designs, *Statist. Sinica*, Vol.17, 1047-1064, 2007.

DiCiccio, T. and Hall, P., Empirical likelihood is bartlett-correctable, *The Annals of Statistics*, Vol.19, 1053-1061, 1991.

Diehl, S. and Stute, W., Kernel density and hazard function estimation in the presence of censoring, *J. Multivariate Anal.*, Vol. 25, 299-310, 1988.

Einmahl, J.H. and McKeague, I.W., Confidence tubes for multiple quantile plots via empirical likelihood, *Ann. Statist.*, Vol. 27, 1348-1367, 1999.

Fleming, T.R. and Harrington, D.P., Counting Processes and Survival Analysis, *Wiley*, 1991.

Giné, E. and Zinn, J., Bootstrapping general empirical measures, *Annals of Probability*, Vol. 18, 851-869, 1990.

Gong, Y., Peng, L. and Qi, Y., Smoothed jackknife empirical likelihood method for ROC curve, *Journal of Multivariate Analysis*, Vol. 70, 1520-1531, 2010.

Hjort, N.L., McKeague, I. W. and Van Keilegom, I., Extending the scope of empirical likelihood, *Annals of Statistics*, Vol. 37, 1079-1111, 2007.

Hsieh, F. and Turnbull, B.W., Non-parametric and semi-parametric estimation of the receiver operating characteristic curve, *The Annals of Statistics*, Vol. 24, 25-40, 1996.

Kosorok, M., Bootstrapping the Grenander estimator, *Institute of Mathematical Statistics*, Vol. 1, 282-292, 2008.

Li, G. and Van Keilegom, I., Likelihood ratio confidence bands in nonparametric regression with censored data, *Scand. J. Statist.*, Vol. 29, 547–562, 2002.

Little, R. and Rubin, D., Statistical Analysis with Missing Data, *John Wiley, New York*, 2002.

Lloyd, C.J., The use of smoothed ROC curves to summarize and compare diagnostic systems, *Journal of the American Statistical Association*, Vol. 93, 1356-1364, 1998.

Lloyd, C.J. and Yong, Z., Kernel estimators for the ROC curve are better than empirical, *Statistics and Probability Letters*, Vol. 44, 221-228, 1999.

Metz, C.E., Herman J., B.A. and Shen, H., Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data, *Statistics in Medicine*, Vol. 17, 1033-1053, 1998.

Mojirsheibani, M., The Glivenko-Cantelli theorem based on data with randomly imputed missing values, *Statistics and Probability Letters*, Vol.55, 385-396, 2001.

Owen, A., Empirical likelihood ratio confidence intervals for a single functional, *Biometrika*, Vol. 75, 237-249, 1988.

Owen, A., Empirical likelihood and confidence regions, *Ann. Statist.*, Vol. 18, 90-120, 1990.

Owen, A., *Empirical Likelihood*, Chapman & Hall, New York, 2001.

Pepe, M.S., A regression modeling framework for ROC curves in medical diagnostic testing, *Biometrika*, Vol.84, 595-608, 1997.

Pepe, M.S., The Statistical Evaluation of Medical Tests for Classification and Prediction, *Oxford University Press*, 2003.

Qin, J. and Lawless, J., Empirical likelihood and general estimating equations, *Annals of Statistics*, Vol. 22, 300-325, 1997.

Qin, Y.S. and Qian, Y.J., Empirical likelihood confidence intervals for the differences of quantiles with missing data, *Acta Mathematicae Applicatae Sinica*, Vol. 25, 105-116, 2009.

Qin, Y. and Zhang, S., Empirical likelihood confidence intervals for differences between two datasets with missing data, *Pattern Recognition Letters*, Vol. 29, 803-812, 2008.

Qin, Y.S. and Zhao, L.C., Empirical likelihood ratio intervals for various differences of two populations, *Chinese Systems Sci. Math. Sci.*, Vol. 13, 23-30, 2000.

Sen, P. K and Singer, J. M., Large sample methods in statistics, *Chapman and Hall Inc*, 2000.

Singh, K., On the asymptotic accuracy of Efron's bootstrap, *The Annals of Statistics*, Vol. 9, 1187-1195, 1981.

Stute W., On almost sure convergence of conditional empirical distribution functions, *Annals of Probability*, Vol.14, 891-901, 1986.

Swets, I.A. and Pickett, R.M., Evaluation of Diagnostic Systems: Methods from Signal Detection Theory, *New York: Academic Press*, 1982.

Tosteson, A.N. and Begg, C.B., A general regression methodology for ROC curve estimation, *Medical Decision Making*, Vol. 8, 205-215, 1988.

van der Vaart, A. W., Asymptotic Statistics, *Cambridge University Press*, 1998.

van der Vaart, A. W. and Wellner, J., Weak Convergence and Empirical Processes: With Applications to Statistics, *Springer Series in Statistics*, 1996.

Wang, Q.H. and Rao, J.N.K., Empirical likelihood for linear regression models under imputation for missing responses, *Canad. J. Statist.*, Vol. 29, 597-608, 2001.

Wang, Q.H. and Rao, J.N.K., Empirical likelihood-based inference in linear models with missing data, *Scand. J. Statist.*, Vol. 29, 563-576, 2002.

Wellner, J., Bootstrap limit theorems: a partial survey, *Elsevier Science Pulishers*, 313-329, 1992.

Zou, K.H., Hall, W.J. and Shapiro, D.E., Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests, *Stat Med.*, Vol. 16, 2143-2156, 1997.

Appendix: Proofs of Theorems

Lemma A.1. *Under conditions in Theorem 2.1, as $n \longrightarrow \infty$, we have*

$$\sqrt{m+n}\left\{\hat{R}_{m,n}(p) - R(p)\right\} \xrightarrow{\mathfrak{D}} N(0, \sigma_1^2(p)),$$

where $\sigma_1^2(p)$ is defined in Theorems 2.1 and $R(p)$ is the true ROC curve at point $p \in (a, b)$.

Proof. Since $R^{''}$ is continuous at $p \in (a, b)$, $R^{'}$ and $R^{''}$ are bounded in (a, b). Denote $\sigma - algebra$ $\mathcal{B}_{r_1} = \{\sigma(x_i, \delta_{xi}, i \in S_{r_1})\}$ and $\mathcal{B}_m = \{\sigma(x_i, \delta_{xi}, i = 1, ..., m)\}$. Because x_i^* are only dependent on $\mathcal{B}_{r_1}$, from Qin and Zhang (2008), we have

$$E(I(x_i^* \leq x)|\mathcal{B}_{r_1}) = E(I(x_i^* \leq x)|\mathcal{B}_m) = \frac{1}{r_1}\sum_{i\in S_{r_1}} I(x_i \leq x).$$

and

$$\begin{aligned}\sqrt{m}\{F_m(x) - F(x)\} = & \frac{\sqrt{m}}{\sqrt{r_1}}\frac{1}{\sqrt{r_1}}\sum_{i\in S_{r_1}}\{I(x_i \leq x) - F(x)\} \\ & + \frac{\sqrt{m_1}}{\sqrt{m}}\frac{1}{\sqrt{m_1}}\sum_{i\in S_{m_1}}\{I(x_i^* \leq x) - E(I(x_i^* \leq x)|\mathcal{B}_{r_1})\}.\end{aligned}$$

Denote the first term as $V_m(x)$ and the second term as $U_m(x)$. The response rate $P_1 > 0$ assures the $r_1 \to \infty$ and $m_1 \to \infty$ when $m \to \infty$. We define the empirical distribution

$F_{r1}(x) = 1/r_1 \sum_{i \in S_{r_1}} I(x_i \leq x)$ of $x_1, \ldots, x_m$ and $x_i^*, \quad i \in S_{m_1}$ with the distribution function $F_{r1}(x)$. Denote $F_{r_1,m_1}(x) = 1/m_1 \sum_{i \in S_{m_1}} \{I(x_i^* \leq x)\}$. $F_{r_1,m_1}(x)$ is the empirical distribution of $x_1^*, \ldots, x_{m_1}^*$ and the bootstrapped version of $F_{r1}(x)$ with weighting mechanism $\mathcal{M}_{m_1}$ which is independent of $\mathcal{B}_{r_1}$ since we can rewrite that $F_{r_1,m_1}(x) = 1/m_1 \sum_{i \in S_{r_1}} \left\{M_{m_1,i}^* I(x_i \leq x)\right\}$ from the equation (4.4) in Wellner (1992), where the weight, $\mathcal{M}_{m_1} = \{M_{m_1,1}^*, \ldots, M_{m_1,m_1}^*\}$, follows multinomial distribution. By Theorem 4.1 of Bickel and Freedman (1981), we have $\sqrt{m_1}\{F_{r_1,m_1}(x) - F_{r1}(x)\} \Longrightarrow B(F(x))$, where $B(\cdot)$ is the Brownian bridge on $[0,1]$. Hence, $E(U_m(s)U_m(t)) \xrightarrow{P} (1-P_1)\{F(\min(s,t)) - F(s)F(t)\}$ and $E(U_m(x)|\mathcal{B}_{r_1}) = 0$. By Donsker's theorem and multivariate central limit theorem from Theorem 19.3 of van der Vaart (1998), $\sqrt{r_1}\,\{F_{r1}(x) - F(x)\} \Longrightarrow B(F(x))$ and $B(F(x))$ is tight. $E(V_m(s)V_m(t)) \xrightarrow{P} P_1^{-1}\{F(\min(s,t)) - F(s)F(t)\}$, where $B(\cdot)$ is the Brownian bridge on $[0,1]$. Then, we consider

$$(V_m(x), U_m(x)) = \left(\frac{\sqrt{m}}{\sqrt{r_1}} \sqrt{r_1}\,\{F_{r1}(x) - F(x)\}\,, \frac{\sqrt{m_1}}{\sqrt{m}} \sqrt{m_1}\{F_{r_1,m_1}(x) - F_{r1}(x)\} \right).$$

We know that Brownian bridge $B(F(x))$ is tight and $V_m(x)$ and $U_m(x)$ marginally converge to Brownian bridge, i.e., $V_m(x) \Longrightarrow \sqrt{1-P_1} B(F(x))$ and $U_m(x) \Longrightarrow \sqrt{P_1^{-1}} B(F(x))$, respectively. From the equation (3.2) in Giné and Zinn (1990), we know that $U_m(x) \underset{W}{\overset{P}{\rightsquigarrow}} \sqrt{P_1^{-1}} B(F(x))$, where weak convergence $\underset{W}{\overset{P}{\rightsquigarrow}}$ is defined as follows by Kosorok (2008),

$$\sup_{h \in BL_1(\mathcal{F})} \|E_{\cdot|\mathcal{B}_{r_1}} h\{U_m(x)\} - Eh\{\sqrt{P_1^{-1}} B(F(x))\}\| \to 0.$$

Note $\mathcal{M}_{m_1}$ is measurable conditional on $\mathcal{B}_{r_1}$. $V_m(s)$ and $U_m(s)$ are uncorrelated since

$$E(V_m(s)U_m(s)) = E(V_m(s)E(U_m(s)|\mathcal{B}_{r_1})) = 0.$$

By p.180 in van der Vaart and Wellner (1996) and Theorem 2.2 in Kosorok (2008),

$$(V_m(x), U_m(x)) \Longrightarrow \left(\sqrt{P_1^{-1}} \tilde{B}_1(F(x)), \sqrt{1-P_1} \tilde{B}_2(F(x)) \right),$$

where $\tilde{B}_1(F(x))$ and $\tilde{B}_2(F(x))$ are independent copies of $B(F(x))$. The sequence converges jointly in distribution to two independent Brownian bridges, which implies that

$$W_1(x) = \sqrt{m}\{F_m(x) - F(x)\} = V_m(x) + U_m(x) \Longrightarrow \sqrt{1 - P_1 + P_1^{-1}}B(F(x)). \quad \text{(A.1)}$$

Similarly, we have

$$W_2(y) = \sqrt{n}\{G_n(x) - G(x)\} \Longrightarrow \sqrt{1 - P_2 + P_2^{-1}}B(G(x)).$$

Then, we consider the uniform convergence of empirical distribution function after hot deck imputation. Mojirsheibani (2001) derived the Glivenko-Cantelli Theorem with completely randomly missing data.

$$\sup_{x\in\mathcal{R}} |F_m(x) - F(x)| \longrightarrow 0 \quad a.s. \quad and \quad \sup_{y\in\mathcal{R}} |G_n(y) - G(y)| \longrightarrow 0 \quad a.s. \quad \text{(A.2)}$$

We write

$$\begin{aligned}
&1 - \frac{1}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G(x_{I,j})}{h}\right) - R(p) \\
&= F(G^{-1}(1-p)) - \int_{-\infty}^{\infty} K\left(\frac{1-p-G(x)}{h}\right) \mathrm{d}F_m(x) \\
&= F(G^{-1}(1-p)) - K\left(\frac{1-p-G(x)}{h}\right) F_m(x)|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} F_m(x)\mathrm{d}K\left(\frac{1-p-G(x)}{h}\right)
\end{aligned}$$

$$= F\{G^{-1}(1-p)\} - K\left(\frac{-p}{h}\right) - \int_{-\infty}^{\infty} F_m(x) w\left(\frac{1-p-G(x)}{h}\right) h^{-1} \mathrm{d}G(x)$$
$$= F\{G^{-1}(1-p)\} - K\left(\frac{-p}{h}\right) - \int_{-1}^{1} F_m\{G^{-1}(1-p-xh)\} w(x) \mathrm{d}x$$
$$= F\{G^{-1}(1-p)\} - F_m(G^{-1}(1-p)) - \int_{-1}^{1} \left[F_m\{G^{-1}(1-p-xh)\} - F_m\{G^{-1}(1-p)\}\right] w(x) \mathrm{d}x$$
$$= F\{G^{-1}(1-p)\} - F_m\{G^{-1}(1-p)\} - \int_{-1}^{1} \left[F\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p)\}\right] w(x) \mathrm{d}x$$
$$- \int_{-1}^{1} ([F_m\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p-xh)\}]$$
$$- \left[F_m\{G^{-1}(1-p)\} - F\{G^{-1}(1-p)\}\right]) w(x) \mathrm{d}x$$

Because $-p/h$ is beyond the support of kernel function K as $h \to 0$, $K(-p/h) = 0$ when $p \in (a, b)$.

$$\begin{aligned}
&\int_{-\infty}^{\infty} [F\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p)\}] w(x) \mathrm{d}x \\
&= -\int_{-p/h}^{(1-p)/h} R^{'}(p) xh w(x) \mathrm{d}x - \frac{1}{2} \int_{-p/h}^{(1-p)/h} R^{''}(p^*)(xh)^2 w(x) \mathrm{d}x \\
&= -\frac{1}{2} \int_{-1}^{1} R^{''}(p^*)(xh)^2 w(x) \mathrm{d}x \\
&= O(h^2),
\end{aligned} \tag{A.3}$$

where p^* is between p and $p + xh$. Because $p \in (a, b)$, we have $F_m\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p-xh)\} - m^{-1/2}\sqrt{1 - P_2 + P_2^{-1}} B[F\{G^{-1}(1-p-xh)\}] = o_p(m^{-1/2})$, for any $x \in [-1, 1]$. Using the conditions on h and the continuity of $B_F(x)$,

$$\int_{-p/h}^{(1-p)/h} (F_m\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p-xh)\} - [F_m\{G^{-1}(1-p)\}$$
$$- F\{G^{-1}(1-p)\}]) w(x) \mathrm{d}x$$

$$
\begin{aligned}
&= \int_{-1}^{1} F_m\{G^{-1}(1-p-xh)\} - F\{G^{-1}(1-p-xh)\} \\
&\quad - m^{-1/2}\sqrt{1-P_2+P_2^{-1}}B[F\{G^{-1}(1-p-xh)\}]w(x)\mathrm{d}x \\
&\quad - \int_{-1}^{1} \left\{ F_m(G^{-1}(1-p)) - F(G^{-1}(1-p)) - m^{-1/2}\sqrt{1-P_2+P_2^{-1}}B(F(G^{-1}(1-p))) \right\} w(x)\mathrm{d}x \\
&\quad + \sqrt{1-P_2+P_2^{-1}} \int_{-1}^{1} \left\{ m^{-1/2}B[F\{G^{-1}(1-p-xh)\}] - m^{-1/2}B[F\{G^{-1}(1-p)\}] \right\} w(x)\mathrm{d}x \\
&= o_p(m^{-1/2}). \qquad \text{(A.4)}
\end{aligned}
$$

Hence, by (A.1), (A.3), (A.5) and (A.6), we have

$$
\begin{aligned}
&\sqrt{m}\left[1 - \frac{1}{m}\sum_{j=1}^{m} K\left\{\frac{1-p-G(x_{I,j})}{h}\right\} - R(p)\right] \\
&= \sqrt{m}[F\{G^{-1}(1-p)\} - F_m\{G^{-1}(1-p)\}] + o_p(m^{-1/2}m^{1/2}) + O(m^{1/2}h^2) \\
&\xrightarrow{\mathfrak{D}} N(0, (1-P_1+P_1^{-1})F\{G^{-1}(1-p)\}\{1-F[G^{-1}(1-p)]\}) \\
&= N(0, (1-P_1+P_1^{-1})R(p)\{1-R(p)\}. \qquad \text{(A.5)}
\end{aligned}
$$

Write

$$
\begin{aligned}
&\frac{\sqrt{n}}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) - \frac{\sqrt{n}}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G(x_{I,j})}{h}\right) \\
&= \int_{-\infty}^{\infty} K\left(\frac{1-p-G_n(x)}{h}\right)\mathrm{d}\sqrt{n}F_m(x) - \int_{-\infty}^{\infty} K\left(\frac{1-p-G(x)}{h}\right)\mathrm{d}\sqrt{n}F_m(x). \qquad \text{(A.6)}
\end{aligned}
$$

Notice that

$$
\begin{aligned}
&\int_{-\infty}^{\infty} K\left(\frac{1-p-G_n(x)}{h}\right)\mathrm{d}\sqrt{m}\{F_m(x)-F(x)\} - \int_{-\infty}^{\infty} K\left(\frac{1-p-G(x)}{h}\right)\mathrm{d}\sqrt{m}\{F_m(x)-F(x)\} \\
&= K\left(\frac{1-p-G_n(x)}{h}\right)\sqrt{m}\{F_m(x)-F(x)\} \mid_{-\infty}^{\infty} - K\left(\frac{1-p-G(x)}{h}\right)\sqrt{m}\{F_m(x)-F(x)\} \mid_{-\infty}^{\infty} \\
&- \int_{-\infty}^{\infty} W_1(x)\mathrm{d}K\left(\frac{1-p-G_n(x)}{h}\right) + \int_{-\infty}^{\infty} W_1(x)\mathrm{d}K\left(\frac{1-p-G(x)}{h}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{h}\int_{-\infty}^{\infty} W_1(x) w\left(\frac{1-p-G(x)}{h}\right) \mathrm{d}G(x) - \frac{1}{h}\int_{-\infty}^{\infty} W_1(x) w\left(\frac{1-p-G_n(x)}{h}\right) \mathrm{d}G_n(x)\\
&= \int_{-1}^{1} W_1\{G^{-1}(1-p-hu)\} w(x)\mathrm{d}u - \int_{-1}^{1} W_1\{G_n^{-1}(1-p-hu)\} w(x)\mathrm{d}u\\
&= \sqrt{1-P_2+P_2^{-1}}\int_{-1}^{1} B[F\{G^{-1}(1-p-hu)\}] - B[F\{G_n^{-1}(1-p-hu)\}] w(x)\mathrm{d}u + o_p(1)\\
&= o_p(1),
\end{aligned}
$$

because of the continuity of $B(F(x))$ and the proof in P. 1525 of Gong et al. (2010). Thus, we can adjust the (A.7) as follows

$$
\begin{aligned}
&\int_{-\infty}^{\infty} K\left(\frac{1-p-G_n(x)}{h}\right) d\sqrt{n}F_m(x) - \int_{-\infty}^{\infty} K\left(\frac{1-p-G(x)}{h}\right) \mathrm{d}\sqrt{n}F_m(x)\\
&= \sqrt{n}\int_{-\infty}^{\infty} K\left(\frac{1-p-G_n(x)}{h}\right) - K\left(\frac{1-p-G(x)}{h}\right) \mathrm{d}F(x)\\
&= \sqrt{n}\int_{-\infty}^{\infty} \left(\frac{G(x)-G_n(x)}{h}\right) w\left(\frac{1-p-G(x)}{h}\right) \mathrm{d}F(x)\\
&+ \sqrt{n}\int_{-\infty}^{\infty} \frac{1}{2}\left(\frac{G(x)-G_n(x)}{h}\right)^2 w^{'}\left(\frac{1-p-G(x)+\xi_x}{h}\right) \mathrm{d}F(x). \qquad \text{(A.7)}
\end{aligned}
$$

Denote $R^{'}(p)$ as the first derivative of $R(p)$. The Brownian bridge $B_1(G(x))$ and $B_2(G(x))$ are uniformly bounded for $x \in (a, b)$. Also, we have the continuities of Brownian bridge $B(\cdot)$ and $R^{'}(p)$. Thus, $B(x)$ is uniformly bounded. We have

$$
\begin{aligned}
&\sqrt{n}\int_{-\infty}^{\infty} \frac{1}{2}\left(\frac{G(x)-G_n(x)}{h}\right)^2 w^{'}\left(\frac{1-p-G(x)+\xi_x}{h}\right) \mathrm{d}F(x)\\
=&\frac{1}{2\sqrt{n}h^2}\int_{-\infty}^{\infty} \left\{\sqrt{n}(G(x)-G_n(x))\right\}^2 w^{'}\left(\frac{1-p-G(x)+\xi_x}{h}\right) \mathrm{d}F(x)\\
=&\frac{1}{2\sqrt{n}h^2}\int_{-1}^{1} \left[W_2\{G^{-1}(1-p-uh+\xi_x)\}\right]^2 w^{'}(u)\mathrm{d}F\{G^{-1}(1-p-uh+\xi_x)\}\\
=&\frac{1}{2\sqrt{n}h}\int_{-1}^{1} \left[W_2\{G^{-1}(1-p-uh+\xi_x)\}\right]^2 w^{'}(u)R^{'}(p+uh+\xi_x)\mathrm{d}u\\
=&\frac{\sqrt{1-P_2+P_2^{-1}}}{2\sqrt{n}h}\int_{-1}^{1} \left\{B[G\{G^{-1}(1-p-uh+\xi_x)\}]\right\}^2 w^{'}(u)R^{'}(p+uh+\xi_x)\mathrm{d}u + o_p(1)\\
=&\frac{\sqrt{1-P_2+P_2^{-1}}}{2\sqrt{n}h}\int_{-1}^{1} \left\{B(1-p-uh+\xi_x)\right\}^2 w^{'}(u)R^{'}(p+uh+\xi_x)\mathrm{d}u + o_p(1)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\sqrt{1-P_2+P_2^{-1}}}{2\sqrt{n}h}\int_{-1}^{1}\{B(1-p)\}^2 w^{'}(u)R^{'}(p)\mathrm{d}u + o_p(1) \\
&= o_p(1). \qquad (A.8)
\end{aligned}
$$

Recall that

$$
\begin{aligned}
&\sqrt{n}\int_{-\infty}^{\infty}\left(\frac{G(x)-G_n(x)}{h}\right) w\left(\frac{1-p-G(x)}{h}\right)\mathrm{d}F(x) \\
&= \int_{-1}^{1} W_2\{G^{-1}(1-p-hu)\}w(u)h^{-1}dF(G^{-1}(1-p-hu)) \\
&= \int_{-1}^{1}\sqrt{1-P_2+P_2^{-1}}B[G\{G^{-1}(1-p-hu)\}]w(u)(R^{'}(p+uh))du + o_p(1) \\
&= \sqrt{1-P_2+P_2^{-1}}R^{'}(p)B(1-p)\int_{-1}^{1} w(u)du + o_p(1) \\
&\xrightarrow{\mathfrak{D}} N(0,(1-P_2+P_2^{-1})p(1-p)R^{'2}(p)). \qquad (A.9)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&\sqrt{m+n}\{\hat{R}_{m,n}(p)-R(p)\} \\
&= \frac{\sqrt{m+n}}{\sqrt{n}}\sqrt{n}\left\{\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G(x_{I,j})}{h}\right)-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\} \\
&+ \frac{\sqrt{m+n}}{\sqrt{m}}\sqrt{m}\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G(x_{I,j})}{h}\right)-R(p)\right\}.
\end{aligned}
$$

Combining (A.6), (A.8), (A.9) and (A.10) and the independence of first term and second term, we can obtain the conclusion as follows,

$$
\sqrt{m+n}\{\hat{R}_{m,n}(p)-R(p)\} = N(0,\sigma_1^2(p)).
$$

□

Lemma A.2. *Under conditions in Theorem 2.1, for $p \in (a,b)$, as $n \longrightarrow \infty$, we have*

$$\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(p) - R(p)\right\} \xrightarrow{\mathfrak{P}} N(0,\sigma_1^2(p)),$$

where $\sigma_1^2(p)$ is defined in Lemma A.1.

Proof. From the definition of $\hat{V}_i(p)$, we have

$$\begin{aligned}
&\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(p)\\
&= \frac{1}{m+n}\sum_{i=1}^{m+n}\{(m+n)\hat{R}_{m,n}(p) - (m+n-1)\hat{R}_{m,n,i}(p)\}\\
&= \frac{1}{m+n}\sum_{i=1}^{m}[(m+n)\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\\
&\quad -(m+n-1)\left\{1-\frac{1}{m-1}\sum_{j=1,j\neq i}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}]\\
&\quad +\frac{1}{m+n}\sum_{i=m+1}^{m+n}[(m+n)\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\\
&\quad -(m+n-1)\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right)\right\}]\\
&= \frac{1}{m+n}\sum_{i=1}^{m}1-\frac{m+n}{m}\left\{\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)-\sum_{j=1,j\neq i}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\\
&\quad +\left(\frac{m+n-1}{m-1}-\frac{m+n}{m}\right)\sum_{j=1,j\neq i}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\\
&\quad +\frac{1}{m+n}\sum_{i=1}^{n}[1+\frac{m+n-1}{m}\{\sum_{j=1}^{m}K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right)\\
&\quad -\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\}+\left(\frac{m+n-1}{m}-\frac{m+n}{m}\right)\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)]
\end{aligned}$$

$$= \frac{1}{m+n}(m+n-\frac{m+n}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)$$
$$+\frac{m+n-1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\})$$
$$+\frac{1}{m+n}\{\left(\frac{m+n-1}{m-1}-\frac{m+n}{m}\right)(m-1)$$
$$+\left(\frac{m+n-1}{m}-\frac{m+n}{m}\right)n\}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)$$
$$= \frac{1}{m+n}(m+n-\frac{m+n}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)$$
$$+\frac{m+n-1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}). \quad (A.10)$$

Write

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{G_{n,i}(x_{I,j})-G_n(x_{I,j})}{h}w\left(\frac{1-p-G_n(x_{I,j})}{h}\right)$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_{n,i}(x_{I,j})-G_n(x_{I,j})}{h}\right\}^2 w^{'}\left(\frac{1-p-\xi_{n,i,j}}{h}\right)$$
$$=\sum_{j=1}^{m}\left\{\sum_{i=1}^{n}\frac{G_{n,i}(x_{I,j})-G_n(x_{I,j})}{h}\right\}w\left(\frac{1-p-G_n(x_{I,j})}{h}\right)$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_{n,i}(x_{I,j})-G_n(x_{I,j})}{h}\right\}^2 w^{'}\left(\frac{1-p-\xi_{n,i,j}}{h}\right)$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{1}{2}\left\{\frac{G_{n,i}(x_{I,j})-G_n(x_{I,j})}{h}\right\}^2 w^{'}\left(\frac{1-p-\xi_{n,i,j}}{h}\right), \quad (A.11)$$

where $\xi_{n,i,j}$ is between the $G_n(x_{I,j})$ and $G_{n,i}(x_{I,j})$,

$$G_n(x_{I,j})-G_{n,i}(x_{I,j})=\frac{1}{n-1}\{G_n(x_{I,j})-I(Y_{I,i}\le x_{I,j})\}=O_p\left(\frac{1}{n-1}\right), \quad (A.12)$$

and

$$\sum_{i=1}^{n}\{G_{n,i}(x_{I,j})-G_n(x_{I,j})\}=0,$$

because

$$
\begin{aligned}
&G_n(x_{I,j}) - G_{n,i}(x_{I,j}) \\
&= \frac{1}{n}\sum_{k=1}^{n} I(y_{I,i} \le x_{I,j}) - \frac{1}{n-1}\sum_{i=k,k\neq i}^{n} I(y_{I,i} \le x_{I,j}) \\
&= \left(\frac{1}{n} - \frac{1}{n-1}\right)\sum_{k=1}^{n} I(y_{I,i} \le x_{I,j}) - \frac{1}{n-1}\{\sum_{k=1}^{n} I(y_{I,i} \le x_{I,j}) - \sum_{i=k,k\neq i}^{n} I(y_{I,i} \le x_{I,j})\} \\
&= \frac{1}{n-1}\{G_n(x_{I,j}) - I(y_{I,i} \le x_{I,j})\}.
\end{aligned}
$$

By similar steps in (A.12) and (A.13), we have

$$
\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right) - K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\} = O_p\left(\frac{mn}{(n-1)^2h}\right). \tag{A.13}
$$

Combining (A.11), (A.14) and Lemma A.1, we have

$$
\begin{aligned}
&\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i(p) - R(p)\right\} \\
&= \sqrt{m+n}\left\{1 - \frac{1}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) + O_p\left(\frac{m+n-1}{(m+n)m}\frac{mn}{h(n-1)^2}\right) - R(p)\right\} \\
&= \sqrt{m+n}\left\{\hat{R}_{m,n}(p) - R(p) + O_p\left(\frac{(m+n-1)n}{(m+n)(n-1)^2h}\right)\right\} \\
&\xrightarrow{\mathfrak{D}} N(0, \sigma_1^2(p)).
\end{aligned}
$$

□

Lemma A.3. *Under conditions in Theorem 2.1, for $p \in (a,b)$, as $n \longrightarrow \infty$, we have*

$$
\frac{1}{m+n}\sum_{i=1}^{m+n}\left\{\hat{V}_i(p) - R(p)\right\}^2 \xrightarrow{\mathcal{P}} \sigma_2^2(p),
$$

where $\sigma_2^2(p)$ is defined in Theorem 2.2.

Proof. For $1 \le i \le m$,

$$\begin{aligned}
\hat{V}_i(p) &= 1 - \frac{m+n}{m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) + \frac{m+n-1}{m-1}\sum_{j=1,j\neq i}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) \\
&= 1 + \left(\frac{m+n-1}{m-1} - \frac{m+n}{m}\right)\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) \\
&\quad - \frac{m+n-1}{m-1}\left\{\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) - \sum_{j=1,j\neq i}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\} \\
&= 1 + \frac{n}{(m-1)m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) - \frac{m+n-1}{m-1} K\left(\frac{1-p-G_n(x_{I,i})}{h}\right),
\end{aligned}$$

and

$$\begin{aligned}
\hat{V}_i^2(p) = &\left\{1 - \frac{m+n-1}{m-1} K\left(\frac{1-p-G_n(x_{I,i})}{h}\right)\right\}^2 + \left\{\frac{n}{(m-1)m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}^2 \\
&+ 2\left[1 - \frac{m+n-1}{m-1} K\left(\frac{1-p-G_n(x_{I,i})}{h}\right)\left\{\frac{n}{(m-1)m}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\right],
\end{aligned}$$

which implies that

$$\begin{aligned}
\sum_{i=1}^{m}\hat{V}_i^2(p) = m &+ \frac{(m+n-1)^2}{(m-1)^2}\sum_{i=1}^{m} K^2\left(\frac{1-p-G_n(x_{I,i})}{h}\right) \\
&- \frac{2(m+n-1)}{m-1}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,i})}{h}\right) + m\left\{\frac{n}{(m-1)m}\Sigma_{i=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}^2 \\
&+ \frac{2n}{(m-1)m}\left\{m - \frac{m+n-1}{m-1}\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,i})}{h}\right)\right\}\left\{\sum_{j=1}^{m} K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}.
\end{aligned} \tag{A.14}$$

Since K^2 is a distribution function, from Gong et al. (2010) and (A.3), we have that

$$\frac{1}{m}\sum_{i=1}^{m} K^2\left(\frac{1-p-G_n(x_{I,i})}{h}\right) \xrightarrow{\mathcal{P}} F\{G^{-1}(1-p)\}.$$

Hence, by (A.13) and Lemma A.1,

$$\begin{aligned}
\frac{1}{m+n}\sum_{i=1}^{m}\hat{V}_i^2(p) &= \frac{m}{m+n} + \frac{(m+n-1)^2}{(m+n)(m-1)^2}\sum_{i=1}^{m}K^2\left(\frac{1-p-G_n(x_{I,i})}{h}\right) \\
&- \frac{2(m+n-1)}{(m-1)(m+n)}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,i})}{h}\right) \\
&+ \frac{n^2}{(m-1)^2 m(m+n)}\left\{\sum_{i=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}^2 \\
&+ \frac{2n}{(m+n)(m-1)m}\left\{m - \frac{m+n-1}{m-1}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,i})}{h}\right)\right\} \\
&\left\{\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\} \\
&\xrightarrow{\mathfrak{D}} \frac{r}{r+1} - 2F\{G^{-1}(1-p)\} + \frac{r+1}{r}F\{G^{-1}(1-p)\} + \frac{1}{r(r+1)}F^2\{G^{-1}(1-p)\} \\
&+ \frac{2}{r+1}F\{G^{-1}(1-p)\}[1 - \frac{r+1}{r}F\{G^{-1}(1-p)\}] \\
&= \frac{r+1}{r}R(p) - \frac{2r+1}{r(r+1)}R^2(p).
\end{aligned} \tag{A.15}$$

Next, for $m+1 \le i \le m+n$, we can write that

$$\begin{aligned}
\hat{V}_i(p) = 1 &- \frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right) \\
&+ \frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right) - K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\},
\end{aligned}$$

and

$$\begin{aligned}
\hat{V}_i^2(p) = &\left\{1 - \frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}^2 \\
&+ \left[\frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right) - K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\right]^2
\end{aligned}$$

$$
\begin{aligned}
&+2\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\\
&\left[\frac{m+n-1}{m}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\right].
\end{aligned}
\tag{A.16}
$$

For the corresponding third terms of $1/(m+n)\sum_{i=m+1}^{m+n}\hat{V}_i^2(p)$, by (A.14), we have

$$
\begin{aligned}
&2\left\{1-\frac{1}{m}\sum_{j=1}^{m}K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\\
&\left[\frac{m+n-1}{m}\frac{1}{m+n}\sum_{i=m+1}^{m+n}\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,m-i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\right]\\
&=O_p((nh)^{-1}).
\end{aligned}
\tag{A.17}
$$

Define $A_i=\left[\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,i})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,i})}{h}\right)\right\}\right]^2$. By (A.13), we have

$$
\begin{aligned}
A_i&=\left[\sum_{j=1}^{m}\left\{K\left(\frac{1-p-G_{n,i}(x_{I,j})}{h}\right)-K\left(\frac{1-p-G_n(x_{I,j})}{h}\right)\right\}\right]^2\\
&=\left\{\int_{-\infty}^{\infty}mK\left(\frac{1-p-G_{n,i}(x)}{h}\right)dF_m(x)-\int_{-\infty}^{\infty}mK\left(\frac{1-p-G_n(x)}{h}\right)\mathrm{d}F_m(x)\right\}^2\\
&=\{m\int_{-\infty}^{\infty}\left(\frac{G_{n,i}(x)-G_n(x)}{h}\right)w\left(\frac{1-p-G_n(x)}{h}\right)\mathrm{d}F_m(x)\\
&+m\int_{-\infty}^{\infty}\frac{1}{2}\left(\frac{G_{n,i}(x)-G_n(x)}{h}\right)^2w^{'}\left(\frac{1-p-G_n(x)+\xi}{h}\right)\mathrm{d}F_m(x)\}^2+o_p(1)\\
&=\left\{\int_{-\infty}^{\infty}m\left(\frac{G_{n,i}(x)-G_n(x)}{h}\right)w\left(\frac{1-p-G_n(x)}{h}\right)\mathrm{d}F_m(x)\right\}^2+o_p(1).
\end{aligned}
$$

By (A.1), (A.2), (A.3) and the continuity of R' and Assumptions A.4 and A.5

$$
\begin{aligned}
\frac{1}{m+n}\sum_{i=1}^{n}A_i=\frac{m^2}{m+n}\sum_{i=1}^{n}\{&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{G_{n,i}(x_1)-G_n(x_1)}{h}\right)\left(\frac{G_{n,i}(x_2)-G_n(x_2)}{h}\right)\\
&w\left(\frac{1-p-G_n(x_1)}{h}\right)w\left(\frac{1-p-G_n(x_2)}{h}\right)\mathrm{d}F_m(x_1)\mathrm{d}F_m(x_2)\}+o_p(1)
\end{aligned}
$$

$$= \frac{m^2}{(m+n)(n-1)^2h^2} \sum_{i=1}^{n} [\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{G_n(x_1) - I(Y_{I,i} \leq x_1)\}\{G_n(x_2) - I(Y_{I,i} \leq x_2)\}$$
$$w\left(\frac{1-p-G_n(x_1)}{h}\right) w\left(\frac{1-p-G_n(x_2)}{h}\right) \mathrm{d}F_m(x_1)\mathrm{d}F_m(x_2)] + o_p(1)$$
$$= \frac{m^2}{(m+n)(n-1)^2h^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{\sum_{i=1}^{n} \{G_n(x_1)G_n(x_2) + I(Y_{I,i} \leq x_1)I(Y_{I,i} \leq x_2)$$
$$- I(Y_{I,i} \leq x_1)G_n(x_2) - G_n(x_1)I(Y_{I,i} \leq x_2)\}$$
$$w\left(\frac{1-p-G_n(x_1)}{h}\right) w\left(\frac{1-p-G_n(x_2)}{h}\right) \mathrm{d}F_m(x_1)\mathrm{d}F_m(x_2)\} + o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2h^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\{G_n(x_1 \wedge x_2) - G_n(x_1)G_n(x_2)\}$$
$$w\left(\frac{1-p-G_n(x_1)}{h}\right) w\left(\frac{1-p-G_n(x_2)}{h}\right) \mathrm{d}F_m(x_1)\mathrm{d}F_m(x_2)] + o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2h^2} \int_{-1}^{1} \int_{-1}^{1} [G_n\{F_m^{-1}(v_1) \wedge F_m^{-1}(v_2)\} - G_n\{F_m^{-1}(v_1)\}G_n\{F_m^{-1}(v_2)\}]$$
$$w\left(\frac{1-p-G_n(F_m^{-1}(v_1))}{h}\right) w\left(\frac{1-p-G_n(F_m^{-1}(v_2))}{h}\right) \mathrm{d}v_1\mathrm{d}v_2 + o_p(1).$$

From the similar proof Lemma A.1 of Gong et al. (2010), the above equation is

$$= \frac{nm^2}{(m+n)(n-1)^2h^2} \int_{-1}^{1} \int_{-1}^{1} [G\{F^{-1}(v_1) \wedge F^{-1}(v_2)\} - G\{F^{-1}(v_1)\}G\{F^{-1}(v_2)\}]$$
$$w\left(\frac{1-p-G\{F^{-1}(v_1)\}}{h}\right) w\left(\frac{1-p-G\{F^{-1}(v_2)\}}{h}\right) \mathrm{d}v_1\mathrm{d}v_2 + o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2} \int_{-1}^{1} \int_{-1}^{1} [G\{G^{-1}(1-p-hu_1) \wedge G^{-1}(1-p-hu_2)\} - G(G^{-1}(1-p-hu_2))$$
$$G_n\{G^{-1}(1-p-hu_2)\}]w(u_1)w(u_2)\mathrm{d}F\{G^{-1}(1-p-hu_2)\}\mathrm{d}F\{G^{-1}(1-p-hu_1)\} + o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2} \int_{-1}^{1} \int_{-1}^{1} [G\{G^{-1}(1-p-hu_1) \wedge G^{-1}(1-p-hu_2)\}$$
$$- G\{G^{-1}(1-p-u_2)\}G_n\{G^{-1}(1-p-u_2)\}] \quad w(u_1)w(u_2)R^{'}(p+hu_2)R^{'}(p+hu_1)\mathrm{d}(u_2)\mathrm{d}(u_1) + o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2} \int_{-1}^{1} \int_{-1}^{1} [\{(1-p-hu_1) \wedge (1-p-hu_2) - (1-p-hu_1)(1-p-hu_2)\}$$
$$w(u_1)w(u_2)R^{'}(p)R^{'}(p)]\mathrm{d}(u_2)\mathrm{d}(u_1) + o_p(1)$$

$$= \frac{nm^2}{(m+n)(n-1)^2}\{(1-p)\wedge(1-p)-(1-p)^2\}R^{'2}(p)\int_{-1}^{1}w(u_1)\mathrm{d}(u_1)\int_{-1}^{1}w(u_2)\mathrm{d}(u_2)+o_p(1)$$
$$= \frac{nm^2}{(m+n)(n-1)^2}\{(1-p)-(1-p)^2\}R^{'2}(p)+o_p(1)$$
$$\xrightarrow{P} \frac{r^2}{(1+r)}p(1-p)R^{'2}(p). \tag{A.18}$$

By (A.17), (A.18), (A.19) and Lemma A.1, we have

$$\frac{1}{m+n}\sum_{i=m+1}^{m+n}\hat{V}_i^2(p) \xrightarrow{\mathcal{P}} \frac{1}{1+r}R^2(p)+(r+1)p(1-p)R^{'2}(p). \tag{A.19}$$

Hence, it follows from (A.16), (A.20) and Lemma A.2 that

$$\frac{1}{m+n}\sum_{i=1}^{m+n}\{\hat{V}_i(p)-R(p)\}^2$$
$$= \frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i^2(p)+R^2(p)-\frac{2}{m+n}R(p)\sum_{i=1}^{m+n}\hat{V}_i(p)$$
$$\xrightarrow{P} \frac{1+r}{r}R(p)-\frac{2r+1}{r(r+1)}R^2(p)+\frac{1}{1+r}R^2(p)+(r+1)p(1-p)R^{'2}(p)-2R^2(p)+R^2(p)$$
$$= (1+\frac{1}{r})R(p)(1-R(p))+(r+1)p(1-p)R^{'2}(p)$$
$$= \sigma_2^2(p).$$

□

Proof of Theorem 2.1 It follows directly from Lemmas A.2 and A.3. □

Proof of Theorem 2.2 Throughout let $\theta = R(p)$. Define

$$g(\lambda) = \frac{1}{m+n}\sum_{i=1}^{m+n}\frac{\hat{V}_i(p)-\theta}{1+\lambda(\hat{V}_i(p)-\theta\}}.$$

It it easy to check that

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

where $S_{m+n} = 1/(m+n)\sum_{i=1}^{m+n}\{\hat{V}_i(p)-\theta\}^2$ and $Z_{m+n} = \max_{1\leq i\leq m+n}\left|\hat{V}_i(p)-\theta\right|$. Because Z_{m+n} is bounded, $S_{m+n} \xrightarrow{\mathfrak{P}} \sigma_2^2$ and $1/(m+n)\sum_{i=1}^{m+n}\{\hat{V}_i(p)-\theta\} = O_p((m+n)^{-1/2})$, we have

$$
|\lambda| = O_p((m+n)^{-1/2}). \tag{A.20}
$$

Define $\gamma_i = \lambda\{\hat{V}_i(p)-\theta\}$. We obtain that

$$
\max_{1\leq i\leq m+n}|\gamma_i| = o_p(1). \tag{A.21}
$$

Using (A.20) and (A.21), we have

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

which imply that

$$
\lambda = S_{m+n}^{-1}\frac{1}{m+n}\sum_{i=1}^{m+n}\{\hat{V}_i(p)-\theta\} + \beta_n, \tag{A.22}
$$

where $\beta_n = O_p((m+n)^{-1})$. Note that $\eta_i = o(\gamma_i^2)$, $i = 1, ..., m+n$. Using Taylor expansion, it follows from (A.20), (A.22), Lemma A.1 and Lemma A.2 that

$$
\begin{aligned}
& l\{R(p), p\} \\
= & \, 2\sum_{i=1}^{m+n} \log[1 + \lambda\{\hat{V}_i(p) - \theta\}] \\
= & \, \frac{2(m+n)}{S_{m+n}} \left[\frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\} \right]^2 + 2(m+n)\beta_n \frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\} \\
& - (m+n) \left[\frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\} + \beta_n S_{m+n} \right] \left[S_{m+n}^{-1} \frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\} + \beta_n \right] + 2\sum_{i=1}^{m+n} \eta_i \\
= & \, \frac{(\sqrt{m+n}[\frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\}])^2}{\frac{1}{m+n} \sum_{i=1}^{m+n} \{\hat{V}_i(p) - \theta\}^2} + o_p(1) \\
\xrightarrow{\mathfrak{D}} & \, \frac{\sigma_1^2(p)\chi_1^2}{\sigma_2^2(p)}. \qquad \square
\end{aligned}
$$

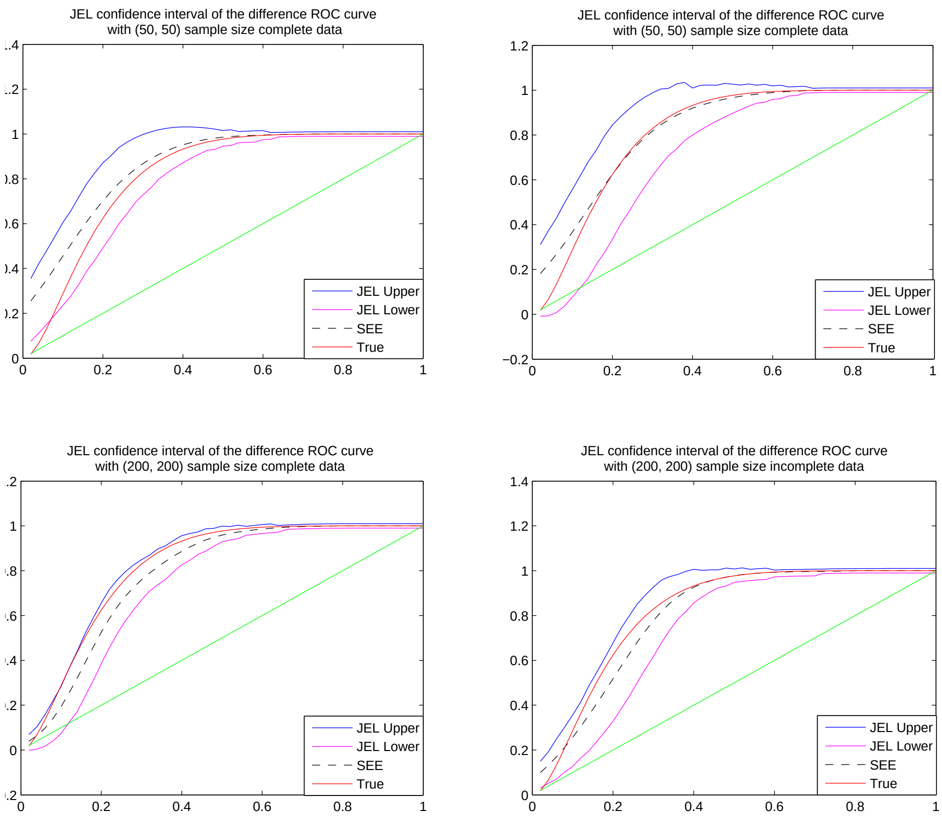

Figure 1: 95% point-wise jackknife empirical likelihood confidence interval for ROC curves from data A, where JEL Upper indicates the upper bound of jackknife empirical likelihood confidence interval, JEL Lower indicates the lower bound of jackknife empirical likelihood confidence interval, SEE means smoothed empirical estimator and True means the true value of ROC curve.

Table 1: Coverage probability of 95% confidence interval for $ROC(p) = 1 - F(G^{-1}(1-p))$.

n1	n2	P	P1	P2	(A) JEL	(A) SEL	(B) JEL	(B) SEL	(C) JEL	(C) SEL
50	50	0.1	0.7	0.6	0.918	0.942	0.951	0.787	0.932	0.773
100	100	0.1	0.7	0.6	0.923	0.940	0.953	0.813	0.917	0.789
200	150	0.1	0.7	0.6	0.937	0.952	0.946	0.804	0.927	0.814
50	50	0.3	0.7	0.6	0.928	0.950	0.968	0.812	0.946	0.856
100	100	0.3	0.7	0.6	0.928	0.948	0.955	0.848	0.925	0.903
200	150	0.3	0.7	0.6	0.942	0.953	0.955	0.853	0.932	0.908
50	50	0.1	0.9	0.8	0.908	0.960	0.951	0.822	0.919	0.821
100	100	0.1	0.9	0.8	0.923	0.953	0.951	0.859	0.929	0.847
200	150	0.1	0.9	0.8	0.933	0.962	0.958	0.869	0.938	0.877
50	50	0.3	0.9	0.8	0.926	0.962	0.941	0.832	0.930	0.894
100	100	0.3	0.9	0.8	0.929	0.960	0.955	0.865	0.940	0.918
200	150	0.3	0.9	0.8	0.932	0.941	0.947	0.905	0.933	0.928

Table 2: Average length of 95% confidence interval for $ROC(p) = 1 - F(G^{-1}(1-p))$.

n1	n2	P	P1	P2	(A) JEL	(A) SEL	(B) JEL	(B) SEL	(C) JEL	(C) SEL
50	50	0.1	0.7	0.6	0.42492	0.40518	0.32284	0.21781	0.35455	0.22543
100	100	0.1	0.7	0.6	0.31733	0.31807	0.23647	0.18237	0.26238	0.19084
200	150	0.1	0.7	0.6	0.25668	0.27594	0.19612	0.15152	0.20800	0.16759
50	50	0.3	0.7	0.6	0.45558	0.44693	0.35349	0.24882	0.42148	0.33445
100	100	0.3	0.7	0.6	0.33850	0.36403	0.25595	0.21040	0.30840	0.28424
200	150	0.3	0.7	0.6	0.27029	0.31737	0.21029	0.17784	0.24542	0.24410
50	50	0.1	0.9	0.8	0.35404	0.36128	0.27093	0.20495	0.29431	0.22238
100	100	0.1	0.9	0.8	0.26807	0.28567	0.20001	0.17431	0.21982	0.18420
200	150	0.1	0.9	0.8	0.21693	0.25042	0.16592	0.15093	0.17548	0.16502
50	50	0.3	0.9	0.8	0.38486	0.41028	0.29694	0.22772	0.35025	0.31367
100	100	0.3	0.9	0.8	0.28426	0.33791	0.21581	0.19633	0.25848	0.26212
200	150	0.3	0.9	0.8	0.22810	0.28021	0.17743	0.17743	0.20477	0.22474