Open Journal of Statistics, 2012, 2, 21-27
http://dx.doi.org/10.4236/ojs.2012.21003 Published Online January 2012 (http://www.SciRP.org/journal/ojs)

Smoothed Empirical Likelihood Inference for ROC Curves with Missing Data

Yueheng An
Department of Mathematics and Statistics, Georgia State University, Atlanta, USA
Email: yan3@student.gsu.edu

Received September 25, 2011; revised November 4, 2011; accepted November 16, 2011

ABSTRACT

The receiver operating characteristic (ROC) curve has been widely used in scientific research fields. After using the random hot deck imputation, we propose the smoothed empirical likelihood ratio statistic for the ROC curve with missing data. Its asymptotic distribution is a scaled chi-square distribution and empirical likelihood confidence intervals for ROC curves are constructed. The simulation study shows that the proposed interval estimates perform well based on the coverage probability for different sample sizes and response rates.

Keywords: Confidence Interval; Missing Data; ROC Curve; Smoothed Empirical Likelihood

1. Introduction

The receiver operating characteristic (ROC) curve has been extensively used to evaluate the diagnostic tests. The ROC curve is usually defined as a graphical plot of the *sensitivity* vs (1-*specificity*). It is clear that ROC curve is an appealing method to summarize the accuracy of predictions in the diagnostic test. In recent years, ROC curves have been widely applied to medical research, diagnostic medicine and many other scientific research fields (Zhou, McClish and Obuchowski [1], Pepe [2]).

Empirical likelihood (EL) is a useful nonparametric statistical inference method which does not need to assume a known family of distributions (see Owen [3]). Owen [4,5] originally proposed EL confidence regions for the population mean parameter in the complete data setting. Chen and Hall [6] introduced smoothed EL confidence intervals for quantiles. To improve the performance of normal approximation methods for the ROC curve for small sample sizes, the EL based method has been used to estimate the ROC curve. Claeskens, Jing, Peng and Zhou [7] developed smoothing EL confidence intervals for ROC curves and Su *et al.* [8] proposed plug-in EL for the ROC curve. Liang and Zhou [9] developed semi-parametric EL confidence intervals for ROC curves with right censoring.

In recent years, missing data problem received much attention in biomedical studies, population survey andmany other related fields. Some of the responses may not be obtained due to information loss (see Qin and Qian [10], Wang and Rao [11]). There is no inference procedure about the ROC curve with missing data. Recently Qin and Qian [10] proposed smoothing EL interval estimation for the difference of two quantiles with missing data. Motivated by their idea, we propose empirical likelihood ratio for the ROC curve with missing data and prove that the resulting EL ratio has a scaled chi-squared limiting distribution. This approach is a natural extension of Claeskens *et al*. [7] to missing data.

The rest of the paper is organized as follows. In Section 2, adopting Qin and Qian [10]'s approach, which was also from Claeskens *et al*. [7], we propose the smoothed empirical likelihood ratio statistic, derive its limiting-distribution and construct the empirical likelihood confidence interval for the ROC curve. In Section 3, we conduct a simulation study to evaluate the finite sample performance of the empirical likelihood interval estimation. The conclusion is given in Section 4. The proofs are given in the Appendix.

2. Main Results

In the following, we adopt the same notations and terminologies as those in Qin and Qian [10]. Suppose there are two independent populations (x,δ_x) and (y,δ_y), where $\delta_x=0$, if x is missing, $\delta_x=1$, otherwise; $\delta_y=0$, if y is missing, $\delta_y=1$, otherwise. We assume that x, y are missing completely at random, *i.e.*, $P(\delta_x=1|x)=P_1$ and $P(\delta_y=1|y)=P_2$. We consider i.i.d. samples of missing data (x_i,δ_{x_i}), $i=1,2,\cdots,m$; (y_j,δ_{y_j}), $j=1,2,\cdots,n$. Let $r_x=\sum_1^m \delta_{x_i}$, $r_y=\sum_1^n \delta_{y_j}$, $m_x=m-r_x$ and $m_y=n-r_y$. Qin and Qian [10] use s_{r_x} and s_{r_y} to denote the sets of respondents with respect to x and y, and use s_{m_x} and s_{m_y} to denote the sets of

non-respondents, respectively. Like Qin and Qian [10], we let x_i^* and y_j^* be the imputed values for the missing data with respect to x and y. Use random hot deck imputation to select a $x_i^* = x_k$ for some $k \in s_{r_x}$ (Little and Rubin [12]). Similarly, we obtain y_j^*. One can obtain complete data as follows.

$$x_{I,i} = \delta_{x_i} x_i + (1-\delta_{x_i}) x_i^*,$$
$$y_{I,j} = \delta_{y_j} y_j + (1-\delta_{y_j}) y_j^*,$$
$$i = 1,2,\cdots,m,\ j = 1,2,\cdots,n.$$

It is of interest to study two populations, one with disease and another one with non-disease. Suppose that the distribution function of disease population X is $F(t)$ and the distribution function of non-disease population Y is $G(t)$. The sensitivity and specificity for a continuous-scale diagnostic test are 1-$F(t)$ and $G(t)$ at a threshold t. At a given level $q = (1-\text{specificity})$, the ROC curve is expressed as

$$\Gamma = 1 - F\left(G^{-1}(1-q)\right),\ for\ 0 < q < 1,$$

where $G^{-1}(q) = \inf\{t : G(t) \ge q\}$. As Qin and Qian [10], let the bandwidth $a = a_m \to 0$ as $m \to \infty$, $b = b_n \to 0$ as $n \to \infty$, and the kernel functions $K_1(\cdot)$ and $K_2(\cdot)$. Define

$$F(t) = \int_{-\infty}^{t/a} K_1(u)\,\mathrm{d}u, G(t) = \int_{-\infty}^{t/b} K_2(u)\,\mathrm{d}u.$$

We adopt the smoothed EL approach of Qin and Qian [10] and define the profile EL ratio statistic at Γ:

$$R(\Gamma) = \sup_{\theta} R(\Gamma,\theta),$$

where

$$R(\Gamma,\theta) = \sup_{p_i, i=1,\cdots,m, q_j, j=1,\cdots,n} \left\{ \sum_{i=1}^{m} \log(mp_i) + \sum_{j=1}^{n} \log(nq_j) \right\}$$

and p_i, q_j satisfy the following equations:

$$\sum_{i=1}^{m} p_i = 1, \sum_{i=1}^{m} p_i F(\theta - x_{I,i}) = q,\ p_i > 0,\ i = 1,\cdots,m,$$

$$\sum_{j=1}^{n} q_j = 1, \sum_{j=1}^{n} q_j G(\theta - y_{I,j}) = q,\ q_j > 0,\ j = 1,\cdots,n,$$

and

$$\nu_1(x_{I,i},\theta,\Gamma) = F_1(\theta - x_{I,i}) - 1 + \Gamma,\ i = 1,\cdots,m,$$

$$\nu_2(y_{I,j},\theta,\Gamma) = F_2(\theta - y_{I,j}) - 1 + q,\ j = 1,\cdots,n.$$

We set $\partial R(\Gamma,\theta)/\partial\theta = 0$. By using the Lagrange multipliers method, we have that

$$R(\Gamma) = -\sum_{i=1}^{m} \log\{1 + \lambda_1(\theta)\nu_1(x_{I,i},\theta,\Gamma)\} - \sum_{j=1}^{n} \log\{1 + \lambda_2(\theta)\nu_2(y_{I,j},\theta,\Gamma)\},$$

where $\lambda_j(\theta)$, $j = 1,2$, and θ satisfy the following score equations:

$$U_{1,m,n}(\theta,\lambda_1,\lambda_2) = \frac{1}{m}\sum_{i=1}^{m} \frac{\nu_1(x_{I,i},\theta,\Gamma)}{1+\lambda_1(\theta)\nu_1(x_{I,i},\theta,\Gamma)} = 0,$$

$$U_{2,m,n}(\theta,\lambda_1,\lambda_2) = \frac{1}{n}\sum_{j=1}^{n} \frac{\nu_2(y_{I,j},\theta,\Gamma)}{1+\lambda_2(\theta)\nu_2(y_{I,j},\theta,\Gamma)} = 0,$$

$$U_{3,m,n}(\theta,\lambda_1,\lambda_2) = \frac{1}{m}\lambda_1(\theta)\sum_{i=1}^{m} \frac{\beta_1(x_{I,i},\theta,\Gamma)}{1+\lambda_1(\theta)\nu_1(x_{I,i},\theta,\Gamma)} + \frac{1}{n}\lambda_2(\theta)\sum_{j=1}^{n} \frac{\beta_2(y_{I,j},\theta,\Gamma)}{1+\lambda_2(\theta)\nu_2(y_{I,j},\theta,\Gamma)} = 0,$$

where

$$\beta_1(x_{I,i},\theta,\Gamma) = \frac{1}{a}K_1\left(\frac{\theta - x_{I,i}}{a}\right)$$

and

$$\beta_2(y_{I,j},\theta,\Gamma) = \frac{1}{b}K_2\left(\frac{\theta - y_{I,j}}{b}\right).$$

Suppose that θ_0 is the true value of θ.In this paper, we assume the same regularity conditions (i)-(v) in Qin and Qian [10] with condition (ii) modified as follows:

(ii) Denote $f(t) = \partial F(t)/\partial t$ and $g(t) = \partial G(t)/\partial t$. For some $t_0 \ge 2$, suppose that $f^{t_0-1}(t)$ and $g^{t_0-1}(t)$ exist and are uniformly continuous and bounded in a neighborhood of θ_0. Assume that $f(\theta_0)g(\theta_0) > 0$.

Recall that condition (iii): $n/m \to k$ *as* $m + n \to \infty$. Then we state the main result about confidence intervals for the ROC curve.

Theorem 1. Under the regularity conditions (i)-(v), as $m + n \to \infty$, there exists a root $\theta_{m,n}$ of Equation (1) such that $R(\Gamma,\theta)$ attains its local maximum at $\theta_{m,n}$,

$$-2R(\Gamma,\theta_{m,n}) \xrightarrow{d} \frac{kg^2(\theta_0)(1-P_1+P_1^{-1})\Gamma(1-\Gamma) + q(1-q)(1-P_2+P_2^{-1})f^2(\theta_0)}{c_0}\chi_1^2$$

where

$$c_0 = q(1-q)f(\theta_0)^2 + kg(\theta_0)^2\Gamma(1-\Gamma)_1^2.$$

We know that k is estimated by n/m, and P_1 and P_2 can be consistently estimated by

$$\hat{P}_1 = \frac{1}{m}\sum\nolimits_{i=1}^{m} \delta_{x_i},$$
$$\hat{P}_2 = \frac{1}{n}\sum\nolimits_{j=1}^{n} \delta_{y_j}$$

respectively. Qin and Qian [10] showed that,

$$\hat{f}(\theta_0)=\frac{1}{ma}\sum_{i=1}^{m}K_1\left(\frac{\theta_{m,n}-x_{I,i}}{a}\right)$$

$$\hat{g}(\theta_0)=\frac{1}{nb}\sum_{j=1}^{n}K_2\frac{\theta_{m,n}-y_{I,j}}{b}$$

are consistent estimators of $f(\theta_0)$ and $g(\theta_0)$. Put

$$b_0(\Gamma)=\frac{\left[k\Gamma(1-\Gamma)\left(1-P_1+P_1^{-1}\right)g(\theta_0)^2+q(1-q)\left(1-P_2+P_2^{-1}\right)f(\theta_0)^2\right]}{c_0}.$$

Plugging in the above consistent estimators, we obtain a consistent estimator $\hat{b}_0(\Gamma)$ of $b_0(\Gamma)$. Let $\chi_1^2(\alpha)$ be the upper α-quantile of χ_1^2. Thus, it follows from Theorem 1 that the $100(1-\alpha)\%$ EL confidence interval for Γ is given by

$$\mathcal{R}=\left\{\Gamma:-2\hat{b}_0^{-1}(\Gamma)R(\Gamma,\theta_{m,n})\le\chi_1^2(\alpha)\right\}$$

Remark: The asymptotic distribution of the EL statistic is a standard χ_1^2 distribution for complete data since $P_1=P_2=1$ and it coincides with the conclusion of Claeskens *et al*. (2003).

3. Simulation Studies

In this section, we carry out extensive simulation studies to evaluate the performance of the EL method for the ROC curve in terms of coverage probability and average length of confidence intervals with different response rates and sample sizes.

The simulation setting is similar to Qin and Qian [10]. The diseased population X is distributed as $N(1, 1)$, while the non-diseased population Y follows $N(0, 1)$. We choose kernel function

$$K_1(u)=K_2(u)=\left(\sqrt{2\pi}\right)^{-1}\exp\left(-u^2/2\right)$$

and the bandwidths $a=c_1m^{-1/3}$, $b=c_2m^{-1/3}$. We draw random samples x and y from the populations X and Y. The response rates for x and y are chosen as $(p_1,p_2)=(0.7,0.6),(0.8,0.7),(0.9,0.8)$. The sample sizes are chosen to be $(m, n)=(50, 50)$, $(75, 75)$,$(100, 100)$, $(200, 150)$. We generate 1000 random samples of the data. The proposed EL confidence intervals for the ROC curve are constructed at q = 0.1, 0.3, 0.5, and 0.7. The nominal level of the confidence intervals is selected as $1-\alpha=95\%$.

From **Tables 1-4** we have the following findings:

1) Note the response rate is higher and larger than 0.6 in the simulation study. For each fixed response rate and sample size, the coverage probability of confidence intervals for the ROC curve is close to the nominal level 95%. In the simulation, when sample size $(50,50)$ is small, the coverage probability is still good.

2) For almost all the cases in the simulation study, when the response rates increase, the coverage probabilities of confidence intervals are closer to 95%, *i.e.*, they are more accurate, and the average length of the confidence intervals decreases, because larger response rates provide more information for the data.

3) Similarly, when the sample sizes increase, the coverage probabilities of confidence intervals are more accurate, and the average length of the confidence intervals decreases.

4) For different q = 0.1, 0.3, 0.5, and 0.7, the EL confidence intervals maintain good coverage probability, and it is very stable.

Table 1. Empirical likelihood confidence intervals for the ROC curve at q = 0.1, (Γ = 0.3891).

(P_1, P_2)	(c_1, c_2)	(m, n)	CP (%)	LE	RE	AL
(0.7, 0.6)	(1.3, 1.3)	(50, 50)	95.6	0.2205	0.6042	0.3837
		(75, 75)	94.7	0.2442	0.5724	0.3281
		(100, 100)	95.3	0.2570	0.5469	0.2899
		(200, 150)	95.3	0.2832	0.5075	0.2242
(0.8, 0.7)	(1.5, 1.5)	(50, 50)	93.5	0.1214	0.5276	0.4062
		(75, 75)	94.7	0.1335	0.5071	0.3736
		(100, 100)	94.5	0.1503	0.4999	0.3496
		(200, 150)	95.5	0.1701	0.4761	0.3060
(0.9, 0.8)	(1.2, 1.2)	(50, 50)	93.4	0.1274	0.5320	0.4046
		(75, 75)	94.7	0.1429	0.5097	0.3668
		(100, 100)	94.2	0.1584	0.4991	0.3407
		(200, 150)	94.9	0.1756	0.4735	0.2980

CP(%): coverage probability; LE: the average left endpoint; RE: the average right endpoint; AL: the average length of the interval.

Table 2. Empirical likelihood confidence intervals for the ROC curve at q = 0.3, (Γ = 0.6828).

(P_1, P_2)	(c_1, c_2)	(m, n)	CP (%)	LE	RE	AL
(0.7,0.6)	(1.3,1.3)	(50,50)	94.8	0.3031	0.8237	0.5206
		(75,75)	96.2	0.3295	0.8055	0.4760
		(100,100)	94.4	0.3374	0.8133	0.4758
		(200,150)	94.8	0.3412	0.7897	0.4484
(0.8,0.7)	(1.5,1.5)	(50, 50)	94.4	0.3207	0.8092	0.4885
		(75, 75)	96.2	0.3359	0.8002	0.4643
		(100,100)	94.8	0.3409	0.7960	0.4552
		(200,150)	95.9	0.3414	0.7857	0.4443
(0.9,0.8)	(1.2,1.2)	(50, 50)	95.7	0.3271	0.8133	0.4862
		(75, 75)	94.1	0.3384	0.7986	0.4603
		(100,100)	95.4	0.3409	0.7910	0.4501
		(200,150)	94.6	0.3414	0.7836	0.4422

Table 3. Empirical likelihood confidence intervals for the ROC curve at q = 0.5, (Γ = 0.8413).

(P_1, P_2)	(c_1, c_2)	(m, n)	CP(%)	LE	RE	AL
(0.7,0.6)	(1.3,1.3)	(50,50)	95.5	0.4167	0.9285	0.5118
		(75,75)	95.2	0.4202	0.9166	0.4964
		(100,100)	94.4	0.4207	0.9110	0.4903
		(200,150)	95.0	0.4207	0.8986	0.4780
(0.8,0.7)	(1.5,1.5)	(50, 50)	94.5	0.4204	0.9213	0.5009
		(75, 75)	94.5	0.4207	0.9110	0.4903
		(100,100)	95.3	0.4207	0.9037	0.4831
		(200,150)	96.4	0.4207	0.8917	0.4710
(0.9,0.8)	(1.2,1.2)	(50, 50)	95.8	0.4205	0.9186	0.4981
		(75, 75)	94.1	0.4207	0.9092	0.4885
		(100,100)	95.0	0.4207	0.9043	0.4836
		(200,150)	95.1	0.4207	0.8933	0.4726

Table 4. Empirical likelihood confidence intervals for the ROC curve at q = 0.7, (Γ = 0.9363).

(P_1, P_2)	(c_1, c_2)	(m, n)	CP (%)	LE	RE	AL
(0.7,0.6)	(1.3,1.3)	(50,50)	93.9	0.4679	0.9758	0.5079
		(75,75)	95.2	0.4681	0.9697	0.5016
		(100,100)	95.8	0.4681	0.9669	0.4988
		(200,150)	94.5	0.4681	0.9627	0.4945
(0.8,0.7)	(1.5,1.5)	(50, 50)	93.3	0.4681	0.9703	0.5021
		(75, 75)	94.6	0.4681	0.9652	0.4971
		(100,100)	96.0	0.4681	0.9628	0.4946
		(200,150)	93.7	0.4681	0.9615	0.4933
(0.9,0.8)	(1.2,1.2)	(50, 50)	93.7	0.4681	0.9702	0.5021
		(75, 75)	95.0	0.4681	0.9654	0.4973
		(100,100)	94.3	0.4681	0.9632	0.4951
		(200,150)	94.9	0.4681	0.9611	0.4930

4. Discussion

In this paper, we developed the smoothing empirical likelihood method for the ROC curve with missing data-which is a natural extension of Claeskens *et al*. [7].The key technique used to impute the missing data is the random hot deck imputation procedure. Under imputation, the proposed smoothed EL statistic converges to a scaled chi-square distribution. In addition, we carry out the simulation studies to evaluate the finite sample performance of the proposed EL interval estimation for the ROC curve. For either smaller or larger q, the EL confidence intervals for the ROC curve have good coverage probabilities which are close to the nominal level. In summary, the proposed EL interval estimation is a reliable and useful tool for the ROC curve analysis with missing data. In the future, we will use other imputation methods to achieve better interval estimation and improve the performance.

5. Acknowledgements

The author acknowledges partial support under a FY09 Research Initiation Grant in Georgia State University. The author would like to thank Dr. Yichuan Zhao for his supervision.

REFERENCES

[1] X.-H. Zhou, D. K. McClish and N. A. Obuchowski, "Statistical Methods in Diagnostic Medicine," Wiley, New York, 2002.

[2] M. S. Pepe, "The Statistical Evaluation of Medical Tests for Classification and Prediction," Oxford University Press, Oxford, 2003.

[3] A. B. Owen, "Empirical Likelihood," Chapman & Hall Ltd, London, 2001. doi:10.1201/9781420036152

[4] A. B. Owen, "Empirical Likelihood Ratio Confidence Intervals for a Single Functional," *Biometrika*, Vol. 75, No. 2, 1988, pp. 237-249. doi:10.1093/biomet/75.2.237

[5] A. B. Owen, "Empirical Likelihood Ratio Confidence Regions. The Annals of Statistics," *Biometrika*, Vol. 18, 1990, pp. 90-120.

[6] S. X. Chen and P. Hall, "Smoothed Empirical Likelihood Confidence Intervals for Quantiles," *The Annals of Statistics*, Vol. 21, No. 3, 1993, pp. 1166-1181. doi:10.1214/aos/1176349256

[7] G. Claeskens, B.-Y. Jing, L. Peng and W. Zhou, "Empirical Likelihood Confidence Regions for Comparison Distributions and ROC Curves," *The Canadian Journal of Statistics*, Vol. 31, 2003, pp. 173-190.

[8] H. Su, Y. Qin and H. Liang, "Empirical Likelihood-Based Confidence Interval of ROC Curves," *Statistics in Biopharmaceutical Research*, Vol. 1, No. 4, 2009, pp. 407-414. doi:10.1198/sbr.2009.0044

[9] H. Liang and Y. Zhou, "Semiparametirc Inference for ROC Curves with Censoring," *Scandinavian Journal of Statistics*, Vol. 35, No. 2, 2008, pp. 212-227. doi:10.1111/j.1467-9469.2007.00580.x

[10] Y. S. Qin and Y. J. Qian, "Empirical Likelihood Confidence Intervals for the Differences of Quantiles with Missing Data," *Acta Mathematicae Applicatae Sinica* (*English Series*), Vol. 25, No. 1, 2009, pp. 105-116. doi:10.1007/s10255-006-6116-0

[11] Q. Wang and J. N. K. Rao, "Empirical Likelihood-Based Inference under for Missing Response Data," *The Annals of Statistics*, Vol. 30, No. 3, 2002, pp. 896-924. doi:10.1214/aos/1028674845

[12] R. J. A. Little and D. B. Rubin, "Statistical Analysis with Missing Data," 2nd Edition, Wiley & John Sons, New York, 2002.

Appendix. Proof of Theorem 1

To prove Theorem 1, we need some additional lemmas-similar to those in Qin and Qian [10]. We only give an outline of the proofs since they follow similar arguments as Qin and Qian [10].

Lemma A.1. *Under the regularity conditions of Theorem* 1, *as* $m+n\to\infty$, *we have*

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\nu_1\left(x_{I,i},\theta_0,\Gamma\right)\xrightarrow{d}N\left(0,\sigma_1^2\right)$$

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\nu_2\left(y_{I,j},\theta_0,\Gamma\right)\xrightarrow{d}N\left(0,\sigma_2^2\right),$$

$$\frac{1}{m}\sum_{i=1}^{m}\nu_1^2\left(x_{I,i},\theta_0,\Gamma\right)=\Gamma\left(1-\Gamma\right)+o_p\left(1\right)$$

$$\frac{1}{n}\sum_{j=1}^{n}\nu_2^2\left(y_{I,j},\theta_0,\Gamma\right)=q\left(1-q\right)+o_p\left(1\right)$$

where

$$\sigma_1^2=\left(1-P_1+P_1^{-1}\right)\Gamma\left(1-\Gamma\right),$$

$$\sigma_2^2=\left(1-P_2+P_2^{-1}\right)q\left(1-q\right).$$

Proof of Lemma A.1. We follow the similar lines as Qin and Qian [10]. Let $\bar{\nu}_{1\mathrm{r}}=\frac{1}{r_x}\sum_{i\in S_{rx}}\nu_1(x_i,\theta_0,\Gamma)$, and $\mathcal{B}_{\mathrm{m}}=\sigma\left(\left(\delta_{xi},x_i\right),i=1,\cdots,m\right)$. It follows that

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\nu_1\left(x_{I,i},\theta_0,\Gamma\right)$$

$$=\sqrt{m}\,\bar{\nu}_{1\mathrm{r}}+\frac{1}{\sqrt{m}}\sum_{i\in S_{m_x}}\left\{\nu_1\left(x_i^*,\theta_0,\Gamma\right)-E\left(\nu_1\left(x_i^*,\theta_0,\Gamma\right)|\mathcal{B}_m\right)\right\}$$

Like Qin and Qian [10], we have that

$$V_m\xrightarrow{d}N\left(0,P_1^{-1}\Gamma\left(1-\Gamma\right)\right).$$

We have

$$\sup_t\left|P\left(\sigma_{2m}^{-1}U_m\le t|\ \mathcal{B}_m\right)-\Phi\left(t\right)\right|=\mathrm{o}_p\left(1\right),$$

where $\sigma_{2m}^2=\left(1-P_1\right)\mathrm{E}\nu_1\left(x,\theta_0,\Gamma\right)=\left(1-P_1\right)\Gamma\left(1-\Gamma\right)$ and $\Phi(t)$ is the cumulative distribution function of $N(0,1)$. By Lemma A.1 of Qin and Qian [10], we have

$$\frac{1}{\sqrt{m}}\sum_i\nu_1\left(x_{I,i},\theta_0,\Gamma\right)\xrightarrow{d}N\left(0,\sigma_1^2\right).$$

As Qin and Qian [10], we have that,

$$\frac{1}{m_x}\sum\nolimits_{\mathrm{i}\in\mathrm{S}_{\mathrm{rx}}}\nu_1^2\left(x_i^*,\theta_0,\Gamma\right)=E\nu_1^2\left(x,\theta_0,\Gamma\right)+o_p\left(1\right)$$

$$\frac{1}{m}\sum\nolimits_{i=1}^{m}\nu_1^2\left(x_{I,i},\theta_0,\Gamma\right)$$

$$=P_1\mathrm{E}\nu_1^2\left(x,\theta_0,\Gamma\right)+o_p\left(1\right)$$

$$+\left(\frac{m_x}{m}\right)\left(\frac{1}{m_x}\right)\sum\nolimits_{i\in S_{mx}}\nu_1^2\left(x_i^*,\theta_0,\Gamma\right)$$

$$=\mathrm{E}\nu_1^2\left(x,\theta_0,\Gamma\right)+o_p\left(1\right)=\Gamma\left(1-\Gamma\right)+o_p\left(1\right)$$

The rest of Lemma A.1 can be proved following same lines. It is omitted.

Lemma A.2. (*Qin and Qian* [10]). *Assume that* $1/3<\eta<1/2$. *Under the regularity conditions* (*i*)-(*v*),

$$\lambda_1\left(\theta\right)=O_p\left(n^{-\eta}a^{-1}+a^{t_0}\right),\quad\lambda_2\left(\theta\right)=O_p\left(n^{-\eta}b^{-1}+b^{t_0}\right),$$

uniformly for $\theta\in\left\{\theta:\left|\theta-\theta_0\right|\le cn^{-\eta}\right\}$ *as* $m+n\to\infty$, *where c is a positive constant.*

Proof of Lemma A.2. We follow the same arguments as Qin and Qian [10]. The proof is omitted.

Lemma A.3. (*Qin and Qian* [10]). *Assume that* $1/3<\eta<1/2$. *Under the regularity conditions* (*i*)-(*v*), *in probability there exists a root* $\theta_{m,n}$ *of Equation* (1) *such that*,

$$\left|\theta_{m,n}-\theta_0\right|=O_p\left(n^{-\eta}\right),$$

as $m+n\to\infty$, *and* $R\left(\Gamma,\theta\right)$ *attains its local maximum value at* $\theta_{m,n}$.

Proof of Lemma A.3. We follow the similar lines as Qin and Qian [10]. The proof is omitted.

Lemma A.4. *Assume that the regularity conditions are satisfied. Then, as* $m+n\to\infty$

$$\sqrt{m}\left(\theta_{m,n}-\theta_0\right)\xrightarrow{d}N\left(0,\frac{q^2\left(1-q\right)^2f^2\left(\theta_0\right)\sigma_1^2+k\Gamma^2\left(1-\Gamma\right)^2g^2\left(\theta_0\right)\sigma_2^2}{c_0^2}\right)$$

$$\lambda_1\left(\theta_{m,n}\right)=-\frac{kg\left(\theta_0\right)}{f\left(\theta_0\right)}\lambda_2\left(\theta_{m,n}\right)+O_p\left(n^{-\frac{1}{2}}\right),$$

$$\sqrt{m}\lambda_2\left(\theta_{m,n}\right)\xrightarrow{d}N\left(0,\frac{f^2\left(\theta_0\right)\left\{g^2\left(\theta_0\right)\sigma_1^2+k^{-1}f^2\left(\theta_0\right)\sigma_2^2\right\}}{c_0^2}\right)$$

where $\sigma_j^2,j=1,2$, *and* c_0 *are defined in Lemma A.*1 *and Theorem* 1.

Proof of Lemma A.4. We follow the similar lines as Qin and Qian [10]. Let $\lambda_1=\lambda_1\left(\theta\right)$, $\lambda_{E1}=\lambda_1\left(\theta_{m,n}\right)$, $\lambda_2=\lambda_2\left(\theta\right)$, $\lambda_{E2}=\lambda_2\left(\theta_{m,n}\right)$. Using the Taylor expansion, Lemma A.2 and Lemma A.3, we have

$$
\begin{aligned}
0 &= U_{i,m,n}\left(\theta_{m,n},\lambda_{E1},\lambda_{E2}\right)\\
&= U_{i,m,n}\left(\theta_0,0,0\right)+\frac{\partial U_{i,m,n}\left(\theta_0,0,0\right)}{\partial\theta\left(\theta_{m,n}-\theta_0\right)}\\
&\quad+\frac{\partial U_{i,m,n}\left(\theta_{0,0},0\right)}{\partial\lambda_1}\lambda_{\mathrm{E1}}+\frac{\partial U_{i,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_2}\lambda_{E2}\\
&\quad+o_p\left(\in_n\right),
\end{aligned}
$$

where $i = 1, 2, 3$, $\in_n = |\theta_{m,n}-\theta_0|+|\lambda_{E1}|+|\lambda_{E2}|$. As Lemma 4.5 of Qin and Qian [10], we can show that

$$
\begin{aligned}
&\frac{\partial U_{1,m,n}\left(\theta_0,0,0\right)}{\partial\theta} = f\left(\theta_0\right)+O_p\left(1\right),\\
&\frac{\partial U_{1,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_1} = -\Gamma\left(1-\Gamma\right)+O_p\left(1\right),\\
&\frac{\partial U_{1,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_2} = 0\\
&\frac{\partial U_{2,m,n}\left(\theta_0,0,0\right)}{\partial\theta} = g\left(\theta_0\right)+O_p\left(1\right),\\
&\frac{\partial U_{2,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_1} = 0,\\
&\frac{\partial U_{2,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_2} = -q\left(1-q\right)+O_p\left(1\right),\\
&\frac{\partial U_{3,m,n}\left(\theta_0,0,0\right)}{\partial\theta} = 0,\\
&\frac{\partial U_{3,\mathrm{m,n}}\left(\theta_0,0,0\right)}{\partial\lambda_1} = f\left(\theta_0\right)+O_p\left(1\right),\\
&\frac{\partial U_{3,m,n}\left(\theta_0,0,0\right)}{\partial\lambda_2} = kg\left(\theta_0\right)+o_p\left(1\right)
\end{aligned}
$$

Thus

$$
\begin{pmatrix}\theta_{m,n}-\theta_0\\ \lambda_{E1}\\ \lambda_{E2}\end{pmatrix} = W^{-1}\begin{pmatrix}-U_{1,m,n}\left(\theta_{0,0,0}\right)\\ -U_{2,m,n}\left(\theta_{0,0,0}\right)\\ 0\end{pmatrix}
$$

where

$$
\mathrm{W} = \begin{pmatrix}f\left(\theta_0\right) & -\Gamma\left(1-\Gamma\right) & 0\\ g\left(\theta_0\right) & 0 & -q\left(1-q\right)\\ 0 & f\left(\theta_0\right) & kg\left(\theta_0\right)\end{pmatrix}
$$

It follows that

$$
\theta_{m,n}-\theta_0 = -\frac{1}{c_0}\left\{q\left(1-q\right)f\left(\theta_0\right)U_{1m,n}(\theta_0,0,0)+k\Gamma\left(1-\Gamma\right)g\left(\theta_0\right)U_{2,m,n}\left(\theta_0,0,0\right)\right\}+O_p\left(n^{-1/2}\right),
$$

$$
\lambda_{E1} = \frac{kg\left(\theta_0\right)}{c_0}\left\{g\left(\theta_0\right)U_{1,m,n}\left(\theta_0,0,0\right)-f\left(\theta_0\right)U_{2,m,n}\left(\theta_0,0,0\right)\right\}+O_p\left(n^{-1/2}\right),
$$

$$
\lambda_{E2} = \frac{f\left(\theta_0\right)}{c_0}\left\{g\left(\theta_0\right)U_{1,m,n}\left(\theta_0,0,0\right)-f\left(\theta_0\right)U_{2,m,n}\left(\theta_0,0,0\right)\right\}+O_p\left(n^{-1/2}\right),
$$

From Lemma A.1, we have

$$
\sqrt{m}\begin{pmatrix}U_{1,m,n}\left(\theta_{0,0,0}\right)\\ U_{2,m,n}\left(\theta_{0,0,0}\right)\end{pmatrix}\xrightarrow{d} N\left(0,\begin{pmatrix}\sigma_1^2 & 0\\ 0 & k^{-1}\sigma_2^2\end{pmatrix}\right)
$$

thus Lemma A.4 is proved.

Proof of Theorem 1. It is similar to the proof of Theorem 1 in Qin and Qian [10]. The proof of Theorem 1 is omitted.

