Strong approximations for resample quantile processes and application to ROC methodology

Jiezhun Gu [1], **Subhashis Ghosal**[2] [3]

Abstract The receiver operating characteristic (ROC) curve is defined as true positive rate versus false positive rate obtained by varying a decision threshold criterion. It has been widely used in medical science for its ability to measure the accuracy of diagnostic or prognostic tests. Mathematically speaking, ROC curve is the composition of survival function of one population with the quantile function of another population. In this paper, we study strong approximation for the quantile processes of bootstrap and the Bayesian bootstrap resampling distributions, and use this result to study strong approximations for the empirical ROC estimator, the corresponding bootstrap, and the Bayesian versions in terms of two independent Kiefer processes. The results imply asymptotically accurate coverage probabilities for the confidence bands for the ROC curve and confidence intervals for the area under the curve functional of the ROC constructed using bootstrap and the BB method.

Key words: Bootstrap; Bayesian bootstrap; Kiefer process; ROC curve; Strong approximation; Quantile process.

1 Introduction

Originally introduced in the context of electronic signal detection (Green and Swets, 1966), the receiver operating characteristic (ROC) curve, which is a plot of the true positive rate versus the false positive rate by varying all the possible threshold values, has become an indispensable tool for measuring the accuracy of diagnostic tests since the 1970s (Metz, 1978).

[1]Duke Clinical Research Institute, Durham, NC, 27705. Email: jiezhun.gu@duke.edu.

[2]Department of Statistics, North Carolina State University, Raleigh, NC 27695. Email: ghosal@stat.ncsu.edu.

[3]Partially supported by NSF grant DMS-0349111.

The main attractions of ROC curve may be described by the following properties: (1) One ROC curve can display all the decision characteristic information of a diagnostic test in a unit graph which can be used for classification purpose; (2) Different diagnostic tests can be compared by their corresponding ROC curves. For example, the data set analyzed by Wieand et al. (1989) was based on 51 patients as control group diagnosed as pancreatitis and 90 patients as cases group diagnosed as pancreatic cancer by two biomarkers, cancer antigen (CA 125) and a carbohydrate antigen (CA 19-9), respectively. The purpose was to decide which biomarker would better distinguish case group from control group on a certain range of false positive rate (0, 0.2). One of the solutions is to estimate and compare two ROC curves and their corresponding confidence bands for the two biomarkers. If one lies above another, then we can conclude that one is better than another. The area under the curve (AUC) functional of ROC can be interpreted as the probability that the diagnostic value of a randomly chosen patient with the positive condition (usually referring to a disease) is greater than the diagnostic value of a randomly chosen patient without the positive condition (Bamber, 1975). A nonparametric estimator of ROC may be obtained by substituting into the empirical counterparts, and its variability may be estimated by bootstrap (Pepe, 2003). More recently, Gu et al. (2006) proposed a smoother estimator and related confidence bands by using the Bayesian bootstrap (BB) method. These authors treated the BB distribution as the posterior distribution of the unknown quantities and make inference based the "BB posterior distribution", which actually corresponds to the posterior distribution with respect to Dirichlet prior with precision parameters converging to zero. Interestingly, the smoothness of their estimator comes from resampling and does not involve selecting a bandwidth. The BB estimator of ROC curve and its AUC performs well compared to some existing semiparametric and nonparametric methods.

In this paper, we will first obtain strong approximations for quantile processes of bootstrap the BB resampling distributions by a sequence of appropriate Gaussian processes (more specifically, Kiefer processes). The reason of employing the strong approximation theory for this study instead of the weak convergence theory is discussed in Section 4 Remark 2. The existing strong approximation theories primarily include these findings:

- Komlós, Major and Tusnády (1975) showed that the uniform empirical process can be strongly approximated by a sequence of Brownian bridges obtained from a single Kiefer process;

- Csörgő and Révész (1978) proved that under suitable conditions quantile process can be strongly approximated by a Kiefer process;

- Lo (1987) studied the strong approximation theory for the cumulative distribution function (c.d.f.) of bootstrap and BB processes.

- Hsieh and Turnbull (1996) studied the strong approximation for the empirical ROC estimate on a fixed subinterval of $(0, 1)$.

The ROC function is $R(t) = \bar{G}(\bar{F}^{-1}(t))$, where $\bar{F}(x) = 1 - F(x)$ and $\bar{G}(y) = 1 - G(y)$ are the survival functions of independent variables $X \sim F$ and $Y \sim G$, respectively. Combining these results, we shall develop strong approximations for bootstrap and the BB version of the ROC process, its AUC and other functionals. In particular, these results imply Gaussian weak limits for the processes and asymptotically valid coverage probabilities for the resulting confidence bands and intervals. Interestingly, it will be seen that the forms of these Gaussian approximations are identical, and therefore both bootstrap and the BB distribution of the ROC function, conditioned on the samples, are identical to the Gaussian approximation of the empirical ROC estimate. This means that frequentist variability of the empirical estimate of ROC can be asymptotically accurately estimated by resampling variability of either bootstrap or the BB procedures, given that the samples and the "posterior" mean (i.e, mean of either bootstrap or the BB distribution of the ROC curve) is asymptotically equivalent to that of the empirical estimate up to the first order ($O(N^{-1/2})$), where N is the total sample size. Also, the result implies that for functionals like AUC, the empirical estimator is asymptotically normal and asymptotically equivalent to the BB estimator, and the corresponding confidence intervals have asymptotic frequentist validity. Graphical displays of these findings will be shown in the simulation studies in Section 5.

Our strong approximation results imply that the estimator proposed by Gu et al. (2006), namely, the mean of the BB distribution, is asymptotically equivalent to the empirical estimator of ROC, and converges to the true ROC curve at the rate $N^{-1/2}$. This, along with the smoothness of the estimator, provides strong justifications in favor of the estimator proposed there. Further, our results establish that confidence bands for the ROC function and confidence intervals for the area under the curve and other functionals of the ROC constructed using the BB method have asymptotically correct coverage probabilities. Moreover, our result on strong approximation of the quantile function of the BB and bootstrap distribution completes a gap in strong approximation theory, and may be useful in various contexts other than the ROC methodology.

2 Notation

For $X_1, \ldots, X_n \sim$ i.i.d. F, define the domain of X as $[a, b]$: $a = \sup\{x : F(x) = 0\}$, $b = \inf\{x : F(x) = 1\}$; the order statistic $X_{j:n}$, $j = 0, 1, \ldots, n+1$, where $X_{0:n} = a$ and $X_{n+1:n} = b$, similarly define $V_{j:n}$ and $U_{j:n}$ for independent uniform samples $V_1, \ldots, V_n$ and $U_1, \ldots, U_n$; the quantile function $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $t \in [0, 1]$; for a given $0 < y < 1$, $k = \lceil ny \rceil$, where $\lceil \cdot \rceil$ is the ceiling function, that is, the smallest integer greater than or equal to x; $U(0,1)$ denotes the uniform distribution on [0,1]; abbreviate almost sure convergence by $a.s.$. Condition A (Csörgő and Révész, 1978) and Condition B are assumed, which hold for a large class of distribution functions.

Condition A: The continuous distribution function F is twice differentiable on (a,b), $F' = f \neq 0$ on (a,b) and for some $\gamma > 0$, $\sup_{a<x<b} F(x)(1-F(x))\left|\frac{f'(x)}{f^2(x)}\right| \leq \gamma$.

Condition B: Let F and G satisfy Condition A and let the following conditions hold:

$$\sup_{a<x<b} F(x)(1-F(x))\left|\frac{g^{'}(x)}{f^2(x)}\right|, \quad \sup_{a<x<b} F(x)(1-F(x))\left|\frac{g(x)}{f(x)}\right| \text{ are bounded.}$$

Define the following orders: $\delta_n = 25n^{-1}\log\log n$, $\delta_n^* = \delta_n + 2n^{-1/2}(\log\log n)^{1/2}$, $\delta_n^{\#} = \delta_n + n^{-1/2}(\log\log n)^{1/2}$. $l_n = n^{-1/4}(\log\log n)^{1/4}(\log n)^{1/2}$, $\tau_m = m^{-3/4}(\log\log m)^{1/4}(\log m)^{1/2}$, $\gamma_n = n^{-1}\log^2 n$. $\alpha_m^{-1} = O(m^{1/4-\epsilon})$ for some $\epsilon \in (0, 1/4)$, $\alpha_m^* = \alpha_m + 2m^{-1/2}(\log\log m)^{1/2}$, $\alpha_m^{\#} = \alpha_m + (m-1)^{-1/2}(\log\log(m-1))^{1/2}$. The symbol $\Box$ is used at the end of the proof.

We define the following empirical, quantile functions and their processes:

$$\text{empirical function: } \mathbb{F}_n(x) = j/n, \quad \text{if } X_{j:n} \leq x < X_{j+1:n}, j = 0, 1, \ldots, n.$$

$$\text{empirical process: } \mathbb{J}_n(x) = \sqrt{n}(\mathbb{F}_n(x) - F(x))$$

$$\text{quantile function: } \mathbb{F}_n^{-1}(y) = X_{k:n} = F^{-1}(U_{k:n}), \text{ where } k = \lceil ny \rceil$$

$$\text{quantile process: } \mathbb{Q}_n(y) = \sqrt{n}(\mathbb{F}_n^{-1}(y) - F^{-1}(y))$$

Based on samples $X_1, \ldots, X_n$, bootstrap resamples, which means samples generated from bootstrap resampling distribution, are given by $\{\mathbb{F}_n^{-1}(V_{j:n}), j = 1, \ldots, n;\ V_1, \ldots, V_n \sim$ i.i.d. $U(0,1)$ independent of X_i's $\}$. The empirical and quantile functions of bootstrap resampling distribution are denoted as $\mathbb{F}_n^*(x)$ and $\mathbb{F}_n^{*-1}(y)$, respectively, based on the bootstrap resamples. Bootstrap empirical and quantile processes are defined as

$$\mathbb{J}_n^*(x) = \sqrt{n}(\mathbb{F}_n^*(x) - \mathbb{F}_n(x))$$

$$\mathbb{Q}_n^*(y) = \sqrt{n}(\mathbb{F}_n^{*-1}(y) - \mathbb{F}_n^{-1}(y)) = \sqrt{n}(\mathbb{F}_n^{-1}(V_{k:n}) - \mathbb{F}_n^{-1}(y)) = \sqrt{n}(F^{-1}(U_{k^*:n}) - F^{-1}(U_{k:n})),$$

where $k = \lceil ny \rceil, k^* = \lceil nV_{k:n} \rceil$, $\mathbb{F}_n^{*-1}(y) = \mathbb{F}_n^{-1}(V_{k:n})$.

Based on sample $X_1, ..., X_n$, BB's c.d.f. is defined as $\mathbb{F}_n^{\#}(x) = \sum_{1 \le j \le n} \Delta_{j:n} 1(X_{j:n} \le x)$, $x \in \mathbb{R}$, where $V_1, \ldots, V_{n-1} \sim$ i.i.d. $U(0,1)$, independent of X_i's, $\Delta_{j:n} = V_{j:n-1} - V_{j-1:n-1}$, the quantile function, empirical and quantile processes of the BB resampling distributions are:

$$\text{quantile function of BB: } \mathbb{F}_n^{\#-1}(y) = \begin{cases} X_{j:n}, & V_{j-1:n-1} < y \le V_{j:n-1}, \\ X_{0:n}, & y = 0, \ j = 1, 2, \ldots, n. \end{cases}$$

$$\text{empirical process of BB: } \mathbb{J}_n^{\#}(x) = \sqrt{n}(\mathbb{F}_n^{\#}(x) - \mathbb{F}_n(x)),$$

$$\text{quantile process of BB: } \mathbb{Q}_n^{\#}(y) = \sqrt{n}(\mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(y)).$$

Set $\bar{\mathbb{F}}_n^{\#}(x) = 1 - \mathbb{F}_n^{\#}(x)$ and $\bar{\mathbb{F}}_n^{\#-1}(y) = \mathbb{F}_n^{\#-1}(1-y)$, and similarly for G. Let $\mathbb{U}_n(x)$, $\mathbb{H}_n(x)$, $\mathbb{U}_n^{-1}(y)$, $\mathbb{W}_n(y)$, $\mathbb{H}_n^{\#}(x)$, $\mathbb{W}_n^{\#}(y)$, respectively stand for $\mathbb{F}_n(x)$, $\mathbb{J}_n(x)$, $\mathbb{F}_n^{-1}(y)$, $\mathbb{Q}_n(y)$, $\mathbb{J}_n^{\#}(x)$, $\mathbb{Q}_n^{\#}(y)$ when F is replaced by $U(0,1)$.

Table 1: Notation for empirical and quantile processes

Definitions	General notation	Uniform case notation
Empirical function	$\mathbb{F}_n(x)$	$\mathbb{U}_n(x)$
Empirical process	$\mathbb{J}_n(x)$	$\mathbb{H}_n(x)$
Quantile function	$\mathbb{F}_n^{-1}(y)$	$\mathbb{U}_n^{-1}(y)$
Quantile process	$\mathbb{Q}_n(y)$	$\mathbb{W}_n(y)$
Bootstrap empirical process	$\mathbb{J}_n^*(x)$	$\mathbb{H}_n^*(x)$
Bootstrap quantile process	$\mathbb{Q}_n^*(y)$	$\mathbb{W}_n^*(y)$
BB empirical process	$\mathbb{J}_n^{\#}(x)$	$\mathbb{H}_n^{\#}(x)$
BB quantile process	$\mathbb{Q}_n^{\#}(y)$	$\mathbb{W}_n^{\#}(y)$

3 Strong approximations for bootstrap and the BB quantile processes

Theorem 3.1 (Strong approximation for bootstrap quantile process) Let $X_1, X_2, \ldots \sim$ i.i.d. F satisfying Condition A. Then the quantile process of bootstrap $\mathbb{Q}_n^*(y)$ can be strongly approximated by a Kiefer process K, in the sense that

$$\sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid f(F^{-1}(y))\mathbb{Q}_n^*(y) - n^{-1/2}K(y,n) \mid =_{a.s.} O(l_n). \tag{1}$$

To prove this theorem, we need the following lemma whose proof is given in Section 6.

Lemma 3.1 Let $X_1, \ldots, X_n \sim$ i.i.d. F satisfying Condition A, and let $V_1, \ldots, V_n \sim$ i.i.d. $U(0,1)$, independent of $X_1, ..., X_n$. Then for the quantile process $\mathbb{Q}_n(y)$ of X_i's, there exists a Kiefer process K such that

$$\sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid (y - V_{k:n})\mathbb{Q}_n(V_{k:n}) f'(F^{-1}(\xi))/f(F^{-1}(\xi))) \mid =_{a.s.} O(l_n), \tag{2}$$

$$\sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid f(F^{-1}(y))\mathbb{Q}_n(V_{k:n}) - n^{-1/2}K(y,n) \mid =_{a.s.} O(l_n), \tag{3}$$

where $k = \lceil ny \rceil$ and for any ξ lies between y and $V_{k:n}$.

Proof of Theorem 3.1 :

When $\delta_n^* \le y \le 1 - \delta_n^*$, $[\delta_n^*, 1-\delta_n^*] \subset [\delta_n, 1-\delta_n]$. Let $k^* = \lceil nV_{k:n} \rceil$. Then

$$\begin{aligned} \mathbb{Q}_n^*(y) &= \sqrt{n}(F^{-1}(U_{k^*:n}) - F^{-1}(U_{k:n})) \\ &= \sqrt{n}(F^{-1}(U_{k^*:n}) - F^{-1}(V_{k:n})) + \sqrt{n}(F^{-1}(V_{k:n}) - y) - \sqrt{n}(F^{-1}(U_{k:n}) - y) \\ &= \mathbb{Q}_n(V_{k:n}) + \tilde{\mathbb{Q}}_n(y) - \mathbb{Q}_n(y), \end{aligned} \tag{4}$$

where $\mathbb{Q}_n(y)$ and $\tilde{\mathbb{Q}}_n(y)$ are independent quantile processes for X_i's and $\tilde{X}_i$'s respectively, $X_i = F^{-1}(U_i)$, $\tilde{X}_i = F^{-1}(V_i)$, $U_i \sim$ i.i.d. $U(0,1)$, $V_i \sim$ i.i.d. $U(0,1)$, and U_i's and V_i's are independent, $i = 1, \ldots, n$. By Theorem B of Appendix applied to $\mathbb{Q}_n(y)$ and $\tilde{\mathbb{Q}}_n(y)$, respectively, there exist independent Kiefer processes K and $\tilde{K}$, satisfying

$$\begin{aligned} &\sup_{\delta_n \le y \le 1-\delta_n} \mid f(F^{-1}(y))\mathbb{Q}_n(y) - n^{-1/2}K(y,n) \mid =_{a.s.} O(l_n), \\ &\sup_{\delta_n \le y \le 1-\delta_n} \mid f(F^{-1}(y))\tilde{\mathbb{Q}}_n(y) - n^{-1/2}\tilde{K}(y,n) \mid =_{a.s.} O(l_n). \end{aligned} \tag{5}$$

By applying Lemma 3.1 and (4), (5), we obtain $\sup_{\delta_n^* \le y \le 1-\delta_n^*} |f(F^{-1}(y))$
$\mathbb{Q}_n^*(y) - n^{-1/2}\tilde{K}(y,n)| =_{a.s.} O(l_n)$. $\square$

Theorem 3.2 (Strong approximation for the BB quantile process) Let $X_1, ..., X_n \sim$ i.i.d. F satisfying Condition A. Then the quantile process of BB resampling distribution $\mathbb{Q}_n^{\#}(y)$ can be strongly approximated by a Kiefer process K,

$$\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \mid f(F^{-1}(y))\mathbb{Q}_n^{\#}(y) - n^{-1/2}K(y,n) \mid =_{a.s.} O(l_n). \tag{6}$$

where $\delta_n^{\#} = \delta_n + n^{-1/2}(\log\log n)^{1/2}$, $\delta_n = 25n^{-1}\log\log n$.

The proof requires the following lemma whose proof is deferred to Section 6.

Lemma 3.2 Let $X_1, \ldots, X_n \sim$ i.i.d. F satisfying Condition A, $V_1, \ldots, V_{n-1} \sim$ i.i.d. $U(0,1)$, independent of $X_1, \ldots, X_n$, $V_1, \ldots, V_{n-1}$'s empirical function denoted as $\tilde{\mathbb{U}}_{n-1}(x)$. Then

$$\sup_{0<y<1} \sqrt{n} \mid \mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) \mid =_{a.s} O(n^{-1/2}\log n), \tag{7}$$

In addition, there exists a Kiefer process K, such that

$$\sup_{\delta_n^{\#} \leq y \leq 1-\delta_n^{\#}} | \sqrt{n} f(F^{-1}(y))(\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - \mathbb{F}_n^{-1}(y)) - n^{-1/2} K(y,n) | =_{a.s.} O(l_n). \quad (8)$$

Proof of Theorem 3.2 :

1. We may represent $X_i = F^{-1}(U_i)$, where U_i's $\sim$ i.i.d. $U(0,1)$, $i = 1, \dots, n$, $V_1, \dots, V_{n-1}$ $\sim$ i.i.d. $U(0,1)$, independent of X's. Then we have (cf. See Table 2 for Values of $\mathbb{F}_n^{\#-1}(y)$ and $\tilde{\mathbb{U}}_{n-1}(y)$)

$$\mathbb{F}_n^{\#-1}(y) = \begin{cases} \mathbb{F}_n^{-1}(\frac{1+(n-1)\tilde{\mathbb{U}}_{n-1}(y)}{n}), & V_{k:n-1} < y < V_{k+1:n-1}, k = 0, \dots, n-1, \\ \mathbb{F}_n^{-1}(\frac{(n-1)\tilde{\mathbb{U}}_{n-1}(y)}{n}), & y = V_{k:n-1}, k = 1, \dots, n-1. \end{cases}$$

Table 2: Values of $\mathbb{F}_n^{\#-1}(y)$ and $\tilde{\mathbb{U}}_{n-1}(y)$

y	$\mathbb{F}_n^{\#-1}(y)$	$\tilde{\mathbb{U}}_{n-1}(y)$
$0 \leq y < V_{1:n-1}$	$X_{1:n}$	0
$= V_{1:n-1}$	$X_{1:n}$	$1/(n-1)$
$\cdots$	$\cdots$	$\cdots$
$V_{k:n-1} < y < V_{k+1:n-1}$	$X_{k+1:n}$	$k/(n-1)$
$y = V_{k+1:n-1}$	$X_{k+1:n}$	$(k+1)/(n-1)$
$\cdots$	$\cdots$	$\cdots$
$V_{n-2:n-1} < y < V_{n-1:n-1}$	$X_{n-1:n}$	$(n-2)/(n-1)$
$y = V_{n-1:n-1}$	$X_{n-1:n}$	1
$V_{n-1:n-1} < y \leq 1$	$X_{n:n}$	1

2. The BB quantile process can be written as

$$\mathbb{Q}_n^{\#}(y) = \sqrt{n}(\mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y))) + \sqrt{n}(\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - \mathbb{F}_n^{-1}(y)). \quad (9)$$

Then by applying Lemma 3.2, (9) and Condition A, Theorem 3.2 holds for a Kiefer process K satisfying (8). $\square$

4 Limit theorem for bootstrap and the BB version of the ROC process and functionals

Let $X_1, \ldots, X_m \sim$ i.i.d. F, and $Y_1, \ldots, Y_n \sim$ i.i.d. G. For example, in medical contexts F may stand for the c.d.f. of a population without disease and G for the c.d.f. of a population with disease. Assume F and G satisfy Condition A and B and X's and Y's are independent. Let $N = m + n$, when $N \to +\infty$, $\frac{m}{N} \to \lambda$, $0 < \lambda < 1$.

Recall Kiefer process defined as follows: $\{K(x, y) : 0 \le x \le 1, 0 \le y < \infty\}$ is determined by $K(x, y) = W(x, y) - xW(1, y)$, where $W(x, y)$ be a two-parameter standard Wiener process. Its covariance function is $E[K(x_1, y_1)K(x_2, y_2)] = (x_1 \wedge x_2 - x_1x_2)(y_1 \wedge y_2)$, where $\wedge$ stands for the minimum.

4.1 Functional limit theorems for ROC curve

The ROC curve is defined as $\{(\mathrm{P}(X > c_t), \mathrm{P}(Y > c_t)) : X \sim F\ , Y \sim G\ , c_t \in \mathbb{R}\}$, or alternatively as $R(t) = \bar{G}(\bar{F}^{-1}(t))$, where t =P$(X > c_t)$; its derivative is $R'(t) = g(\bar{F}^{-1}(t))/f(\bar{F}^{-1}(t))$. Empirical estimate of ROC can be obtained by plugging in their empirical counterparts. Based on the given samples, BB estimate can be obtained by doing the following two steps: (1) get one realization from the posterior distribution of ROC curves by plugging the c.d.f. and quantile function of the BB resampling distributions into the ROC functional form, respectively (2) get BB posterior mean by averaging over many realizations. The BB version of the ROC curve described above can be equivalently regarded by putting two independent non-informative Dirichlet priors on F and G. Hence, our BB estimate of ROC is essentially a non-parametric Bayesian estimate. Let $\mathbb{R}(t) = \mathbb{R}_{m,n}(t) = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t))$, $\mathbb{R}^*_{m,n}(t) = \bar{\mathbb{G}}^*_n(\bar{\mathbb{F}}_m^{*-1}(t))$, and $\mathbb{R}^{\#}_{m,n}(t) = \bar{\mathbb{G}}^{\#}_n(\bar{\mathbb{F}}_m^{\#-1}(t))$.

Theorem 4.1 (Functional limit theorem for empirical, bootstrap and BB version of ROC curves.) Let $X_1, \ldots, X_m \sim$ i.i.d. F, $X_i = F^{-1}(U_i)$ and $Y_1, \ldots, Y_n \sim$ i.i.d. G, X's and Y's are independent. Assume F and G satisfy Condition A and B. Then

$$\mathbb{R}_{m,n}(t) = R(t) + R'(t)\frac{K_1(t,m)}{m} + \frac{K_2(R(t),n)}{n} + O(\alpha_m^{-1}\tau_m), \qquad t \in (\alpha_m, 1-\alpha_m) \tag{10}$$

$$\mathbb{R}^*_{m,n}(t) = \mathbb{R}_{m,n}(t) + R'(t)\frac{K_1(t,m)}{m} + \frac{K_2(R(t),n)}{n} + O(\alpha_m^{-1}\tau_m),\ t \in (\alpha^*_m, 1-\alpha^*_m) \tag{11}$$

$$\mathbb{R}^\#_{m,n}(t) = \mathbb{R}_{m,n}(t) + R'(t)\frac{K_1(t,m)}{m} + \frac{K_2(R(t),n)}{n} + O(\alpha_m^{-1}\tau_m),\ t \in (\alpha^\#_m, 1-\alpha^\#_m) \tag{12}$$

in the sense of $a.s.$ for (11) and (12), where K_1 and K_2 are independent generic Kiefer processes (not identical in each appearance).

Proof of Theorem 4.1 :

1. **Proof of (10) of Theorem 4.1**: By Theorem A and B of Appendix, when $t \in (\alpha_m, 1-\alpha_m) \subset (\delta_m, 1-\delta_m)$, the empirical ROC curve estimator $\mathbb{R}_{m,n}(t)$ can be approximated by the following:

$$\begin{aligned}\mathbb{R}_{m,n}(t) &= \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) = \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)) + \frac{1}{n}K_2(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)), n) + O(\gamma_n)\\ &= \bar{G}(\bar{F}^{-1}(t)) + \frac{1}{n}K_2(\bar{G}(\bar{F}^{-1}(t)), n) + I_1 + I_2 + O(\gamma_n),\end{aligned}$$

where

$$I_1 = \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)) - \bar{G}(\bar{F}^{-1}(t)) = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m) \tag{13}$$

$$I_2 = n^{-1}K_2(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)), n) - n^{-1}K_2(\bar{G}(\bar{F}^{-1}(t)), n) = O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m) \tag{14}$$

Proof of (13) : By Theorem B of Appendix, we have

$$I_1 = g(\bar{F}^{-1}(t))\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))} - g'(\bar{F}^{-1}(\xi))\left(\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))}\right)^2$$

and ξ lies between t and $U_{\lceil mt\rceil:m}$, which ensures $f(\bar{F}^{-1}(\xi))/f(\bar{F}^{-1}(t)) \le 10^{\gamma}$ by Theorem N of Appendix. By Condition B, we get

$$\mid R'(t)O(\tau_m) \mid = \left| t(1-t)\frac{g(\bar{F}^{-1}(t))}{f(\bar{F}^{-1}(t))}\frac{O(\tau_m)}{t(1-t)} \right| \le O(\alpha_m^{-1}\tau_m) \tag{15}$$

$$\left|\xi(1-\xi)\frac{g'(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(\xi))}\frac{f^2(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))\xi(1-\xi)}\right|(m^{-1}K_1(t,m) - O(\tau_m))^2$$

$$\le O(\alpha_m^{-1}m^{-1}\log\log m) \quad a.s. \ \square \tag{16}$$

By combining (15) and (16), we finish the proof of (13). □

Proof of (14) : By modulus of continuity for Brownian motion on bounded interval, we get

$$\begin{aligned}
\mid I_2 \mid &= \mid \frac{1}{n}K_2(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)), n) - \frac{1}{n}K_2(\bar{G}(\bar{F}^{-1}(t)), n) \mid \\
&\le n^{-1/2} \mid I_1 \mid^{1/2} (\log(1/\mid I_1 \mid))^{1/2} \\
&\le n^{-1/2}\left|\frac{R'(t)}{m}K_1(t,m) + O(\alpha_m^{-1}\tau_m)\right|^{1/2} (\log m)^{1/2} \\
&\le n^{-1/2} \mid O(\alpha_m^{-1}m^{-1/2}(\log\log m)^{1/2}) + O(\alpha_m^{-1}\tau_m) \mid^{1/2} (\log m)^{1/2} \\
&= O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m) \ \square
\end{aligned} \tag{17}$$

Therefore, by combining the results (13) and (14) , we complete the proof of (10). □

2. **Proof of (11) of Theorem 4.1**: The idea of the proof of (11) is similar to (10), except for the following: (1) We can expand I_{12}^* in Taylor's series instead of I_1^* ($I_1^* =$

$I^*_{12}+I^*_{11}$); (2) Because the bootstrap version is expanded around the empirical estimator, the Taylor's series expansion is evaluated at the empirical point $\bar{\mathbb{F}}^{*-1}_m(t)$ instead of at the true point $\bar{F}^{-1}(t)$. Measuring these difference enlarges the range of domain from $t \in (\alpha_m, 1-\alpha_m)$ to $t \in (\alpha^*_m, 1-\alpha^*_m)$, which allows the limiting distribution based on bootstrap resampling distribution to be the same as that of the empirical one.

By Theorem 3.1 and Theorem C of Appendix, there exist two independent Kiefer processes K_1 and K_2, such that when $t \in (\alpha^*_m, 1-\alpha^*_m) \subset (\delta^*_m, 1-\delta^*_m)$, the strong approximation for bootstrap version of ROC curve is given by

$$\mathbb{R}^*_{m,n}(t) = \bar{\mathbb{G}}^*_n(\bar{\mathbb{F}}^{*-1}_m(t)) = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{-1}_m(t)) + I^*_1 + n^{-1}K_2(\bar{G}(\bar{F}^{-1}(t)), n) + I^*_2 + O(\tau_n)$$

where $I^*_1 = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{*-1}_m(t)) - \bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{-1}_m(t)) = I^*_{11} + I^*_{12}$, $I^*_{12} = \bar{G}(\bar{\mathbb{F}}^{*-1}_m(t)) - \bar{G}(\bar{\mathbb{F}}^{-1}_m(t))$, $I^*_{11} = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{*-1}_m(t)) - \bar{G}(\bar{\mathbb{F}}^{*-1}_m(t)) - (\bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{-1}_m(t)) - \bar{G}(\bar{\mathbb{F}}^{-1}_m(t)))$;

$I^*_2 = I^*_{21} + I^*_{22}$, $I^*_{21} = n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{*-1}_m(t)), n) - n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{-1}_m(t)), n)$,

$I^*_{22} = n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}^{-1}_m(t)), n) - n^{-1}K_2(\bar{G}(\bar{F}^{-1}(t)), n)$. We can show

$$I^*_{12} = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m) \tag{18}$$

$$I^*_{11} = O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m) \tag{19}$$

$$I^*_{21} = O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m) \tag{20}$$

$$I^*_{22} = O(\tau_m) \tag{21}$$

Proof of (18) : By taking Taylor's series expansion of I^*_{12}, we have

$$I^*_{12} = g(\bar{\mathbb{F}}^{-1}_m(t))\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))} - g'(\bar{F}^{-1}(\xi))\left(\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))}\right)^2 \tag{22}$$

where ξ lies between $U_{k:m}$ and $U_{k^*:m}$, $k = \lceil mt \rceil$, $k^* = \lceil mV_{k:m} \rceil$, $\bar{\mathbb{F}}_m^{*-1}(t) = \bar{F}^{-1}(U_{k^*:m})$, $\bar{\mathbb{F}}_m^{-1}(t) = \bar{F}^{-1}(U_{k:m})$.

- Because the first term of (22) can be split as

$$g(\bar{F}^{-1}(t))\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))} + \frac{g(\bar{\mathbb{F}}_m^{-1}(t)) - g(\bar{F}^{-1}(t))}{f(\bar{F}^{-1}(t))}(m^{-1}K_1(t,m) - O(\tau_m)) \tag{23}$$

  By (15), the first term of (23) can be majorized by $\frac{R'(t)}{m}K_1(t,m) + O(\alpha_m^{-1}\tau_m)$. The second term of (23) can be bounded by

$$\begin{aligned} &\left|\frac{g'(\bar{F}^{-1}(\xi^*))}{f(\bar{F}^{-1}(t))}\right| \mid \bar{\mathbb{F}}_m^{-1}(t) - \bar{F}^{-1}(t) \mid\mid m^{-1}K_1(t,m) - O(\tau_m) \mid \\ &\leq \left|\xi^*(1-\xi^*)\frac{g'(\bar{F}^{-1}(\xi^*))}{f^2(\bar{F}^{-1}(\xi^*))}\right| \frac{f^2(\bar{F}^{-1}(\xi^*))}{f^2(\bar{F}^{-1}(t))\xi^*(1-\xi^*)} O(m^{-1}\log\log m) \end{aligned}$$

  which is $O(\alpha_m^{-1}m^{-1}\log\log m)$, where ξ^* lies between $U_{k:m}$ and t. This condition ensures the boundedness of $\frac{f^2(\bar{F}^{-1}(\xi^*))}{f^2(\bar{F}^{-1}(t))}$ by Theorem N. Also, because $t \in (\alpha_m^*, 1-\alpha_m^*)$, by Theorem K, $\mid \frac{1}{\xi^*} \mid \leq O(\alpha_m^{-1})$. Therefore the first term of (22) can be bounded by $m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m)$.

- The second term of (22) can be majorized as follows:

$$\begin{aligned} &\left|\xi(1-\xi)\frac{g'(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(\xi))}\right| \frac{f^2(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))\xi(1-\xi)} O(m^{-1}\log\log m) \\ &\leq O(\alpha_m^{-1})O(m^{-1}\log\log m) \end{aligned}$$

By combining the results of these two parts, we prove the statement (18). □

Proof of (19) : By Theorem A, there exists Kiefer process K, such that I_{11}^* can be majorized by

$$
\begin{aligned}
&\left| \frac{K(\bar{G}(\bar{\mathbb{F}}_m^{*-1}(t)),n) - K(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)),n)}{n} \right| + O(n^{-1}\log^2 n) \\
&\le n^{-1/2} \mid I_{12}^* \mid^{1/2} (\log(1/ \mid I_{12}^* \mid))^{1/2} + O(n^{-1}\log^2 n) \qquad (24) \\
&= O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m) \text{ (following the similar argument in (17))} \quad \square
\end{aligned}
$$

By combining the results of I_{11}^* and I_{12}^*, we have

$$
I_1^* = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m), \qquad (25)
$$

Proof of (20) : By following the same argument of (17) using (25), the result (20) is immediate.

Proof of (21) : By applying (10), we have

$$
\begin{aligned}
&\mid I_{22}^* \mid \le n^{-1/2} \mid \mathbb{R}_{m,n} - R(t) \mid^{1/2} (\log(\mid \mathbb{R}_{m,n} - R(t) \mid)^{-1})^{1/2} \\
&\le n^{-1/2}\left| \frac{R'(t)}{m}K_1(t,m) + \frac{1}{n}K_2(R(t),n) + O(\alpha_m^{-1}\tau_m) \right|^{1/2} O((\log N)^{1/2}) \\
&\le n^{-1/2}O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m + n^{1/2}\tau_n + \alpha_m^{-1/2}\tau_m^{-1/2}(\log N)^{1/2}) = O(\tau_m). \ \square \qquad (26)
\end{aligned}
$$

Hence, by combining (18), (19), (20) and (21), we complete the proof of (11). □

3. **Proof of (12) of Theorem 4.1**: The proof of (12) follows the same arguments in the proof of (11) with the following modifications: (1) In order to use Taylor's series expansion and the boundedness result of the ratio of $\frac{f(\bar{F}^{-1}(\xi))}{f(\bar{F}^{-1}(t))}$, where ξ lies between t and $\tilde{\mathbb{U}}_{m-1}(t)$, we will do Taylor's series expansion on $I_{122}^{\#}$ instead of $I_{12}^{\#}$. That is, the decomposition $I_{12}^{\#} = I_{121}^{\#} + I_{122}^{\#}$ is essential; (2) Similarly, one consequence of

measuring the remainder term is to shrink the range of domain from $t \in (\alpha_m^*, 1-\alpha_m^*)$ to $t \in (\alpha_m^{\#}, 1-\alpha_m^{\#})$.

By Theorem 3.2 and Theorem C of Appendix, there exist two independent Kiefer processes K_1 and K_2, such that when $t \in (\alpha_m^{\#}, 1-\alpha_m^{\#}) \subset (\delta_m^{\#}, 1-\delta_m^{\#})$, the strong approximation for the BB version of ROC curve is given by

$$\mathbb{R}_{m,n}^{\#}(t) = \bar{\mathbb{G}}_n^{\#}(\bar{\mathbb{F}}_m^{\#-1}(t)) = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) + I_1^{\#} + n^{-1}K_2(\bar{G}(\bar{F}^{-1}(t)), n) + I_2^{\#} + O(\tau_n)$$

where

$$\begin{aligned}
I_1^{\#} &= \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) = I_{11}^{\#} + I_{12}^{\#} \\
I_{11}^{\#} &= \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{\#-1}(t)) - (\bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t))) \\
I_{12}^{\#} &= \bar{G}(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)) = I_{121}^{\#} + I_{122}^{\#} \\
I_{121}^{\#} &= \bar{G}(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t))) \\
I_{122}^{\#} &= \bar{G}(\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t))) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)) \\
I_2^{\#} &= \{n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{\#-1}(t)), n) - n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)), n)\} \\
&+ \{n^{-1}K_2(\bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)), n) - n^{-1}K_2(\bar{G}(\bar{F}^{-1}(t)), n)\},
\end{aligned}$$

and $\tilde{\mathbb{U}}_{m-1}(t)$ is the empirical function of $V_1, \dots, V_{m-1} \sim$ i.i.d. $U(0,1)$.

First, we will show

$$I_{122}^{\#} = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m) \tag{27}$$

Proof of (27) :

$$I_{122}^{\#} = \bar{G}(\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t))) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t))) = \{-g(\bar{\mathbb{F}}_m^{-1}(t))[\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t)) - \bar{\mathbb{F}}_m^{-1}(t)]\}$$
$$+ \{-g'(\bar{\mathbb{F}}_m^{-1}(\xi))[\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t))) - \bar{\mathbb{F}}_m^{-1}(t)]^2\}, \tag{28}$$

where ξ lies between t and $\tilde{\mathbb{U}}_{m-1}(t)$. The first term of (28) can be strongly approximated a.s. by $m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m)$, because $-g(\bar{\mathbb{F}}_m^{-1}(t))[\bar{\mathbb{F}}_m^{-1}\tilde{\mathbb{U}}_{m-1}(t) - \bar{\mathbb{F}}_m^{-1}(t)]$ can be split as

$$\{-g(\bar{\mathbb{F}}_m^{-1}(t))[\bar{\mathbb{F}}_m^{\#-1}(t) - \bar{\mathbb{F}}_m^{-1}(t)]\} + \{-g(\bar{\mathbb{F}}_m^{-1}(t))[\bar{\mathbb{F}}_m^{-1}\tilde{\mathbb{U}}_{m-1}(t) - \bar{\mathbb{F}}_m^{\#-1}(t)]\} \tag{29}$$

The first term of (29) can be strongly approximated by $m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m)$ due to Theorem 3.2, and by an argument similar to the proof of the bootstrap case; the second term of (29) can be bounded *a.s.* by $O(\frac{\log m}{m})$ due to Condition A and Lemma 3.2.

The second term of (28) can be bounded *a.s.* by $O(\alpha_m^{-1}\tau_m)$, because

$$\begin{aligned}
&| -g'(\bar{\mathbb{F}}_m^{-1}(\xi)) \mid [\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t)) - \bar{\mathbb{F}}_m^{-1}(t)]^2 \\
&\leq 2 \mid -g'(\bar{\mathbb{F}}_m^{-1}(\xi)) \mid \{[\bar{\mathbb{F}}_m^{\#-1}(t) - \bar{\mathbb{F}}_m^{-1}(t)]^2 + [\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t)) - \bar{\mathbb{F}}_m^{\#-1}(t)]^2\} \\
&\leq 2 \mid -g'(\bar{\mathbb{F}}_m^{-1}(\xi)) \mid [\bar{\mathbb{F}}_m^{\#-1}(t) - \bar{\mathbb{F}}_m^{-1}(t)]^2 + O((\frac{\log m}{m})^2) \text{ (by Lemma 3.2 and Condition A)} \\
&\leq 2 \mid -g'(\bar{\mathbb{F}}_m^{-1}(\xi)) \mid \left(\frac{m^{-1}K_1(t,m) - O(\tau_m)}{f(\bar{F}^{-1}(t))}\right)^2 + O((m^{-1}\log m)^2) \\
&\leq O(\frac{g'(\bar{\mathbb{F}}_m^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))}) \cdot O(m^{-1}\log\log m) + O((m^{-1}\log m)^2) \\
&\leq O(\alpha_m^{-1}\tau_m).
\end{aligned} \tag{30}$$

The statement (30) holds because

$$\left|\frac{g'(\bar{\mathbb{F}}_m^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))}\right| = \left|\frac{g'(\bar{\mathbb{F}}_m^{-1}(\xi))}{f^2(\bar{\mathbb{F}}_m^{-1}(\xi))}\right|\left|\frac{f^2(\bar{\mathbb{F}}_m^{-1}(\xi))}{f^2(\bar{F}^{-1}(\xi))}\right|\left|\frac{f^2(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))}\right| \tag{31}$$
$$\leq O(\alpha_m^{-1})$$

where the second and third term of (31) can be bounded by a constant by Theorem N; the first term of (31) can be bounded by $O(\alpha_m^{-1})$ due to Condition B and

$$|\ \xi - t\ | \leq (m-1)^{-1/2}(\log\log(m-1))^{1/2} \text{ by Theorem J}$$
$$t \in (\alpha_m^{\#}, 1-\alpha_m^{\#}) \qquad \square$$

In addition, $I_{121}^{\#}$ is bounded *a.s.* by $O(\frac{\log m}{m})$ by Condition A and Lemma 3.2. Hence $I_{12}^{\#} = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m)$. $I_{11}^{\#}$ can be bounded *a.s.* by $O((m/n)^{1/2}\alpha_m^{-1/2}\tau_m)$ followed the same argument of (24) incorporated with the result of $I_{12}^{\#}$. Therefore $I_1^{\#} = m^{-1}R'(t)K_1(t,m) + O(\alpha_m^{-1}\tau_m)$. The same argument used in the proof of the bootstrap case can be used to prove $I_2^{\#} = O(\tau_m)$. Hence, we complete the proof of (12). $\square$

Remark 1: The similar result of (10) can be found in Hsieh and Turnbull (1996). The difference between our strong approximation for $\mathbb{R}_{m,n}(t)$ given by (10) and their's is a different trade-off between interval of validity and the approximation error. They used any fixed subinterval of [0,1] and obtained a better approximation error. From practical point of view, inclusion of levels near 0 is important. Our results treats the domain of $R(t)$ as the whole of [0,1] in the limiting sense.

4.2 Implication of the functional limit theorems for ROC curves

From (10), then as $L^{\infty}[0,1]$ valued random function,

$$\begin{aligned} &\sqrt{N}(\mathbb{R}(t) - R(t))1(\{\alpha_m < t < 1-\alpha_m\}) \\ &= \left\{ \tfrac{\sqrt{N}}{\sqrt{m}} R'(t) \tfrac{K_1(t,m)}{\sqrt{m}} + \tfrac{\sqrt{N}}{\sqrt{n}} \tfrac{K_2(R(t),n)}{\sqrt{n}} + O(\sqrt{N}\alpha_m^{-1}\tau_m) \right\} 1(\{\alpha_m < t < 1-\alpha_m\}) \\ &\rightsquigarrow \tfrac{1}{\sqrt{\lambda}} R'(t) B_1(t) + \tfrac{1}{\sqrt{1-\lambda}} B_2(R(t)), \end{aligned} \tag{32}$$

in the sense of generalized weak convergence on non-separable spaces (Van der Vaart & Wellner, 1996), where B_1 and B_2 are independent Brownian bridges. To see that (32) holds, observe that the realizations of $\frac{1}{\sqrt{\lambda}} R'(t) B_1(t) + \frac{1}{\sqrt{1-\lambda}} B_2(R(t))$ are continuous and are tield to 0 at the two end points 0 and 1. Hence, as $m \to \infty$,

$$\frac{1}{\sqrt{\lambda}} R'(t) B_1(t) + \frac{1}{\sqrt{1-\lambda}} B_2(R(t)) 1(\{\alpha_m < t < 1-\alpha_m\}) \to \frac{1}{\sqrt{\lambda}} R'(t) B_1(t) + \frac{1}{\sqrt{1-\lambda}} B_2(R(t))$$

uniformly on $t \in [0,1]$. Similarly, from (12), we can get, *a.s.*, conditionally on samples,

$$\sqrt{N}(\mathbb{R}^{\#}(t) - \mathbb{R}(t))1(\{\alpha_m^{\#} < t < 1-\alpha_m^{\#}\} \rightsquigarrow \frac{1}{\sqrt{\lambda}} R'(t) B_1(t) + \frac{1}{\sqrt{1-\lambda}} B_2(R(t)). \tag{33}$$

Remark 2: Li et al. (1996) obtained a functional weak convergence result to a Gaussian limit for the empirical estimate ROC process using Donsker's theorem and functional Delta method. Although the weak convergence approach is very elegant, the functional delta method requires fixed mean and hence does not apply directly in the BB or bootstrap distribution. Further, the approach based on strong approximation lets us work with real valued random variables and ordinary Taylor's series expansion, derive a stronger representation, and treat the whole domain $t \in [0,1]$ in the limit. In contrast, the weak convergence approach must leave out some neighborhoods of 0 and 1 in the limit.

Result 1 If A is a continuity set for the process on the right hand side of (32), then

$$\Pr{}^{\#}\{\sqrt{N}(\mathbb{R}^{\#}(\cdot) - \mathbb{R}(\cdot)) \in A\} - \Pr\{\sqrt{N}(\mathbb{R}(\cdot) - R(\cdot)) \in A\} \to 0. \tag{34}$$

where $\Pr^{\#}$ stands for BB resampling distribution conditional on samples.

Result 2 [Equivalence of empirical and BB estimator of ROC] Conditionally on samples,

$$\sqrt{N} \parallel \mathrm{E}(\mathbb{R}^{\#}(\cdot) \mid X_1, \ldots, X_m, Y_1, \ldots, Y_n) - \mathbb{R}(\cdot) \parallel_{L^\infty} \to 0. \ \ a.s. \tag{35}$$

4.3 Limit theorems for functional and AUC

It is not difficult to show the following corollaries.

Corollary 4.2 Let $\psi : \mathbb{R} \to \mathbb{R}$ be a bounded function with continuous second derivatives. Then for any $q > 0$ being the continuity point of the limiting random variable

$$\sup_{t\in(0,1)} \{\psi'(R(t))[\lambda^{-1/2}R'(t)B_1(t) + (1-\lambda)^{-1/2}B_2(R(t))]\},$$

we have that almost surely for the given samples,

$$\begin{aligned} &\mid \Pr{}^{*}\{ \sup_{t\in(\alpha_m^*,1-\alpha_m^*)} \sqrt{N} \mid \psi(\mathbb{R}^*(t)) - \psi(\mathbb{R}(t)) \mid \leq q\} \\ &\quad - \Pr\{ \sup_{t\in(\alpha_m^*,1-\alpha_m^*)} \sqrt{N} \mid \psi(\mathbb{R}(t)) - \psi(R(t)) \mid \leq q\} \mid \to 0 \end{aligned} \tag{36}$$

$$\begin{aligned} &\mid \Pr{}^{\#}\{ \sup_{t\in(\alpha_m^{\#},1-\alpha_m^{\#})} \sqrt{N} \mid \psi(\mathbb{R}^{\#}(t)) - \psi(\mathbb{R}(t)) \mid \leq q\} \\ &\quad - \Pr\{ \sup_{t\in(\alpha_m^{\#},1-\alpha_m^{\#})} \sqrt{N} \mid \psi(\mathbb{R}(t)) - \psi(R(t)) \mid \leq q\} \mid \to 0 \end{aligned} \tag{37}$$

where $\Pr^*$ stands for bootstrap resampling distribution conditional on the given samples.

Remark 3: Equivalently, a metric d (such as Levy's metric) which can characterize weak convergence can be also used to quantify the nearness of the distributions.

Corollary 4.3 Let $\phi(\cdot) : C[0,1] \to \mathbb{R}$ be a bounded linear functional. Then conditionally on the samples, we have for all q

$$\Pr^*\{\sqrt{N}(\phi(\mathbb{R}^*) - \phi(\mathbb{R})) \le q\} - \Pr\{\sqrt{N}(\phi(\mathbb{R}) - \phi(R)) \le q\} \to 0 \;\; a.s. \tag{38}$$

$$\Pr^\#\{\sqrt{N}(\phi(\mathbb{R}^\#) - \phi(\mathbb{R})) \le q\} - \Pr\{\sqrt{N}(\phi(\mathbb{R}) - \phi(R)) \le q\} \to 0 \;\; a.s. \tag{39}$$

Corollary 4.4 [For partial AUC (pAUC)] Conditionally on the samples, a.s. for any $(e_1, e_2) \subset (0,1)$, we have $\sqrt{N}(\mathbb{A}^\#(e_1^*, e_2^*) - \hat{\mathbb{A}}(e_1^*, e_2^*)) \to_d N(0, \sigma^2(e_1, e_2))$, where $e_1^* = \max(e_1, \alpha_m^\#)$, $e_2^* = \min(e_2, 1 - \alpha_m^\#)$

$$\sigma^2(e_1, e_2) = \lambda^{-1} \int_{e_1}^{e_2} \int_{e_1}^{e_2} R'(t)R'(s)(s \wedge t - st)dtds$$
$$+(1-\lambda)^{-1} \int_{e_1}^{e_2} \int_{e_1}^{e_2} (R(t) \wedge R(s) - R(t)R(s))dtds,$$
$$\mathbb{A}^\#(e_1^*, e_2^*) = \int_{e_1^*}^{e_2^*} \mathbb{R}^\#(t)dt, \quad \hat{\mathbb{A}}(e_1^*, e_2^*) = \int_{e_1^*}^{e_2^*} \mathbb{R}(t)dt.$$

Remark 4: (1) It can be easily shown that Condition B ensures that $\sigma^2(0,1) < \infty$. (2) Asymptotic form for AUC ($e_1 = 0, e_2 = 1$) can be obtained from Corollary 4.3. (3) Similar results holds for the empirical estimate and the bootstrap resampling distribution. Further, it follows that

$$\sup\{|\Pr^\#\{\sqrt{N}(\mathbb{A}^\#(e_1^*, e_2^*) - \hat{\mathbb{A}}(e_1^*, e_2^*)) \le q\} - \Pr\{\sqrt{N}(\hat{\mathbb{A}}(e_1^*, e_2^*) - A(e_1^*, e_2^*)) \le q\}| : q \in \mathbb{R}\} \to 0 \;\; a.s.$$

where $A(e_1^*, e_2^*) = \int_{e_1^*}^{e_2^*} R(t)dt$.

5 Simulation study

5.1 Comparison of the frequentist variability and the BB resampled variability

To investigate the asymptotic equivalence of frequentist variability of the empirical estimate of AUC and resampled variability of AUC estimate based on the BB procedures, we simulated 6000 data sets of $X_1, \ldots, X_m \sim$ i.i.d. F = Normal (0,1), $Y_1, \ldots, Y_n \sim$ i.i.d. G = Normal$(1.868, 1.5^2)$ to get 6000 frequentist samples of $\sqrt{N}(\hat{\mathbb{A}} - A)$, where $\hat{\mathbb{A}} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} 1(Y_j > X_i)$, A being the true AUC value, $N = m + n$, m and n chosen to be 50 and 300, shown in Figure 1 (a) and (b) respectively; based on one simulated data set defined above, BB's one random realization of AUC denoted as $\mathbb{A}^{\#}$ can be obtained based on that BB's random realization of the ROC curve denoted as $\mathbb{R}^{\#}(t)$; hence, one BB resample of $\sqrt{N}(\mathbb{A}^{\#} - \hat{\mathbb{A}})$ can be generated. Let the resample size be 6000. The density plot (Figure 1) of the asymptotic normal distribution $\mathrm{N}(0, \sigma^2(0,1))$, where σ can be calculated by Corollary 4.3, overlays those of $\sqrt{N}(\hat{\mathbb{A}} - A)$ and $(\sqrt{N}(\mathbb{A}^{\#} - \hat{\mathbb{A}}) \mid \text{data})$. The grid points on [0,1] are chosen with equal interval length 0.001 to calculate AUC.

The accuracy of the BB distribution of AUC in estimating the sampling distribution of the empirical estimate of the AUC is studied. The accuracy of the normal limit is also studied for the same data. It is observed that both estimates are fairly close to the actual distribution obtained by simulation, but the BB estimate is somewhat more accurate than the normal limit.

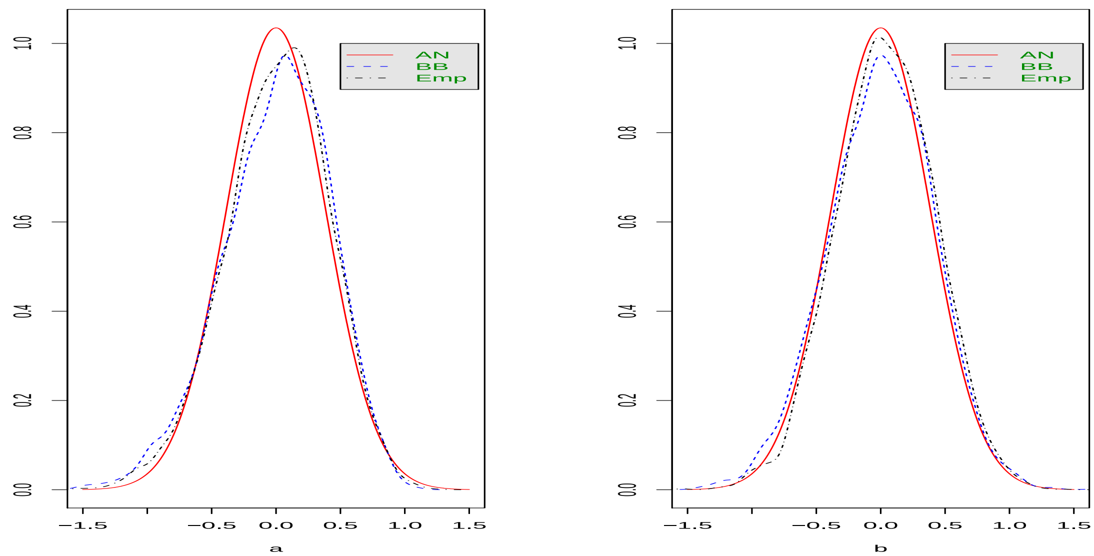

Figure 1: Comparison of density plots obtained by empirical $\sqrt{N}(\hat{\mathbb{A}} - A)$ based on 6000 simulated data sets, the BB's $(\sqrt{N}(\mathbb{A}^{\#} - \hat{\mathbb{A}})|\text{data})$ based on one simulated data set and corresponding 6000 resamples, and the asymptotic normal distribution $\mathrm{N}(0, \sigma^2(0,1))$ where σ can be calculated by Corollary 4.3.

5.2 Comparison of the empirical and the BB version of ROC

By simulating the data sets described above, with sample size $m = n = 50$ and $\psi = R(0.5)$, we compare the estimate of the ROC at the point $t = 0.5$. We obtain $P(\hat{\mathbb{R}}(0.5) \leq q)$ based on 1000 simulated data sets, and $P^{\#}(\mathbb{R}^{\#}(0.5) \leq q \mid \text{data})$ based on 1000 resamples. The grid points on [0,1] are chosen with equal interval length 0.001 to calculate ROC. The plot $\hat{P}(\hat{\mathbb{R}}(0.5) \leq q)$ and the interval of 25–75th percentile of $P^{\#}(\mathbb{R}^{\#}(0.5) \leq q \mid \text{data}\ )$ are given in Figure 2 .

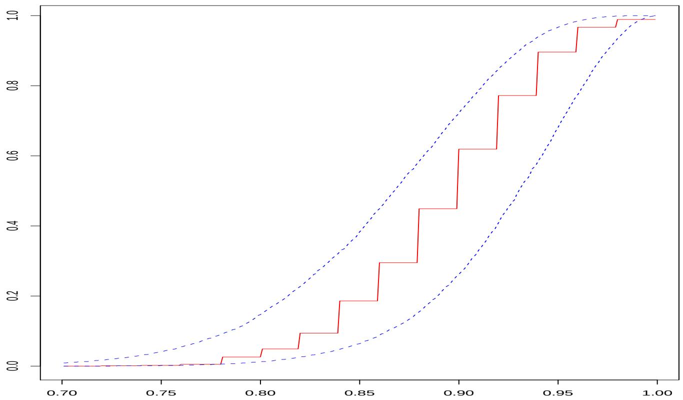

Figure 2: Comparison c.d.f. of Empirical (solid curve) and the BB's interval of 25–75th percentile (dotted curves) the cdf of the empirical estimate of $R(0.5)$. Based on sample size $m = n = 50$, 1000 simulated data sets, 1000 resample size, the grid points on [0,1] with equal interval 0.001.

6 Proof of the Lemmas

6.1 Proof of the Lemma 3.1

For any $y \in [\delta_n^*, 1-\delta_n^*]$, let $k = \lceil ny \rceil$, ξ lies between y and $V_{k:n}$. Note that $\mid y - k/n \mid \leq 1/n$.

6.1.1 Proof of the statement (2) of the Lemma 3.1

1. For any $y \in [\delta_n, 1-\delta_n]$, $V_{k:n} = y + n^{-1/2}\tilde{\mathbb{W}}_n(y)$, where $\tilde{\mathbb{W}}_n(\cdot)$ is the uniform quantile process for V_i's, $i = 1, \ldots, n$, then

$$\mid y - V_{k:n} \mid \leq 4n^{-1/2}(y(1-y))^{1/2}(\log\log n)^{1/2} \text{ by Theorem K of Appendix.} \quad (40)$$

2. When n is large, $(y(1-y))^{-1/2} \leq 2(\delta_n^*)^{-1/2}$. Then by (48), (50) of Theorem N of

Appendix and Condition A,

$$\begin{aligned}
&\sup_{\delta_n^*\le y\le 1-\delta_n^*}\sqrt{y(1-y)}\frac{\mid f'(F^{-1}(\xi))\mid}{f(F^{-1}(\xi))f(F^{-1}(V_{k:n}))}\\
&=\sup_{\delta_n^*\le y\le 1-\delta_n^*}[(y(1-y))^{-1/2}]\cdot\frac{y(1-y)}{\xi(1-\xi)}\cdot\xi(1-\xi)\frac{\mid f'(F^{-1}(\xi))\mid}{f^2(F^{-1}(\xi))}\cdot\frac{f(F^{-1}(\xi))}{f(F^{-1}(V_{k:n}))} \quad (41)\\
&\le_{a.s.} 2(2n^{-1/2}(\log\log n)^{1/2})^{-1/2}5\gamma 10^{\gamma} =_{a.s.} O(n^{1/4}(\log\log n)^{-1/4}).
\end{aligned}$$

3. δ_n^* ensures $V_{k:n}\in[\delta_n, 1-\delta_n]$ $a.s.$. By Theorem D of Appendix, we have

$$\sup_{\delta_n^*\le y\le 1-\delta_n^*}\mid f(F^{-1}(V_{k:n}))\mathbb{Q}_n(V_{k:n})-\mathbb{W}_n(V_{k:n})\mid\le 40\gamma 10^{\gamma}n^{-1/2}\log\log n,$$

$$\sup_{\delta_n^*\le y\le 1-\delta_n^*}\mid \mathbb{W}_n(V_{k:n})\mid\le 2(\log\log n)^{1/2}.$$

Hence

$$\sup_{\delta_n^*\le y\le 1-\delta_n^*}\mid f(F^{-1}(V_{k:n}))\mathbb{Q}_n(V_{k:n})\mid =_{a.s.} O((\log\log n)^{1/2}). \quad (42)$$

Combining results (40), (41) and (42), we now have

$$\begin{aligned}
&\sup_{\delta_n^*\le y\le 1-\delta_n^*}\left|(y-V_{k:n})\mathbb{Q}_n(V_{k:n})\frac{f'(F^{-1}(\xi))}{f(F^{-1}(\xi))}\right|\\
&\le \sup_{\delta_n^*\le y\le 1-\delta_n^*}\frac{\mid y-V_{k:n}\mid}{\sqrt{y(1-y)}}\cdot\left\{\sqrt{y(1-y)}\frac{\mid f'(F^{-1}(\xi))\mid}{f(F^{-1}(\xi))f(F^{-1}(V_{k:n}))}\right\}\\
&\quad\cdot\mid f(F^{-1}(V_{k:n}))\mathbb{Q}_n(V_{k:n})\mid =_{a.s.} O(n^{-1/4}(\log\log n)^{3/4})\le_{a.s.} O(l_n).\ \Box \quad (43)
\end{aligned}$$

6.1.2 Proof of the statement (3) of the Lemma 3.1

By Condition A,

$$f(F^{-1}(y)) = f(F^{-1}(V_{k:n})) + (y-V_{k:n})f'(F^{-1}(\xi))/f(F^{-1}(\xi)), \quad (44)$$

where ξ lies between y and $V_{k:n}$. Note that $y \in [\delta_n^*, 1-\delta_n^*]$ ensures that $V_{k:n} \in [\delta_n, 1-\delta_n]$ *a.s.*. Let K be the Kiefer process approximation to $\mathbb{Q}_n$ as in Theorem B of Appendix. Then we have

$$\begin{aligned}
&\sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid f(F^{-1}(y))\mathbb{Q}_n(V_{k:n}) - n^{-1/2}K(y,n) \mid \\
&\le \sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid f(F^{-1}(V_{k:n}))\mathbb{Q}_n(V_{k:n}) - n^{-1/2}K(V_{k:n},n) \mid \\
&\quad + \sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid (y - V_{k:n}) f'(F^{-1}(\xi))/f(F^{-1}(\xi))\mathbb{Q}_n(V_{k:n}) \mid \\
&\quad + \sup_{\delta_n^* \le y \le 1-\delta_n^*} \mid n^{-1/2}(K(V_{k:n},n) - K(y,n)) \mid . \qquad (45)
\end{aligned}$$

The first two terms of (45) are $O(l_n)$ *a.s.* respectively by Theorem B of Appendix and (2). The third term of (45) can be bounded by $\sup_{\delta_n^* \le y \le 1-\delta_n^*}\{\mid n^{-1/2}K(V_{k:n},n) - n^{-1/2}K(k/n,n) \mid + \mid n^{-1/2}K(k/n,n) - n^{-1/2}K(y,n) \mid\}$, which are $O(l_n)$, respectively, by Theorem E and Theorem F of Appendix (let $h_n = n^{-1}$). □

6.2 Proof of the Lemma 3.2

6.2.1 Proof of the statement (7) of Lemma 3.2

By Theorem M, we have

$$\begin{aligned}
&\sup_{0<y<1} \sqrt{n} \mid \mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) \mid \\
&\le \sup_{0<y<1} \sqrt{n} \left| \mathbb{F}_n^{-1}\left(\frac{1+(n-1)\tilde{\mathbb{U}}_{n-1}(y)}{n}\right) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) \right| \\
&\quad + \sup_{0<y<1} \sqrt{n} \left| \mathbb{F}_n^{-1}\left(\frac{(n-1)\tilde{\mathbb{U}}_{n-1}(y)}{n}\right) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) \right| \\
&\le \sup_{0\le k\le n} 2\sqrt{n} \mid X_{k+1:n} - X_{k:n} \mid =_{a.s.} O(n^{-1/2}\log n) \quad \square
\end{aligned}$$

6.2.2 Proof of the statement (8) of Lemma 3.2

Because $y \in [\delta_n^{\#}, 1-\delta_n^{\#}] \subset [\epsilon_n, 1-\epsilon_n]$, where $\epsilon_n = 0.236 n^{-1} \log\log n$, let $k = \lceil ny \rceil$, ξ lies between y and $\tilde{\mathbb{U}}_{n-1}(y)$. The following four steps are needed to finish the proof.

1. By similar argument of (43), we have

$$\begin{aligned}
&\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \left| (\tilde{\mathbb{U}}_{n-1}(y) - y) \frac{f'(F^{-1}(\xi))}{f(F^{-1}(\xi))} \mathbb{Q}_n(\tilde{\mathbb{U}}_{n-1}(y)) \right| \\
&\le \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \left\{ \frac{\mid \tilde{\mathbb{U}}_{n-1}(y) - y \mid}{\sqrt{y(1-y)}} \right\} \left\{ \sqrt{y(1-y)} \frac{\mid f'(F^{-1}(\xi)) \mid}{f(F^{-1}(\xi)) f(F^{-1}(\tilde{\mathbb{U}}_{n-1}(y)))} \right\} \\
&\quad \cdot \mid f(F^{-1}(\tilde{\mathbb{U}}_{n-1}(y))) \mathbb{Q}_n(\tilde{\mathbb{U}}_{n-1}(y)) \mid \\
&=_{a.s.} O(n^{-1/4} (\log\log n)^{3/4}) \le_{a.s} O(l_n)
\end{aligned}$$

2. By Theorem B of Appendix, $\mid \sqrt{n} f(F^{-1}(y))(\mathbb{F}_n^{-1}(y) - F^{-1}(y)) - n^{-1/2} K(y,n) \mid =_{a.s.} O(l_n)$.

3. We have

$$\begin{aligned}
&\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \mid \sqrt{n} f(F^{-1}(y))[F^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - F^{-1}(y)] - n^{-1/2} K(y,n) \mid \\
&\quad \le \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \mid \sqrt{n} f(F^{-1}(y))(\tilde{\mathbb{U}}_{n-1}(y) - y)/f(F^{-1}(y) - n^{-1/2} K(y,n) \mid \\
&\quad\quad + \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \left| (\tilde{\mathbb{U}}_{n-1}(y) - y)^2 \frac{f'(F^{-1}(\xi))}{f^2(F^{-1}(\xi))} \frac{f(F^{-1}(y))}{f(F^{-1}(\xi))} \right| \\
&\quad \le \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \mid \sqrt{n}(\tilde{\mathbb{U}}_{n-1}(y) - y) - n^{-1/2} K(y,n) \mid \\
&\quad\quad + \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \left| (\tilde{\mathbb{U}}_{n-1}(y) - y)^2 \frac{f'(F^{-1}(\xi))}{f^2(F^{-1}(\xi))} \frac{f(F^{-1}(y))}{f(F^{-1}(\xi))} \right| \qquad (46) \\
&\quad =_{a.s.} O(l_n)
\end{aligned}$$

The reasons for (46) are as follows:

(a) By Theorem A of Appendix, we have

$$\sup_{0<y<1} \sqrt{\frac{n}{n-1}} \left| \sqrt{n-1}(\tilde{\mathbb{U}}_{n-1}(y) - y) - \frac{K(y, n-1)}{\sqrt{n-1}} \right| =_{a.s.} O(\frac{\sqrt{n}\log^2(n-1)}{n-1})$$

By Theorems G and H of Appendix, we have

$$\begin{aligned}
&\sup_{0<y<1} \left| \frac{\sqrt{n}}{n-1} K(y, n-1) - n^{-1/2} K(y, n) \right| \\
&\leq \sup_{0<y<1} \left\{ \frac{\sqrt{n}}{n-1} \mid K(y, n-1) - K(y, n) \mid + \mid (\frac{\sqrt{n}}{n-1} - \frac{1}{\sqrt{n}}) K(y, n) \mid \right\} \\
&=_{a.s.} O\left(\sqrt{\frac{n}{n-1}} l_n + \frac{(\log \log n)^{\frac{1}{2}}}{n-1} \right) =_{a.s.} O(l_n)
\end{aligned}$$

Therefore, the first term of (46) can be majorized by

$$\begin{aligned}
&\sup_{0<y<1} \sqrt{\frac{n}{n-1}} \left| \sqrt{n-1}(\tilde{\mathbb{U}}_{n-1}(y) - y) - \frac{1}{\sqrt{n-1}} K(y, n-1) \right| \\
&+ \sup_{0<y<1} \left| \frac{\sqrt{n}}{n-1} K(y, n-1) - n^{-1/2} K(y, n) \right| =_{a.s.} O(l_n),
\end{aligned}$$

(b) By Theorem J of Appendix, we have $(\tilde{\mathbb{U}}_{n-1}(y)-y)^2 \leq_{a.s} 4(n-1)^{-1} y(1-y) \log\log(n-1)$. By following a similar argument to that of Theorem 3 (Csörgő and Révész, 1978), we get $\frac{f(F^{-1}(y))}{f(F^{-1}(\xi))} \leq 10^{\gamma}$. Also by Condition A, the second term of (46) can be majorized by

$$\begin{aligned}
&\sup_{\delta_n^{\#} \leq y \leq 1-\delta_n^{\#}} [4(n-1)^{-1} \log\log(n-1)] \left\{ \frac{y(1-y)}{\xi(1-\xi)} \right\} \left\{ \xi(1-\xi) \frac{f'(F^{-1}(\xi))}{f^2(F^{-1}(\xi))} \right\} \left\{ \frac{f(F^{-1}(y))}{f(F^{-1}(\xi))} \right\} \\
&\leq_{a.s.} O(l_n)
\end{aligned}$$

4. By Theorems G and I of Appendix, we have

$$\begin{aligned}
&\sup_{0<y<1} \frac{1}{\sqrt{n-1}} \mid K(\tilde{\mathbb{U}}_{n-1}(y), n) - K(y, n) \mid \\
&\leq \sup_{0<y<1} \frac{1}{\sqrt{n-1}} \mid K(\tilde{\mathbb{U}}_{n-1}(y), n) - K(\tilde{\mathbb{U}}_{n-1}(y), n-1) \mid \\
&\quad + \sup_{0<y<1} \frac{1}{\sqrt{n-1}} \mid K(y, n) - K(y, n-1) \mid \\
&\quad + \sup_{0<y<1} \frac{1}{\sqrt{n-1}} \mid K(\tilde{\mathbb{U}}_{n-1}(y), n-1) - K(y, n-1) \mid =_{a.s.} O(l_n)
\end{aligned}$$

Notice that $\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - \mathbb{F}_n^{-1}(y)$ can be split as

$$[\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - F^{-1}(\tilde{\mathbb{U}}_{n-1}(y))] - [\mathbb{F}_n^{-1}(y) - F^{-1}(y)] + [F^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - F^{-1}(y)] \qquad (47)$$

By combining parts (1)- (4) above and (47), we finish the proof of (8). □

7 Appendix

The following theorems were used:

Theorem A: [Implied by Komlós, Major and Tusnády, 1975, Theorem 4]: Let $X_1, X_2, \ldots \sim$ i.i.d. F, where F is any arbitrary continuous distribution function and $\mathbb{J}_n(x)$ be the empirical process. Then there exists a Kiefer process $\{K(y,t); 0 \leq y \leq 1, t \geq 0\}$, such that $\sup_x \left|\mathbb{J}_n(x) - n^{-1/2}K(F(x), n)\right| =_{a.s.} O(n^{-1/2}\log^2 n)$.

Theorem B: [Csörgő and Révész, 1978, Theorem 6]: Let $X_1, X_2, \ldots \sim$ i.i.d. F, satisfying Condition A. Then the quantile process $\mathbb{Q}_n(x)$ of X can be approximated by a Kiefer process $\{K(y,t); 0 \leq y \leq 1, t \geq 0\}$ as $\sup_{\delta_n \leq y \leq 1-\delta_n} \mid f(F^{-1}(y))\mathbb{Q}_n(y) - n^{-1/2}K(y, n) \mid =_{a.s.} O(l_n)$.

Theorem C: [Lo, 1987, Lemma 6.3]: There exists a Kiefer process $\{K(s,t); 0 \leq s \leq 1, t \geq 0\}$ independent of X such that $\sup_x \mid \mathbb{J}_n^{\#}(x) - n^{-1/2}K(\mathbb{F}_n(x), n) \mid =_{a.s.} O(l_n)$.

Theorem D: [Csörgő and Révész, 1978, Theorem 3]: Let $X_1, X_2, ...$ be i.i.d. F satisfying Condition A. Then $\limsup_{n\to\infty} \frac{\sqrt{n}}{\log\log n} \sup_{\delta_n \le y \le 1-\delta_n} \mid f(F^{-1}(y))\mathbb{Q}_n(y) - \mathbb{W}_n(y) \mid \le 40\gamma 10^{\gamma}$ $a.s.$ (γ is defined in Condition A)

Theorem E: [Csörgő and Révész, 1981, Lemma 4.5.1]: Let $U_1, U_2, \dots$ be i.i.d. $U(0,1)$ random variables, and let Kiefer process $\{K(y,t); 0 \le y \le 1, 0 \le t\}$ defined on the same probability space. Then $\sup_{1\le k\le n} n^{-1/2} \mid K(U_{k:n}, n) - K(k/n, n) \mid =_{a.s.} O(l_n)$.

Theorem F: [Csörgő and Révész, 1981,Theorem 1.15.2]: Let h_n be a sequence of positive numbers for which $\lim_{n\to\infty} \frac{\log h_n^{-1}}{\log\log n} = \infty$, $\gamma_n = (2nh_n \log h_n^{-1})^{-1/2}$. Then

$$\lim_{n\to\infty} \sup_{0\le t\le 1-h_n} \sup_{0\le s\le h_n} \gamma_n \mid K(t+s,n) - K(t,n) \mid =_{a.s} 1.$$

Theorem G: Let $K(s,n)$ be a Kiefer process, $\sup_{0<s<1} \left| \frac{K(s,n-1) - K(s,n)}{\sqrt{n-1}} \right| =_{a.s.} O(l_n)$. The result is contained in the proof of Lemma 6.3 of Lo (1987).

Theorem H: [Law of Iterated Logarithm (LIL)]: Let $K(y,n)$ be Kiefer process,

$$\limsup_{n\to\infty} \sup_{0\le y\le 1} \frac{\mid K(y,n) \mid}{(2n\log\log n)^{\frac{1}{2}}} =_{a.s.} \frac{1}{2}.$$

Theorem I: [Special case of Lemma 6.2 (Lo, 1987) when $F = U(0,1)$]: Let $U_1, \cdots, U_m \sim$ i.i.d. $U(0,1)$, $\mathbb{U}_m(s)$ is the empirical function of U_i's. For any Kiefer process K independent of $U_1, \cdots, U_m$, then $\sup_{0<s<1} \left| \frac{K(\mathbb{U}_m(s), m) - K(s,m)}{\sqrt{m}} \right| =_{a.s.} O(l_m)$.

Theorem J: [Csörgő and Révész, 1978, Theorem D] : For $\mathbb{H}_n(x)$ the uniform empirical process and $\epsilon_n = 0.236\frac{\log\log n}{n}$, we have $\limsup_{n\to\infty} \sup_{\epsilon_n \le x \le 1-\epsilon_n} (x(1-x)\log\log n)^{-1/2} |\mathbb{H}_n(x)| =_{a.s.} 2$.

Theorem K: [Csörgő and Révész, 1978, Theorem 2]: Let $\mathbb{W}_n(y)$ be the uniform quantile process, then $\limsup_{n\to\infty} \sup_{\delta_n \le y \le 1-\delta_n} (y(1-y)\log\log n)^{-1/2} \mid \mathbb{W}_n(y) \mid =_{a.s.} 4$.

Theorem L: [Csörgő and Révész, 1978, Lemma 1]: Under Condition A,

$$\frac{f(F^{-1}(y_1))}{f(F^{-1}(y_2))} \le \left\{ \frac{y_1 \vee y_2}{y_1 \wedge y_2} \cdot \frac{1 - y_1 \wedge y_2}{1 - y_1 \vee y_2} \right\}^{\gamma} \text{ for any pair } y_1, y_2 \in (0,1).$$

Theorem M: [Slud, 1978]: Let $U_1, \ldots, U_n \sim$ i.i.d. $U(0,1)$ and $M_n = \max_{0 \le k \le n}(U_{k+1:n} - U_{k:n})$. Then $nM_n / \log n \to_{a.s} 1$.

Theorem N: Under condition A, for any $y \in [\delta_n^*, 1 - \delta_n^*] \subset [\delta_n, 1 - \delta_n]$, let $k = \lceil ny \rceil$, ξ lies between y and $V_{k:n}$. Then

$$\frac{y(1-y)}{\xi(1-\xi)} \le 5 \tag{48}$$

$$\frac{\xi}{V_{k:n}} \le 5, \quad \frac{1-\xi}{1-V_{k:n}} \le 5 \tag{49}$$

$$\sup_{\delta_n^* \le y \le 1-\delta_n^*} f(F^{-1}(\xi))/f(F^{-1}(V_{k:n})) \le 10^{\gamma} \tag{50}$$

Remark: When $y \in [\delta_n^{\#}, 1 - \delta_n^{\#}]$, ξ lies between y and $\tilde{\mathbb{U}}_{n-1}(y)$. By Theorem J of Appendix, the inequalities (48), (49) and (50) hold for y, ξ and $\tilde{\mathbb{U}}_{n-1}(y)$ by the same arguments, yielding

$$\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \frac{y(1-y)}{\xi(1-\xi)} \le 5 \tag{51}$$

$$\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \frac{\xi}{\tilde{\mathbb{U}}_{n-1}(y)} \le 5, \qquad \sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} \frac{1-\xi}{1-\tilde{\mathbb{U}}_{n-1}(y)} \le 5 \tag{52}$$

$$\sup_{\delta_n^{\#} \le y \le 1-\delta_n^{\#}} f(F^{-1}(\xi))/f(F^{-1}(\tilde{\mathbb{U}}_{n-1}(y))) \le 10^{\gamma} \tag{53}$$

The results are contained in the proof of Theorem 3 of Csörgő and Révész (1978).

8 References

1. Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating graph. *Journal of Mathematical Psychology* . **12**, 387–415.

2. Csörgő, M. and Révész, P. (1978). Strong approximations of the quantile process. *The Annals of Statistics* **6**, 882–894.

3. Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic press, New York.

4. Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* John Wiley & Sons, New York.

5. Gu, J. , Ghosal, S. and Roy A. (2006). Non-parametric estimation of ROC curve. *Institute of Statistics Mimeo Series* **2592**, North Carolina State University.

6. Hsieh, F. S. and Turnbull, B. W. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *The Annals of Statistics* **24**, 25-40.

7. Komlós, J. , Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. Z. *Wahrsch. verw. Gebiete* **32**, 111–131.

8. Li, G. Tiwari, R. C. and Wells, M. T. (1996). Quantile comparison functions in two-sample problems: With applications to comparisons of diagnostic markers. *Journal of the American Statistical Association* **91**, 689–698.

9. Lo, A. Y. (1987). A large sample study of the Bayesian bootstrap. *The Annals of Statistics* **15**, 360–375.

10. Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine* **8**, 283–298.

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

12. Slud, E. (1978). Entropy and maximal spacings for random partitions. *Z. Wahrsch. Verw. Gebiete* **41**, 341–352.

13. Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York.

14. Wieand, S., Gail, M. H., James, B. R. and James, K. L. (1989). A family of non-parametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika* **76**, 585–592.