

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 a composition of survival function of one population to 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 bootstrap and the Bayesian bootstrap confidence bands, and accurate frequentist coverage probabilities of bootstrap and the Bayesian bootstrap confidence intervals for the area under the curve functional of the ROC.

Key words: Bayesian bootstrap; Bootstrap; Empirical process; Kiefer process; ROC curve; Strong approximations.

¹Department of Statistics, North Carolina State University, Raleigh, NC 27695. Email: jgu@unity.ncsu.edu.

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

³Both authors are partially supported by NSF grant DMS-0349111.

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, has become a popular method for measuring the accuracy of diagnostic tests since the 1970s (Metz, 1978). The true positive rate and the false positive rate can be obtained by varying the threshold criterion. The main attractions of ROC curve may be described by the following properties: (1) it can display the trade-off between the true positive rate and false positive rate by varying the decision threshold values in a unit graph; (2) it can be compared with ROC curves of other diagnostic tests, even with different measurements of the diagnostic variables. 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 disease) is greater than the diagnostic value of a randomly chosen patient without the positive condition. A nonparametric estimator of ROC may be obtained by substituting into the empirical distributions, 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. In this paper, we develop strong approximations of bootstrap and BB quantile process by a sequence of appropriate Gaussian processes (more specifically, Kiefer processes). Combining these results with some earlier results from strong approximation theory in an appropriate way, we shall develop strong approximations for the empirical, bootstrap and BB version of the ROC process, its AUC and other functionals. In particular, these results imply a Gaussian weak limit for the processes and asymptotically valid coverage probabilities for the resulting confidence bands and intervals.

The existing strong approximation theory primarily includes 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.

In this paper, we will first obtain strong approximations for quantile processes of resampling processes based on the bootstrap and the BB. 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 $Y \sim G$. As a consequence, we will obtain the strong approximations for the empirical estimate of $R(t)$ and the corresponding bootstrap, and the Bayesian versions of $R(t)$. Interestingly, it will be seen that the forms of these Gaussian approximations are identical, and therefore the distribution of the ROC function, conditioned on the samples, is 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 resampled variability of bootstrap and BB procedures, given that the samples and the “posterior” mean (i.e, mean of ROC under BB distribution) are 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.

2 Notation

2.1 Preliminary notation

Before introducing the notations for empirical and quantile processes, we will define some commonly used notation:

1. Define the domain of X as $[a, b]$: $a = \sup\{x : F(x) = 0\}$, $b = \inf\{x : F(x) = 1\}$; 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.*; Define inverse of a general c.d.f. as follows: $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $t \in [0, 1]$; $X_1, \dots, X_n \sim \text{i.i.d. } F$, equivalently, $X_j = F^{-1}(U_j)$, where $U_1, \dots, U_n \sim \text{i.i.d. } U(0, 1)$. $X_{j:n}$ denote the j th largest order statistic based on $\{X_1, \dots, X_n\}$, $j = 0, 1, \dots, n + 1$, where $X_{0:n} = a, X_{n+1:n} = b$. We define $V_{j:n}, U_{j:n}$ in the same way in this paper.

2. **Condition A** (Csörgő and Révész, 1978): Let X_1, \dots be i.i.d. random variables with a continuous distribution function F which is twice differentiable on (a, b) , where $F' = f \neq 0$ on (a, b) . For some $\gamma > 0$, $\sup_{a < x < b} F(x)(1 - F(x)) \left| \frac{f'(x)}{f^2(x)} \right| \leq \gamma$.

3. **Condition B** Let F and G satisfy Condition A and the following conditions:

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

For example, if $F = \text{Normal}(0, 1)$, $G = \text{Normal}(1, 1)$, then Condition A and B hold since $\lim_{x \rightarrow +\infty} \int_x^\infty x e^{-(t^2 - x^2)/2} dt = 2$.

4. $\delta_n = 25n^{-1} \log \log n$, $\delta_n^* = \delta_n + 2n^{-1/2}(\log \log n)^{1/2}$, $\delta_n^\# = \delta_n + n^{-1/2}(\log \log n)^{1/2}$.
 $l_n = n^{-1/4}(\log \log n)^{1/4}(\log n)^{1/2}$, $\tau_m = m^{-3/4}(\log \log m)^{1/4}(\log m)^{1/2}$, $\gamma_n = n^{-1} \log^2 n$.

$\alpha_m^{-1} = o(m^{1/4})$, $\alpha_m^* = \alpha_m + 2m^{-1/2}(\log \log m)^{1/2}$, $\alpha_m^\# = \alpha_m + (m-1)^{-1/2}(\log \log(m-1))^{1/2}$. The symbol \square is used at the end of the proof.

2.2 Empirical and quantile functions:

For $X_1, \dots, X_n \sim \text{i.i.d. } F$,

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

$$\text{empirical process: } \mathbb{J}_n(x) = \sqrt{n}(\mathbb{F}_n(x) - F(x)) \quad (2)$$

$$\text{quantile function: } \mathbb{F}_n^{-1}(y) = X_{k:n} = F^{-1}(U_{k:n}), \text{ where } k = \lceil ny \rceil \quad (3)$$

$$\text{quantile process: } \mathbb{Q}_n(y) = \sqrt{n}(\mathbb{F}_n^{-1}(y) - F^{-1}(y)) \quad (4)$$

In particular, we will study bootstrap and BB resampling quantile processes.

2.2.1 Resampling distribution under bootstrap

$X_1, \dots, X_n \sim \text{i.i.d. } F$, bootstrap resample is given by $\{\mathbb{F}_n^{-1}(V_{j:n}), j = 1, \dots, n; V_1, \dots, V_n \sim \text{i.i.d. } U(0, 1) \text{ independent of } X_i\text{'s}\}$. 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

$$\mathbb{J}_n^*(x) = \sqrt{n}(\mathbb{F}_n^*(x) - \mathbb{F}_n(x)) \quad (5)$$

$$\mathbb{Q}_n^*(y) = \sqrt{n}(\mathbb{F}_n^{*-1}(y) - \mathbb{F}_n^{-1}(y)) = \sqrt{n}(\mathbb{F}_n^{-1}(V_{k^*:n}) - \mathbb{F}_n^{-1}(y)) = \sqrt{n}(F^{-1}(U_{k^*:n}) - F^{-1}(U_{k:n})), \quad (6)$$

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

2.2.2 Resampling distribution under BB

For $X_1, \dots, X_n \sim \text{i.i.d. } F$, BB distribution (c.d.f.) is defined as follows:

$$\mathbb{F}_n^\#(x) = \sum_{1 \leq j \leq n} \Delta_{j:n} 1(X_{j:n} \leq x) \quad x \in \mathbb{R},$$

where $V_1, \dots, V_{n-1} \sim \text{i.i.d. } U(0, 1)$, independent of X_i 's, $\Delta_{j:n} = V_{j:n-1} - V_{j-1:n-1}$, $j = 1, \dots, n$.

The BB quantile function is defined as follows:

$$\mathbb{F}_n^{\#-1}(y) = \begin{cases} X_{k:n}, & V_{k-1:n-1} < y \leq V_{k:n-1}, \quad k = 1, 2, \dots, n, \\ X_{0:n}, & y = 0. \end{cases}$$

BB empirical and quantile processes are defined as follows:

$$\mathbb{J}_n^\#(x) = \sqrt{n}(\mathbb{F}_n^\#(x) - \mathbb{F}_n(x)), \tag{7}$$

$$\mathbb{Q}_n^\#(y) = \sqrt{n}(\mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(y)). \tag{8}$$

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 BB quantile processes

Theorem 3.1 (Strong approximation for bootstrap quantile process) Let $X_1, X_2, \dots \sim$ 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

$$\sup_{\delta_n^* \leq y \leq 1 - \delta_n^*} | f(F^{-1}(y))\mathbb{Q}_n^*(y) - n^{-1/2}K(y, n) | =_{a.s.} O(l_n). \quad (9)$$

To prove this theorem, we need the following lemma whose proof is given in Section 5.

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

$$\sup_{\delta_n^* \leq y \leq 1 - \delta_n^*} | f(F^{-1}(y))\mathbb{Q}_n(V_{k:n}) - n^{-1/2}K(y, n) | =_{a.s.} O(l_n), \quad (10)$$

where $k = \lceil ny \rceil$.

Proof of Theorem 3.1 :

When $\delta_n^* \leq y \leq 1 - \delta_n^*$, $[\delta_n^*, 1 - \delta_n^*] \subset [\delta_n, 1 - \delta_n]$. Let $k = \lceil ny \rceil$ and $k^* = \lceil nV_{k:n} \rceil$. Then $\mathbb{Q}_n^*(y)$ can be split into summation of three parts as $\mathbb{Q}_n(V_{k:n}) + \tilde{\mathbb{Q}}_n(y) - \mathbb{Q}_n(y)$, 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, $i = 1, \dots, n$. By Theorem B of Appendix and Lemma 3.1, Theorem 3.1 is immediate. \square

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

by a Kiefer process K , in the sense that

$$\sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} |f(F^{-1}(y))\mathbb{Q}_n^\#(y) - n^{-1/2}K(y, n)| =_{a.s.} O(l_n). \quad (11)$$

The proof requires the following lemma whose proof is deferred to Section 5.

Lemma 3.2 Let $X_1, \dots, X_n \sim$ i.i.d. F satisfying Condition A, $V_1, \dots, V_{n-1} \sim$ i.i.d. $U(0, 1)$, $\mathbb{F}_n^{-1}(y)$, $\tilde{\mathbb{U}}_{n-1}(x)$ and $\mathbb{F}_n^{\#-1}(y)$ denote the quantile function of X 's, empirical function of V 's and quantile function of the Bayesian bootstrap respectively. Then

$$\sup_{0 < y < 1} \sqrt{n} | \mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) | =_{a.s.} O(n^{-1/2} \log n), \quad (12)$$

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

$$\sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} | \sqrt{n} f(F^{-1}(y)) (\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - \mathbb{F}_n^{-1}(y)) - n^{-1/2} K(y, n) | =_{a.s.} O(l_n). \quad (13)$$

Proof of Theorem 3.2 :

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 $\tilde{\mathbb{U}}_{n-1}(y)$)

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

Also, the BB quantile process $\mathbb{Q}_n^\#(y)$ can be split as $\sqrt{n}(\mathbb{F}_n^{\#-1}(y) - \mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y))) + \sqrt{n}(\mathbb{F}_n^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - \mathbb{F}_n^{-1}(y))$. Then Theorem 3.2 follows by applying Lemma 3.2 and Condition A. \square

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 |
| $= V_{1:n-1}$ | $X_{1:n}$ | $1/(n-1)$ |
| \dots | \dots | \dots |
| $V_{k:n-1} < y < V_{k+1:n-1}$ | $X_{k+1:n}$ | $k/(n-1)$ |
| $y = V_{k+1:n-1}$ | $X_{k+1:n}$ | $(k+1)/(n-1)$ |
| \dots | \dots | \dots |
| $V_{n-2:n-1} < y < V_{n-1:n-1}$ | $X_{n-1:n}$ | $(n-2)/(n-1)$ |
| $y = V_{n-1:n-1}$ | $X_{n-1:n}$ | 1 |
| $V_{n-1:n-1} < y \leq 1$ | $X_{n:n}$ | 1 |

4 Functional limit theorems for ROC curves and their AUCs

4.1 Functional limit theorems for ROC curves

Let $X_1, X_2, \dots, X_m \sim \text{i.i.d. } F$, and $Y_1, Y_2, \dots, Y_n \sim \text{i.i.d. } G$, which are c.d.f.'s of populations of interest. 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.

The ROC curve is defined as $\{(P(X > t), P(Y > t)) : X \sim F, Y \sim G, t \in \mathbb{R}\}$, or alternatively as $R(t) = \bar{G}(\bar{F}^{-1}(t))$; its derivative is $R'(t) = g(\bar{F}^{-1}(t))/f(\bar{F}^{-1}(t))$.

Theorem 4.1 (Functional limit theorem for empirical, bootstrap and BB's ROC curve estimators, denoted as $\mathbb{R}_{m,n}(t)$, $\mathbb{R}_{m,n}^*(t)$, $\mathbb{R}_{m,n}^\#(t)$ respectively.) Let $X_1, \dots, X_m \sim \text{i.i.d. } F$, $X_i = F^{-1}(U_i)$ and $Y_1, \dots, Y_n \sim \text{i.i.d. } G$. Assume F and G satisfy Condition A and B,

$N = m + n$, $\frac{m}{N} \rightarrow \lambda$, $\frac{n}{N} \rightarrow 1 - \lambda$. Then

$$\mathbb{R}_{m,n}(t) = R(t) + R'(t) \frac{K_1(t, m)}{m} + \frac{K_2(R(t), n)}{n} + O(\alpha_m^{-1} \tau_m), \quad t \in (\alpha_m, 1 - \alpha_m), \quad (14)$$

$$\mathbb{R}_{m,n}^*(t) = \mathbb{R}_{m,n}(t) + R'(t) \frac{K_1(t, m)}{m} + \frac{K_2(R(t), n)}{n} + O(\alpha_m^{-1} \tau_m), \quad t \in (\alpha_m^*, 1 - \alpha_m^*), \quad (15)$$

$$\mathbb{R}_{m,n}^\#(t) = \mathbb{R}_{m,n}(t) + R'(t) \frac{K_1(t, m)}{m} + \frac{K_2(R(t), n)}{n} + O(\alpha_m^{-1} \tau_m), \quad t \in (\alpha_m^\#, 1 - \alpha_m^\#). \quad (16)$$

in the sense of *a.s.* conditionally on the observed samples, where K_1 and K_2 are independent generic Kiefer processes (not identical in each appearance).

Proof of Theorem 4.1 :

1. **Proof of (14) 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

$$I_1 = \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)) - \bar{G}(\bar{F}^{-1}(t)) = m^{-1} R'(t) K_1(t, m) + O(\alpha_m^{-1} \tau_m), \quad (17)$$

$$I_2 = n^{-1} K_2(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)), n) - n^{-1} K_2(\bar{G}(\bar{F}^{-1}(t)), n) = O((m/n)^{1/2} \alpha_m^{-1/2} \tau_m) \quad (18)$$

Proof of (17) : 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 $f(\bar{F}^{-1}(\xi))/f(\bar{F}^{-1}(t)) \leq 10^\gamma$. By Condition B, we get

$$|R'(t)O(\tau_m)| = \left| t(1-t) \frac{g(\bar{F}^{-1}(t))}{f(\bar{F}^{-1}(t))} \frac{O(\tau_m)}{t(1-t)} \right| \leq O(\alpha_m^{-1} \tau_m) \quad (19)$$

$$\begin{aligned} & \left| \xi(1-\xi) \frac{g'(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(\xi))} \frac{f^2(\bar{F}^{-1}(\xi))}{f^2(\bar{F}^{-1}(t))\xi(1-\xi)} \right| (m^{-1}K_1(t, m) - O(\tau_m))^2 \\ & \leq O(\alpha_m^{-1} m^{-1} \log \log m) \quad a.s. \quad \square \end{aligned} \quad (20)$$

By combining (19) and (20), we finish the proof of (17). \square

Proof of (18) : By modulus of continuity for Brownian motion on bounded interval, we get

$$\begin{aligned} |I_2| &= \left| \frac{1}{n} K_2(\bar{G}(\bar{\mathbb{F}}_m^{-1}(t)), n) - \frac{1}{n} K_2(\bar{G}(\bar{F}^{-1}(t)), n) \right| \\ &\leq n^{-1/2} |I_1|^{1/2} (\log(1/|I_1|))^{1/2} \\ &\leq n^{-1/2} \left| \frac{R'(t)}{m} K_1(t, m) + O(\alpha_m^{-1} \tau_m) \right|^{1/2} (\log m)^{1/2} \\ &\leq n^{-1/2} \left| O(\alpha_m^{-1} m^{-1/2} (\log \log m)^{1/2}) + O(\alpha_m^{-1} \tau_m) \right|^{1/2} (\log m)^{1/2} \\ &= O((m/n)^{1/2} \alpha_m^{-1/2} \tau_m) \quad a.s. \quad \square \end{aligned} \quad (21)$$

The proof of (14) is now complete. \square

2. Proof of (15) of Theorem 4.1: The idea of the proof of (15) is similar to (14), except for the following: (1) A major difference is that Theorem 3.1 in Section 3 is used instead of Theorem B of Appendix; (2) We can expand I_{12}^* in Taylor's series instead of I_1^* , where $I_1^* = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{*-1}(t)) - \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) = I_{11}^* + I_{12}^*$, $I_{12}^* = \bar{G}(\bar{\mathbb{F}}_m^{*-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t))$, $I_{11}^* = \bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{*-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{*-1}(t)) - (\bar{\mathbb{G}}_n(\bar{\mathbb{F}}_m^{-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t)))$; (3) Because the bootstrap

estimator is expanded around the empirical estimator, the Taylor's series expansion is evaluated at the empirical point $\bar{\mathbb{F}}_m^{*-1}(t)$ 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 of bootstrap estimator to be the same as the empirical estimator's. \square

3. Proof of (16) of Theorem 4.1: The proof of (16) follows the same lines as the proof of (15) with the following modifications: (1) Theorem 3.2 is applied instead of Theorem 3.1 in Section 3; (2) 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, where $I_{12}^\# = \bar{G}(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t))$, $I_{121}^\# = \bar{G}(\bar{\mathbb{F}}_m^{\#-1}(t)) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t)))$, $I_{122}^\# = \bar{G}(\bar{\mathbb{F}}_m^{-1}(\tilde{\mathbb{U}}_{m-1}(t))) - \bar{G}(\bar{\mathbb{F}}_m^{-1}(t))$; (3) Similarly, one consequence of measuring the remainder term is to shrink the range of domain from $t \in (\alpha_m^*, 1 - \alpha_m^*)$ to $t \in (\alpha_m^\#, 1 - \alpha_m^\#)$. \square

4.2 Implication of the functional limit theorems for ROC curves

From (14), let $N = m + n$ and assume that $\frac{m}{N} \rightarrow \lambda$, $\frac{n}{N} \rightarrow 1 - \lambda$, then as $L^\infty[0, 1]$ valued random function,

$$\sqrt{N}(\mathbb{R}_{m,n}(t) - R(t)) \rightsquigarrow \frac{1}{\sqrt{\lambda}}R'(t)B_1(t) + \frac{1}{\sqrt{1-\lambda}}B_2(R(t)), \quad (22)$$

in the sense of general weak convergence (Van der Vaart & Wellner, 1996), where B_1 and B_2 are independent Brownian bridges. Similarly, from (15) and (16), we can get, *a.s.*,

conditionally on samples,

$$\sqrt{N}(\mathbb{R}_{m,n}^*(t) - \mathbb{R}_{m,n}(t)) \rightsquigarrow \frac{1}{\sqrt{\lambda}}R'(t)B_1(t) + \frac{1}{\sqrt{1-\lambda}}B_2(R(t)), \quad (23)$$

$$\sqrt{N}(\mathbb{R}_{m,n}^\#(t) - \mathbb{R}_{m,n}(t)) \rightsquigarrow \frac{1}{\sqrt{\lambda}}R'(t)B_1(t) + \frac{1}{\sqrt{1-\lambda}}B_2(R(t)). \quad (24)$$

Remark: We suspect that a possibly shorter approach to the distribution of (22), (23), (24) may be based on Donsker theorem, conditional Donsker theorem for bootstrap and Hadamard differentiability (see Section 3.9 of Van der Vaart & Wellner, 1996). However, calculation of Hadamard derivative of $(F, G) \mapsto \bar{G} \circ \bar{F}^{-1}$ seems very challenging. 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. From practical point of view, inclusion of levels near 0 is important.

Result 1 For any continuity set $A \subset L^\infty[0, 1]$ of the process on the right hand side of (22), interpreting probabilities as outer probabilities, then

$$\Pr^*\{\sqrt{N}(\mathbb{R}_{m,n}^*(\cdot) - \mathbb{R}_{m,n}(\cdot)) \in A\} - \Pr\{\sqrt{N}(\mathbb{R}_{m,n}(\cdot) - R(\cdot)) \in A\} \rightarrow 0. \quad (25)$$

where \Pr^* stands for bootstrap probability conditional on samples.

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

$$\sqrt{N} \|\mathbb{E}(\mathbb{R}_{m,n}^\#(\cdot) \mid X_1, \dots, X_m, Y_1, \dots, Y_n) - \mathbb{R}_{m,n}(\cdot)\|_{L^\infty} \rightarrow 0. \quad a.s. \quad (26)$$

4.3 Functional limit theorems for AUC

It is not difficult to show the following corollaries.

Corollary 1 Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded function with continuous second derivatives, $N = m + n$ and assume that $\frac{m}{N} \rightarrow \lambda$, $\frac{n}{N} \rightarrow 1 - \lambda$. Then as $N \rightarrow \infty$, for any $q > 0$ being the continuity point of the limiting random variable $\sup_{t \in (0,1)} \{\psi'(R(t))[\lambda^{-1/2}R'(t)B_1(t) + (1 - \lambda)^{-1/2}B_2(R(t))]\}$, interpreting probabilities as outer probabilities,

$$\begin{aligned} & | \Pr^* \left\{ \sup_{t \in (\alpha_m^*, 1 - \alpha_m^*)} \sqrt{N} \left| \psi(\mathbb{R}_{m,n}^*(t)) - \psi(\mathbb{R}_{m,n}(t)) \right| \leq q \right\} \\ & - \Pr \left\{ \sup_{t \in (\alpha_m^*, 1 - \alpha_m^*)} \sqrt{N} \left| \psi(\mathbb{R}_{m,n}(t)) - \psi(R(t)) \right| \leq q \right\} | \rightarrow 0, \end{aligned} \quad (27)$$

$$\begin{aligned} & | \Pr^\# \left\{ \sup_{t \in (\alpha_m^\#, 1 - \alpha_m^\#)} \sqrt{N} \left| \psi(\mathbb{R}_{m,n}^\#(t)) - \psi(\mathbb{R}_{m,n}(t)) \right| \leq q \right\} \\ & - \Pr \left\{ \sup_{t \in (\alpha_m^\#, 1 - \alpha_m^\#)} \sqrt{N} \left| \psi(\mathbb{R}_{m,n}(t)) - \psi(R(t)) \right| \leq q \right\} | \rightarrow 0, \end{aligned} \quad (28)$$

where $\Pr^\#$ stands for BB probability conditional on samples.

Remark: 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 2 Let $\phi(\cdot) : C[0,1] \rightarrow \mathbb{R}$ be a bounded linear functional, $N = m + n$ and assume that $\frac{m}{N} \rightarrow \lambda$, $\frac{n}{N} \rightarrow 1 - \lambda$. Then as $N \rightarrow \infty$, conditionally on the samples, we have for all q

$$\Pr^* \{ \sqrt{N}(\phi(\mathbb{R}_{m,n}^*) - \phi(\mathbb{R}_{m,n})) \leq q \} - \Pr \{ \sqrt{N}(\phi(\mathbb{R}_{m,n}) - \phi(R)) \leq q \} \rightarrow 0 \quad a.s. \quad (29)$$

$$\Pr^\# \{ \sqrt{N}(\phi(\mathbb{R}_{m,n}^\#) - \phi(\mathbb{R}_{m,n})) \leq q \} - \Pr \{ \sqrt{N}(\phi(\mathbb{R}_{m,n}) - \phi(R)) \leq q \} \rightarrow 0 \quad a.s. \quad (30)$$

Corollary 3 For partial AUC (denoted as $A(\alpha, \beta)$): Let $N = m + n$ and assume that $\frac{m}{N} \rightarrow \lambda$, $\frac{n}{N} \rightarrow 1 - \lambda$, then as $N \rightarrow \infty$, conditionally on the samples, for any $(\alpha, \beta) \subset (0, 1)$,

we have

$$\sqrt{N}(\hat{\mathbb{A}}_{m,n}(\alpha, \beta) - A(\alpha, \beta)) \rightarrow N(0, \sigma^2(\alpha, \beta)), \quad (31)$$

where $\sigma^2(\alpha, \beta) = \lambda^{-1} \int_{\alpha}^{\beta} \int_{\alpha}^{\beta} R'(t)R'(s)(s \wedge t - st)dtds + (1 - \lambda)^{-1} \int_{\alpha}^{\beta} \int_{\alpha}^{\beta} (R(t) \wedge R(s) - R(t)R(s))dtds$, $\hat{\mathbb{A}}_{m,n}(\alpha, \beta) = \int_{\alpha}^{\beta} \mathbb{R}_{m,n}(t)dt$, $A(\alpha, \beta) = \int_{\alpha}^{\beta} R(t)dt$.

Remark: It can be easily shown that Condition B ensures the integrals in the expression for $\sigma(\alpha, \beta)$ as $\alpha \rightarrow 0$ and $\beta \rightarrow 1$.

5 Proof of the Lemmas in the sense of *a.s.*

5.1 Proof of the Lemma 3.1

For any $y \in [\delta_n^*, 1 - \delta_n^*]$, let $k = \lceil ny \rceil$. By applying Theorem D, K and N of Appendix, we can get

$$\sup_{\delta_n^* \leq y \leq 1 - \delta_n^*} |(y - V_{k:n})\mathbb{Q}_n(V_{k:n})f'(F^{-1}(\tau))/f(F^{-1}(\tau))| =_{a.s.} O(l_n), \quad (32)$$

for any τ lies between y and $V_{k:n}$. By Condition A, we have

$$f(F^{-1}(y)) = f(F^{-1}(V_{k:n})) + (y - V_{k:n})f'(F^{-1}(\xi))/f(F^{-1}(\xi)), \quad (33)$$

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 by plugging (33) into $f(F^{-1}(y))$ and applying Theorem B, E and F, Lemma 3.1 holds. \square

5.2 Proof of the Lemma 3.2

5.2.1 Proof of the statement (12) of Lemma 3.2

By Theorem M, we have

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

5.2.2 Proof of the statement (13) of Lemma 3.2

Because $y \in [\delta_n^\#, 1 - \delta_n^\#] \subset [\epsilon_n, 1 - \epsilon_n]$, where $\epsilon_n = 0.236n^{-1} \log \log n$, 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 the similar argument as the proof of (32), we have $\sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} \left| (\tilde{\mathbb{U}}_{n-1}(y) - y) \frac{f'(F^{-1}(\xi))}{f(F^{-1}(\xi))} \mathbb{Q}_n(\tilde{\mathbb{U}}_{n-1}(y)) \right| =_{a.s.} O(l_n)$.
2. By Theorem B of Appendix, $\left| \sqrt{n} f(F^{-1}(y)) (\mathbb{F}_n^{-1}(y) - F^{-1}(y)) - n^{-1/2} K(y, n) \right| =_{a.s.} O(l_n)$.
3. $\sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} \left| \sqrt{n} f(F^{-1}(y)) [F^{-1}(\tilde{\mathbb{U}}_{n-1}(y)) - F^{-1}(y)] - n^{-1/2} K(y, n) \right|$ can be bounded by

$$\begin{aligned}
& \sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} \left| \sqrt{n} (\tilde{\mathbb{U}}_{n-1}(y) - y) - n^{-1/2} K(y, n) \right| \\
& \quad + \sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} \left| (\tilde{\mathbb{U}}_{n-1}(y) - y)^2 \frac{f'(F^{-1}(\xi)) f(F^{-1}(y))}{f^2(F^{-1}(\xi)) f(F^{-1}(\xi))} \right|, \quad (34)
\end{aligned}$$

which is $O(l_n)$ *a.s.*. The reasons for (34) are as follows:

(a) By Theorem A of Appendix, we have $\sup_{0 < y < 1} \sqrt{\frac{n}{n-1}} \left| \sqrt{n-1}(\tilde{U}_{n-1}(y) - y) - \frac{K(y, n-1)}{\sqrt{n-1}} \right| =_{a.s.} O\left(\frac{\sqrt{n} \log^2(n-1)}{n-1}\right)$. By Theorems G and H of Appendix, we have $\sup_{0 < y < 1} \left| \frac{\sqrt{n}}{n-1} K(y, n-1) - n^{-1/2} K(y, n) \right| =_{a.s.} O(l_n)$. Therefore, the first term of (34) can be majorized by $\sup_{0 < y < 1} \sqrt{\frac{n}{n-1}} \left| \sqrt{n-1}(\tilde{U}_{n-1}(y) - y) - \frac{1}{\sqrt{n-1}} K(y, n-1) \right| + \sup_{0 < y < 1} \left| \frac{\sqrt{n}}{n-1} K(y, n-1) - n^{-1/2} K(y, n) \right| =_{a.s.} O(l_n)$.

(b) By Theorem J of Appendix, we have $(\tilde{U}_{n-1}(y) - y)^2 \leq_{a.s.} 4(n-1)^{-1} y(1-y) \log \log(n-1)$. 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 (34) can be majorized by $\sup_{\delta_n^\# \leq y \leq 1 - \delta_n^\#} [4(n-1)^{-1} \log \log(n-1)] \frac{y(1-y)}{\xi(1-\xi)} \xi(1-\xi) \frac{f'(F^{-1}(\xi))}{f^2(F^{-1}(\xi))} \frac{f(F^{-1}(y))}{f(F^{-1}(\xi))} \leq_{a.s.} O(l_n)$.

4. By Theorems G and I of Appendix, we have $\sup_{0 < y < 1} \frac{1}{\sqrt{n-1}} \left| K(\tilde{U}_{n-1}(y), n) - K(y, n) \right| =_{a.s.} O(l_n)$.

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

$$[\mathbb{F}_n^{-1}(\tilde{U}_{n-1}(y)) - F^{-1}(\tilde{U}_{n-1}(y))] - [\mathbb{F}_n^{-1}(y) - F^{-1}(y)] + [F^{-1}(\tilde{U}_{n-1}(y)) - F^{-1}(y)] \quad (35)$$

By combining parts (1)- (4) above and plugging (35) into $\mathbb{F}_n^{-1}(\tilde{U}_{n-1}(y)) - \mathbb{F}_n^{-1}(y)$, we finish the proof of (13). \square

6 Appendix

Kiefer process: $\{K(x, y) : 0 \leq x \leq 1, 0 \leq y < \infty\}$ is determined by $K(x, y) = W(x, y) - xW(1, y)$, where $W(x, y)$ be a two-parameter Wiener process.

Kiefer process's covariance function: $E[K(x_1, y_1)K(x_2, y_2)] = (x_1 \wedge x_2 - x_1 x_2)(y_1 \wedge y_2)$, where \wedge stands for the minimum.

The following theorems were used:

Theorem A: (Strong approximation for arbitrary empirical process) [Corollary of strong approximation for uniform empirical process: Komlós, Major and Tusnády, 1975]: Let $X_1, X_2, \dots \sim$ i.i.d. F , where F is any arbitrary continuous distribution function and $\mathbb{G}_n(x)$ be the empirical process. Then there exists a Kiefer process $\{K(y, t); 0 \leq y \leq 1, t \geq 0\}$, such that

$$\sup_x |\mathbb{J}_n(x) - n^{-1/2}K(F(x), n)| =_{a.s.} O(n^{-1/2} \log^2 n).$$

Theorem B: (Strong approximation for arbitrary quantile process) [Csörgő and Révész, 1978, Theorem 6]: Let $X_1, X_2, \dots \sim$ i.i.d. F , satisfying Condition A. Then the quantile process $\mathbb{Q}_n(x)$ of X can be approximated by a Kiefer process $\{K(y, t); 0 \leq y \leq 1, t \geq 0\}$

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

Theorem C: (Strong approximations for empirical processes of bootstrap and the Bayesian Bootstrap) [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

$$\begin{aligned} \sup_x |\mathbb{J}_n^*(x) - n^{-1/2}K(\mathbb{F}_n(x), n)| &=_{a.s.} O(l_n), \\ \sup_x |\mathbb{J}_n^\#(x) - n^{-1/2}K(\mathbb{F}_n(x), n)| &=_{a.s.} O(l_n). \end{aligned}$$

Theorem D: [Csörgő and Révész, 1978, Theorem 3]: Let X_1, X_2, \dots be i.i.d. F satisfying

Condition A. Then

$$\limsup_{n \rightarrow \infty} \frac{\sqrt{n}}{\log \log n} \sup_{\delta_n \leq y \leq 1 - \delta_n} |f(F^{-1}(y))\mathbb{Q}_n(y) - \mathbb{W}_n(y)| \leq 40\gamma 10^\gamma \quad a.s.$$

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 $\{K(y, t); 0 \leq y \leq 1, 0 \leq t\}$ defined on the same probability space. Then

$$\sup_{1 \leq k \leq n} n^{-1/2} |K(U_{k:n}, n) - K(k/n, n)| =_{a.s.} O(l_n).$$

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

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{0 \leq t \leq 1 - h_n} \gamma_n |K(t + h_n, n) - K(t, n)| &=_{a.s.} 1, \\ \lim_{n \rightarrow \infty} \sup_{0 \leq t \leq 1 - h_n} \sup_{0 \leq s \leq h_n} \gamma_n |K(t + s, n) - K(t, n)| &=_{a.s.} 1. \end{aligned}$$

Theorem G: Let $K(s, n)$ be a Kiefer process,

$$\sup_{0 < s < 1} \frac{1}{\sqrt{n-1}} \left| K(s, n-1) - K(s, n) \right| =_{a.s.} O(l_n).$$

Remark: Because $B_n(s) = K(s, n) - K(s, n-1)$ is a series of independent Brownian bridges, this result follows immediately from the result of Lo (1987).

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

$$\limsup_{n \rightarrow \infty} \sup_{0 \leq y \leq 1} \frac{|K(y, n)|}{(2n \log \log n)^{\frac{1}{2}}} =_{a.s.} \frac{1}{2}. \quad (36)$$

Theorem I: [Special case of Lemma 6.2 (Lo, 1987) when $F = U(0, 1)$]: Let $U_1, \dots, U_m \sim$ 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, \dots, U_m ,

$$\sup_{0 < s < 1} \frac{1}{\sqrt{m}} | K(\mathbb{U}_m(s), m) - K(s, m) | =_{a.s.} O(l_m).$$

Theorem J: [Csörgő and Révész, 1978, Theorem D] : Let $\mathbb{H}_n(x)$ be the uniform empirical process, $\epsilon_n = 0.236n^{-1} \log \log n$, then

$$\limsup_{n \rightarrow \infty} \sup_{\epsilon_n \leq x \leq 1 - \epsilon_n} (x(1-x) \log \log n)^{-1/2} | \mathbb{H}_n(x) | =_{a.s.} 2.$$

Theorem K: [Csörgő and Révész, 1978, Theorem 2]: Let $\mathbb{W}_n(y)$ be the uniform quantile process, then

$$\limsup_{n \rightarrow \infty} \sup_{\delta_n \leq y \leq 1 - \delta_n} (y(1-y) \log \log n)^{-1/2} | \mathbb{W}_n(y) | =_{a.s.} 4.$$

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

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

Theorem M: [Slud, 1978]: Let $U_1, \dots, U_n \sim$ i.i.d. $U(0, 1)$ and define maximal uniform space M_n as $M_n = \max_{0 \leq k \leq n} U_{k+1:n} - U_{k:n}$. Then

$$nM_n / \log n \rightarrow_{a.s.} 1.$$

Theorem N: Important inequalities [Proof follows Csörgő and Révész, 1978, Theorem 3]:

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

$$\frac{y(1-y)}{\xi(1-\xi)} \leq 5 \quad (37)$$

$$\frac{\xi}{V_{k:n}} \leq 5, \quad \frac{1-\xi}{1-V_{k:n}} \leq 5 \quad (38)$$

$$\sup_{\delta_n^* \leq y \leq 1-\delta_n^*} f(F^{-1}(\xi))/f(F^{-1}(V_{k:n})) \leq 10^\gamma \quad (39)$$

Remark: When $y \in [\delta_n^\#, 1 - \delta_n^\#]$, ξ lies between y and $\tilde{U}_{n-1}(y)$. By Theorem J of Appendix, the inequalities (37), (38) and (39) hold for y , ξ and $\tilde{U}_{n-1}(y)$ by the same arguments, yielding

$$\sup_{\delta_n^\# \leq y \leq 1-\delta_n^\#} \frac{y(1-y)}{\xi(1-\xi)} \leq 5, \quad (40)$$

$$\sup_{\delta_n^\# \leq y \leq 1-\delta_n^\#} \frac{\xi}{\tilde{U}_{n-1}(y)} \leq 5, \quad \sup_{\delta_n^\# \leq y \leq 1-\delta_n^\#} \frac{1-\xi}{1-\tilde{U}_{n-1}(y)} \leq 5, \quad (41)$$

$$\sup_{\delta_n^\# \leq y \leq 1-\delta_n^\#} f(F^{-1}(\xi))/f(F^{-1}(\tilde{U}_{n-1}(y))) \leq 10^\gamma. \quad (42)$$

7 References

1. Csörgő, M. and Révész, P. (1978). Strong approximations of the quantile process. *The Annals of Statistics* **6**, 882–894.
2. Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic press, New York.
3. Green, D. M. and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. John Wiley & Sons, New York.
4. Gu, J. , Ghosal, S. and Roy A. (2006). Non-parametric estimation of ROC curve.

Institute of Statistics Mimeo Series 2592, NCSU.

5. 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.
6. Lo, A. Y. (1987). A large sample study of the Bayesian bootstrap. *The Annals of Statistics* **15**, 360–375.
7. Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine* **8**, 283–298.
8. Pepe, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford Statistical Science Series, Oxford University Press.
9. Slud, E. (1978). Entropy and maximal spacings for random partitions. *Z. Wahrsch. Verw. Gebiete* **41**, 341–352.
10. Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.