A Linear Regression Framework for the Receiver Operating Characteristic (ROC) Curve Analysis

Zheng Zhang* and Ying Huang

1Department of Biostatistics and Center for Statistical Sciences, Brown University, Providence, RI 02912, USA
2Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

Abstract

The receiver operating characteristic (ROC) curve has been a popular statistical tool for characterizing the discriminating power of a classifier, such as a biomarker or an imaging modality for disease screening or diagnosis. It has been recognized that the accuracy of a given procedure may depend on some underlying factors, such as subject’s demographic characteristics or disease risk factors, among others. Non-parametric- or parametric-based methods tend to be either inefficient or cumbersome when evaluating effect of multiple covariates is the main focus. Here we propose a semi-parametric linear regression framework to model covariate effect. It allows the estimation of sensitivity at given specificity to vary according to the covariates and provides a way to model the area under the ROC curve indirectly. Estimation procedure and asymptotic theory are presented. Extensive simulation studies have been conducted to investigate the validity of the proposed method. We illustrate the new method on a diagnostic test dataset.

Keywords: AUC; Covariate effect; Linear regression; ROC curve; Sensitivity

Introduction

New advances in medical technology have produced an array of potentially powerful tools to screen for and diagnose various medical conditions. Effective screening and accurate diagnosis can ensure optimal treatment and improve prognosis. Before a screening or diagnostic test can be applied in a clinical setting, rigorous statistical assessment of its performance in discriminating the diseased state from the non-diseased state is required. For tests measured on the continuous scale, the receiver operating characteristic (ROC) curve is a common statistical tool for describing the performance of such tests [1]. Let D be disease status (1 for disease and 0 otherwise) and Y the test result with positivity defined whenever Y ≥ c. Define the true positive fraction (TPF(c)) and the false positive fraction (FPF(c)) as

\[\text{TPF}(c) = \frac{\sum n_D}{n_D}, \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where TPF and FPF are also called test’s sensitivity and specificity. The ROC curve is a plot of TPF(c) versus FPF(c) when the threshold c ranges from -∞ to ∞. Alternatively, the ROC curve can be written as a function of t, by defining

\[S_D(t) = \frac{\sum n_D}{n_D}, \quad S_Y(t) = \frac{\sum n_Y}{n_Y},\]

where SD and SY indicate the diseased and non-diseased population, then

\[\text{ROC}(t) = \frac{S_D(t)}{S_Y(t)}, \quad t \in (0,1) \quad (1)\]

Different approaches have been proposed to estimate an ROC curve. Nonparametrically, we can obtain the empirical ROC curve based on

\[\text{TPF}(c) = \frac{\sum n_D}{n_D} \quad \text{and} \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where n_D and n_Y are number of diseased and non-diseased observations respectively. Alternatively, we can estimate the ROC curve parametrically by assuming a distributional form for SD and SY and then calculating the induced ROC curve with equation (1). The derived ROC curve, however, is not invariant to transformation of the test results. Pepe [2] notes that ROC curve describes only the relationship between the distributions of SD and SY, not the distributions themselves. Semiparametric estimators that directly model the ROC curve as a parametric function without specifying the underlying distributions of SD and SY provides a desirable alternative.

Abstract

The receiver operating characteristic (ROC) curve has been a popular statistical tool for characterizing the discriminating power of a classifier, such as a biomarker or an imaging modality for disease screening or diagnosis. It has been recognized that the accuracy of a given procedure may depend on some underlying factors, such as subject’s demographic characteristics or disease risk factors, among others. Non-parametric- or parametric-based methods tend to be either inefficient or cumbersome when evaluating effect of multiple covariates is the main focus. Here we propose a semi-parametric linear regression framework to model covariate effect. It allows the estimation of sensitivity at given specificity to vary according to the covariates and provides a way to model the area under the ROC curve indirectly. Estimation procedure and asymptotic theory are presented. Extensive simulation studies have been conducted to investigate the validity of the proposed method. We illustrate the new method on a diagnostic test dataset.

Keywords: AUC; Covariate effect; Linear regression; ROC curve; Sensitivity

Introduction

New advances in medical technology have produced an array of potentially powerful tools to screen for and diagnose various medical conditions. Effective screening and accurate diagnosis can ensure optimal treatment and improve prognosis. Before a screening or diagnostic test can be applied in a clinical setting, rigorous statistical assessment of its performance in discriminating the diseased state from the non-diseased state is required. For tests measured on the continuous scale, the receiver operating characteristic (ROC) curve is a common statistical tool for describing the performance of such tests [1]. Let D be disease status (1 for disease and 0 otherwise) and Y the test result with positivity defined whenever Y ≥ c. Define the true positive fraction (TPF(c)) and the false positive fraction (FPF(c)) as

\[\text{TPF}(c) = \frac{\sum n_D}{n_D}, \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where TPF and FPF are also called test’s sensitivity and specificity. The ROC curve is a plot of TPF(c) versus FPF(c) when the threshold c ranges from -∞ to ∞. Alternatively, the ROC curve can be written as a function of t, by defining

\[S_D(t) = \frac{\sum n_D}{n_D}, \quad S_Y(t) = \frac{\sum n_Y}{n_Y},\]

where SD and SY indicate the diseased and non-diseased population, then

\[\text{ROC}(t) = S_D(S_Y(t)), \quad t \in (0,1) \quad (1)\]

Different approaches have been proposed to estimate an ROC curve. Nonparametrically, we can obtain the empirical ROC curve based on

\[\text{TPF}(c) = \frac{\sum n_D}{n_D} \quad \text{and} \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where n_D and n_Y are number of diseased and non-diseased observations respectively. Alternatively, we can estimate the ROC curve parametrically by assuming a distributional form for SD and SY and then calculating the induced ROC curve with equation (1). The derived ROC curve, however, is not invariant to transformation of the test results. Pepe [2] notes that ROC curve describes only the relationship between the distributions of SD and SY, not the distributions themselves. Semiparametric estimators that directly model the ROC curve as a parametric function without specifying the underlying distributions of SD and SY provides a desirable alternative.

Abstract

The receiver operating characteristic (ROC) curve has been a popular statistical tool for characterizing the discriminating power of a classifier, such as a biomarker or an imaging modality for disease screening or diagnosis. It has been recognized that the accuracy of a given procedure may depend on some underlying factors, such as subject’s demographic characteristics or disease risk factors, among others. Non-parametric- or parametric-based methods tend to be either inefficient or cumbersome when evaluating effect of multiple covariates is the main focus. Here we propose a semi-parametric linear regression framework to model covariate effect. It allows the estimation of sensitivity at given specificity to vary according to the covariates and provides a way to model the area under the ROC curve indirectly. Estimation procedure and asymptotic theory are presented. Extensive simulation studies have been conducted to investigate the validity of the proposed method. We illustrate the new method on a diagnostic test dataset.

Keywords: AUC; Covariate effect; Linear regression; ROC curve; Sensitivity

Introduction

New advances in medical technology have produced an array of potentially powerful tools to screen for and diagnose various medical conditions. Effective screening and accurate diagnosis can ensure optimal treatment and improve prognosis. Before a screening or diagnostic test can be applied in a clinical setting, rigorous statistical assessment of its performance in discriminating the diseased state from the non-diseased state is required. For tests measured on the continuous scale, the receiver operating characteristic (ROC) curve is a common statistical tool for describing the performance of such tests [1]. Let D be disease status (1 for disease and 0 otherwise) and Y the test result with positivity defined whenever Y ≥ c. Define the true positive fraction (TPF(c)) and the false positive fraction (FPF(c)) as

\[\text{TPF}(c) = \frac{\sum n_D}{n_D}, \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where TPF and FPF are also called test’s sensitivity and specificity. The ROC curve is a plot of TPF(c) versus FPF(c) when the threshold c ranges from -∞ to ∞. Alternatively, the ROC curve can be written as a function of t, by defining

\[S_D(t) = \frac{\sum n_D}{n_D}, \quad S_Y(t) = \frac{\sum n_Y}{n_Y},\]

where SD and SY indicate the diseased and non-diseased population, then

\[\text{ROC}(t) = S_D(S_Y(t)), \quad t \in (0,1) \quad (1)\]

Different approaches have been proposed to estimate an ROC curve. Nonparametrically, we can obtain the empirical ROC curve based on

\[\text{TPF}(c) = \frac{\sum n_D}{n_D} \quad \text{and} \quad \text{FPF}(c) = \frac{\sum n_Y}{n_Y},\]

where n_D and n_Y are number of diseased and non-diseased observations respectively. Alternatively, we can estimate the ROC curve parametrically by assuming a distributional form for SD and SY and then calculating the induced ROC curve with equation (1). The derived ROC curve, however, is not invariant to transformation of the test results. Pepe [2] notes that ROC curve describes only the relationship between the distributions of SD and SY, not the distributions themselves. Semiparametric estimators that directly model the ROC curve as a parametric function without specifying the underlying distributions of SD and SY provides a desirable alternative.
\( \Phi^{-1}(FPF) \) have a linear relationship. Hence we propose to fit a linear line through the points of \((\Phi^{-1}(FPF), \Phi^{-1}(TPF))\) and obtain estimates of \(a_2\) and \(a_2\) by the least squares technique, and consequently, estimate the ROC curve itself. This method is intuitive, conceptually easy to understand, and very easy to implement. More importantly, it can be readily extended to allow for covariate \((Z)\) effects on the ROC curve by writing

\[
\Phi^{-1}(ROC_{c}(t)) = a_0 + a_1 \Phi^{-1}(t) + \alpha Z. \tag{3}
\]

Our linear regression approach provided a simple and widely accessible algorithm for fitting such models.

Hsieh and Turnbull [11] described a weighted least squares approach to estimate the binormal ROC curve. For continuous data, their approach groups the data into a pre-determined number (independent of the number of observations) of categories, and the largest they chose is 12. Our framework is similar in spirit but allows significant improvement in efficiency by eliminating the need of grouping, as will be shown later in the next section. Moreover, our framework offers the flexibility to model additional covariate effects of grouping, as will be shown later in the next section. Moreover, our approach to estimate the binormal ROC curve. For continuous data, their approach groups the data into a pre-determined number \((\alpha, \beta)\) and density \(f_D\) functions

\[
\alpha, \beta
\]

understand, and very easy to implement. More importantly, it can be

Theorem 1

\[
\sqrt{n_\alpha - \alpha_0} \to \mathcal{N}(0, \Sigma),
\]

meaning that the process converges to a mean zero Gaussian process with variance-covariance function

\[
\Sigma(s, t) = 1/\{\phi(a_0 + a_1 \Phi^{-1}(s))\phi(a_0 + a_1 \Phi^{-1}(t))\{2s\Sigma(s, s) + \Sigma(s, t)\},
\]

where \(s, t\) is \(ROC(s)/ROC(t)\), \(ROC\) is the vector of the estimators \(\hat{\alpha}_0, \hat{\alpha}_1\) and the vector

\[
M^* = \begin{pmatrix}
1 & \cdots & 1 \\
\Phi^{-1}(t_1) & \cdots & \Phi^{-1}(t_n)
\end{pmatrix}
\]

and the vector

\[
\Phi^{-1}(ROC(t_j)) = (\Phi^{-1}(ROC(t_1)), \ldots, \Phi^{-1}(ROC(t_n)));
\]

5. Calculate the ordinary least squares(OLS) estimator of \(a\) as

\[
\hat{a} = \begin{pmatrix}
\hat{\alpha}_0 \\
\hat{\alpha}_1
\end{pmatrix} = (MM^*)^{-1}M^*\Phi^{-1}(ROC(t));
\]

6. Then the estimated ROC curve and its associated AUC are

\[
\hat{ROC}(t) = \Phi(\hat{\alpha}_0 + \hat{\alpha}_1 \Phi^{-1}(t)) \quad \text{and}
\]

\[
\hat{AUC} = \Phi(\hat{\alpha}_0 / \sqrt{1 + \hat{\alpha}_1^2})
\]

**Asymptotic theory**

We develop the asymptotic distribution of the OLS estimator

\[
\hat{a} = (\hat{\alpha}_0, \hat{\alpha}_1)
\]

under the following assumptions: \(\{Y_{n_\alpha}\}\) and \(\{Y_{n_\beta}\}\) are i.i.d. random variables with a survival functions \(S_{\alpha}\) and \(S_{\beta}\) and density functions \(f_{\alpha}\) and \(f_{\beta}\) respectively; \(n_\alpha / n_\beta \to \lambda, 0 < \lambda < \infty, n_\alpha \to \infty\); the slope of the ROC curve, \(f_{\alpha}(S_{\beta}(t))/f_{\beta}(S_{\alpha}(t))\), is bounded on the subinterval \([a,b]\) of \((0,1), 0 < a < b < 1\).

\[
\sqrt{n_{\alpha} - \alpha_0} \to \mathcal{N}(0, \Sigma_A)
\]

\[
A = \left[ \int_a^b 1/\{\Phi^{-1}(t)\} (\Phi^{-1}(t)) d(t) \right]^{-1},
\]

and

\[
\Sigma_A = \begin{pmatrix}
\sigma_0^2 & \sigma_1^2 \\
\sigma_1^2 & \sigma_2^2
\end{pmatrix},
\]

where

\[
\sigma_0^2 = \lambda_0 \int_{a}^{b} J_1(s) J_1(t) [ROC(s) \wedge ROC(t) - ROC(s)ROC(t)] ds dt
\]

\[
+ \lambda_1 \int_{a}^{b} J_2(s) J_2(t) (s \wedge t - st) ds dt,
\]

\[
\sigma_1^2 = \lambda_0 \int_{a}^{b} K_1(s) K_1(t) [ROC(s) \wedge ROC(t) - ROC(s)ROC(t)] ds dt
\]

\[
+ \lambda_2 \int_{a}^{b} K_2(s) K_2(t) (s \wedge t - st) ds dt,
\]

\[
\sigma_2^2 = \lambda_0 \int_{a}^{b} K_1(s) J_1(t) [ROC(s) \wedge ROC(t) - ROC(s)ROC(t)] ds dt
\]

\[
+ \lambda_2 \int_{a}^{b} J_1(s) K_2(t) (s \wedge t - st) ds dt,
\]

\[
\sigma_3^2 = \lambda_0 \int_{a}^{b} K_1(s) K_1(t) [ROC(s) \wedge ROC(t) - ROC(s)ROC(t)] ds dt
\]

\[
+ \lambda_2 \int_{a}^{b} J_1(s) J_1(t) (s \wedge t - st) ds dt,
\]

with \(J_1(s) = (\Phi(\Phi^{-1}(ROC(s)))^{-1})^{-1}, J_2(s) = (\Phi(\Phi^{-1}(s)))^{-1}, K_1(s) = \Phi^{-1}(s) J_1(s)\) and \(K_2(s) = \Phi^{-1}(s) J_2(s)\).

**Proof:** The proof for Theorem 1 can be found in the Appendix II.

**Asymptotic efficiency relative to HT method**

Here we compare the asymptotic efficiency between the OLS estimator and the estimator derived using Hsieh & Turnbull’s method (HT) for estimating \(\alpha_0, \alpha_1, \alpha_{\lambda}\) and \(ROC(t)\) at \(t = 0.2, 0.4, 0.7, \) when \(\lambda\) (the
limit of the ratio of $n_2$ and $n_o$ varies from 0.5 to 2. We choose $(\alpha_0, \alpha_t) = (1.2, 0.45)$ so that the area under the ROC curve (AUC) is 0.863 and the ROC curve takes values of 0.794, 0.861 and 0.924 at $t = 0.2, 0.4, 0.7$, respectively. Note that for $h(t) = \left( \Phi^{-1}(t) \right)^{-1}$, let $\Sigma$ denote the asymptotic variance of $\sqrt{n_2}(\hat{\alpha}_t - \alpha_t)$ from an ROC modeling method (OLS or HT), and the asymptotic variance expression for the corresponding $\widehat{ROC}(t)$ are:

$$\text{var}(\widehat{ROC}(t)) \doteq 1/n_2 \{ (\Phi(\alpha_0 + \alpha_t \Phi^{-1}(t))^2 h(t)^2 \Sigma h(t) \}. $$

Table 1 shows the asymptotic efficiency of HT relative to OLS. For the OLS estimator, the boundary points $(a, b)$ chosen are (0.0001, 0.9999). For the HT estimator, the number of categories chosen is eight. Note that asymptotically, the OLS estimator can lead to substantial efficiency gain compared to the HT estimator. The biggest improvement is seen when estimating $\alpha_t$ with efficiency gain above 50%. The efficiency gain in estimating points on the ROC curve varies from 3% to 20% depending on the point of interest (Table 1).

### Regression Model with Discrete Covariates

Suppose there are $K$ categories that potentially could overlap with each other. For $k = 1, \ldots, K$, let $Z(k)$ be a vector of length $K-1$, with value 1 for the $k-1$th element and zero else where (so $Z(1)$ is a vector of zeros), and let $n_k$ be the number of non-diseased observations in category $k$. Suppose the ROC curve within category $k$ is characterized by

$$\Phi^{-1}(\text{ROC}(t)) = \beta_t + \gamma \Phi^{-1}(t) + \theta^T Z(k)$$

where $\theta = (\beta_2, \ldots, \beta_k)^T$.

That is, for subset 1 (the reference subset)

$$\Phi^{-1}(\text{ROC}(t)) = \beta_1 + \gamma \Phi^{-1}(t)$$

and for subset $k = 2, \ldots, K$,

$$\Phi^{-1}(\text{ROC}(t)) = \beta_1 + \gamma \Phi^{-1}(t) + \beta_k.$$  

Hence $\beta_t$ is the difference in the intercept parameter of the ROC curves between subset $k$ and the reference subset. The parameter of interest here is $\theta^* = (\beta_2, \ldots, \beta_K, \gamma)^T$. The underlying assumptions for equation (5) is that there exists an unknown monotone increasing function $h_1$ such that $h_1(Y_{1n1}) \sim N(0,1)$ and $h_1(Y_{1n0}) \sim N(\beta_1/\gamma, 1/\gamma^2)$. Similarly, for subsequent subset $k, k = 2, 3, \ldots, K$, (6) implies there exists an unknown monotone increasing function $h_k$ such that $h_k(Y_{knk}) \sim N(0,1)$ and $h_k(Y_{kn0}) \sim N(\beta_1/\gamma, 1/\gamma^2)$. Notice that $h_{n0}$ are not required to be the same for different $k$.

Let $\widehat{\text{ROC}}(t)$ be the empirical estimate of $\text{ROC}(t)$ based on data from category $k$. Like in the case of equation (4), $\sqrt{n_k}(\Phi^{-1}(\text{ROC}(t))) - \Phi^{-1}(\text{ROC}(t)))$ converge to a Gaussian process, therefore $\Phi^{-1}(\text{ROC}(t))$ can be approximated by $\Phi^{-1}(\text{ROC}(t))$, which motivates the following estimation procedure:

1. Calculate pairs of $(t, \text{ROC}(t))$ for each subset separately;
2. Let design matrix $D$ be
   $$D = \begin{pmatrix} 1 & O & O & \cdots & O \\ M & M_1 & O & \cdots & O \\ M & O & M_1 & \cdots & O \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ M & O & O & \cdots & M_1 \end{pmatrix} \top$$

   where $M = \begin{pmatrix} 1 & \cdots & 1 \\ \Phi^{-1}(t_1) & \cdots & \Phi^{-1}(t_p) \end{pmatrix}$, $M_1$ is a vector of length $p$ with constant value of one, and $O$ is a zero value matrix.
3. Let $\tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_K)^T$ and $\tilde{Y}_k = (\Phi^{-1}(\text{ROC}_k(t_{p1})), \ldots, \Phi^{-1}(\text{ROC}_k(t_{pl}))$;
4. Our linear model is: for the reference subset, $\Phi^{-1}(\text{ROC}_k(t_p)) = \beta_t + \gamma \Phi^{-1}(t_p) + \epsilon_1$,
   where $\tilde{\beta}_2, \ldots, \tilde{\beta}_K$ is normally distributed with mean 0 and asymmetric covariance matrix $\Sigma_{00}$ and for subset $k, k = 2, 3, \ldots, K$
   $$\Phi^{-1}(\text{ROC}_k(t_p)) = \beta_t + \gamma \Phi^{-1}(t_p) + \beta_k + \epsilon_k,$$
   where $\tilde{\beta}_2, \ldots, \tilde{\beta}_K$ is normally distributed with mean 0 and asymmetric covariance matrix $\Sigma_{k}$;
5. Our OLS estimator for $\theta^*$ is
   $$\hat{\theta} = (DD \top)^{-1} D \tilde{Y}$$

   The above method assumes covariate effects can be explained adequately by the difference in the intercept parameter $(\alpha_t)$. In addition we allow the slope of a binormal ROC curve to be different across covariate categories by assuming

$$\Phi^{-1}(\text{ROC}(t)) = \beta_t + \gamma \Phi^{-1}(t) + (1 \Phi^{-1}(t)) \theta^T Z(k),$$

where $\theta = (\beta_2, \ldots, \beta_K, \gamma, 1/\gamma)^T$ then the parameter of interest is $\theta^* = (\beta_t, \gamma, \beta_2, \ldots, \beta_K, 1/\gamma)^T$. The estimating procedure is similar to the case where only intercept parameters vary, but with $M_1$ replaced by $M$ in (7). Inference for the significance of $\theta^*$ in both settings can be achieved by estimating the variance of $\theta^*$ with the bootstrap resampling method.

### Simulation

#### Estimation of the ROC curve

The estimation procedure specified in the previous section starts with the choice of the false positive set $T$. Although in theory, any

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$a_1$</th>
<th>$a_t$</th>
<th>$R(0.2)$</th>
<th>$R(0.4)$</th>
<th>$R(0.7)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.10</td>
<td>1.72</td>
<td>1.09</td>
<td>1.06</td>
<td>1.20</td>
</tr>
<tr>
<td>1</td>
<td>1.10</td>
<td>1.64</td>
<td>1.05</td>
<td>1.06</td>
<td>1.20</td>
</tr>
<tr>
<td>2</td>
<td>1.10</td>
<td>1.59</td>
<td>1.03</td>
<td>1.06</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Table 1: Asymptotic efficiency of OLS relative to HT (with 8 categories).
chosen set of $T$ would yield estimators with the same asymptotic property, their small sample properties need to be investigated. An obvious starting point is to choose the collection of the observed false positive fractions that fall into the interval $[a, b]$, we call this observed FP (OFP) method. In the case where there are no ties in the test results of non-diseased subjects, this is equivalent to the method selecting the subset of $\{1/n_{\alpha}, 2/n_{\alpha}, \ldots, (n_{\alpha} - 1)/n_{\alpha}\}$ within $[a, b]$, which we call the equal fraction (EF) method. Another possible choice is to divide interval $[a, b]$ into $n_{\alpha} - 1$ equally spaced sub-intervals and choose the midpoints of those subintervals to be the set $T$ (midpoint (MP) method). For the last two methods, the number of points in $T$ always equal to $n_{\alpha} - 1$ regardless of the length of $[a, b]$.

We compare the performance of the OLS estimators between those three different methods of selecting $T$ (OFP, EF and MP). We simulate dataset both with and without ties and vary the values of $[a, b]$ to estimate either a full curve or a partial ROC curve.

First, we generate $Y_n \sim N(0,1)$ and $Y_0 \sim N(a_{\alpha}/\alpha, 1/\alpha^2)$. The resulting ROC curve follows a binormal model: $\text{ROC}(t) = \Phi(\alpha_{\alpha} + \alpha \Phi^{-1}(t))$. The parameters $(\alpha_{\alpha}, \alpha)$ are chosen to be $(1.2, 0.45)$, corresponding to an area under the curve (AUC) value of 0.863 and a partial AUC (pAUC) value of 0.142 for $t \in (0, 0.2)$. With sample size of $n_{\alpha} = n_{\alpha} = 100$, we compare the biases and sampling standard errors in estimating AUC, ROC(0.2), ROC(0.4) and ROC(0.7) in the full curve estimation and pAUC(0.2) and ROC(0.1) in the partial curve estimation. To simulate data with ties, we chose 20% tied values within each population.

From Table 2, when the test results have no ties, midpoint (MP) method has the best performance. When data has ties, observed false positive fraction (OFP) method has the smallest bias when estimating the entire curve, but MP has the best performance for estimating the partial curve. We recommend using either method MP or OFP in practice and we choose MP method for the subsequent simulations (Table 2).

Next we compare performance of the OLS estimator with two existing semiparametric ROC modeling approach (ROC-GLM and PV), which have been shown to have good performance among others. Table 3 summarizes relative biases and standard errors for the three estimators for estimating either a full or a partial ROC curve. We observe that when the full ROC curve is of interest, the OLS and the GLM estimators have comparable performances while the PV estimator has somewhat larger biases. When estimating a partial ROC curve, however, the OLS estimator can have substantially smaller bias compared to other estimators (Table 3).

In Table 4, we demonstrate performance of the OLS estimator relative to the nonparametric ROC estimator and the parametric ROC estimator assuming normal case and control distributions after Box-Cox transformation. We generate data as $Y_n \sim LOGN(0,1)$ and $Y_0 \sim LOGN(\alpha_{\alpha}/\alpha,\alpha^{-1/3})$ so the resulting ROC curve is the same as above. We further modify the data as following: (i) inflating all data points by two fold (scaled log normal distribution); (ii) increasing all data points by 2 units (shifted log normal distribution); (iii) $\exp(Y^{(x)})$ (non Box-Cox). For the parametric method, Box-Cox transformation is applied to $Y_n$ and $Y_0$ separately before the fitting of a normal distribution. For log normal data or scaled log normal data, the OLS estimator and the parametric estimator combined with Box-Cox transformation have similar performances. For shifted log normal and non Box-Cox data, however, the parametric estimator performs poorly with a large bias even after a Box-Cox transformation. Both the OLS estimator and the nonparametric estimator are unbiased in these scenarios with the former substantially more efficient (Table 4).

Lastly, we perform simulation studies to investigate the use of large sample inference for ROC curve and AUC based on the OLS estimator. Table 5 shows the mean estimated asymptotic error and the coverage of 95% confidence intervals. The variance estimate from the asymptotic theory reflects the actual sampling variance and the coverage of 95% confidence intervals is excellent (Table 5).

Application to DPOAE Data Set

The DPOAE data set was first published by Stover et al. [13].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>OFP</th>
<th>EF</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of ties: 0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(a, b) = (0.0001, 0.9999)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(0.2)$</td>
<td>0.794</td>
<td>0.791(0.039)</td>
<td>0.792(0.039)</td>
<td>0.795(0.038)</td>
</tr>
<tr>
<td>$(0.4)$</td>
<td>0.861</td>
<td>0.860(0.031)</td>
<td>0.860(0.031)</td>
<td>0.862(0.031)</td>
</tr>
<tr>
<td>$(0.7)$</td>
<td>0.924</td>
<td>0.923(0.024)</td>
<td>0.923(0.024)</td>
<td>0.924(0.024)</td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.861(0.027)</td>
<td>0.861(0.027)</td>
<td>0.863(0.027)</td>
</tr>
<tr>
<td>$(a, b) = (0.0001, 0.2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(0.1)$</td>
<td>0.733</td>
<td>0.727(0.049)</td>
<td>0.729(0.048)</td>
<td>0.733(0.048)</td>
</tr>
<tr>
<td>pAUC(0.2)</td>
<td>0.142</td>
<td>0.141(0.010)</td>
<td>0.141(0.010)</td>
<td>0.142(0.010)</td>
</tr>
<tr>
<td>Percent of ties: 20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(a, b) = (0.0001, 0.9999)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(0.2)$</td>
<td>0.794</td>
<td>0.794(0.043)</td>
<td>0.805(0.041)</td>
<td>0.798(0.042)</td>
</tr>
<tr>
<td>$(0.4)$</td>
<td>0.861</td>
<td>0.862(0.034)</td>
<td>0.879(0.031)</td>
<td>0.864(0.034)</td>
</tr>
<tr>
<td>$(0.7)$</td>
<td>0.924</td>
<td>0.924(0.026)</td>
<td>0.942(0.022)</td>
<td>0.926(0.025)</td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.863(0.030)</td>
<td>0.875(0.027)</td>
<td>0.865(0.030)</td>
</tr>
<tr>
<td>$(a, b) = (0.0001, 0.2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(0.1)$</td>
<td>0.733</td>
<td>0.728(0.053)</td>
<td>0.739(0.051)</td>
<td>0.734(0.053)</td>
</tr>
<tr>
<td>pAUC(0.2)</td>
<td>0.142</td>
<td>0.141(0.011)</td>
<td>0.143(0.010)</td>
<td>0.143(0.011)</td>
</tr>
</tbody>
</table>

Table 2: Inference of the ROC curve by the semiparametric least squares based method (OLS) with different choices of the false positive sets (OFP: observed false positive fraction, EF: equal fraction or MP: midpoint method). Cases and controls are drawn from normal distributions. $(\alpha_{\alpha}, \alpha) = (1.2, 0.45)$. $n_{\alpha} = 100(100, 100)$. Sample means and sampling standard errors from 1000 simulations are shown.
Table 3: Inference of the ROC curve by the semiparametric least squares based method (OLS), the ROC-GLM method (GLM) and the placement value method (PV). Cases and controls are drawn from normal distributions. \((\alpha_0, \alpha_1) = (1.2, 0.45), (\alpha_0, \alpha_1) = (100, 100)\). Relative bias and sampling standard error from 500 simulations are shown. Relative bias = bias/true value \times 100%.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>OLS</th>
<th>GLM</th>
<th>PV</th>
</tr>
</thead>
<tbody>
<tr>
<td>((a,b) = (0.0001, 0.9999))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a_0)</td>
<td>1.2</td>
<td>1.09(0.150)</td>
<td>1.46(0.152)</td>
<td>2.40(0.161)</td>
</tr>
<tr>
<td>(a_1)</td>
<td>0.45</td>
<td>2.42(0.085)</td>
<td>1.38(0.080)</td>
<td>9.44(0.101)</td>
</tr>
<tr>
<td>R(0.2)</td>
<td>0.794</td>
<td>-0.14(0.040)</td>
<td>0.16(0.038)</td>
<td>-0.53(0.040)</td>
</tr>
<tr>
<td>R(0.4)</td>
<td>0.861</td>
<td>-0.02(0.031)</td>
<td>0.12(0.031)</td>
<td>0.15(0.032)</td>
</tr>
<tr>
<td>R(0.7)</td>
<td>0.924</td>
<td>-0.02(0.023)</td>
<td>-0.002(0.023)</td>
<td>0.38(0.024)</td>
</tr>
<tr>
<td>((a,b) = (0.0001,0.2))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a_0)</td>
<td>1.2</td>
<td>0.18(0.246)</td>
<td>-1.79(0.235)</td>
<td>7.45(0.266)</td>
</tr>
<tr>
<td>(a_1)</td>
<td>0.45</td>
<td>-0.62(0.141)</td>
<td>-6.82(0.130)</td>
<td>14.04(0.159)</td>
</tr>
<tr>
<td>R(0.05)</td>
<td>0.677</td>
<td>0.11(0.051)</td>
<td>1.27(0.049)</td>
<td>-1.00(0.051)</td>
</tr>
<tr>
<td>R(0.1)</td>
<td>0.733</td>
<td>-0.02(0.046)</td>
<td>0.51(0.046)</td>
<td>0.11(0.045)</td>
</tr>
<tr>
<td>R(0.15)</td>
<td>0.768</td>
<td>-0.13(0.046)</td>
<td>0.08(0.045)</td>
<td>0.61(0.045)</td>
</tr>
</tbody>
</table>

Table 4: Inference of the ROC curve by the semiparametric least squares based method (OLS), the nonparametric method (EMP) and the parametric method (RES). Cases and controls are drawn from log-normal distributions and modified as noted. \((\alpha_0, \alpha_1) = (1.2, 0.45), (\alpha_0, \alpha_1) = (100, 100)\). Sample means and sampling standard errors from 1000 simulations are shown.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>Distribution</th>
<th>OLS</th>
<th>EMP</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R(0.2))</td>
<td>0.794</td>
<td>lognormal</td>
<td>0.796(0.039)</td>
<td>0.795(0.044)</td>
<td>0.795(0.036)</td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>0.861</td>
<td>0.863(0.031)</td>
<td>0.864(0.036)</td>
<td>0.862(0.030)</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>0.924</td>
<td>0.926(0.023)</td>
<td>0.926(0.027)</td>
<td>0.863(0.026)</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.864(0.027)</td>
<td>0.864(0.027)</td>
<td>0.863(0.026)</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>0.794</td>
<td>lognormal-scaled</td>
<td>0.795(0.039)</td>
<td>0.796(0.043)</td>
<td>0.794(0.036)</td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>0.861</td>
<td>0.862(0.032)</td>
<td>0.863(0.037)</td>
<td>0.861(0.030)</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>0.924</td>
<td>0.925(0.024)</td>
<td>0.925(0.028)</td>
<td>0.923(0.023)</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.863(0.028)</td>
<td>0.863(0.028)</td>
<td>0.862(0.026)</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>0.794</td>
<td>lognormal-shift</td>
<td>0.794(0.039)</td>
<td>0.795(0.043)</td>
<td>0.818(0.043)</td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>0.861</td>
<td>0.861(0.031)</td>
<td>0.861(0.036)</td>
<td>0.900(0.031)</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>0.924</td>
<td>0.923(0.023)</td>
<td>0.923(0.027)</td>
<td>0.961(0.017)</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.864(0.027)</td>
<td>0.864(0.027)</td>
<td>0.888(0.027)</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>0.794</td>
<td>non Box-Cox</td>
<td>0.796(0.038)</td>
<td>0.796(0.044)</td>
<td>0.764(0.051)</td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>0.861</td>
<td>0.863(0.031)</td>
<td>0.864(0.036)</td>
<td>0.887(0.037)</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>0.924</td>
<td>0.926(0.023)</td>
<td>0.926(0.027)</td>
<td>0.968(0.017)</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.863</td>
<td>0.864(0.027)</td>
<td>0.864(0.027)</td>
<td>0.862(0.029)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Result of 1000 simulations to evaluate the application of inference based on asymptotic theory to finite sample studies. \((\alpha_0, \alpha_1) = (1.2, 0.45)\). Relative bias = bias/true value \times 100%; SSE: sampling standard error; ASE: mean estimated standard error using asymptotic theory; CP: coverage of 95% Wald-confidence intervals using estimated asymptotic error.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>Distribution</th>
<th>OLS</th>
<th>EMP</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R(0.2))</td>
<td>0.794</td>
<td>lognormal</td>
<td>0.163</td>
<td>0.151</td>
<td>94.8%</td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>2.2%</td>
<td>0.085</td>
<td>0.083</td>
<td>95.4%</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>0.4%</td>
<td>0.039</td>
<td>0.039</td>
<td>94.6%</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.3%</td>
<td>0.032</td>
<td>0.031</td>
<td>93.4%</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>0.1%</td>
<td>0.025</td>
<td>0.023</td>
<td>90.8%</td>
<td></td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>0.9%</td>
<td>0.159</td>
<td>0.157</td>
<td>95.0%</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>2.9%</td>
<td>0.088</td>
<td>0.088</td>
<td>94.6%</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.4%</td>
<td>0.041</td>
<td>0.042</td>
<td>94.6%</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>-0.3%</td>
<td>0.033</td>
<td>0.033</td>
<td>94.6%</td>
<td></td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>-0.1%</td>
<td>0.025</td>
<td>0.024</td>
<td>93.8%</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>-0.1%</td>
<td>0.205</td>
<td>0.208</td>
<td>96.4%</td>
<td></td>
</tr>
<tr>
<td>(R(0.2))</td>
<td>1.1%</td>
<td>0.09</td>
<td>0.101</td>
<td>97.2%</td>
<td></td>
</tr>
<tr>
<td>(R(0.4))</td>
<td>-0.2%</td>
<td>0.053</td>
<td>0.053</td>
<td>93.8%</td>
<td></td>
</tr>
<tr>
<td>(R(0.7))</td>
<td>-0.3%</td>
<td>0.030</td>
<td>0.033</td>
<td>94.6%</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Result of 1000 simulations to evaluate the application of inference based on asymptotic theory to finite sample studies. \((\alpha_0, \alpha_1) = (1.2, 0.45)\). Relative bias = bias/true value \times 100%; SSE: sampling standard error; ASE: mean estimated standard error using asymptotic theory; CP: coverage of 95% Wald-confidence intervals using estimated asymptotic error.
DPOAE stands for distortion product otoacoustic emission, which is an audiology test used to separate normal-hearing from hearing-impaired ears.

The test is administered under nine different auditory stimulus conditions with three levels of frequency (1001, 1416 and 2002 Hz) and three levels of intensity (55, 60 and 65 dB SPL). A total of 210 subjects were included in the study. The subjects were considered cases with hearing impairment at a given frequency if their audiometric threshold exceeds 20dB HL measured by a behavior test (gold standard). Each subject was tested in only one ear. Test result is the negative signal-to-noise ratio. SNR, with higher value being more indicative of hearing impairment.

The objective of the analysis is to determine the optimal noise ratio, -SNR, with higher value being more indicative of hearing impairment. The test setting for the clinical use of DPOAE to separate normal from hearing-impaired ears, but bear in mind an ear may be determined to be hearing impaired or normal at different frequencies.

We partition the data into nine subsets, corresponding to the nine test settings. Data is analyzed by the method specified in (8) where both intercept and slope parameters vary across subsets. The set T is chosen by the midpoint method from (0.0001, 0.9999). We choose the reference subset (subset 1) to be the test setting with frequency value of 1001Hz and intensity value of 55 dB SPL. Let (β̂k, γ̂k) be the intercept and slope estimates for the ROC curve for setting (1001, 55) and the subsequent β̂k and γ̂k, with k = 2,…,9 represent the differences in the intercept and slope parameters of the ROC curves between subset k and subset 1. None of the γ̂k is statistically significantly different from 0 at α = 0.05 α = 0.05 level (data are not shown).

We also develop a χ² test statistic $\hat{\chi}^2 = \hat{\Sigma}^{-1}/y$, where y = (γ̂2,…,γ̂9) and $\hat{\Sigma}$ is the estimated covariance matrix of $\hat{y}$. Write the null hypothesis as $H_0: y = 0$ and compare the above statistic with a χ² distribution with 8 degrees of freedom gives a P-value of 0.84, suggesting insignificant slope terms consistent with the result from testing the significance of each γ̂k separately.

We re-analyze the data with the slope terms omitted and Table 6 summarizes the results. We can see that among the nine test settings, the setting (1416, 55) (β̂5) has the largest intercept estimate and the difference from the intercept of the reference setting (1001, 55) is statistically significant (p=0.0041). The P values for other parameters ($\hat{\beta}_{10}$, $\hat{\beta}_{11}$ to $\hat{\beta}_{10}$) are not significant. The estimated ROC curve for setting (1416, 55) is $ROC(t) = \Phi(2.54 + 0.82\Phi^{-1}(t))$ with estimated AUC value of 0.975, which is the largest among all settings. The performance of the test declines with increasing intensity at each fixed frequency value. This analysis suggests (1416Hz, 55 dB SPL) is a better test setting than the reference setting. It has also shown that at fixed false positive level of 10%, the expected sensitivities for the nine test settings range from 62.5% to 93.2%, demonstrating the importance of choosing an optimal test setting (Table 6).

### Table 6: Covariates effects on the ROC curves estimated by the the OLS method for the DPOAE test.

<table>
<thead>
<tr>
<th>(Frequency, intensity)</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Err</th>
<th>P value</th>
<th>AUC</th>
<th>ROC(0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1001,55)</td>
<td>$\beta_0$</td>
<td>1.73</td>
<td>0.26</td>
<td>&lt; 0.0001</td>
<td>0.910</td>
<td>0.751</td>
</tr>
<tr>
<td>(1001,55)</td>
<td>$\gamma_0$</td>
<td>0.82</td>
<td>0.12</td>
<td>&lt; 0.0001</td>
<td>0.910</td>
<td>0.751</td>
</tr>
<tr>
<td>(1001,60)</td>
<td>$\beta_0$</td>
<td>-0.20</td>
<td>0.14</td>
<td>0.1455</td>
<td>0.882</td>
<td>0.684</td>
</tr>
<tr>
<td>(1001,65)</td>
<td>$\beta_0$</td>
<td>-0.36</td>
<td>0.19</td>
<td>0.0607</td>
<td>0.855</td>
<td>0.625</td>
</tr>
<tr>
<td>(1416,55)</td>
<td>$\beta_0$</td>
<td>0.81</td>
<td>0.28</td>
<td>0.0041*</td>
<td>0.975</td>
<td>0.932</td>
</tr>
<tr>
<td>(1416,60)</td>
<td>$\beta_0$</td>
<td>0.36</td>
<td>0.29</td>
<td>0.2172</td>
<td>0.947</td>
<td>0.851</td>
</tr>
<tr>
<td>(1416,65)</td>
<td>$\beta_0$</td>
<td>0.19</td>
<td>0.28</td>
<td>0.5064</td>
<td>0.931</td>
<td>0.808</td>
</tr>
<tr>
<td>(2002,55)</td>
<td>$\beta_0$</td>
<td>0.47</td>
<td>0.29</td>
<td>0.1052</td>
<td>0.956</td>
<td>0.875</td>
</tr>
<tr>
<td>(2002,60)</td>
<td>$\beta_0$</td>
<td>0.42</td>
<td>0.28</td>
<td>0.1293</td>
<td>0.952</td>
<td>0.864</td>
</tr>
<tr>
<td>(2002,65)</td>
<td>$\beta_0$</td>
<td>0.10</td>
<td>0.27</td>
<td>0.7019</td>
<td>0.921</td>
<td>0.782</td>
</tr>
</tbody>
</table>

In summary, the proposed linear regression framework provides an unified approach for the ROC curve analysis. It can be used to estimate the ROC curve, as well as model covariate effect. The application of ROC curve goes beyond the medical diagnostic field and it can be used for evaluating any discrimination tools. It is, and will continue to be an important and exciting area to engage in research.

### Concluding Remarks

This manuscript proposes a semiparametric OLS method to estimate and compare performance of diagnostic tests and more generally, to assess potential covariate effects on the test performance. The asymptotic distribution theory for the OLS estimator is developed for the ROC curve estimation, in which the estimators are shown to be consistent and asymptotically normally distributed. For modeling covariate effects, we recommend bootstrap resampling for variance estimation.

Our proposed estimator provides useful addition to the field of rank-based semiparametric ROC modeling. Those semiparametric approaches are more robust than parametric approaches, by assuming a functional form on the ROC curve itself but not the test results and thus invariant to monotone transformation of the test results. At the same time, they offer better efficiency compared to nonparametric method. We have done extensive simulations to compare the proposed OLS estimator with two other commonly used semiparametric ROC modeling methods (the ROC-GLM and placement value based method), and found the OLS estimator has comparable performance in general and slightly better performance in some scenarios [12]. The OLS estimator, however, is much more intuitive compared to other estimators and very easy to implement using standard linear regression software, which could make it particularly appealing to clinical audience.

In short, the proposed linear regression framework provides an unified approach for the ROC curve analysis. It can be used to estimate the ROC curve, as well as model covariate effect. The application of ROC curve goes beyond the medical diagnostic field and it can be used for evaluating any discrimination tools. It is, and will continue to be an important and exciting area to engage in research.

### Acknowledgements

The authors would like to acknowledge Margaret S. Pepe for her invaluable contribution to the development of this manuscript. The work is partially funded by NIH grant GM54438 and P30CA015704.
References

Appendix I: Proof of Equation (4)

Hsieh and Turnbull (1996, Theorem 2.1) showed, as \( n \to \infty \),

\[
\sup_{0 \leq s \leq 1} | \widehat{ROC}_c(t) - ROC(t) | \to 0 \text{ a.s.} \tag{9}
\]

They also showed (Theorem 2.2) for independent observations, under the above conditions, there exists a probability space on which one can define sequences of two independent versions of Brownian bridges \( (B_1^{(n)}, B_2^{(n)}, 0 \leq t \leq 1) \), and the following statement holds:

\[
\sqrt{n} \left( \widehat{ROC}_c(t) - ROC(t) \right) = \sqrt{\lambda} B_1^{(n)}(ROC(t)) + \alpha_1 \frac{\phi(\alpha_0 + \alpha_1 \Phi^{-1}(t))}{\phi(\Phi^{-1}(t))} B_2^{(n)}(t) + o(n^{-\frac{1}{2}} \log n) \tag{10}
\]

a.s. uniformly on \([a, b], 0 < a < b < 1\).

The above two theorems stated the strong consistency and strong approximation properties for the ROC curve.

Fix \( t \in [a, b] \), by intermediate value theorem,

\[
\Phi^{-1}(\widehat{ROC}_c(t)) - \Phi^{-1}(ROC(t)) = \phi(\Phi^{-1}(ROC^*(t))) \left( \widehat{ROC}_c(t) - ROC(t) \right) \tag{11}
\]

where

\[
\left| ROC^*(t) - ROC(t) \right| \leq \left| \widehat{ROC}_c(t) - ROC(t) \right| \tag{12}
\]

Therefore, \( ROC^*(t) \to ROC(t) \) a.s. by (9)

From continuous mapping theorem:

\[
\phi(\Phi^{-1}(ROC^*(t))) \to \phi(\Phi^{-1}(ROC(t))) \text{ a.s.} \tag{13}
\]

Notice

\[
\phi(\Phi^{-1}(ROC(t))) = \frac{1}{\phi(\Phi^{-1}(ROC(t)))} = \frac{1}{\phi(\alpha_0 + \alpha_1 \Phi^{-1}(t))} \tag{14}
\]

Then we have

\[
\sqrt{n} \left( \Phi^{-1}(\widehat{ROC}_c(t)) - \Phi^{-1}(ROC(t)) \right) \to \frac{1}{\phi(\alpha_0 + \alpha_1 \Phi^{-1}(t))} \sqrt{n} \left( \widehat{ROC}_c(t) - ROC(t) \right) \tag{15}
\]
\[
\frac{1}{\phi(\Phi^{-1}(ROC(t)))} B_1^{(n)}(ROC(t)) + \alpha \frac{1}{\phi(\Phi^{-1}(t))} B_2^{(n)}(t) + o(n^{-\frac{1}{2}} \log n^2) \quad (16)
\]

a.s. uniformly on \([a, b], 0 < a < b < 1\).

Let

\[
V_n(t) = \sqrt{n} \left( \Phi^{-1}(\widehat{ROC}_t(t)) - \Phi^{-1}(ROC(t)) \right)
\]

\[V(t) = \sqrt{\frac{\lambda}{\phi(\Phi^{-1}(ROC(t)))}} B_1^{(n)}(ROC(t)) + \alpha \frac{1}{\phi(\Phi^{-1}(t))} B_2^{(n)}(t) \quad (18)
\]

Equation (16) implies \(V_n(t) \Rightarrow V(t) \text{ in } (D[a, b], \| . \|_\infty)\).

Equation (4) resulted from the fact that \(V(t)\) is the sum of two independent Brownian Bridges.

**Appendix II: Proof of Theorems 1**

Theorem 1 is a direct consequence of the equation (4), namely, asymptotic normality of the empirical ROC estimates.