Published in final edited form as:

Biom J. 2016 July ; 58(4): 747–765. doi:10.1002/bimj.201500099.

Least Squares Regression Methods for Clustered ROC Data with Discrete Covariates

Liansheng Larry Tang^{1,2}, Wei Zhang³, Qizhai Li³, Xuan Ye¹, and Leighton Chan²

¹Department of Statistics, George Mason University, Fairfax, VA 22030, USA

²Epidemiology and Biostatistics, NIH Clinical Center, Rockville, MD, 20814 USA

³Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

Abstract

The receiver operating characteristic (ROC) curve is a popular tool to evaluate and compare the accuracy of diagnostic tests to distinguish the diseased group from the non-diseased group when test results from tests are continuous or ordinal. A complicated data setting occurs when multiple tests are measured on abnormal and normal locations from the same subject and the measurements are clustered within the subject. Although least squares regression methods can be used for the estimation of ROC curve from correlated data, how to develop the least squares methods to estimate the ROC curve from the clustered data has not been studied. Also, the statistical properties of the least squares methods under the clustering setting are unknown. In this article, we develop the least squares ROC methods to allow the baseline and link functions to differ, and more importantly, to accommodate clustered data with discrete covariates. The methods can generate smooth ROC curves which satisfy the inherent continuous property of the true underlying curve. The least squares methods are shown to be more efficient than the existing nonparametric ROC methods under appropriate model assumptions in simulation studies. We apply the methods to a real example in the detection of glaucomatous deterioration. We also derive the asymptotic properties of the proposed methods.

1 Introduction

A diagnostic test is commonly used to screen or diagnose diseases. For example, a cancer screening test is often used to distinguish a cancer patient from a non-cancer patient. For diagnostic tests that generate binary results, their accuracy can be summarized in terms of sensitivity (the probability of identifying a diseased subject when the disease truly exists) and specificity (the probability of correctly ruling out a non-diseased subject when the disease is truly absent). For diagnostic tests that generate ordinal or continuous results, the receiver operating characteristic (ROC) curve is a standard statistical tool to describe and compare the accuracy of tests (Zhou et al., 2011). An ROC curve combines all possible pairs of sensitivities and one minus specificities from different decision thresholds and thus describes the accuracy of tests apart from decision thresholds. The ROC analysis of single

Conflict of interest

The authors have declared no conflict of interest.

test data has been extensively investigated, see Qin and Zhou (2006), Qin et al. (2011), and Hsieh and Turnbull (1996), since the seminal work by Dorfman and Alf (1969).

Statistical tests are often complicated when used in diagnostic biomarker studies where two or more different diagnostic biomarkers are simultaneously measured on normal and abnormal locations. The results from the abnormal and normal locations are clustered within the same subject. Statistical analysis on this clustered data presents new challenges, because the general statistical methods applied on clustered data without considering the clustering feature could result in biased and incorrect statistical inference. To the best of our knowledge, Li and Zhou (2008) and Tang et al. (2012) are the existing methods to estimate ROC curves from this type of clustered data. Li and Zhou (2008) discussed a nonparametric method for estimating the ROC curve for clustered data. Tang et al. (2012) extended Li and Zhou (2008)'s non-parametric methods to allow the simultaneous comparison among multiple tests. Although Obuchowski (1997) also considered clustered diagnostic test data, her methods dealt with the area under the ROC curve, which can be considered as a special case of Li and Zhou's method. These nonparametric methods generated rough ROC curve estimators, and did not incorporate smoothing techniques in estimation. Hence, they may not be as efficient as parametric or semi-parametric methods, which can incorporate appropriate model assumptions.

The least squares methods by Tang and Zhou (2009) and Tang and Zhou (2012) can yield smooth ROC curves for correlated data and are more efficient than nonparametric methods, but the methods have not been studied for clustered ROC data. This paper focuses on applying parametric and semi-parametric least squares ROC methods to clustered data and procedures for deriving the properties of the estimated ROC parameters. The regression framework for clustered ROC data is similar to that for correlated ROC data, and the main difference lies in the response variables. For correlated data, the standard empirical ROC curve was used as the response variable, while the nonparametric ROC curve was used for clustered data (Li and Zhou, 2008). In addition, simultaneous confidence bands can be obtained using the ROC parameter estimators and their estimated variances from the proposed methods to visualize the uncertainty of the estimated ROC curves.

Furthermore, various discrete covariates may have effects on ROC curve analysis (Zhou et al., 2011). For example, the marker type is a common discrete covariate in ROC comparison analysis. Because of this, we further take into consideration of the discrete covariate in the parametric and semi-parametric least squares method for clustered ROC curves analysis. We allow both methods to have interaction terms between the level of covariates and the false positive rates (FPRs). In addition, the semi-parametric method allows unknown baseline functions.

2 Notations and definitions

Suppose that there are V diagnostic tests, where tests could be combinations of multiple discrete covariates. For example, combining gender and 3 biomarkers produces 3 biomarker levels for each gender, totaling six diagnostic tests. Let $X_{\nu ip}$ denote the test result of the ν th level on the p th abnormal location within subject i , where $\nu = 1, \dots, V$, $i = 1, \dots, m$, and $p =$

$1, \dots, m_{vi}$. m is the total number of subjects involved in the trial and m_{vi} is the number of abnormal locations from the i th subject for the v th level. Similarly, let Y_{viq} denotes the result of the v th level on the q th normal location within subject i , where $v = 1, \dots, V$, $i = 1, \dots, m$, and $q = 1, \dots, n_{vi}$. n_{vi} is the number of normal locations from the i th subject for the v th level.

The data of X_{vip} and Y_{viq} are clustered within the same subject i . The marginal survival functions of X_{vip} and Y_{viq} are $S_{D, v}$ and $S_{\bar{D}, v}$ respectively, where $X_{vip} \sim S_{D, v}(x)$, and $Y_{viq} \sim S_{\bar{D}, v}(y)$, for the p th abnormal location and the q th normal location in subject i respectively, for $p = 1, \dots, m_{vi}$ and $q = 1, \dots, n_{vi}$. This implies the exchangeability of locations within subjects. In other words, for a test, the distribution for normal locations is the same for all subjects, and so is the distribution for abnormal locations. Without loss of generality, we assume that measurements tend to be larger for an abnormal location than that for a normal location. Given a false positive rate (FPR) u , which is 1-specificity, the ROC curve of the v th level at u is given by $R_v(u) = S_{D, v}(S_{\bar{D}, v}^{-1}(u))$. The survival functions $S_{D, v}$ and $S_{\bar{D}, v}$ can be estimated empirically. The extended empirical functions of $S_{D, v}$ and $S_{\bar{D}, v}$ defined in Li and Zhou (2008) are, respectively,

$$\hat{S}_{D, v}(x) = \frac{1}{\sum_{i=1}^m m_{vi}} \sum_{i=1}^m \sum_{p=1}^{m_{vi}} I(X_{vip} > x),$$

and

$$\hat{S}_{\bar{D}, v}(y) = \frac{1}{\sum_{i=1}^m n_{vi}} \sum_{i=1}^m \sum_{q=1}^{n_{vi}} I(Y_{viq} > y).$$

Then the resulting empirical ROC curve for level v is given by

$$\hat{R}_v(u) = \hat{S}_{D, v}(\hat{S}_{\bar{D}, v}^{-1}(u)). \quad (1)$$

These are the simple averages of the indicator random variables for the normal and abnormal locations. Although the computation is straightforward, it is not trivial to obtain the statistical properties of these empirical functions as shown in Li and Zhou (2008). We will use these empirical functions as the response variables in our previously proposed regression models (Tang and Zhou, 2012).

We can model the ROC curves as follows

$$R_v(u) = G^{-1}\{H_0(u) + \alpha_v + \beta_v H_v(u)\}, \quad v = 1, \dots, V. \quad (2)$$

where G and H are known inverse distribution functions, for $\nu = 1, \dots, V$. The commonly used binormal model assumes that $G = H = \Phi^{-1}$. However, this paper allows a more general setting when G and H can differ. Based on the model (2), and the ROC curves can be transformed by the function G , the model (2) becomes

$$G\{R_\nu(u)\} = H_0(u) + (\alpha_\nu + \beta_\nu H_\nu(u))I(\nu \neq 1), \quad \nu = 1, \dots, V. \quad (3)$$

Let H_0 be a continuous function, and for the model identifiability reason, we let $\alpha_1 = \beta_1 = 0$. Here, the parameters α_ν and β_ν can be considered as the location and scale parameters with α_ν measuring the superiority of the ν th ROC curve over the baseline curve and β_ν allowing the different shapes of the ROC curves for the ν th level and the baseline level. We can model the baseline function H_0 nonparametrically or parametrically by letting $H_0(u) = \alpha_1 + \beta_1 H_1(u)$, where H_1 is a known inverse distribution function.

We can choose the range of u to obtain the entire ROC curve or only part of the curve. By choosing u in the range of $(0, 1)$, we can get the entire curve, and by choosing $u \in [t_0, t_1]$ ($0 < t_0 < t_1 < 1$), we have the partial ROC curve from t_0 to t_1 . The partial ROC curve at lower specificities is of partial interest to evaluate screening tests used in a large population. Utilizing the partial ROC curve is particularly important in comparing markers that are developed to screen a large population for specific diseases. A misclassified patient may have to undergo a more invasive test for verification purposes. It is desired to achieve a higher specificity so that the chance of misclassifying healthy patients can be kept at a minimal level. Thus it is more interesting to compare screening markers at a higher range of specificities than in the entire range of $(0, 1)$.

Various authors have considered similar models as the model (2) when there is only one discrete covariate, namely, the test type. Metz et al. (1998) used $G = H_1 = \Phi^{-1}$ for one test, where Φ is the cdf of the standard normal distribution. Zhang and Pepe (2005) proposed to use Φ^{-1} for all inverse distribution functions to estimate ROC curves for multiple tests. The authors also proposed an intuitive parametric least squares (PLS) ROC method that requires no iteration and thus takes much less computation time than the aforementioned methods. Tang and Zhou (2009) further derived the asymptotic covariance of the PLS estimators for correlated ROC data. The baseline function H_0 is assumed to take on a known function with some unknown parameters by these authors.

In this paper we propose novel semi-parametric and parametric least squares methods for clustered continuous ROC data. We further derive the properties of our least squares parameter estimators for clustered ROC data, and apply them to draw inference on the estimated ROC parameters and to visualize the difference among the ROC curves using confidence bands. Due to the clustering structure, the derivation of the properties is considerably more difficult than the correlated data structure considered in Tang and Zhou (2009) and Tang and Zhou (2012).

3 Least Squares Regression ROC Methods

The ROC curve for the ν th level can be estimated using the empirical function in (1). We then need to choose a reference level as the baseline from the V test levels. A reference level among the V ROC curves can be chosen if we are particularly interested in comparing it with other levels of covariate combinations. And we can choose L distinct FPR points, $\mathbf{u} = (u_1, \dots, u_L)^T$, within the range of interest $[t_0, t_1]$ of ROC curves. In the simulation studies and the example, the FPR points are chosen to be equally spaced points in $[0, 1]$ with the increment of 0.02. Accounting for the estimation errors for the ROC curves, the model (3)

can be written in the following $\sum_{\nu=1}^V L_{\nu}$ linear regression equations, where L_{ν} is the number of FPRs we use for the level ν ROC curve.

$$\begin{aligned} G\{\hat{R}_{\nu}(u_1)\} &= H_0(u_1) + (\alpha_{\nu} + \beta_{\nu} H_{\nu}(u_1))I(\nu \neq 1) + \varepsilon_1 \\ &\vdots \\ G\{\hat{R}_{\nu}(u_L)\} &= H_0(u_L) + (\alpha_{\nu} + \beta_{\nu} H_{\nu}(u_L))I(\nu \neq 1) + \varepsilon_L \quad \text{for } \nu=1, \dots, V. \end{aligned} \quad (4)$$

Our parametric least squares (PLS) method assumes that the baseline function has a known form, while our semi-parametric least squares (SLS) method allows an unknown baseline function form of H_0 . Both methods require the empirical ROC estimators at the distinct jump FPR points as the estimators are to be used as the response variables. For PLS, the baseline function could be $H_0(u) = \alpha_1 + \beta_1 H_1(u)$ with unknown parameters α_1, β_1 and a known basis function H_1 . We can then fit the equations above using a simple least squares method. Although the error terms are correlated, we will show the estimator vector $\hat{\boldsymbol{\theta}} = (\hat{\alpha}_1, \hat{\beta}_1, \dots, \hat{\alpha}_V, \hat{\beta}_V)^T$ for $\boldsymbol{\theta} = (\alpha_1, \beta_1, \dots, \alpha_V, \beta_V)^T$ is consistent later. Furthermore, if we let $H_0 = \alpha_1 + \beta_1 \Phi^{-1}$ and $H_2 = \dots = H_V = G = \Phi^{-1}$, the model (4) becomes the multivariate binormal model.

We outline the PLS and SLS procedure as follows. For simplicity, we use the notation of $\mathcal{S}(\mathbf{u}) = (\mathcal{S}(u_1), \mathcal{S}(u_2), \dots, \mathcal{S}(u_L))^T$ be any process or function $\mathcal{S}$ applied on the vector $\mathbf{u}$ for $\nu = 1, \dots, V$. Compared with the PLS method, the SLS method involves an additional step of estimating the unknown baseline function H_0 . And the parameter vector to be estimated is $\boldsymbol{\theta}^* = (\alpha_2, \beta_2, \dots, \alpha_V, \beta_V)^T$, which has two less parameters to be estimated compared to the PLS method.

- *Step 1:* Estimate the ROC curves using the empirical ROC curves described in (1). For SLS, estimate the unknown baseline function H_0 by transforming the empirical reference ROC curve:

$$\hat{H}_0(\mathbf{u}) = G\{\hat{R}_1(\mathbf{u})\}.$$

- *Step 2:* Transform the empirically estimated ROC curves of the ν th level, $\hat{R}_{\nu}$, by G , and define the following vector for PLS

$$Y_\nu = G\{\hat{R}_\nu(u)\},$$

for $\nu = 1, \dots, V$.

Transform the empirically estimated ROC curves of the ν th level, $\hat{R}_\nu(u)$, by G :

$$Y_\nu^* = G\{\hat{R}_\nu(u)\} - \hat{H}_0(u),$$

for $\nu = 2, \dots, V$.

- *Step 3:* For PLS, combine $Y_1, Y_2, \dots, Y_V$ to get the following linear regression equation

$$Y = X\theta + \varepsilon, \quad (5)$$

where $Y = (Y_1^T, \dots, Y_V^T)^T$ is a $(\Sigma_\nu L) \times 1$ vector, and the $(\Sigma_\nu L) \times (2V)$ design matrix X is

$$X = \begin{pmatrix} X_1 & 0 & 0 & \cdots & 0 \\ X_2 & X_2 & 0 & \cdots & 0 \\ X_3 & 0 & X_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_V & 0 & 0 & \cdots & X_V \end{pmatrix},$$

with its $L_\nu \times 2$ submatrices X_ℓ for $\ell = 1, \dots, L$.

$$X_\nu = \begin{pmatrix} 1 & \cdots & 1 \\ H_\ell(u_1) & \cdots & H_\ell(u_L) \end{pmatrix}^T.$$

Also, the error term ε is given by $\varepsilon = (\varepsilon_{11}, \dots, \varepsilon_{1L}, \dots, \varepsilon_{V1}, \dots, \varepsilon_{VL})^T$.

For SLS, perform linear regression on the $L(V-1) \times 1$ vector

$$Y^* = (Y_2^{*T}, \dots, Y_V^{*T})^T:$$

$$Y^* = X^*\theta^* + \varepsilon^*. \quad (6)$$

Here the $L(V-1) \times 2(V-1)$ matrix X^* is given by:

$$\mathbf{X}^* = \begin{pmatrix} \mathbf{X}_2^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_3^* & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_V^* \end{pmatrix}$$

with the $L(V-1) \times 1$ error vector $\mathbf{e}^*$.

- *Step 4:* Obtain the PLS estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ from the regression equations in (5):

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (7)$$

Obtain the SLS estimator $\hat{\boldsymbol{\theta}}^*$ of $\boldsymbol{\theta}$ from the regression equations in (6):

$$\hat{\boldsymbol{\theta}}^* = ((\mathbf{X}^*)^T (\mathbf{X}^*))^{-1} (\mathbf{X}^*)^T \mathbf{Y}^*. \quad (8)$$

The regression model framework is similar as the one for the correlated ROC data proposed in Tang and Zhou (2009) and Tang and Zhou (2012). However, the response variables in the aforementioned models are empirical functions from clustered ROC data. These variables include the ones for the correlated ROC data as a special case when normal locations only come from healthy subjects and abnormal locations only come from diseased subjects. In this correlated ROC case, the correlation arises from the same population which is either the healthy population or the diseased population, while in the clustered ROC case, the correlation arises both from within the population and between the populations. Also, for the SLS method, the estimated reference ROC curve is the empirical curve. And the resulting $\hat{H}_0$ and $\hat{\boldsymbol{\theta}}^*$ can then be plugged in the model (3) to get the semi-parametric ROC curve estimators.

3.1 Theoretical Properties and a Separation Curve

Hsieh and Turnbull (1996) showed that an empirical ROC curve process can be decomposed into two independent Brownian bridge processes. One of the processes is based on observations of the diseased subjects, and the other process is based on observations of the non-diseased subjects. We can not apply their results to derive the large sample properties of our estimators on clustered data. However, since observations on different subjects are independent, we would be able to derive the asymptotic covariance structure of the empirical ROC curves for clustered data, and then obtain the asymptotic distributions of our least squares estimators.

We derive the theoretical properties for the PLS and SLS methods in the following with the general settings of $m_{vi}, n_{vi} \geq 1$. Here we assume that for each level v on each subject i , test data are collected on m_{vi} normal locations and n_{vi} abnormal locations. We also define

$R(u) = (R_1^T(u_1), \dots, R_V^T(u_L))^T$ and its estimator $\hat{R}(u) = (\hat{R}_1^T(u_1), \dots, \hat{R}_V^T(u_L))^T$, where u_v is a L -dimension vector for $v = 1, \dots, V$.

Result 1—Suppose that $S_{D,v}$ and $S_{\bar{D},v}$ have continuous and positive derivatives, $s_{D,v}$ and $s_{\bar{D},v}$, respectively, on $[S_{\bar{D},v}^{-1}(a) - \varepsilon, S_{\bar{D},v}^{-1}(a) + \varepsilon]$, for $\varepsilon > 0$ and some $0 < a < 1$, and assume $n \rightarrow \infty$, n is the total subjects, $n^{-1} \sum_{i=1}^n m_{vi} \rightarrow \lambda_v$ and $n^{-1} \sum_{i=1}^n n_{vi} \rightarrow \gamma_v$ for some positive

constants λ_v and γ_v , $v = 1, 2, \dots, V$. Define $r_v(u_{K_v}) = \frac{s'_{D,v}(S_{\bar{D},v}^{-1}(u_{K_v}))}{s'_{\bar{D},v}(S_{\bar{D},v}^{-1}(u_{K_v}))}$. Then $\sqrt{n}(\hat{R}(u) - R(u))$ converges to a multivariate normal distribution with mean zero and the variance-covariance matrix is $E(WW^T)$ in probability 1, where

$$W = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \begin{pmatrix} \frac{n}{m_1} \sum_{p=1}^{m_1} \{I(X_{1ip} > S_{\bar{D},1}^{-1}(u_{K_1})) - S_{D,1}(S_{\bar{D},1}^{-1}(u_{K_1}))\} \\ \frac{n}{m_2} \sum_{p=1}^{m_2} \{I(X_{2ip} > S_{\bar{D},2}^{-1}(u_{K_2})) - S_{D,2}(S_{\bar{D},2}^{-1}(u_{K_2}))\} \\ \vdots \\ \frac{n}{m_V} \sum_{p=1}^{m_V} \{I(X_{vip} > S_{\bar{D},V}^{-1}(u_{K_V})) - S_{D,V}(S_{\bar{D},V}^{-1}(u_{K_V}))\} \\ r_1(u_{K_1}) \frac{n}{n_1} \sum_{q=1}^{n_1} \{I(Y_{1iq} > S_{\bar{D},1}^{-1}(u_{K_1})) - u_{K_1}\} \\ r_2(u_{K_2}) \frac{n}{n_2} \sum_{q=1}^{n_2} \{I(Y_{2iq} > S_{\bar{D},2}^{-1}(u_{K_2})) - u_{K_2}\} \\ \vdots \\ r_V(u_{K_V}) \frac{n}{n_V} \sum_{q=1}^{n_V} \{I(Y_{viq} > S_{\bar{D},V}^{-1}(u_{K_V})) - u_{K_V}\} \end{pmatrix} \right\}.$$

proof: See Appendix.

Result 2—Under the conditions in Result 1, $(\hat{\theta} - \theta)$ is asymptotically distributed as a $2V$ -dimension multivariate normal random vector with mean zero and the variance-covariance matrix $(X^T \sum_{\varepsilon}^{-1} X)^{-1}$, where $\Sigma_{\varepsilon} = \text{var}(\varepsilon)$. Also, $(\hat{\theta}^* - \theta^*)$ is asymptotically distributed as a $2(V-1)$ -dimension multivariate normal random vector with mean zero and the variance-covariance matrix $((X^*)^T (\sum_{\varepsilon}^*)^{-1} X^*)^{-1}$, where $\sum_{\varepsilon}^* = \text{var}(\varepsilon^*)$.

Proof: See Appendix.

With the large sample properties for the clustered ROC curve estimators, we can use the separation curve method outlined in Tang and Zhou (2012) to identify the region of specificities in which diagnostic tests differ. The separation curve, defined as $\gamma(u) = \alpha_v + \beta_v H_{\gamma}(u)$, can be used to analytically and visually differentiate between the v th and the reference ROC curves, for $v = 2, \dots, V$. Similarly, a separation curve to visualize the difference between the v_1 th ROC curve and the v_2 th ROC curve is defined as

$$\Delta_{v_1, v_2}(u) = (\alpha_{v_1} - \alpha_{v_2}) + \beta_{v_1} H_{v_1}(u) - \beta_{v_2} H_{v_2}(u).$$

The separation curve is estimated by plugging in respective parameter estimators from the LS methods. Following similar lines of proof from Tang and Zhou (2012), we see that the confidence band of the estimated separation curve, $\hat{\nu}(u)$, can be written as

$\Delta_{\nu}(u) \pm \sqrt{\chi^2(1, H_{\nu}(t)) \text{cov}(\hat{\alpha}_{\nu}, \hat{\beta}_{\nu})(1, H_{\nu}(t))'}$. Here $\text{cov}(\hat{\alpha}_{\nu}, \hat{\beta}_{\nu})$ is a 2×2 submatrix from the estimated covariance matrix of the LS parameter estimators. The confidence band for the estimated curve $\nu_{1, \nu_2}(u)$ involves a vector of four parameters, $\theta_{\nu_1, \nu_2} = (\alpha_{\nu_1}, \alpha_{\nu_2}, \beta_{\nu_1}, \beta_{\nu_2} H_{\nu_2}(u))'$. We can obtain $\text{cov}(\hat{\theta}_{\nu_1, \nu_2})$ which is a 4×4 submatrix from the estimated covariance matrix of the LS estimators. The confidence band for comparing the ν_1 th and ν_2 th tests takes the form of

$$\Delta_{\nu_1, \nu_2}(u) \pm \sqrt{\chi^4(1, H_{\nu_1}(t), -1, -H_{\nu_2}(t)) \text{cov}(\hat{\theta}_{\nu_1, \nu_2})(1, H_{\nu_1}(t), -1, -H_{\nu_2}(t))'}$$

4 Simulation Studies

4.1 Simulation Setting

In simulation studies, we investigate the finite sample performance of the proposed PLS and SLS methods on clustered data. We incorporate the correlation structure between abnormal and normal locations within the same subject to simulate clustered data. We consider two discrete covariates including gender and the marker type, and assume that the total number of the normal locations and abnormal locations for every subject is the same in the simulation studies. We consider a combination of each biomarker type and each gender as a test, and the overall number of tests, V , is 2 times the number of biomarker types. We have total 10 tests with gender and 5 biomarker types in the simulation. And the number of total locations (normal and abnormal) for the k th subject is S_k , $k = 1, 2, \dots, m$. We let the total number of the locations for every subject follow a discrete uniform distribution $\text{unif}\{6, 10\}$ so that the five values $\{6, 7, 8, 9, 10\}$ are equally likely to be observed. For the k th subject, the number of abnormal locations for a subject is determined by a binomial distribution, $\text{Binomial}(S_k, p)$ and is denoted as D_k , $k = 1, 2, \dots, m$, and the number of normal locations for the k th subject is $\bar{D}_k = S_k - D_k$.

We generate multivariate normal clustered marker data using a similar setting as in Li and Zhou (2008). First, for the abnormal locations of the k th subject, D_k V -dimension multivariate normal random vectors $T_{D,s} = (T_{s1}, \dots, T_{sV})^T$ with mean vector $\mathbf{0}$ and variance-covariance matrix $\Sigma_{T,s} = \text{cov}(T_{si}, T_{sj})$ are generated. In the $\Sigma_{T,s}$, we set $\text{cov}(T_{si}, T_{sj}) = \rho_1^{|j-i|}$ for $i, j \in \{1, \dots, V\}$ and $s = 1, 2, \dots, D_k$. Similarly, for the normal locations of the k th subject, we generate $\bar{D}_k$ random vectors from the V -dimension multivariate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix $\text{cov}(T_{si}, T_{sj}) = \rho_2^{|j-i|}$, where $i, j = 1, 2, \dots, V$, $s = 1, 2, \dots, \bar{D}_k$. Through simulation we obtain the test results $\tilde{T}_{D,s}$ and $\tilde{T}_{\bar{D},s}$ for the abnormal and normal location, respectively, where $\tilde{T}_{D,s} = T_{D,1} \sqrt{\eta} + T_{D,s+1} \sqrt{1-\eta}$ and $\tilde{T}_{\bar{D},s} = T_{\bar{D},1} \sqrt{\eta} + T_{\bar{D},s} \sqrt{1-\eta}$. So, the test results for the abnormal and normal locations are clustered within the same subject. The test results are $\tilde{T}_{D,s} + (1, 1.05, \dots, 1 + 0.05(v-1), \dots,$

$1.45)^{\tau}$, $v = 2, 3, \dots, V$, for the abnormal locations, and $\tilde{T}_{\tilde{D},s}$ for the normal locations. In this case, different test levels possess different true ROC curves.

We simulate 1000 replicates for sample size $m = 20, 50, 80$, $\rho_1 = 0.5$, $\rho_2 = 0.4$, $\eta = 0.3$, $p = 0.5$, and $V = 10$, to evaluate the finite performance of the proposed parametric least squares method (PLS) and semi-parametric least squares method (SLS) for estimating the ROC parameters from clustered data. Li and Zhou's method is a nonparametric method and is not included in the simulation. Use the common assumption for inverse distribution H and G in both least squares methods, i.e. $H_1 = H_2 = H_3 = \dots = H_V = G = \Phi^{-1}$. In the PLS approach, the true baseline function is given by $H_0 = 1 + \Phi^{-1}$, $\alpha_1 = 1$, $\beta_1 = 1$. Under this setting, we have $\alpha_v = 1 + 0.05(v - 1)$, $\beta_v = 0$, $v = 2, 3, \dots, 10$. In the SLS approach, we first apply the empirical estimate method to estimate the baseline function H_0 , then use the parametric least squares method to estimate the other parameters.

Using our estimators and the existing methods on the simulated data, we investigate the finite sample performance under different sample sizes. The bias and the root of the mean squared error (RMSE) are summarized in Table 1. Both LS methods perform relatively well with small biases and RMSEs. As the sample size increases, the bias and RMSE tend to get smaller for both methods. The PLS method tends to have smaller biases and RMSEs than the SLS method. This is expected since PLS required a stronger model assumption on the reference test.

4.2 Comparing the performance of PLS and SLS methods

Using the previous setting, we see that all the V levels have different true ROC curve, $R_v(u) = \Phi(1 + 0.05(v - 1) + \Phi^{-1}(u))$, $v = 1, 2, \dots, 10$, under these simulation assumptions. Now we try to compare the PLS method, SLS method with Li and Zhou (LZ) method based on the performance of estimating the partial AUC at FPR ranging from 0 to 0.2 and the AUC at FPR ranging from 0 to 1. The bias is calculated as the difference between partial AUCs of 10 estimated ROC curves by three methods with the true AUC in Table 2. The comparison of diagnostic tests at smaller FPRs is of interest in screening studies in a large population. This is due to the efforts to minimize the number of healthy patients who can be incorrectly diagnosed and may later undergo invasive tests. It is clear from the table that the PLS method tends to have the smallest bias among all three methods. The bias decreases as the sample size gets larger for all methods, and the SLS method has slightly larger biases than the LZ method. In addition, Table 3 summarizes the RMSE of estimating the partial AUC at FPR ranging from 0 to 0.2 using three methods. The results show that the PLS method always has smaller RMSE than the other two methods. For example, when $m = 50$, the RMSE of the partial AUC ranging from 0 to 0.2 for the test 5 by the PLS, LZ and SLS method are 0.0125, 0.0128 and 0.0132, respectively. Furthermore, we compare the bias and RMSE of estimating AUC at FPR ranging from 0 to 1 using the three methods. The results are showed in Table 4 and Table 5. Overall, the LZ method has positive biases, while the PLS and SLS method have negative biases. **The proposed methods tend to have larger RMSEs than the LZ method when estimating the AUC. The SLS method tends to have the largest RMSEs among three methods.**

5 Example: Application to the Detection of Glaucomatous Deterioration

In this section, we apply the proposed methods to estimate the ROC parameters associated with diagnostic tests in the detection of glaucomatous deterioration. We chose glaucoma diagnosis for several reasons. First, as Li and Zhou (2008) have noted, glaucoma is a progressive optic neuropathy. The related symptoms include loss of retinal ganglion cells, morphological changes to the optic nerve and retinal nerve fiber layer, and vision loss. The global prevalence of glaucoma for population aged 40–80 years is 3.54%. The number of people with glaucoma worldwide was estimated to be 64.3 million in 2013, which will increase to 76.0 million in 2020 and 111.8 million in 2040 (Tham et al., 2014). If glaucoma is not diagnosed and treated, damage can progress and cause a loss of peripheral vision and may eventually lead to complete sight loss. In fact, glaucoma is one of the leading causes of globally preventable blindness. Glaucoma does not cause symptoms in early stages, which makes it hard to be diagnosed, but an eye exam might detect the signs of glaucoma. The visual field deterioration due to glaucoma can be tested using imaging techniques. But it is challenging to accurately identify the pathologically progressive eyes in glaucoma patients since the noise level of the image data is high. It is of interest to study the diagnostic accuracy of two newly developed Bayesian hierarchical models on classifying the stable and progressive eyes (Li and Zhou, 2008). Second, these data includes 160 patients. Visual field tests were given to these patients over 8 years of follow-up. Some patients were measured on both eyes and others were measured only on one eye. This resulted in visual field testing on 188 eyes in total. Because some data are from both eyes of the same patient, test scores from the hierarchical models calculated from both eyes of the same patients are clustered and traditional ROC methods are not applicable to the data. Third, these are the same data used in Li and Zhou (2008) outlining the LZ strategy, thereby allowing better direct comparisons to our methods. The LZ method generated empirical ROC curves for the two models, and AUC difference was found to be significant between the models in their paper. The computation time to estimate the ROC curve using LZ method with small steps at FPR is not trivial. On the other hand, PLS and SLS method can quickly estimate the whole ROC curve based on empirical ROC estimates.

The jump points we used in our methods are equally-spaced points from 0 to 1 with the increment of 0.02. We let $G = H = \Phi^{-1}$. We followed the steps proposed in Section 3 by first estimating the empirical ROC curves. The values of the response variable for PLS and SLS were obtained by transforming the empirical ROC curves, and the values for the design matrix were obtained by taking $\Phi^{-1}(\cdot)$ on the jump points. Final ROC parameter estimates were least squares estimates from Equation (7) and (8) for PLS and CLS, respectively. We found that the parameter estimates have little changes when the number of jump points is increased beyond 50. Figure 1 shows estimated ROC curves for Model 1 (denoted as ROC 1) and for Model 2 (denoted as ROC 2) by all three methods. The reason that there are only five instead of six ROC curves in Figure 1 is the estimated ROC 1 curves by SLS method and LZ method are both empirical ROC curve for Model 1. Using the PLS method, we obtain the estimated ROC curves of Model 1 and 2 as $\hat{R}_1(u) = \Phi(0.604 + 0.850\Phi^{-1}(u))$ and $\hat{R}_2(u) = \Phi(2.471 + 0.817\Phi^{-1}(u))$. For the SLS method, the estimated ROC curve of Model 1 and 2

are $\tilde{R}_1(u)$ and $\Phi(\Phi^{-1}(\tilde{R}_1(u)) + 1.867 - 0.032\Phi^{-1}(u))$, where $\tilde{R}_1(u)$ is the empirical reference ROC curve for the first test level based on the LZ method.

For the SLS method, we also constructed the separation curve and its confidence bands for the example. By comparing the confidence band with a horizontal line at $(u) = 0$, we can also visualize the difference between each pair of estimated ROC curves. Figure 2 depicts the separation curve and its 95% confidence band discussed in Section 3.1. It is important to note that the separation curve and the confidence band are based on the difference of the transformed ROC curves, $G(R(u))$, which in our case is $\Phi^{-1}(R(u))$. The construction method utilized here is different from the one used in Li and Zhou (2008) which consists of point-wise confidence intervals. Using our method, we can identify, with a significance level of 0.05, the range of FPRs in which diagnostic test 2 is superior to diagnostic test 1 based on ROC curves. In this example, we found that test 2 is superior to test 1 when the FPR is between 0.078 and 0.513. This is an advantage over the LZ method in that it provides a global measure of sampling variability of estimated difference between two ROC curves.

6 Discussion

The clustered ROC data structure presents a challenge for methods that have been previously developed for other types of data. To the best of our knowledge, the PLS method is the first to generate smooth ROC curves for the clustered data. The simulation results show satisfactory performance of our proposed methods when comparing to the existing methods for clustered ROC data. The computer code implementing the proposed methods is included as supplementary material.

The correlation between the normal and abnormal observations within the same subject is the main reason for the difficulty to derive the explicit expression for the variance of estimated ROC curves. Such difficulty is also noted in Li and Zhou (2008). We have developed the large sample properties for the LS parameter estimators. The properties, especially the asymptotic covariance matrix of the estimated parameters, facilitate the computation of the test statistics for comparing two diagnostic tests. Also, with the theoretical results derived in this paper, we can construct the separation curve and its confidence band to visualize the difference between ROC curves for clustered data.

We chose 50 equally-space FPR points in (0,1) in this paper for estimating LS parameters. Based on our limited experience from the simulation studies, the choice of the number of equally spaced points has little effect on the parameter estimation when the number is sufficiently large (30 or larger). Unequally spaced FPR points may be a feasible choice when the test results are highly skewed and the FPR points are chosen based on the unique jump points from the normal observations. It would be interesting to investigate how equally spaced points and unequally spaced points affect the inference in future research.

The proposed methods have potential applications in evaluating diagnostic tests for eye disease as indicated in the example. It also has applications in tests for dental disease. Test results for diseased teeth and healthy teeth from the same subject are likely to be correlated,

and the proposed methods to account for this type of clustering will provide valid statistical tools for dental tests.

As pointed out by a referee, another possible method for estimating the ROC curve from the clustered data is to use a mixed model with fixed effects for diagnostic tests and random effects for subjects. This potential method may allow for more complex correlation structure than the proposed methods, and it is worthwhile to explore as a future topic.

Acknowledgments

The authors thank two referees and the associate editor for their constructive comments. The authors thank Professor Gang Li for providing the Glaucoma dataset. This research was supported by the Intramural Research Program of the National Institutes of Health and the U.S. Social Security Administration. We would like to thank the NIH Library Writing Center for manuscript editing assistance. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Qizhai Li is partially supported by National Natural Science of China, No 11371353, 61134013 and the Strategic Priority Research Program of the Chinese Academy of Sciences.

References

- Dorfman DD, Alf E. Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals-rating method data. *Journal of Mathematical Psychology*. 1969; 6:487–496.
- Hsieh F, Turnbull BW. Non- & semi-parametric estimation of the receiver operating characteristics (ROC) curve. *Annals of Statistics*. 1996; 24:25–40.
- Li G, Zhou K. A Unified Approach to Nonparametric Comparison of Receiver Operating Characteristic Curves for Longitudinal and Clustered Data. *Journal of the American Statistical Association*. 2008; 103:705–713. [PubMed: 20209021]
- Metz CE, Herman BA, Roe CA. Statistical Comparison of Two ROC-curve Estimates Obtained from Partially-paired Datasets. *Medical Decision Making*. 1998; 18:110–121. [PubMed: 9456215]
- Obuchowski NA. Nonparametric Analysis of Clustered ROC Curve Data. *Biometrics*. 1997; 53:567–578. [PubMed: 9192452]
- Qin G, Jin X, Zhou X-H. Non-parametric interval estimation for the partial area under the ROC curve. *Canadian Journal of Statistics*. 2011; 39:17–33.
- Qin G, Zhou X-H. Empirical Likelihood Inference for the Area under the ROC Curve. *Biometrics*. 2006; 62:613–622. [PubMed: 16918927]
- Tang L, Zhou XH. Semiparametric Inferential Procedures for Comparing Multivariate ROC Curves with Interaction Terms. *Statistica Sinica*. 2009; 19:1137–1145.
- Tang LL, Liu A, Schisterman EF, Zhou X-H, Liu CC-I. Homogeneity tests of clustered diagnostic markers with applications to the BioCycle Study. *Statistics in Medicine*. 2012; doi: 10.1002/sim.5391
- Tang LL, Zhou X-H. A Semiparametric Separation Curve Approach for Comparing Correlated ROC Data From Multiple Markers. *Journal of Computational and Graphical Statistics*. 2012; 21:662–676. [PubMed: 23074360]
- Tham Y-C, Li X, Wong TY, Quigley HA, Aung T, Cheng C-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014; 121:2081–2090. [PubMed: 24974815]
- Zhang, Z.; Pepe, MS. A Linear Regression Framework for Receiver Operating Characteristic(ROC) Curve Analysis. Tech rep, UW Biostatistics Working Paper Series, Working Paper 253. 2005. <http://www.bepress.com/uwbiostat/paper253>
- Zhou, X-H.; Obuchowski, NA.; McClish, DK. Statistical methods in diagnostic medicine. Vol. 712. John Wiley & Sons; 2011.

A Appendix

A.1 Proof of Result 1

Proof

For $\mathbf{u} = (u_{K_1}, u_{K_2}, \dots, u_{K_V})^T$, where u_{K_i} , $i = 1, 2, \dots, V$ is a L_T -dimension vector

$$\sqrt{n}(\hat{R}(\mathbf{u}) - R(\mathbf{u})) = \sqrt{n} \begin{pmatrix} \hat{R}_1(u_{K_1}) - R_1(u_{K_1}) \\ \hat{R}_2(u_{K_2}) - R_2(u_{K_2}) \\ \vdots \\ \hat{R}_V(u_{K_V}) - R_V(u_{K_V}) \end{pmatrix} = \sqrt{n} \begin{pmatrix} \hat{S}_{D,1}(\hat{S}_{\bar{D},1}^{-1}(u_{K_1})) - S_{D,1}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \hat{S}_{D,2}(\hat{S}_{\bar{D},2}^{-1}(u_{K_2})) - S_{D,2}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \hat{S}_{D,V}(\hat{S}_{\bar{D},V}^{-1}(u_{K_V})) - S_{D,V}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix}$$

where $\hat{S}_{D,v}(u) = \frac{1}{m_v} \sum_{i=1}^n \sum_{p=1}^{m_{vi}} I(X_{vip} > u)$ and

$\hat{S}_{\bar{D},v}(u) = \frac{1}{n_v} \sum_{i=1}^n \sum_{q=1}^{n_{vi}} I(Y_{viq} > u)$, $m_v = \sum_{i=1}^n m_{vi}$ and $n_v = \sum_{i=1}^n n_{vi}$ for $v = 1, 2, \dots, V$ and $u \in [0, 1]$. According to Lemma 1 of Li and Zhou (2008),

$$\sqrt{n} \begin{pmatrix} \hat{S}_{D,1}(u_{K_1}) - S_{D,1}(u_{K_1}) \\ \hat{S}_{D,2}(u_{K_2}) - S_{D,2}(u_{K_2}) \\ \vdots \\ \hat{S}_{D,V}(u_{K_V}) - S_{D,V}(u_{K_V}) \\ \hat{S}_{\bar{D},1}(u_{K_1}) - S_{\bar{D},1}(u_{K_1}) \\ \hat{S}_{\bar{D},2}(u_{K_2}) - S_{\bar{D},2}(u_{K_2}) \\ \vdots \\ \hat{S}_{\bar{D},V}(u_{K_V}) - S_{\bar{D},V}(u_{K_V}) \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^n V_i(\mathbf{u}),$$

$$\text{where } V_i(\mathbf{u}) = \begin{pmatrix} \frac{n}{m_1} \sum_{p=1}^{m_1} \{I(X_{1ip} > u_{K_1}) - S_{D,1}(u_{K_1})\} \\ \frac{n}{m_2} \sum_{p=1}^{m_2} \{I(X_{2ip} > u_{K_2}) - S_{D,2}(u_{K_2})\} \\ \vdots \\ \frac{n}{m_V} \sum_{p=1}^{m_V} \{I(X_{Vip} > u_{K_V}) - S_{D,V}(u_{K_V})\} \\ \frac{n}{n_1} \sum_{q=1}^{n_1} \{I(Y_{1iq} > u_{K_1}) - S_{\bar{D},1}(u_{K_1})\} \\ \frac{n}{n_2} \sum_{q=1}^{n_2} \{I(Y_{2iq} > u_{K_2}) - S_{\bar{D},2}(u_{K_2})\} \\ \vdots \\ \frac{n}{n_V} \sum_{q=1}^{n_V} \{I(Y_{Vi q} > u_{K_V}) - S_{\bar{D},V}(u_{K_V})\} \end{pmatrix}.$$

$V_i(\mathbf{u})$ are independent identically random vectors. Following the approach of Billingsley (1999) for empirical distribution process, it can be shown that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n V_i(\mathbf{u}) \xrightarrow{d} \mathbf{U}(\mathbf{u}) \text{ in } D(R)^{2V} \quad (1')$$

$\mathbf{U}(\mathbf{u})$ is a Gaussian process in $D(R)^{2^V}$, where

$$\mathbf{U}(u) = (\mathbf{U}_{S_{D,1}}(u_{K_1}), \mathbf{U}_{S_{D,2}}(u_{K_2}), \dots, \mathbf{U}_{S_{D,V}}(u_{K_V}), \mathbf{U}_{S_{\bar{D},1}}(u_{K_1}), \mathbf{U}_{S_{\bar{D},2}}(u_{K_2}), \dots, \mathbf{U}_{S_{\bar{D},V}}(u_{K_V}))^T,$$

the variance-covariance function is the limit of $\frac{1}{n} \sum_{i=1}^n \text{cov}(V_i(\mathbf{u}), V_i(\mathbf{u}))$.

Under (1') and the compact differentiability of the inverse function and the functional delta method (see, Andersen et al. 1993), we have

$$\sqrt{n} \begin{pmatrix} \hat{S}_{D,1}(u_{K_1}) - S_{D,1}(u_{K_1}) \\ \hat{S}_{D,2}(u_{K_2}) - S_{D,2}(u_{K_2}) \\ \vdots \\ \hat{S}_{D,V}(u_{K_V}) - S_{D,V}(u_{K_V}) \\ \hat{S}_{\bar{D},1}^{-1}(u_{K_1}) - S_{\bar{D},1}^{-1}(u_{K_1}) \\ \hat{S}_{\bar{D},2}^{-1}(u_{K_2}) - S_{\bar{D},2}^{-1}(u_{K_2}) \\ \vdots \\ \hat{S}_{\bar{D},V}^{-1}(u_{K_V}) - S_{\bar{D},V}^{-1}(u_{K_V}) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \mathbf{U}_{S_{D,1}}(u_{K_1}) \\ \mathbf{U}_{S_{D,2}}(u_{K_2}) \\ \vdots \\ \mathbf{U}_{S_{D,V}}(u_{K_V}) \\ -\frac{\mathbf{U}_{S_{\bar{D},1}}(S_{\bar{D},1}^{-1}(u_{K_1}))}{S'_{\bar{D},1}(S_{\bar{D},1}^{-1}(u_{K_1}))} \\ -\frac{\mathbf{U}_{S_{\bar{D},2}}(S_{\bar{D},2}^{-1}(u_{K_2}))}{S'_{\bar{D},2}(S_{\bar{D},2}^{-1}(u_{K_2}))} \\ \vdots \\ -\frac{\mathbf{U}_{S_{\bar{D},V}}(S_{\bar{D},V}^{-1}(u_{K_V}))}{S'_{\bar{D},V}(S_{\bar{D},V}^{-1}(u_{K_V}))} \end{pmatrix}. \quad (2)$$

According to the Lemma A.1 of Li et al (1996) and the functional delta method, we have

$$\sqrt{n} \begin{pmatrix} \hat{S}_{D,1}(\hat{S}_{\bar{D},1}^{-1}(u_{K_1})) - S_{D,1}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \hat{S}_{D,2}(\hat{S}_{\bar{D},2}^{-1}(u_{K_2})) - S_{D,2}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \hat{S}_{D,V}(\hat{S}_{\bar{D},V}^{-1}(u_{K_V})) - S_{D,V}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \mathbf{U}_{S_{D,1}}(S_{\bar{D},1}^{-1}(u_{K_1})) - r_1(u_{K_1}) \mathbf{U}_{S_{\bar{D},1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \mathbf{U}_{S_{D,2}}(S_{\bar{D},2}^{-1}(u_{K_2})) - r_2(u_{K_2}) \mathbf{U}_{S_{\bar{D},2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \mathbf{U}_{S_{D,V}}(S_{\bar{D},V}^{-1}(u_{K_V})) - r_V(u_{K_V}) \mathbf{U}_{S_{\bar{D},V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix} \quad (3)$$

where $\mathbf{U}_{S_{D,v}}, \mathbf{U}_{S_{\bar{D},v}}$ are Gaussian process, and $r_v(u_{K_v}) = \frac{S'_{D,v}(S_{\bar{D},v}^{-1}(u_{K_v}))}{S'_{\bar{D},v}(S_{\bar{D},v}^{-1}(u_{K_v}))}, v = 1, 2, \dots, V$. Define

$$\mathbf{W} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \begin{pmatrix} \frac{n}{m_1} \sum_{p=1}^{m_1} \{I(X_{1ip} > S_{\bar{D},1}^{-1}(u_{K_1})) - S_{D,1}(S_{\bar{D},1}^{-1}(u_{K_1}))\} \\ \frac{n}{m_2} \sum_{p=1}^{m_2} \{I(X_{2ip} > S_{\bar{D},2}^{-1}(u_{K_2})) - S_{D,2}(S_{\bar{D},2}^{-1}(u_{K_2}))\} \\ \vdots \\ \frac{n}{m_V} \sum_{p=1}^{m_V} \{I(X_{vip} > S_{\bar{D},V}^{-1}(u_{K_V})) - S_{D,V}(S_{\bar{D},V}^{-1}(u_{K_V}))\} \end{pmatrix} \right. \\ \left. - \begin{pmatrix} r_1(u_{K_1}) \frac{n}{n_1} \sum_{q=1}^{n_1} \{I(Y_{1iq} > S_{\bar{D},1}^{-1}(u_{K_1})) - u_{K_1}\} \\ r_2(u_{K_2}) \frac{n}{n_2} \sum_{q=1}^{n_2} \{I(Y_{2iq} > S_{\bar{D},2}^{-1}(u_{K_2})) - u_{K_2}\} \\ \vdots \\ r_V(u_{K_V}) \frac{n}{n_V} \sum_{q=1}^{n_V} \{I(Y_{viq} > S_{\bar{D},V}^{-1}(u_{K_V})) - u_{K_V}\} \end{pmatrix} \right\} \triangleq \frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{w}^{(1)} + \mathbf{w}^{(2)})$$

Not only we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{w}^{(1)} \xrightarrow{d} \begin{pmatrix} \mathbf{U}_{S_{D,1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \mathbf{U}_{S_{D,2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \mathbf{U}_{S_{D,V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix} \text{ and}$$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{w}^{(2)} \xrightarrow{d} \begin{pmatrix} -r_1(u_{K_1}) \mathbf{U}_{S_{\bar{D},1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ -r_2(u_{K_2}) \mathbf{U}_{S_{\bar{D},2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ -r_V(u_{K_V}) \mathbf{U}_{S_{\bar{D},V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix},$$

but also according to (1'), we have

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{w}^{(1)}, \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{w}^{(2)} \right)^\tau \xrightarrow{d} \begin{pmatrix} \mathbf{U}_{S_{D,1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \mathbf{U}_{S_{D,2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \mathbf{U}_{S_{D,V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \\ -r_1(u_{K_1}) \mathbf{U}_{S_{\bar{D},1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ -r_2(u_{K_2}) \mathbf{U}_{S_{\bar{D},2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ -r_V(u_{K_V}) \mathbf{U}_{S_{\bar{D},V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix},$$

then using the functional delta method, we can get

$$\mathbf{W} \xrightarrow{d} \begin{pmatrix} \mathbf{U}_{S_{D,1}}(S_{\bar{D},1}^{-1}(u_{K_1})) - r_1(u_{K_1}) \mathbf{U}_{S_{\bar{D},1}}(S_{\bar{D},1}^{-1}(u_{K_1})) \\ \mathbf{U}_{S_{D,2}}(S_{\bar{D},2}^{-1}(u_{K_2})) - r_2(u_{K_2}) \mathbf{U}_{S_{\bar{D},2}}(S_{\bar{D},2}^{-1}(u_{K_2})) \\ \vdots \\ \mathbf{U}_{S_{D,V}}(S_{\bar{D},V}^{-1}(u_{K_V})) - r_V(u_{K_V}) \mathbf{U}_{S_{\bar{D},V}}(S_{\bar{D},V}^{-1}(u_{K_V})) \end{pmatrix},$$

Then it follows from the multivariate Lindberg-Levy central limit theorem (serfling,1980) that $\hat{\mathbf{R}}(\mathbf{u})$ is multivariate normal, and the variance-covariance matrix of $\sqrt{n}\hat{\mathbf{R}}(\mathbf{u})$ is given by

$$\text{var}(\sqrt{n}\hat{\mathbf{R}}(\mathbf{u})) \xrightarrow{p} E(\mathbf{W}\mathbf{W}^T).$$

A.2 Proof of Result 2

Proof

Along with Result 1 and the functional delta method, it is obvious that when $n \rightarrow \infty$, $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ converges to a $2V$ -dimension and $(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}^*)$ converges to a $2(V-1)$ -dimension multivariate normal distribution with mean zero, respectively. As given earlier,

$$\mathbf{Y} = (G(\hat{R}_1(u_{K_1})), G(\hat{R}_2(u_{K_2})), \dots, G(\hat{R}_V(u_{K_V})))^T$$

and

$$\mathbf{Y}^* = (G(\hat{R}_2(u_{K_2})) - G(\hat{R}_1(u_{K_1})), \dots, G(\hat{R}_V(u_{K_V})) - G(\hat{R}_1(u_{K_1})))^T.$$

Denote the variance of $\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\theta}$ as $\Sigma_{\boldsymbol{\varepsilon}}$ and the variance of $\boldsymbol{\varepsilon}^* = \mathbf{Y}^* - \mathbf{X}^*\boldsymbol{\theta}^*$, because $\hat{\boldsymbol{\theta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ and $\hat{\boldsymbol{\theta}}^* = ((\mathbf{X}^*)^T(\mathbf{X}^*))^{-1}(\mathbf{X}^*)^T\mathbf{Y}^*$, then we obtain

$$\begin{aligned} \text{var}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) &= (\mathbf{X}^T \Sigma_{\boldsymbol{\varepsilon}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma_{\boldsymbol{\varepsilon}}^{-1} \text{var}(\boldsymbol{\varepsilon}) \Sigma_{\boldsymbol{\varepsilon}}^{-1} \mathbf{X} (\mathbf{X}^T \Sigma_{\boldsymbol{\varepsilon}}^{-1} \mathbf{X})^{-1} = (\mathbf{X}^T \Sigma_{\boldsymbol{\varepsilon}}^{-1} \mathbf{X})^{-1}, \\ \text{var}(\hat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}^*) &= ((\mathbf{X}^*)^T (\Sigma_{\boldsymbol{\varepsilon}}^*)^{-1} \mathbf{X}^*)^{-1} (\mathbf{X}^*)^T (\Sigma_{\boldsymbol{\varepsilon}}^*)^{-1} \text{var}(\boldsymbol{\varepsilon}^*) (\Sigma_{\boldsymbol{\varepsilon}}^*)^{-1} \mathbf{X}^* ((\mathbf{X}^*)^T (\Sigma_{\boldsymbol{\varepsilon}}^*)^{-1} \mathbf{X}^*)^{-1}. \end{aligned}$$

Figure 1.

Estimated ROC curves by PLS, SLS, and LZ methods for two tests. PLS method – solid lines; SLS method – long dashed line; LZ method – dotted lines.

Figure 2.

Estimated separation curve and its confidence band. The solid line is the separation curve, the shaded region is the 95% confidence band, and the dashed line is the horizontal line at $(u) = 0$.

Table 1

Bias and RMSE for Estimating the LS Parameters

	m=20						m=50						m=80					
	PLS			SLS			PLS			SLS			PLS			SLS		
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
α_1	0.00238	0.37559			0.00364	0.31092			0.00431	0.19725								
β_1	-0.02380	0.29845			-0.01578	0.24910			-0.00546	0.15239								
α_2	-0.03907	0.47982	-0.03879	0.52100	-0.00799	0.38587	-0.00680	0.42157	-0.00306	0.24956	-0.00410	0.27102						
β_2	-0.03372	0.39463	-0.03403	0.41940	-0.01405	0.31604	-0.01370	0.33731	-0.00204	0.20587	-0.00243	0.21862						
α_3	-0.02486	0.48059	-0.03562	0.52501	-0.02010	0.40002	-0.01851	0.43642	-0.00380	0.25489	-0.00245	0.27203						
β_3	-0.02358	0.38177	-0.03181	0.41019	-0.01884	0.33385	-0.01734	0.35539	-0.00750	0.20479	-0.00659	0.21532						
α_4	-0.01575	0.47351	-0.02146	0.50356	-0.00518	0.41813	-0.00750	0.44997	-0.00371	0.25112	-0.00578	0.27502						
β_4	-0.01732	0.38430	-0.02191	0.40309	-0.00946	0.33710	-0.01080	0.35729	-0.00481	0.20377	-0.00599	0.21752						
α_5	0.00007	0.48540	-0.00198	0.52212	-0.01241	0.40643	-0.00406	0.43380	0.00010	0.25578	-0.00051	0.27851						
β_5	-0.00835	0.38894	-0.00999	0.41137	-0.00975	0.32085	-0.00395	0.33646	-0.00226	0.20264	-0.00306	0.21599						
α_6	0.01297	0.49639	0.00827	0.53183	-0.00815	0.42582	-0.00328	0.45600	-0.00485	0.25994	-0.00540	0.28470						
β_6	-0.01038	0.38696	-0.01387	0.40939	-0.00877	0.33944	-0.00550	0.35749	-0.00742	0.20337	-0.00784	0.21808						
α_7	-0.01302	0.50037	-0.01514	0.54109	-0.00024	0.40141	0.00124	0.43441	0.00294	0.25870	-0.00024	0.28314						
β_7	-0.01923	0.40388	-0.02095	0.42794	-0.00213	0.32140	-0.00114	0.34051	-0.00151	0.20350	-0.00364	0.21823						
α_8	-0.00361	0.49985	-0.01024	0.53609	0.01172	0.41992	0.01630	0.44665	-0.01043	0.25758	-0.01262	0.27593						
β_8	-0.00956	0.39656	-0.01507	0.41910	-0.00069	0.32837	0.00236	0.34500	-0.01081	0.20566	-0.01234	0.21587						
α_9	-0.00956	0.49268	-0.01474	0.52100	-0.01213	0.41545	-0.01018	0.44544	-0.01555	0.26005	-0.01675	0.28105						
β_9	-0.01771	0.38315	-0.02217	0.40235	-0.01488	0.33452	-0.01358	0.35537	-0.00975	0.20563	-0.01079	0.21630						
α_{10}	0.00237	0.49230	-0.00106	0.53145	-0.01427	0.40654	-0.00823	0.43273	-0.00532	0.26061	-0.00612	0.27920						
β_{10}	-0.00467	0.38238	-0.00708	0.40572	-0.01804	0.33023	-0.01421	0.34515	-0.00301	0.20554	-0.00344	0.21636						

Table 2

Bias for Estimating the Partial AUC

	m=20			m=50			m=80		
	PLS	LZ	SLS	PLS	LZ	SLS	PLS	LZ	SLS
Test 1	0.0031	0.0036	0.0036	0.0023	0.0026	0.0026	0.0010	0.0013	0.0013
Test 2	0.0035	0.0038	0.0046	0.0030	0.0034	0.0038	0.0010	0.0012	0.0013
Test 3	0.0035	0.0040	0.0046	0.0026	0.0029	0.0034	0.0015	0.0016	0.0019
Test 4	0.0036	0.0042	0.0047	0.0028	0.0033	0.0035	0.0012	0.0014	0.0016
Test 5	0.0038	0.0044	0.0049	0.0023	0.0026	0.0031	0.0012	0.0015	0.0017
Test 6	0.0049	0.0055	0.0061	0.0024	0.0028	0.0033	0.0014	0.0016	0.0018
Test 7	0.0039	0.0046	0.0051	0.0023	0.0029	0.0032	0.0013	0.0016	0.0018
Test 8	0.0036	0.0043	0.0048	0.0030	0.0036	0.0039	0.0013	0.0016	0.0018
Test 9	0.0038	0.0045	0.0051	0.0026	0.0031	0.0035	0.0008	0.0011	0.0013
Test 10	0.0033	0.0040	0.0045	0.0028	0.0034	0.0037	0.0008	0.0011	0.0013

Table 3

RMSE for Estimating the Partial AUC

	m=20			m=50			m=80		
	PLS	LZ	SLS	PLS	LZ	SLS	PLS	LZ	SLS
Test 1	0.0140	0.0142	0.0142	0.0116	0.0118	0.0118	0.0073	0.0074	0.0074
Test 2	0.0142	0.0143	0.0150	0.0121	0.0123	0.0128	0.0069	0.0071	0.0074
Test 3	0.0148	0.0150	0.0158	0.0121	0.0124	0.0130	0.0075	0.0077	0.0081
Test 4	0.0151	0.0152	0.0159	0.0122	0.0125	0.0129	0.0073	0.0075	0.0078
Test 5	0.0147	0.0148	0.0156	0.0125	0.0128	0.0132	0.0073	0.0075	0.0078
Test 6	0.0154	0.0157	0.0164	0.0126	0.0129	0.0133	0.0072	0.0074	0.0076
Test 7	0.0159	0.0163	0.0171	0.0121	0.0124	0.0129	0.0077	0.0079	0.0083
Test 8	0.0154	0.0157	0.0166	0.0127	0.0129	0.0135	0.0079	0.0082	0.0084
Test 9	0.0149	0.0153	0.0162	0.0125	0.0129	0.0136	0.0075	0.0077	0.0081
Test 10	0.0156	0.0158	0.0165	0.0127	0.0130	0.0137	0.0078	0.0080	0.0084

Table 4

Bias for Estimating the AUC

	m=20			m=50			m=80		
	PLS	LZ	SLS	PLS	LZ	SLS	PLS	LZ	SLS
Test 1	-0.0150	0.0030	0.0030	-0.0074	0.0030	0.0030	-0.0017	0.0006	0.0006
Test 2	-0.0091	0.0048	-0.0096	-0.0052	0.0036	-0.0066	-0.0035	0.0009	-0.0054
Test 3	-0.0099	0.0057	-0.0097	-0.0067	0.0026	-0.0081	-0.0028	0.0010	-0.0048
Test 4	-0.0097	0.0055	-0.0097	-0.0049	0.0027	-0.0063	-0.0018	0.0016	-0.0038
Test 5	-0.0084	0.0050	-0.0099	-0.0076	0.0023	-0.0088	-0.0038	0.0012	-0.0055
Test 6	-0.0070	0.0057	-0.0076	-0.0044	0.0036	-0.0066	-0.0019	0.0019	-0.0035
Test 7	-0.0075	0.0050	-0.0083	-0.0055	0.0021	-0.0075	-0.0003	0.0020	-0.0016
Test 8	-0.0081	0.0051	-0.0084	-0.0038	0.0041	-0.0051	-0.0017	0.0018	-0.0028
Test 9	-0.0059	0.0052	-0.0071	-0.0048	0.0035	-0.0060	-0.0007	0.0022	-0.0018
Test10	-0.0051	0.0059	-0.0057	-0.0049	0.0028	-0.0065	-0.0015	0.0017	-0.0026

Table 5

RMSE for Estimating the AUC

	m=20			m=50			m=80		
	PLS	LZ	SLS	PLS	LZ	SLS	PLS	LZ	SLS
Test 1	0.0650	0.0363	0.0363	0.0513	0.0293	0.0293	0.0289	0.0186	0.0186
Test 2	0.0614	0.0352	0.0712	0.0497	0.0300	0.0585	0.0292	0.0177	0.0351
Test 3	0.0593	0.0345	0.0704	0.0456	0.0278	0.0538	0.0277	0.0175	0.0335
Test 4	0.0580	0.0348	0.0661	0.0444	0.0282	0.0542	0.0260	0.0172	0.0322
Test 5	0.0552	0.0355	0.0698	0.0468	0.0278	0.0558	0.0262	0.0165	0.0311
Test 6	0.0529	0.0331	0.0615	0.0433	0.0277	0.0529	0.0245	0.0166	0.0292
Test 7	0.0522	0.0334	0.0621	0.0394	0.0261	0.0489	0.0225	0.0156	0.0260
Test 8	0.0496	0.0310	0.0585	0.0365	0.0257	0.0455	0.0222	0.0154	0.0252
Test 9	0.0460	0.0314	0.0576	0.0365	0.0253	0.0467	0.0207	0.0155	0.0243
Test 10	0.0437	0.0298	0.0529	0.0365	0.0247	0.0442	0.0201	0.0149	0.0231