IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 2 Ver. III (Mar. - Apr. 2016), PP 77-86
www.iosrjournals.org

Detection of AUC and Confidence Interval Using Normal-Mixture ROC Curve

Sudesh Pundir[1], Azharuddin[2]
[1,2]*Department of Statistics, Pondicherry University, INDIA*

Abstract: *Receiver Operating Characteristic (ROC) Curve is one of the renowned statistical techniques to assess the accuracy of a diagnostic test to discriminate between healthy and diseased subjects. In this paper, two component Normal-Mixture ROC curve is studied and its properties are discussed to know the characteristics of the ROC Curve. Area under the Curve (AUC) of Normal-Mixture ROC Curve and Confidence Interval of AUC are also derived. The proposed model is validated by using the simulation studies*.

Keywords : *Normal-Mixture distribution, ROC Curve, AUC, Confidence Interval of Area Under the ROC Curve, MLE of AUC of ROC Curve via Expectation Maximization (EM) algorithm and Monte Carlo simulation*.

I. Introduction

The Mixture distribution arises where a statistical population contains two or more sub populations. It takes into account the heterogeneity of the population. For example, in medical diagnosis, if we have a set of patients suffering from cancer disease and we can also do further sub-classification again of the patients suffering from stomach cancer, blood cancer, throat cancer, neck cancer etc. So, in this case the mixture distribution is used for considering the heterogeneity of the population.

The finite mixture distributions are widely used in real life situations as a statistical model. First time, the Normal-Mixture distribution introduced is by Newcomb (1886) who showed the application of normal-mixture distributions as models for outliers. Pearson (1894) discussed the estimation of parameters of two component normal-mixture distribution by the method of moments. In his paper Pearson estimate the five parameter of the two component normal-mixture distribution. The first comprehensive monographs of the mixture distribution are Everitt and Hand (1981) and Titterington *et al.* (1985). They discussed the finite mixture distributions and the different estimation methods of mixture distributions. Mclachlan and Basford (1988) discussed different mixture distributions. Bohning *et al.* (1998) also discussed the application of mixture model in medicine. McLachlan and Peel (2000) discussed the finite mixture distributions. Schlattmann (2009) discussed the application of finite mixture distribution in medicine. Wirjanto and Dinghai (2009) discussed the application of normal mixture distribution in finance. They discussed the maximum likelihood estimate, moment generating function, switching regression model and stochastic regression model by using the normal-mixture distribution. Patrick *et al.* (2011) discussed estimation of finite mixture distribution by the method of moments. Dass and Seong (2012) discussed estimation of multivariate bi-normal mixture ROC Curve. Lee (2012) showed the Gaussian mixture model in pathological voices. Gonen (2013) discussed normal mixture of Receiver Operating Characteristic Curve. He also discussed that the diseased population has higher variability than can be explained by a single distribution. Hughes and Bhattacharya (2013) discussed the symmetric properties of Binormal and Bi-Gamma ROC Curves using the Kullback-Leibler Divergences.

The Receiver Operating Characteristic (ROC) Curve is discovered by the radar engineers during World war II in signal detection theory. Green and Swets (1966) discussed the ROC curve and Egan (1975) discussed the ROC Curve. Karzanowski and Hand (2009) also discussed the ROC Curve for continuous data. The Area under the ROC Curve generally denoted AUC and discussed by Green and Swets (1966). Bamber (1975) discussed the area above and below the ROC curve. Hanley and McNeil (1982) discussed use of the area under the ROC Curve. Bradley (1997) also discussed the use of the Area under the Curve. Pundir S and Amala R (2013) discussed various aspects of ROC Curve in continuous data.

Dorfman and Alf (1969) discussed the maximum likelihood estimation of parameters in signal detection theory and they also discussed how to find the confidence interval. Wirjanto and Xu (2009) discussed the maximum likelihood estimate of normal-mixture distribution. Kamaruzzaman *et al.* (2012) also discussed the maximum likelihood estimate the normal-mixture distribution. They also discussed the Expectation Maximization algorithm for estimating the parameters. For finding the ML estimate of normal-mixture distribution we also used the EM algorithm because the normal-mixture distribution have not closed form. Dempster *et al.* (1977) discussed the Maximum Likelihood Estimate of incomplete data via EM algorithm. Joshua (2013) discussed the EM algorithm in multivariate Gaussian mixture models.

Suppose X is a random variable which follows any continuous distribution having scores of patients into healthy and diseased group of individuals. For classifying the patients into healthy and diseased cases, let us

take a reference standard or a threshold value (t). Let the diseased and healthy individuals are denoted by 1 and 0. The individuals are regarded as diseased cases if the score (s) exceeds the threshold value (t) else they are regarded as healthy cases. To know the performance of classifier, we define four probabilities rates viz. True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR) and False Negative Rate (FNR). These are defined as follows:

True Positive Rate is the probability that an individual from diseased population is correctly classified i.e.,

$$\text{TPR} = p(s > t/1).$$

False Positive Rate is the probability that an individual from healthy population is misclassified i.e.,

$$\text{FPR} = p(s > t/0).$$

True Negative Rate is probability that an individual from healthy population is correctly classified i.e.,

$$\text{TNR} = p(s \le t/0).$$

False Negative Rate is the probability that an individual from Diseased population is misclassified i.e.,

$$\text{FNR} = p(s \le t/1).$$

Moreover the ROC Curve focuses only on the probabilities such that s > t in the two populations, then for all smaller values of t, we must have TPR=1 while FPR varies from 0 to 1 and for all larger values of t we must have FPR=0 while TPR varies from 1 to 0. The coordinates of the area under the curve lying between [0,0], [0,1] and [1,1] correspond to the ROC space. The ROC curve that falls near to [0,1] get maximum accuracy 1. The ROC Curve is a graph that shows True Positive Rate on the vertical axis and False Positive Rate on the horizontal axis, as the classification threshold t varies.

ROC Curve have been extensively used in the evaluation of diagnostic test. We have two groups of individuals namely diseased cases and healthy controls. ROC Curve is defined as

$$y(t) = 1 - G\left[F^{-1}(1 - x(t))\right], \quad 0 \le x(t) \le 1 \tag{1.1}$$

where x (t) denote the 1-specificity or False Positive Rate (FPR) and y(t) denote the sensitivity or True Positive Rate (TPR). Specificity is also defined as the proportion of patients without disease who test negative and Sensitivity is defined as the proportion of patients with disease who test positive. G and F denote the distribution function of the healthy cases and diseased cases. The performance of the curve is measured by the area under the ROC Curve. The area under a ROC Curve assesses the overall ability of the test to discriminate between healthy cases and diseased cases. The mathematical definition of AUC is

$$\mathbf{AUC} = \mathbf{P(y > x)} = \int_0^1 \mathbf{y(t)dt} \tag{1.2}$$

A test having an area of 0.5 is a worthless test and a perfect test has an area of 1.00.

The ROC Curve must satisfy the following properties.

1. ROC Curve and AUC is unaltered with respect to monotone increasing transformation of the test scores.
2. The test values of X are smaller than Y.
3. ROC Curve is monotonically increasing function, i.e. $\frac{dy(t)}{dx(t)} > 0$
4. ROC Curve is said to be convex if $\frac{d^2 y(t)}{dx^2(t)} > 0$ and concave if $\frac{d^2 y(t)}{dx^2(t)} < 0$.
5. The slope of ROC Curve at any operating point is equal to the ratio of PDF of diseased to PDF of healthy at a particular cut-off point t is given by

$$slope = \frac{p(t)}{q(t)} \tag{1.3}$$

6. For symmetric ROC Curve, if p(x) and q(x) are the two continuous probability distribution, the Kullback-Leibler divergence is a non-symmetric measure of two probability distributions of p(x) and q(x). Let p(x) denote the healthy distribution and q(x) denote the diseased distribution then KL(p,q) denote the Kullback-Leibler (K-L) divergence between the distributions of diseased and healthy group with p(x) as the comparison distribution and q(x) as the comparison reference. The

$$KL(p,q) = \int_D p(x) \ln\left(\frac{p(x)}{q(x)}\right) dx \tag{1.4}$$

Similarly, let KL(q,p) denote the K-L divergence between the distribution of healthy and diseased group with q(x) as the comparison distribution and p(x) as the reference distribution, then

$$KL(q,p)=\int_D q(x)\ln\left(\frac{q(x)}{p(x)}\right)dx \tag{1.5}$$

where D is the common range of p and q. It should be noted that if KL(p,q) and KL(q,p) are positive and KL(p,q) = KL(q,p) = 0, if and only if p(x) = q(x). these two measure gives the asymmetry of ROC Curve about the diagonal. If KL(p,q) < KL(q,p), then the ROC Curve is said to be True Positive Rate asymmetric and if KL(p,q) > KL(q,p), then the ROC Curve is said to be True Negative Rate asymmetric.

In this paper, there are six sections. In Section 2, we discussed the Normal Mixture ROC (NMROC) Curve and its properties are discussed to know the behavior of the normal-mixture ROC Curve. To know the accuracy of ROC Curve, Area under the normal-mixture ROC curve is also found. In Section 3, we found the maximum likelihood estimator of AUC of normal-mixture ROC curve. In Section 4, confidence interval of Area Under the NMROC Curve is derived. We discussed the simulation studies by taking a numerical example for the validation of the model in Section 5. Section 6 consists of the conclusion.

II. Normal Mixture ROC Curve

Let X be a random variable which follows normal-mixture distribution and it comes from healthy controls. It has probability density function

$$f(x)=pf\left(x_i,\mu_{10},\sigma_{10}^2\right)+(1-p)f\left(x_i,\mu_{20},\sigma_{20}^2\right) \tag{2.1}$$

where, μ_{10}, μ_{20}, σ_{10}^2, and σ_{20}^2 are the means and variances of the normal mixture distribution. First subscript in parameter shows that the it is coming from 1st and 2nd normal density and second subscript 0 shows that it is coming from healthy controls. p and 1-p are the weights of the mixture distribution. The sum of weight should be equal to 1 in mixture distribution.

Similarly, Y be a random variable which follows normal-mixture distribution and it comes from diseased cases. It has probability density function

$$g(y)=pf\left(y_i,\mu_{11},\sigma_{11}^2\right)+(1-p)f\left(y_i,\mu_{21},\sigma_{21}^2\right) \tag{2.2}$$

where μ_{11}, μ_{21}, σ_{11}^2, and σ_{21}^2 are the means and variances of normal mixture distribution. First subscript shows that it is coming from 1st and 2nd normal density and second subscript 1 means that it is coming from diseased case. p and 1-p are the weights of the mixture distribution.

The two component normal-mixture ROC (NMROC) curve is defined as

$$y(t)=p\Phi\left(\frac{\mu_{11}}{\sigma_{11}}+\frac{1}{\sigma_{11}}\Phi^{-1}(x(t))\right)+(1-p)\Phi\left(\frac{\mu_{21}}{\sigma_{21}}+\frac{1}{\sigma_{21}}\Phi^{-1}(x(t))\right) \tag{2.3}$$

where, Φ is the cdf of normal mixture distribution and

$$x(t)=p\Phi\left(\frac{\mu_{10}-t}{\sigma_{10}}\right)+(1-p)\Phi\left(\frac{\mu_{20}-t}{\sigma_{20}}\right),\ -\infty<t<\infty,\ -\infty<\mu_{ij}<\infty,\ \sigma_{ij}>0,\ i=1,2,\ j=1,0.$$

Before plotting ROC curve, one should check the following assumptions

1. The mean of the healthy controls which follows normal mixture distribution should be greater than the mean of the disease cases which also follows normal mixture distribution in normal-mixture ROC model.
2. If $\mu_{10}>\mu_{20}$, $\sigma_{10}^2>\sigma_{20}^2$, and $\mu_{11}>\mu_{21}$, $\sigma_{11}^2>\sigma_{21}^2$, higher the Area under the NMROC curve.
3. All properties of NMROC Curve should be satisfied.
4. Attach more weights to the density function which have high mean and high variance.

Properties

1. NMROC Curve is monotonically increasing function.

Proof : A function is said to be monotonically increasing function, if the first derivative of the function is positive. Differentiating (2.3) with respect to x, we get

$$\frac{dy(t)}{dx(t)}=\frac{1}{\phi\left[\Phi^{-1}(x(t))\right]}\left[\frac{p}{\sigma_{11}}\Phi\left(\frac{\mu_{11}}{\sigma_{11}}+\frac{1}{\sigma_{11}}\Phi^{-1}(x(t))\right)+\frac{(1-p)}{\sigma_{21}}\Phi\left(\frac{\mu_{21}}{\sigma_{21}}+\frac{1}{\sigma_{21}}\Phi^{-1}(x(t))\right)\right]>0$$

$p>0$, $-\infty<x<\infty$, $-\infty<\mu_{ij}<\infty$ and $\sigma_{ij}>0$ i=1,2. (2.4)

where ϕ is the density function and Φ is the distribution function of normal-mixture distribution. Hence, NMROC Curve is monotonically increasing function.

2. NMROC Curve is concave.

Proof : A function is said to be concave if the second derivative is negative. Differentiating (2.3) with respect to 'x' i.e.

$$\frac{d^2y(t)}{dx^2(t)} = \frac{1}{\left[\varphi\left(\Phi^{-1}(x(t))\right)\right]^2}\left[\begin{array}{l}\left(\frac{p}{\sigma_{11}^2}\Phi\left(\frac{\mu_{11}}{\sigma_{11}}+\frac{1}{\sigma_{11}}\Phi^{-1}(x(t))\right)+\frac{(1-p)}{\sigma_{21}^2}\Phi\left(\frac{\mu_{21}}{\sigma_{21}}+\frac{1}{\sigma_{21}}\Phi^{-1}(x(t))\right)\right)\\ -\Phi^{-1}(x(t))\left(\frac{p}{\sigma_{11}}\Phi\left(\frac{\mu_{11}}{\sigma_{11}}+\frac{1}{\sigma_{11}}\Phi^{-1}(x(t))\right)+\frac{(1-p)}{\sigma_{21}}\Phi\left(\frac{\mu_{21}}{\sigma_{21}}+\frac{1}{\sigma_{21}}\Phi^{-1}(x(t))\right)\right)\end{array}\right] < 0$$

$p > 0$, $-\infty < x < \infty$, $-\infty < \mu_{ij} < \infty$ and $\sigma_{ij} > 0$. i=1,2. (2.5)

3. The slope of the NMROC Curve at the threshold t is given by

$$\frac{dy(t)}{dx(t)} = \frac{\sigma_{10}\sigma_{20}\left[p\sigma_{21}\exp\left\{-\frac{(t-\mu_{11})^2}{2\sigma_{11}^2}\right\}+(1-p)\sigma_{11}\exp\left\{-\frac{(t-\mu_{21})^2}{2\sigma_{21}^2}\right\}\right]}{\sigma_{11}\sigma_{21}\left[p\sigma_{20}\exp\left\{-\frac{(t-\mu_{11})^2}{2\sigma_{10}^2}\right\}+(1-p)\sigma_{10}\exp\left\{-\frac{(t-\mu_{20})^2}{2\sigma_{20}^2}\right\}\right]} \quad (2.6)$$

4. NMROC curve is invariant with respect to monotonically increasing transformation of the test scores.

5. The NMROC Curve is TNR asymmetric.

Proof: The K-L divergence between the distribution of diseased and healthy group with p(x) as the comparison distribution and q(x) as the reference distribution is given as

$$KL(p,q) = \left(\begin{array}{l}\ln\sigma_{10}+\ln\sigma_{20}+\ln\left[p\sigma_{21}\exp\left\{-\frac{(x-\mu_{11})^2}{2\sigma_{11}^2}\right\}+(1-p)\sigma_{11}\exp\left\{-\frac{(x-\mu_{21})^2}{2\sigma_{21}^2}\right\}\right]-\ln\sigma_{11}-\ln\sigma_{21}-\\ \ln\left[p\sigma_{20}\exp\left\{-\frac{(x-\mu_{10})^2}{2\sigma_{10}^2}\right\}+(1-p)\sigma_{10}\exp\left\{-\frac{(x-\mu_{20})^2}{2\sigma_{20}^2}\right\}\right]\end{array}\right)\left(\frac{p}{\sigma_{11}\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_{11})^2}{2\sigma_{11}^2}\right\}+\frac{(1-p)}{\sigma_{21}\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_{21})^2}{2\sigma_{21}^2}\right\}\right)$$

(2.7)

Similarly, the K-L divergence between the distribution of healthy and diseased group with q(x) as the comparison distribution and p(x) as the reference distribution has been given as

$$KL(q,p) = \left(\begin{array}{l}\ln\sigma_{21}+\ln\sigma_{11}+\ln\left[p\sigma_{20}\exp\left\{-\frac{(x-\mu_{10})^2}{2\sigma_{10}^2}\right\}+(1-p)\sigma_{10}\exp\left\{-\frac{(x-\mu_{20})^2}{2\sigma_{20}^2}\right\}\right]-\ln\sigma_{10}-\ln\sigma_{20}-\\ \ln\left[p\sigma_{21}\exp\left\{-\frac{(x-\mu_{11})^2}{2\sigma_{11}^2}\right\}+(1-p)\sigma_{11}\exp\left\{-\frac{(x-\mu_{21})^2}{2\sigma_{21}^2}\right\}\right]\end{array}\right)\left(\frac{p}{\sigma_{10}\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_{10})^2}{2\sigma_{10}^2}\right\}+\frac{(1-p)}{\sigma_{20}\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_{20})^2}{2\sigma_{20}^2}\right\}\right)$$

(2.8)

It is found that KL(q,p) > KL(p,q). Hence, NMROC Curve is TNR asymmetric.

Optimal cut-off value

In medical diagnosis, the optimal cut-off value, i.e., t, decides about the condition of patient whether he is healthy or have disease. First time, it is introduced by the Fluss *et al*. (2005). He applied the optimal cut-off value in the youden index which is obtained by taking the maximum difference between F(t) and G(t) where F(t) represents the distribution function of healthy controls and G(t) represents distribution function of disease cases. The optimal cut-off value of NMROC Curve is given as

$$t = \max_t\{F(t)-G(t)\}$$

$$= p\max_t\left(\frac{\left(2\mu_{10}\sigma_{11}^2-2\mu_{11}\sigma_{10}^2\right)\pm 2\sigma_{11}\sigma_{10}\sqrt{\left(\mu_{11}^2-\mu_{10}^2\right)+\left(\sigma_{11}^2+\sigma_{10}^2\right)\ln\left(\frac{\sigma_{11}}{\sigma_{10}}\right)}}{2\left(\sigma_{11}^2-\sigma_{10}^2\right)}\right)+(1-p)\max_t\left(\frac{\left(2\mu_{20}\sigma_{21}^2-2\mu_{21}\sigma_{20}^2\right)\pm 2\sigma_{21}\sigma_{20}\sqrt{\left(\mu_{21}^2-\mu_{20}^2\right)+\left(\sigma_{21}^2+\sigma_{20}^2\right)\ln\left(\frac{\sigma_{21}}{\sigma_{20}}\right)}}{2\left(\sigma_{21}^2-\sigma_{20}^2\right)}\right)$$

(2.9)

The AUC of two component NMROC curve is derived as

$$\mathrm{AUC} = \mathrm{p}\Phi\left(\frac{\mu_{11}-\mu_{10}}{\sqrt{\sigma_{11}^2+\sigma_{10}^2}}\right)+(1-\mathrm{p})\Phi\left(\frac{\mu_{21}-\mu_{20}}{\sqrt{\sigma_{21}^2+\sigma_{20}^2}}\right).$$

(2.10)

III. Maximum Likelihood Estimate of parameters of NMROC curve via EM algorithm

For assessing the diagnostic accuracy, there are three approaches of estimating the ROC Curve- Parametric, non-parametric and semi-parametric approaches. Here, we discuss the parametric approach for estimating the ROC Curve. The probability density function of the normal-mixture distribution is

$$f(x)=\sum_{j=1}^{k} p_j N\left(\mu_j,\sigma_j^2\right), \quad \sum_{j=1}^{k} p_j = 1, \ 0 \le p_j \le 1.$$

and the CDF of the normal-mixture distribution is

$$F(x)=\sum_{j=1}^{k} p_j \Phi\left(\frac{x-\mu_j}{\sigma_j}\right)$$

For k=2, the pdf of two component normal-mixture distribution is given as

$$\mathrm{f}\left(\mathrm{x},\mathrm{p},\mu_1,\mu_2,\sigma_1^2,\sigma_2^2\right)=\frac{\mathrm{p}}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(\mathrm{x}-\mu_1)^2}{2\sigma_1^2}\right]+\frac{(1-\mathrm{p})}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(\mathrm{x}-\mu_2)^2}{2\sigma_2^2}\right]. \tag{3.1}$$

The likelihood function of the two component normal-mixture distribution is

$$\mathrm{L}=\prod_{\mathrm{i}=1}^{\mathrm{n}}\frac{\mathrm{p}}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(\mathrm{x_i}-\mu_1)^2}{2\sigma_1^2}\right]+\frac{(1-\mathrm{p})}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(\mathrm{x_i}-\mu_2)^2}{2\sigma_2^2}\right]. \tag{3.2}$$

The log-likelihood function of the (3.2) is as follows

$$\ln \mathrm{L}=\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathrm{n}\left(\frac{\mathrm{p}}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(\mathrm{x_i}-\mu_1)^2}{2\sigma_1^2}\right]+\frac{(1-\mathrm{p})}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(\mathrm{x_i}-\mu_2)^2}{2\sigma_2^2}\right]\right). \tag{3.3}$$

For maximizing the log-likelihood function, differentiating (3.3) with respect to p, $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ and equate them to zero, we get

$$\frac{\partial \ln L}{\partial p}=\sum_{i=1}^{n}\left(\frac{1}{\alpha}\left(\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x_i-\mu_1)^2}{2\sigma_1^2}\right]-\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_2^2}\right]\right)\right)=0 \tag{3.4}$$

$$\frac{\partial \ln L}{\partial \mu_1}=\sum_{i=1}^{n}\left(\frac{1}{\alpha}\frac{p}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_1^2}\right]\left(\frac{x_i-\mu_1}{\sigma_1^2}\right)\right)=0 \tag{3.5}$$

$$\frac{\partial \ln L}{\partial \mu_2}=\sum_{i=1}^{n}\left(\frac{1}{\alpha}\frac{(1-p)}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_2^2}\right]\left(\frac{x_i-\mu_2}{\sigma_2^2}\right)\right)=0 \tag{3.6}$$

$$\frac{\partial \ln L}{\partial \sigma_1^2}=\sum_{i=1}^{n}\left(\frac{1}{\alpha}\left(-\frac{p}{2\sqrt{2\pi}}\left(\sigma_1^2\right)^{-\frac{3}{2}}\exp\left[-\frac{(x_i-\mu_1)^2}{2\sigma_1^2}\right]+\frac{p}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x_i-\mu_1)^2}{2\sigma_1^2}\right]\left(\frac{(x_i-\mu_1)^2}{2\left(\sigma_1^2\right)^2}\right)\right)\right)=0 \tag{3.7}$$

$$\frac{\partial \ln L}{\partial \sigma_2^2}=\sum_{i=1}^{n}\left(\frac{1}{\alpha}\left(-\frac{(1-p)}{2\sqrt{2\pi}}\left(\sigma_2^2\right)^{-\frac{3}{2}}\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_2^2}\right]+\frac{(1-p)}{\sqrt{2\pi}\sigma_2}\right)\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_2^2}\right]\left(\frac{x_i-\mu_2}{2\left(\sigma_2^2\right)^2}\right)\right)=0 \tag{3.8}$$

where $\alpha = \frac{p}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x_i-\mu_1)^2}{2\sigma_1^2}\right] + \frac{(1-p)}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{(x_i-\mu_2)^2}{2\sigma_2^2}\right].$

We observe that the likelihood equation of two component normal-mixture distribution is not in a closed form. If the equation is not found in a closed form, we will use the Expectation-Maximization (EM) algorithm to estimate the parameters. In this method, there are two steps to get the estimates using EM algorithm.

(i) Take expectation of the likelihood function and
(ii) Maximize the log likelihood function.

Suppose X is a mixture data with N observations, the likelihood of the data assuming that x_i are independently distributed is given as

$$f(X \mid \Theta) = L(\Theta \mid X) = \prod_{i=1}^{N}\sum_{k=1}^{K} p_k f(x_i \mid \vartheta_k) \tag{3.9}$$

The problem of mixture estimation from data X can be formulated as to find the set of parameters Θ that gives the maximum likelihood estimate (MLE) solution

$$\Theta^* = \arg\max_{\Theta} L(\Theta \mid X). \tag{3.10}$$

The summation inside the product in (3.9) checks the possibility of analytical solutions. One alternative is to maximize the complete likelihood in an expectation-maximization (EM) algorithm. So, after maximization of the of (3.4)-(3.8), we get the estimates as follows

$$\hat{\mu}_{10} = \frac{\sum_{i-1}^{n_{10}}(1-\Delta_{i0})x_i}{\sum_{i-1}^{n_{10}}(1-\Delta_{i0})},\ \hat{\mu}_{20} = \frac{\sum_{i=1}^{n_{20}}\Delta_{i0}x_i}{\sum_{i=1}^{n_{20}}\Delta_{i0}},\ \hat{\mu}_{11} = \frac{\sum_{i=1}^{m_{11}}(1-\omega_{i1})y_i}{\sum_{i=1}^{m_{11}}(1-\omega_{i1})},$$

$$\hat{\mu}_{21} = \frac{\sum_{i=1}^{m_{21}}\omega_{i1}y_i}{\sum_{i=1}^{m_{21}}\omega_{i1}},\ \hat{\sigma}_{10}^2 = \frac{\sum_{i=1}^{n_{10}}(1-\Delta_{i0})(x_i-\mu_{10})^2}{\sum_{i=1}^{n_{10}}(1-\Delta_{i0})},\ \hat{\sigma}_{20}^2 = \frac{\sum_{i=1}^{n_{20}}\Delta_{i0}(x_i-\mu_{20})^2}{\sum_{i=1}^{n_{20}}\Delta_{i0}}$$

$$\hat{\sigma}_{11}^2 = \frac{\sum_{i=1}^{m_{11}}(1-\omega_{i1})(y_i-\mu_{11})^2}{\sum_{i=1}^{m_{11}}(1-\omega_{i1})} \quad \text{and} \quad \hat{\sigma}_{21}^2 = \frac{\sum_{i=1}^{m_{21}}\omega_{i1}(y_i-\mu_{21})^2}{\sum_{i=1}^{m_{21}}\omega_{i1}}. \tag{3.11}$$

For healthy control, we used weight $p = \sum_{i=1}^{n}\frac{\Delta_i}{n}$ and for diseased cases, the weight is $p = \sum_{i=1}^{m}\frac{\omega_i}{m}$ where

$$\Delta_{i0} = \frac{p\phi(x_i;\mu_{20},\sigma_{20}^2)}{p\phi_1(x_i;\mu_{10},\sigma_{10}^2)+(1-p)\phi_2(x_i;\mu_{20},\sigma_{20}^2)} \text{ and } \omega_{i1} = \frac{p\phi(y_i;\mu_{21},\sigma_{21}^2)}{p\phi_1(y_i;\mu_{11},\sigma_{11}^2)+(1-p)\phi_2(y_i;\mu_{21},\sigma_{21}^2)}.$$

Substituting maximum likelihood estimates from (3.11) in (2.10), we can get the estimate of AUC.

IV. Confidence Interval of AUC of NMROC Curve

The estimated AUC of two component NMROC is given as

$$\hat{AUC} = p\Phi\left(\frac{\hat{\mu}_{11}-\hat{\mu}_{10}}{\sqrt{\hat{\sigma}_{11}^2+\hat{\sigma}_{10}^2}}\right) + (1-p)\Phi\left(\frac{\hat{\mu}_{21}-\hat{\mu}_{20}}{\sqrt{\hat{\sigma}_{21}^2+\hat{\sigma}_{20}^2}}\right) \tag{4.1}$$

$$= p\Phi(\hat{\delta}_1) + (1-p)\Phi(\hat{\delta}_2)$$

where $\hat{\delta}_1 = \dfrac{\hat{\mu}_{11} - \hat{\mu}_{10}}{\sqrt{\hat{\sigma}_{11}^2 + \hat{\sigma}_{10}^2}}$ and $\hat{\delta}_2 = \dfrac{\hat{\mu}_{21} - \hat{\mu}_{20}}{\sqrt{\hat{\sigma}_{21}^2 + \hat{\sigma}_{20}^2}}$ (4.2)

The confidence interval of AUC of NMROC Curve cannot be found directly because it is a mixture of two component CDF function. First we will find the variance of $\hat{\delta}_1$ and $\hat{\delta}_2$ because $\Phi\left(\hat{\delta}_1\right)$ and $\Phi\left(\hat{\delta}_2\right)$ are monotonically increasing functions of $\hat{\delta}_1$ and $\hat{\delta}_2$.

$$\begin{aligned}\mathrm{V}(\hat{\delta}_1) = &\left(\frac{\partial\delta_1}{\partial\mu_{1y}}\right)^2 \mathrm{V}(\hat{\mu}_{11}) + \left(\frac{\partial\delta_1}{\partial\mu_{1x}}\right)^2 \mathrm{V}(\hat{\mu}_{1x}) + \left(\frac{\partial\delta_1}{\partial\sigma_{1y}}\right)^2 \mathrm{V}(\hat{\sigma}_{1y}^2) + \left(\frac{\partial\delta_1}{\partial\sigma_{1x}}\right)^2 \mathrm{V}(\hat{\sigma}_{1x}^2) + \\ &2C\,\mathrm{ov}(\hat{\mu}_{1y}, \hat{\mu}_{1x})\left(\frac{\partial\delta_1}{\partial\mu_{1y}}\right)\left(\frac{\partial\delta_1}{\partial\mu_{1x}}\right) + 2C\,\mathrm{ov}(\hat{\mu}_{1y}, \hat{\sigma}_{1y}^2)\left(\frac{\partial\delta_1}{\partial\mu_{1y}}\right)\left(\frac{\partial\delta_1}{\partial\sigma_{1y}^2}\right) + 2C\,\mathrm{ov}(\hat{\mu}_{1y}, \hat{\sigma}_{1x}^2) \\ &\left(\frac{\partial\delta_1}{\partial\mu_{1y}}\right)\left(\frac{\partial\delta_1}{\partial\sigma_{1x}^2}\right) + 2C\,\mathrm{ov}(\hat{\mu}_{1x}, \hat{\sigma}_{1x}^2)\left(\frac{\partial\delta_1}{\partial\mu_{1x}}\right)\left(\frac{\partial\delta_1}{\partial\sigma_{1x}^2}\right) + 2C\,\mathrm{ov}(\hat{\sigma}_{1x}^2, \hat{\sigma}_{1y}^2)\left(\frac{\partial\delta_1}{\partial\sigma_{1x}^2}\right)\left(\frac{\partial\delta_1}{\partial\sigma_{1y}^2}\right).\end{aligned} \tag{4.3}$$

$$\begin{aligned}\mathrm{V}(\hat{\delta}_2) = &\left(\frac{\partial\delta_2}{\partial\mu_{2y}}\right)^2 \mathrm{V}(\hat{\mu}_{2y}) + \left(\frac{\partial\delta_2}{\partial\mu_{2x}}\right)^2 \mathrm{V}(\hat{\mu}_{2x}) + \left(\frac{\partial\delta_2}{\partial\sigma_{2y}^2}\right)^2 \mathrm{V}(\hat{\sigma}_{2y}^2) + \left(\frac{\partial\delta_2}{\partial\sigma_{2x}^2}\right)^2 \mathrm{V}(\hat{\sigma}_{2x}^2) + \\ &2\mathrm{C\,ov}(\hat{\mu}_{2y}, \hat{\mu}_{2x})\left(\frac{\partial\delta_2}{\partial\mu_{2y}}\right)\left(\frac{\partial\delta_2}{\partial\mu_{2x}}\right) + 2\mathrm{C\,ov}(\hat{\mu}_{2y}, \hat{\sigma}_{2y}^2)\left(\frac{\partial\delta_2}{\partial\mu_{2y}}\right)\left(\frac{\partial\delta_2}{\partial\sigma_{2y}^2}\right) + 2\mathrm{C\,ov}(\hat{\mu}_{2y}, \hat{\sigma}_{2x}^2) \\ &\left(\frac{\partial\delta_2}{\partial\mu_{2y}}\right)\left(\frac{\partial\delta_2}{\partial\sigma_{2x}^2}\right) + 2\mathrm{C\,ov}(\hat{\mu}_{2x}, \hat{\sigma}_{2x}^2)\left(\frac{\partial\delta_2}{\partial\mu_{2x}}\right)\left(\frac{\partial\delta_2}{\partial\sigma_{2x}^2}\right) + 2\mathrm{C\,ov}(\hat{\sigma}_{2y}^2, \hat{\sigma}_{2x}^2)\left(\frac{\partial\delta_2}{\partial\sigma_{2y}^2}\right)\left(\frac{\partial\delta_2}{\partial\sigma_{2x}^2}\right)\end{aligned} \tag{4.4}$$

Now, on partially differentiating δ_1 and δ_2 with respect to μ_{1y}, μ_{1x}, σ_{1y}^2, σ_{1x}^2 an μ_{2y}, μ_{2x}, σ_{2y}^2, σ_{2x}^2 respectively, we get

$$\begin{aligned}&\frac{\partial\delta_1}{\partial\mu_{1y}} = \frac{1}{(\sigma_{1y}^2 + \sigma_{1x}^2)^{1/2}}, \qquad \frac{\partial\delta_1}{\partial\mu_{1x}} = -\frac{1}{(\sigma_{1y}^2 + \sigma_{1x}^2)^{1/2}}, \\ &\frac{\partial\delta_1}{\partial\sigma_{1y}^2} = -\frac{(\mu_{1y} - \mu_{1x})}{2(\sigma_{1y}^2 + \sigma_{1x}^2)^{3/2}}, \qquad \frac{\partial\delta_1}{\partial\sigma_{1x}^2} = -\frac{(\mu_{1y} - \mu_{1x})}{2(\sigma_{1y}^2 + \sigma_{1x}^2)^{3/2}}, \\ &\frac{\partial\delta_2}{\partial\mu_{2y}} = \frac{1}{(\sigma_{2y}^2 + \sigma_{2x}^2)^{1/2}}, \qquad \frac{\partial\delta_2}{\partial\mu_{2x}} = -\frac{1}{(\sigma_{2y}^2 + \sigma_{2x}^2)^{1/2}}, \\ &\frac{\partial\delta_2}{\partial\sigma_{2y}^2} = -\frac{(\mu_{2y} - \mu_{2x})}{2(\sigma_{2y}^2 + \sigma_{2x}^2)^{3/2}}, \qquad \frac{\partial\delta_2}{\partial\sigma_{2x}^2} = -\frac{(\mu_{2y} - \mu_{2x})}{2(\sigma_{2y}^2 + \sigma_{2x}^2)^{3/2}}.\end{aligned} \tag{4.5}$$

The estimated variance of μ_{1y}, μ_{1x}, σ_{1y}^2 and σ_{1x}^2 are given as

$$\begin{aligned}&\mathrm{V}(\hat{\mu}_{1y}) = \frac{\sigma_{1y}^2}{n_{1y}}, \qquad \mathrm{V}(\hat{\mu}_{1x}) = \frac{\sigma_{1x}^2}{n_{1x}}, \qquad \mathrm{V}(\hat{\sigma}_{1y}^2) = \frac{2\sigma_{1y}^2}{n_{1y} - 1}, \qquad \mathrm{V}(\hat{\sigma}_{1x}^2) = \frac{2\sigma_{1x}^2}{n_{1x} - 1} \\ &\mathrm{V}(\hat{\mu}_{2y}) = \frac{\sigma_{2y}^2}{n_{2y}}, \mathrm{V}(\hat{\mu}_{2x}) = \frac{\sigma_{2x}^2}{n_{2x}}, \qquad \mathrm{V}(\hat{\sigma}_{2y}^2) = \frac{\sigma_{2y}^2}{n_{2y} - 1}, \qquad \mathrm{V}(\hat{\sigma}_{2x}^2) = \frac{\sigma_{2x}^2}{n_{2x} - 1}\end{aligned} \tag{4.6}$$

The covariance terms are all zero due to identically independently distributed random variables. Substituting all the expressions from (4.5) and (4.6) in (4.3) and (4.4), we get

$$V(\hat{\delta}_1) = \frac{1}{(\sigma_{1y}^2 + \sigma_{1x}^2)}\left[\frac{\sigma_{1y}^2}{n_{1y}} + \frac{\sigma_{1x}^2}{n_{1x}}\right] + \frac{(\mu_{1y} - \mu_{1x})^2}{2(\sigma_{1y}^2 + \sigma_{1x}^2)^3}\left[\frac{\sigma_{1y}^2}{n_{1y} - 1} + \frac{\sigma_{1x}^2}{n_{1x} - 1}\right] \tag{4.7}$$

$$V(\hat{\delta}_2) = \frac{1}{(\sigma_{2y}^2 + \sigma_{2x}^2)}\left[\frac{\sigma_{2y}^2}{n_{2y}} + \frac{\sigma_{2x}^2}{n_{2x}}\right] + \frac{(\mu_{2y} - \mu_{2x})^2}{2(\sigma_{2y}^2 + \sigma_{2x}^2)^3}\left[\frac{\sigma_{2y}^2}{n_{2y} - 1} + \frac{\sigma_{2x}^2}{n_{2x} - 1}\right]. \tag{4.8}$$

Using (4.7) and (4.8), the confidence interval of $\hat{AUC}$ is given as follows

$$p\Phi\left[\hat{\delta}_1 - Z_{\frac{\alpha}{2}}\sqrt{V\left(\hat{\delta}_1\right)}\right] + (1-p)\Phi\left[\hat{\delta}_2 - Z_{\frac{\alpha}{2}}\sqrt{V\left(\hat{\delta}_2\right)}\right], \Phi\left[\hat{\delta}_1 + Z_{\frac{\alpha}{2}}\sqrt{V\left(\hat{\delta}_1\right)}\right] + (1-p)\Phi\left[\hat{\delta}_2 + Z_{\frac{\alpha}{2}}\sqrt{V\left(\hat{\delta}_2\right)}\right] \tag{4.9}$$

where $Z_{\frac{\alpha}{2}}$ is the critical value of Z for a two tailed test and α is the level of significance.

V. Simulation Studies

Monte Carlo simulation is an estimating algorithm that is based on repeated random sampling to obtain numerical results. Anderson (1986) discussed the Metropolis algorithm and Monte Carlo Simulation. The following steps are used in Monte Carlo simulation.

1) Define a range of possible inputs.
2) Generate random numbers from a probability distribution over the range.
3) Execute a deterministic computation from the inputs.
4) Aggregate the results.

In this section, we discuss a numerical example to observe how the confidence interval of AUC behaves by using Monte Carlo simulation. We check the behavior of Area under the NMROC Curve (AUC).

Let the sample sizes for healthy and diseased cases are n=10, 20, 30, 100, 200, 300 and the mixture of healthy cases with population parameters p=0.7, $\mu_{10} = 6, \sigma_{10} = 2, \mu_{20} = 3, \sigma_{20} = 1.$ Similarly, the mixture of disease cases with population parameters, p=0.7, $\mu_{11} = 10, \sigma_{11} = 3, \mu_{21} = 7, \sigma_{21} = 2.$ The weights are same for healthy and disease cases. Using these values of parameters, the maximum likelihood estimates of parameters, estimated AUC and 95% confidence interval are given in Table 5.1.

Let the fixed values of population parameters of healthy controls as $\mu_{10} = 1.5,\ \mu_{20} = 1,\ \sigma_{10} = 0.6$ and $\sigma_{20} = 0.4$ and fixed values of $\sigma_{11} = 2.1$ and $\sigma_{21} = 1$ for disease cases for all sample sizes. The mean of disease cases are $\mu_{11} = (4, 5, 6, 9, 14)$ and $\mu_{21} = (1.5, 1.8, 2.1, 3.2, 4.3)$ for all sample sizes. The weight of healthy controls and disease cases are fix p=0.7 for all. The sample sizes n=(10, 20, 30, 40, 50) for all healthy and disease cases. Using these values of parameters, the maximum likelihood estimates of parameters, estimated AUC and 95% confidence interval are given in Table 5.2.

It is observed from Tables 5.1 and 5.2 that if we increase the sample size, the estimated value of mean and variance become closer to the population parameters. In table 5.2, we see that as the difference between the mean of disease cases increases with sample sizes, the AUC of NMROC Curve increases and difference between the UCL and LCL also decreases.

VI. Conclusion

When a population has the heterogeneity and it is further divided into subpopulations, then mixture distribution is used for considering the heterogeneity of the population. In this case, the two component normal-mixture ROC curve gives higher accuracy of diagnostic test as compared to the Binormal ROC curve. Its properties are discussed and it is found that Normal mixture ROC curve is monotonically increasing, concave, TNR asymmetric and is invariant under the monotonically increasing transformation. The confidence interval of AUC is also derived. Hence, it is observed that whenever heterogeneity is present in the population, one should use two component normal-mixture ROC curve instead of Binormal ROC curve.

References

[1]. Newcomb S. (1886), A generalized theory of the combination of observations so as to obtain the best result, Amer. J. Math., 8, 343-346.
[2]. Pearson K (1894), Contribution to the mathematical theory of evolution, Phil. Trans. Roy. Soc. A, 185, 71-110.
[3]. Everitt B. and Hand D.J. (1981), Finite Mixture Distributions, Chapman and Hall. New York.
[4]. Titterington, D.M., Smith, A.F.M. and Makov U.E. (1985), Statistical Analysis of Finite Mixture Distributions, John Wiley and Sons Ltd.
[5]. McLachlan G.J. and Basford K.E. (1988), Mixture Models: Inference and Applications to Clustering, Marcel Dekker, New York.
[6]. Bohning D., Dietz E. and Schlattmann P. (1998), Recent Developments in Computer Assisted Analysis of Mixtures, Biometrics, 54, 525-536.
[7]. McLachlan, G. and Peel D. (2000), Finite Mixture Models, John Wiley, New York.
[8]. Schlattmann P. (2009), Medical Application of Finite Mixture Models, Springer-Verlag Berlin Heidelberg.
[9]. Wirjanto Tony S. and Dinghai Xu (2009), The Application of Mixtures of Normal Distributions in Empirical Finance: A Selected Survey,University of Waterloo, Department of Statistics 2009.
[10]. Patrick J.F., Ehsanes Saleh A.K.M. and Zhang Z. (2011), Methods of Moments Estimation in Finite Mixtures, Sankhya: The Indian Journal of Statistics, 73-A, part 2, 218-230.
[11]. Dass S.C. and Seong W.K. (2012), A Semi-parameteric Approach to Estimation of ROC Curves for Multivariate Binormal Mixtures. www.stt.msu.edu/Links/Research_Memoranda/RM/RM_683.pdf.
[12]. Lee J.Y. (2012), A two-stage approach using Gaussian mixture models and higher-order statistics for a classification of normal and pathological voices, Journal on Advances in Signal Processing , 252, 1-8.
[13]. Gonen M. (2013), Mixtures of Receiver Operating Characteristic Curves, Memorial Sloan-Katter in Cancer. http//works.bepress.com/mithat_gonen/23.
[14]. Hughes, G. and Bhattacharya, B. (2013), Symmetry Properties of Bi-Normal and Bi-Gamma Receiver Operating Curves are Described by Kullback-Leibler Divergences, Entropy, 15(4), 1342-1356.
[15]. Green D.M. and Swets J.A. (1966), Signal Detection Theory and Psychophysics, Wiley, New York.
[16]. Egan J.P. (1975), Signal Detection Theory and ROC Analysis, Academic Press, New York.
[17]. Krzanowski, W.J. and Hand, D.J. (2002), ROC curves for continuous data, Monographs on Statistics and Applied Probability, CRC Press, Taylor and Francis Group, New York.
[18]. Bamber D. (1975), The area above the ordinal dominance graph and area below the receiver operating characteristic graph, Journal of Mathematical Psychology, 12, 387-415.
[19]. Hanley J.A. and McNeil B.J. (1982), The meaning and use of the area under an ROC curve, Radiology, 143, 29-36.
[20]. Bradley A.P. (1997), The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition, 30, 1145-1159.
[21]. Pundir, S. and Amala, R. (2013), Receiver Operating Characteristic modeling for continuous data: A glance, Model Assisted Statistics and Application, 9(2), 121-135.
[22]. Dorfman, D.D. and Alf, E.J., (1969), Maximum likelihood estimation of parameters of signal detection theory and determination of confidence interval – rating method data, Journal of Mathematical Psychology, 6, 487-496.
[23]. Kamaruzzaman Z.A., Ziadi I and Ismail M.T. (2012), Mixture of Normal Distributions: Application to Bursa Malaysia Stock Market Indices, World Applied Science Journal, 16(6), 781-790.
[24]. Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977), Maximum- Likelihood from incomplete data, via the EM algorithm, Journal of Royal Statistics Society Ser. B (Methodological), 39 (1), 1-38.
[25]. Joshua H.P. (2013), The EM Algorithm in Multivariate Gaussian Mixture Models using Anderson Acceleration. A Project Report, Worcester Polytechnic Institute.
[26]. Fluss R., Faraggi D. and Reiser B (2005), Estimation of the Youden Index and its associated cutoff point, Biometrical Journal, 47(4), 458-472.
[27]. Anderson H.L. (1986) Metropolis, Monte Carlo and the MANIAC, Los Alamos Science.

Table 5.1: MLE, $\hat{AUC}$ and 95% confidence interval of $\hat{AUC}$ for different sample sizes

n	$\hat{p}_0$	$\hat{\mu}_{11}$	$\hat{\mu}_{10}$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{10}$	$\hat{p}_1$	$\hat{\mu}_{21}$	$\hat{\mu}_{20}$	$\hat{\sigma}_{21}$	$\hat{\sigma}_{20}$	$\hat{AUC}$	Confidence Interval
10	0.70 (0.0)	11.818 (1.818)	6.718 (0.718)	2.402 (-0.598)	1.944 (-0.056)	0.70 (0.0)	4.580 (-2.42)	1.903 (-1.097)	0.128 (-1.872)	0.179 (-0.821)	0.965 (0.069)	[0.786, 0.992]
20	0.85 (0.15)	10.006 (0.006)	5.313 (-0.687)	1.874 (-1.126)	1.558 (-0.442)	0.90 (0.2)	3.225 (-3.775)	1.956 (-1.044)	0.211 (-1.789)	0.649 (-0.351)	0.972 (0.077)	[0.911, 0.992]
30	0.60 (-0.1)	9.461 (-0.539)	7.384 (1.384)	3.670 (0.670)	1.068 (-0.932)	0.77 (0.07)	8.536 (1.536)	3.532 (0.532)	0.276 (-1.724)	0.650 (-0.35)	0.774 (-0.121)	[0.671, 0.858]
100	0.67 (-0.03)	11.983 (1.983)	7.234 (1.234)	2.896 (-0.104)	1.332 (-0.668)	0.80 (0.10)	7.512 (0.512)	3.698 (0.698)	1.916 (-0.084)	1.261 (0.261)	0.935 (0.040)	[0.905, 0.958]
200	0.88 (0.18)	11.079 (1.079)	7.005 (1.005)	2.439 (-0.561)	2.075 (0.075)	0.59 (-0.11)	7.692 (0.692)	4.918 (1.918)	2.425 (0.425)	0.635 (-0.365)	0.884 (-0.010)	[0.854, 0.910]
300	0.70 (0.0)	10.487 (0.487)	6.326 (0.326)	2.855 (-0.145)	1.722 (-0.278)	0.62 (-0.08)	6.837 (-0.163)	2.833 (-0.167)	1.974 (-0.026)	1.103 (0.103)	0.919 (0.024)	[0.900, 0.935]

( ): bias of the estimates

Table 5.2 MLE, $\hat{AUC}$ and 95% confidence interval of $\hat{AUC}$ for different sample sizes

n	$\hat{p}_0$	$\hat{\mu}_{11}$	$\hat{\mu}_{10}$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{10}$	$\hat{p}_1$	$\hat{\mu}_{21}$	$\hat{\mu}_{20}$	$\hat{\sigma}_{21}$	$\hat{\sigma}_{20}$	$\hat{AUC}$	Confidence Interval
10	0.72 (0.02)	5.685 (1.685)	1.169 (-0.331)	0.350 (-1.75)	0.466 (-0.134)	0.59 (-0.11)	1.989 (0.489)	0.359 (-0.641)	1.392 (0.392)	0.036 (-0.364)	0.950 (0.135)	[0.830, 0.988]
20	0.54 (-0.15)	3.651 (-1.349)	1.611 (0.111)	2.403 (0.303)	0.170 (-0.43)	0.85 (0.15)	0.719 (-1.081)	0.764 (-0.236)	0.371 (-0.629)	0.213 (-0.169)	0.749 (-0.144)	[0.599, 0.862]
30	0.44 (-0.25)	6.550 (0.55)	1.662 (0.162)	1.164 (-0.936)	0.356 (-0.244)	0.52 (-0.18)	2.148 (0.048)	0.669 (-0.331)	1.440 (0.44)	0.291 (-0.109)	0.924 (-0.016)	[0.869, 0.961]
40	0.23 (-0.46)	9.138 (0.138)	2.230 (0.73)	2.210 (0.11)	0.352 (-0.248)	0.79 (0.09)	2.736 (-0.464)	0.948 (-0.052)	0.852 (-0.148)	0.465 (0.065)	0.992 (-0.001)	[0.977, 0.997]
50	0.20 (-0.49)	13.991 (-0.009)	2.029 (0.529)	1.766 (-0.334)	0.691 (0.091)	0.72 (0.02)	4.226 (-0.074)	1.246 (0.246)	1.012 (0.012)	0.504 (0.104)	0.998 (-0.001)	[0.944, 0.999]