Special Proceedings: ISBN #: 978-81-950383-0-5
23rd Annual Conference, 24-28 February 2021; pp 113–123

Inferential Procedures to Compare Parallel, Superior and Crossover Multivariate ROC Curves

Vishnu Vardhan Rudravaram
Department of Statistics, Ramanujan School of Mathematical Sciences
Pondicherry University, Puducherry

Received: 13 June 2021; Revised: 22 July 2021; Accepted: 24 July 2021

Abstract

Receiver Operating Characteristic (ROC) curve is widely used and accepted tool to assess the performance of a classifier or procedure. Along with this, comparing the diagnostic test procedures or ROC curves is also of major concern and interest. A multivariate extension of ROC (MROC) curve considers a linear combination of several markers for classification. In this work, some inferential procedures are given to compare MROC curves that are parallel, superior and crossover using the scores of MROC curve and also using mean vectors and dispersion matrices. Further, a modified version of AUC (mAUC) under MROC setup is proposed to address the case of crossover MROC curves. It is also shown that mAUC performs better than AUC. The performance of mAUC in the aspect of crossover curves is supported by a real dataset and simulation studies at different sample sizes.

Key words: MROC Curve; Modified AUC and Crossover curves

1. Introduction

In theory of Statistics, there are several tools and techniques available and being developed to address various practical issues in diversified areas. One such prominent area is *Classification Scenario*. The term *"Classification"* indicates the method of allocating or assigning a group of objects/individuals into one of the predefined classes or populations. The prominent fields of research where the logical thinking and analytical processing of classification techniques can flourish are *Diagnostic Medicine, Life Sciences, Experimental Psychology* etcetera.

In general, there are two major objectives in classification problems: the first one is to define a *classifier rule* and the second is to determine an *optimal cutoff*. These two objectives are to be met in such a way that it should minimize the rate of misclassification. The classifier rule will help in generating a decision matrix (usually referred as *confusion matrix*) with probabilities of correct and incorrect classifications. The techniques available to handle such classification problems are Logistic Regression, Discriminant Analysis and Receiver Operating Characteristic (ROC) curve analysis. All these techniques are branched from the hub of *Statistical Decision Theory* (SDT). The first two techniques meet the above mentioned criteria of obtaining a classifier rule and optimal cutoff. In addition to the mentioned

Corresponding Author: Vishnu Vardhan Rudravaram Email: vrstatsguru@gmail.com

objectives, ROC curve has yet another feature of providing the accuracy of a classifier.

ROC curve took its origin during World War II and was first used in Signal Detection Theory for analyzing radar images. This technique has its applications in wide variety of fields such as Medicine, Experimental Psychology, Banking, Finance and many more. However, the promising area for the theoretical development of ROC curve is *Diagnostic Medicine*. Apart from providing a classifier rule and optimal threshold another important advantage of ROC curve is in assessing the performance of diagnostic test and in choosing a better one when there are two tests for a particular scenario. Over the years, application of ROC curve analysis has been observed in many fields and a few to mention are Experimental Psychology, Diagnostic Medicine and Radiology (Krzanowski and Hand, 2009), Machine learning (Provost *et al.*, 1998).

Let there exist two populations denoted by '0' (normal or healthy) and '1' (abnormal or diseased) where 'c' be the cutoff. The individual/object is said to belong to population '1' if the score 'S' is greater than c otherwise belongs to population '0'. Four probabilities and their associated classification rates can be defined as

- The probability that an individual/object from '1' is correctly classified as '1'

$$\text{i.e, True Positive Rate } TPR = P(S > c|1)$$

- The probability that an individual/object from '0' is misclassified as '1'

$$\text{i.e., False Positive Rate } FPR = P(S > c|0)$$

- The probability that an individual/object from '0' is correctly classified as '0'

$$\text{i.e., True Negative Rate } TNR = P(S \leq c|0)$$

- The probability that an individual/object from '1' is misclassified as '0'

$$\text{i.e., False Negative Rate } FNR = P(S \leq c|1)$$

The ROC curve underpins an unknown monotonic transformation and it can be defined as the tradeoff between two intrinsic measures namely *1- specificity (FPR)* and *sensitivity (TPR)*. It is a unit square plot ranging from (0, 0) to (1, 1) and a line connecting these points is called the chance diagonal. Sensitivity is the probability that the test result is positive when the condition is present and Specificity is the probability that the test result is negative when the condition is absent.

$$\text{Sensitivity}(S_n) = P(S > c|1); \;\; \text{Specificity}(S_p) = P(S \leq c|0) \tag{1}$$

The typical forms of ROC curve are depicted in Figure 1. The figure constitutes of three cases of ROC curves: best, moderate and worst case. Each case indicates the extent of classification that can be performed using the marker. The curve for best case reaches the top left corner of the graph, the moderate case lies between the top left corner and chance diagonal and the worst case runs parallel to the chance diagonal.

A widely used summary measure of the ROC curve is the Area under the Curve (AUC) which is the probability that a randomly chosen individual/object from population '1' has a higher score than a randomly chosen individual/object from population '0'. It depicts the amount of correct classification that can be achieved using the cutoff of a marker under study. Probabilistically,

$$AUC = P(S_1 > S_0) \tag{2}$$

where S_0 and $S_1 \in S$ are the test scores of populations '0' and '1' respectively. A practical lower bound for AUC is 0.5 and any test with AUC = 0.5 is said to have random classification. As the value of AUC gets closer to 1, better the performance of a test. The mathematical formulation of *Binormal ROC model* is given in detail in the next section.

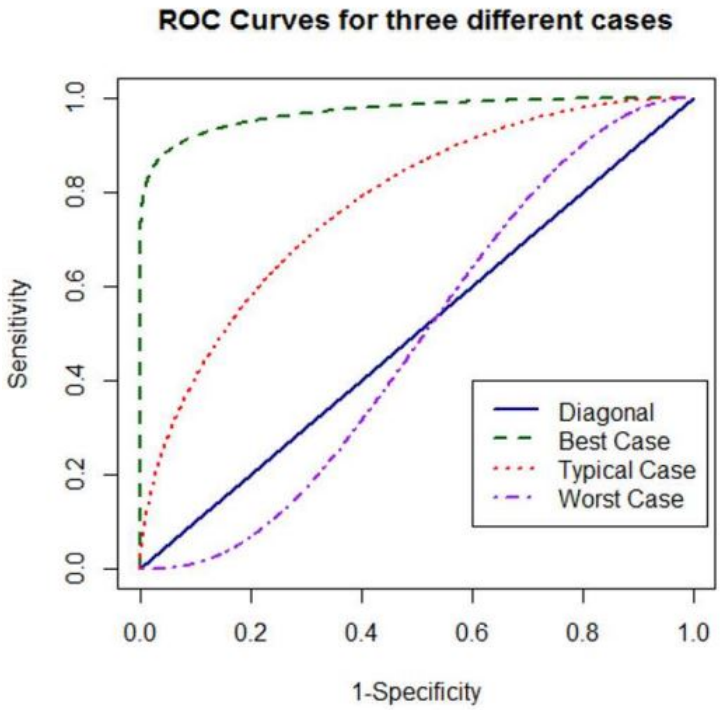


Figure 1: Typical forms of ROC curve

2. Binormal ROC Curve and its Ramifications

Let S_0 and S_1 be two random variables which denote scores from populations '0' and '1' respectively. These are assumed to follow normal distribution with means μ_0 and μ_1 and standard deviations σ_0 and σ_1 respectively (Green and Swets, 1966). i.e., $S_0 \sim N(\mu_0, \sigma_0^2)$ and $S_1 \sim N(\mu_1, \sigma_1^2)$. It is assumed that the mean of population '1' is greater than population '0' ($\mu_1 > \mu_0$) but no restrictions are posed on standard deviations. The intrinsic measures of ROC curve, TPR and FPR at a cutoff 'c' are defined as

$$FPR = x(c) = \Phi\left(\frac{\mu_0 - c}{\sigma_0}\right); \;\; TPR = y(c) = \Phi\left(\frac{\mu_1 - c}{\sigma_1}\right) \tag{3}$$

using FPR, we obtain $c = \mu_0 - \sigma_0 \Phi^{-1}[x(c)]$, then TPR can be written as

$$TPR = y(c) = \Phi\left(\frac{\mu_1 - \mu_0}{\sigma_1} + \frac{\sigma_0}{\sigma_1}\Phi^{-1}[x(c)]\right) \tag{4}$$

The expression for Binormal ROC curve becomes

$$y(c) = \Phi(a + b\ \Phi^{-1}[x(c)]) \tag{5}$$

where $a = \frac{(\mu_1 - \mu_0)}{\sigma_1}$, $b = \frac{\sigma_0}{\sigma_1}$ and $\Phi(.)$ is the standard normal deviate.

The AUC of the ROC curve can also be defined as the average true positive rate over all possible false positive rates in the range (0, 1).

$$AUC = \int_0^1 y(c)\ dx(c) \tag{6}$$

The expression for the AUC of Binormal ROC model is given as

$$AUC = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right) \tag{7}$$

3. ROC Models with Multiple Markers

One of the problems in Obstetrics and Gynecology is to identify a better procedure which helps in studying the blood flow from womb of the mother to baby for identifying the baby's growth. The study has multiple markers that need to be considered to identify whether there is a sufficient blood flow which in turn helps in classifying the subjects into one of the two groups: with adequate and inadequate blood flow. Another situation pertaining to Ecology is also observed where there is a need to identify the species type of a bug as well as to distinguish the gender of a particular species based on the features/characteristics of that particular bug. Hence, one cannot always depend on a single marker to judge the individual's/object's status. This scenario creates a necessity to develop a model that considers more than one marker for classification.

Su and Liu (1993) proposed best linear combinations where both healthy and diseased populations follow multivariate normal distribution by considering two cases, one with proportional covariance matrices and the other with no restriction on covariance matrices. In the first case with proportional covariance matrices the linear combination, is said to maximize sensitivity over a range of specificities. In the case of populations with unequal covariance matrices, the linear combination is the one that maximizes AUC among all possible combinations. Further confidence intervals were developed for AUC of the Su and Liu model by Reiser and Faraggi (1997) and named it as *Generalized ROC model*. Schisterman *et.al.* (2004) discussed covariates effects on the generalized ROC model and provided approximate confidence intervals for the measure AUC.

Liu *et al.* (2005) proposed methods to estimate the best linear combinations by maximizing sensitivity at a fixed specificity for Su and Liu model. They proved that the linear combinations proposed outperform Su and Liu model when there exists heterogeneity among covariance structures.

Countable articles were found in literature for the above multivariate classification in the context of ROC curves. The model proposed by Su and Liu (1993) was used as a base model for a considerable number of these articles. However, Su and Liu model has mathematical complexity when dealing with non-proportional covariance structures. Several authors provided improvisations on the model but none of them suggested a single linear combination that can accommodate both equal and unequal covariance structures. This motivated to the development of a new ROC model that can accommodate equal and unequal covariance structures by linearly combining multiple markers at hand (Sameera *et al.*, 2016).

4. The Multivariate Receiver Operating Characteristic (MROC) Curve

Let $X = (x_1, x_2, \ldots, x_k)$ be the 'k' markers involved in the study. Let π_0 and π_1 be two independent populations (groups) assumed to follow multivariate normal distribution with mean vectors μ_0, μ_1; covariance matrices Σ_0, Σ_1 and sample sizes n_0, n_1 respectively and $n = n_0 + n_1$. Then the probability density function for $\pi_i, i = 0, 1$ is given by

$$f(X|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{k}{2}}|\Sigma|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}(X-\mu_i)^T\Sigma_i^{-1}(X-\mu_i)\right\}$$

where Σ is positive definite.

Let $x(c)$ denote the false positive rate (FPR) and $y(c)$ denote the true positive rate (TPR) where 'c' is the cutoff. The expressions for FPR, c and TPR are defined as

$$FPR = x(c) = P(U > c|\pi_0) = 1 - \Phi\left(\frac{c - b^T\mu_0}{\sqrt{b^T\Sigma_0 b}}\right) \tag{8}$$

where $b(\neq 0)$ be a $kX1$ vector and 'U' is the test score. Using (8), the expression for 'c' is given as

$$c = b^T\mu_0 + \sqrt{b^T\Sigma_0 b}\ \Phi^{-1}(1 - x(c)) \tag{9}$$

where $\phi^{-1}(.)$ is the inverse function of $\Phi(.)$

$$TPR = y(c) = P(U > c|\pi_1) = \Phi\left(\frac{b^T\mu_1 - c}{\sqrt{b^T\Sigma_1 b}}\right) \tag{10}$$

substituting (9) in (10) we get

$$ROC(c) = \Phi\left[\frac{b^T(\mu_1 - \mu_0) - \sqrt{b^T\Sigma_0 b}\ \Phi^{-1}\left(1 - FPR\right)}{\sqrt{b^T\Sigma_1 b}}\right] \tag{11}$$

The expression in (11) is the form of MROC curve.

The linear combination is defined as $U = b^T X = b_1x_1 + b_2x_2 + \ldots + b_kx_k$, where the vector 'b' is obtained using Minimax procedure (Anderson and Bahadur ,1962) as

$$b = [t\Sigma_1 + (1-t)\Sigma_0]^{-1}(\mu_1 - \mu_0) \tag{12}$$

here 't' is a constant which lies in the interval (0, 1) and its value is determined by trial and error method. The cutoff 'c' at each 't' can be obtained through Minimax procedure by

equating TPR and FPR. On solving, we obtain

$$c = \frac{b^T\mu_1\sqrt{b^T\Sigma_0 b} + b^T\mu_0\sqrt{b^T\Sigma_1 b}}{\sqrt{b^T\Sigma_1 b} + \sqrt{b^T\Sigma_0 b}} \tag{13}$$

The AUC of MROC curve can be obtained by integrating (11) over [0, 1]. However, the expression of AUC can also be derived using probabilistic notations. Let U_0 and U_1 be the test scores randomly taken from π_0 and π_1 populations respectively, $(U_1 > U_0)$.

$$\therefore AUC = P(U_1 > U_0)$$

$$AUC = P(U_1 - U_0 > 0)$$

The test scores U_0 and U_1 are independent and follow normal distribution

$$i.e., U_0 \sim N(b^T\mu_0, b^T\Sigma_0 b) \text{ and } U_1 \sim N(b^T\mu_1, b^T\Sigma_1 b), \text{ then}$$

$$U_1 - U_0 \sim N(b^T\mu_1 - b^T\mu_0, b^T\Sigma_0 b + b^T\Sigma_1 b).$$

Hence, if 'z' denotes standard normal variable then,

$$AUC = \Phi\left(\frac{b^T(\mu_1 - \mu_0)}{\sqrt{b^T(\Sigma_1 + \Sigma_0)b}}\right) \tag{14}$$

Higher the AUC lower the overlapping area of two populations and vice versa.

The two intrinsic measures of ROC curve that are used in plotting the curve, Sensitivity (S_n) and Specificity (S_p) defined as abilities of correct identification of the two groups '1' and '0' respectively and are given as follows

$$S_n = P(U > c|\pi_1) = \Phi\left(\frac{b^T\mu_1 - c}{\sqrt{b^T\Sigma_1 b}}\right); \quad S_p = P(U < c|\pi_0) = \Phi\left(\frac{c - b^T\mu_0}{\sqrt{b^T\Sigma_0 b}}\right) \tag{15}$$

5. Introduction to Crossover Curves

In classification, attention is required for those reference values (cutoff) of markers which provide at least a moderate amount of correct classification with a greater susceptibility. In usual context of assessing the performance of a test, scores which are nearer to reference value are given same amount of weightage as that of the scores farther from reference value. The AUC so computed will be contaminated and the true accuracy or the actual performance will be masked. This misleads the interpretation of the measures of ROC as well as the optimal cutoff and leads to high amount of misclassification. Let us consider two tests A and B for better identification of a particular abnormality in individuals. Suppose that the curves of tests A and B cross each other and have at most similar accuracies. Under these circumstances, it is very difficult to notify a better test which has more ability to distinguish individuals.

To fix this issue, a solution to compare two crossover curves by means of a modified version of AUC (mAUC) of MROC curve by eliciting the importance of mAUC over AUC. Numerical illustrations are given using real data sets.

5.1. Modified Area under the Curve (mAUC)

AUC is the probability that an individual/object from group '1' has a score greater than individual/object from group '0'. One small drawback with this definition is that, it does not take into account the amount by which the scores of group '1' and group '0' differ. To overcome this, a small weight is assigned to those scores where the difference between scores is comparatively small (Figure 2).

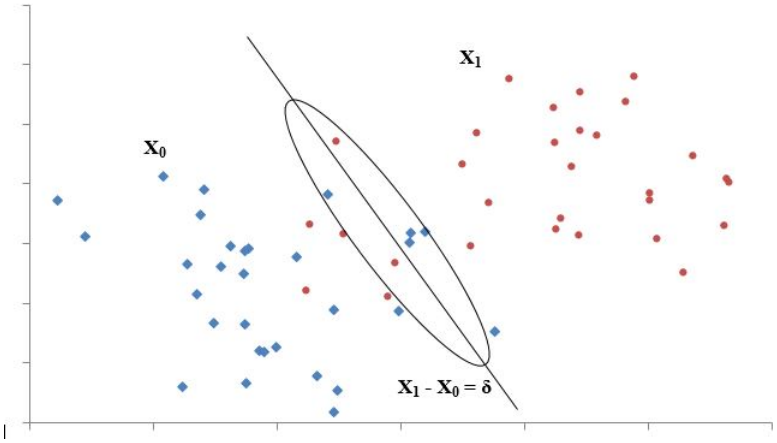


Figure 2: Hypothetical distribution of scores of two populations

mAUC was defined probabilistically by Yu *et al.* (2014) under univariate setup as weighted sum of two AUC's.

$$i.e.,\ mAUC = P(X_1 - X_0 > \delta) + (1-\lambda)P(0 < X_1 - X_0 \leq \delta)$$

$$mAUC = (1-\lambda)P(X_1 > X_0) + \lambda P(X_1 > X_0 + \delta) \tag{16}$$

The first part of this mAUC represents the conventional AUC with $(1-\lambda)$ *as its weight and the second part constitutes an additional parameter '*δ*' with '*λ*' as its weight.* The main role of 'δ' is to magnify the true status of scores of the individuals that are nearer to the reference value. Once the 'δ' value is imposed, a clear identification can be made about those scores that can be treated as true positives, which is the criterion of interest. This supports in giving out an accuracy which can be considered to be better than the conventional AUC. Using the above probabilistic notations, mAUC is derived for MROC model and is given as

$$mAUC = (1-\lambda)\ \Phi\left(\frac{b^T(\mu_1-\mu_0)}{[b^T(\Sigma_0+\Sigma_1)b]^{\frac{1}{2}}}\right) + \lambda\ \Phi\left(\frac{(b^T(\mu_1-\mu_0)-\delta)}{[b^T(\Sigma_0+\Sigma_1)b]^{\frac{1}{2}}}\right) \tag{17}$$

In equation (17), the values of parameters 'λ' and 'δ' are to be chosen in such a way that the true accuracy of a test can be extracted by minimizing the effect of nearby points of the threshold. If 'λ' value is taken to be 0, mAUC reduces to AUC and if it is taken as 1, the probability $P(X_1 > X_0+\delta)$ is only taken into account. Any value of 'λ' greater than 1 would result in making the probability $P(0 < X_1 - X_0 \leq \delta)$ value a penalty. Hence, a reasonable choice for 'λ' lies in the range (0, 1), larger the 'λ' value lower the importance on AUC. In order to choose a 'δ' that is meaningful and reasonable the following result is used.

Result: The confidence interval for mean vector of multivariate normal distribution is

$$\left(b^T \bar{X} - \sqrt{\frac{k(n-1)}{n(n-k)}\, F_{k,(n-k)}(\alpha)\, b^T S b},\;\; b^T \bar{X} + \sqrt{\frac{k(n-1)}{n(n-k)}\, F_{k,(n-k)}(\alpha)\, b^T S b} \right)$$

where n is the number of samples; k is the number of markers and $b^T S b$ is the quadratic form. (Result 5.3, Johnson and Wichern (2007), p.225)

Using the above result, the upper bound of population '0' can be written as

$$b^T \bar{X}_0 - \sqrt{\frac{k(n_0-1)}{n_0(n_0-k)}\, F_{k,(n_0-k)}(\alpha)\, b^T S_0 b}$$

and parameter 'δ' can be chosen as

$$\sqrt{\frac{k(n_0-1)}{n_0(n_0-k)}\, F_{k,(n_0-k)}(\alpha)\, b^T S_0 b}$$

. The main reason for this choice of 'δ' is that it is part of the upper bound for the mean vector of population '0'. If an observed score is larger than this upper bound, then individual's status can be affirmatively called as true positive. The variance of mAUC expression cannot be derived explicitly and hence the concept of bootstrapping is used. If 'B' bootstraps are generated from the dataset, then the estimate and variance of mAUC is given as

$$\hat{mAUC}_{bs} = \frac{1}{B}\sum_{b=1}^{B} mAUC_b \;;\;\; Var(\hat{mAUC}_{bs}) = \frac{1}{B-1}\sum_{b=1}^{B}(mAUC_b - \hat{mAUC}_{bs})^2 \qquad (18)$$

5.2. Comparing two Crossover MROC curves

Two crossover curves can be compared using their mAUC values. The testing procedure proposed to test the hypothesis

$$H_0 : mAUC_{(1)} = mAUC_{(2)} \sim H_1 : mAUC_{(1)} \neq mAUC_{(2)}$$

for identifying the difference between two cross over MROC curves is defined as

$$Z = \frac{\hat{mAUC}_{bs(1)} - \hat{mAUC}_{bs(2)}}{var(\hat{mAUC}_{bs(1)}) - var(\hat{mAUC}_{bs(2)})} \qquad (19)$$

where $\hat{mAUC}_{bs(i)}$ and $var(\hat{mAUC}_{bs(i)});\; i = 1, 2$ can be estimated using equation (18). The Z statistic follows standard normal distribution asymptotically. The bootstrapped confidence interval for mAUC can be obtained using

$$\hat{mAUC}_{bs} \pm Z_{\left(1-\frac{\alpha}{2}\right)}\, var(\hat{mAUC}_{bs})$$

6. Results and Discussions

The proposed methodology is supported using real datasets namely, IUGRFDS and ILP datasets. For illustration purposes, all computations of mAUC and its confidence intervals under real datasets are given at $\lambda = 0.3, 0.5, 0.8$.

IUGRFDS dataset

The dataset IUGRFDS contains data collected from two independent diagnostic procedures CPR and MCA which exhibit a moderate amount of classification. Here, comparison is to be made between CPR and MCA procedures in order to find out which procedure is better in identifying the sufficient blood flow from the mother to baby. The AUC's and mAUC's of CPR and MCA along with their corresponding Z statistic value are computed and reported in Table 1. The crossover MROC curves for CPR and MCA procedures are shown in Figure 3.

Table 1: Comparison between CPR and MCA using mAUC

Measure	CPR (LL, UL)	MCA (LL, UL)	Z value (p-value)
$mAUC_{0.3}$	0.6551 (0.5196, 0.7815)	0.5902 (0.5034, 0.7254)	0.8008 (0.212^{NS})
$mAUC_{0.5}$	0.6369 (0.5106, 0.7631)	0.5774 (0.4453, 0.7090)	0.7070 (0.239^{NS})
$mAUC_{0.8}$	0.6095 (0.4817, 0.7577)	0.5581 (0.4558, 0.7032)	0.5777 (0.282^{NS})
AUC	0.6824 (0.5702, 0.7906)	0.6095 (0.5329, 0.7139)	0.9536 (0.170^{NS})

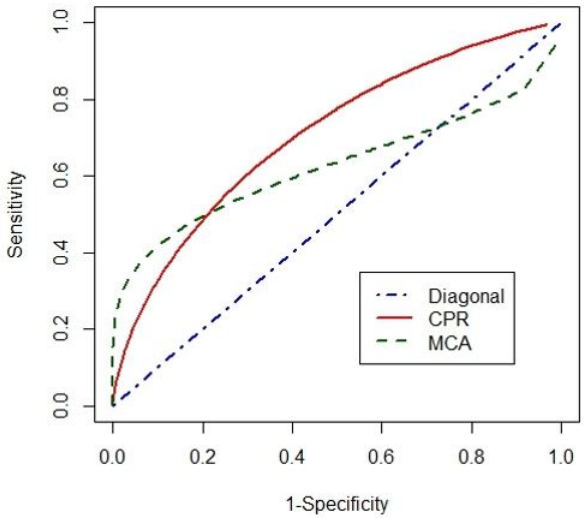


Figure 3: Crossover MROC curves for IUGRFDS dataset

Here NS = Not Significant, LL = Lower Limit, UL = Upper Limit.

For three values of λ, mAUC values are lower than that of AUC values. This is due to the fact that mAUC expression takes value of λ into account which results in assigning an appropriate weight to those scores that are closer to the threshold for extracting the true accuracy of a diagnostic procedure. However, the choice of λ should be in such a way that the accuracy is not too low. The results portrayed in Table 1 depict that both the procedures; CPR and MCA are equally effective in identifying the blood flow from mother to baby.

ILP dataset

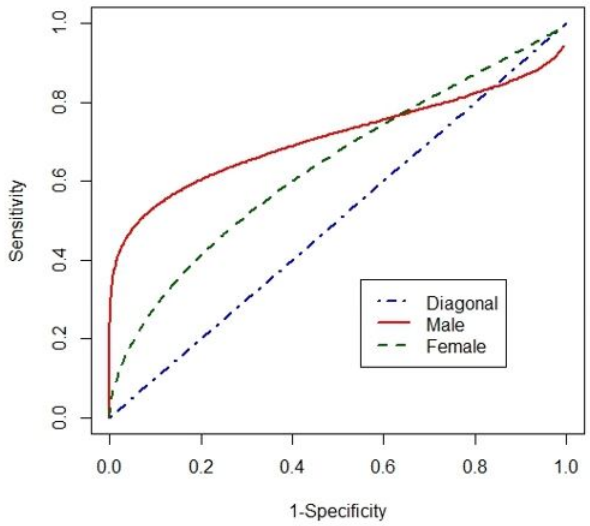
The ILP dataset is divided into two sets based on gender as males and females to check whether disease identification is identical in both the genders. MROC curves of males and females are then compared to check if classification is better in one gender compared to the other. The mAUC and AUC values are calculated for both datasets and placed in Table 2 along with their Z values and significance. The MROC curves obtained for males and females can be seen in Figure 4.

Table 2: Comparison between Males and Females using mAUC

Measure	Males (LL, UL)	Females (LL, UL)	Z value (p-value)
$mAUC_{0.3}$	0.6989 (0.6721, 0.7241)	0.6116 (0.5252, 0.7025)	1.8327 (0.033^{NS})
$mAUC_{0.5}$	0.6908 (0.6597, 0.7230)	0.5912 (0.5118, 0.6697)	2.0605 (0.019*)
$mAUC_{0.8}$	0.6788 (0.6571, 0.7124)	0.5606 (0.4512, 0.6598)	2.3824 (0.009*)
AUC	0.7109 (0.6759, 0.7297)	0.6422 (0.5591, 0.7203)	1.4726 (0.070^{NS})

Here NS = Not significant, * = significant, LL = Lower Limit, UL = Upper Limit.

Figure 4: Crossover MROC curves for ILP dataset



A better classification is seen in males than females when mAUC's obtained at $\lambda = 0.5$ and $\lambda = 0.8$. However, the Z value obtained for AUC's shows no difference between the curves indicating that the influence of scores close to the threshold is high and masking the true information. The result obtained at $\lambda = 0.3$ depicts an insignificant value stating that the weight assigned is not sufficient to extract true information from the scores of markers that are nearer to cutoff.

Observations

In this work, detailed discussion is made on a new summary measure for the MROC curve namely modified AUC is proposed. Inferential procedure is developed for this modified AUC in order to identify the true difference between MROC curves that cross each other.

References

Anderson, T. W. and Bahadur, R. R. (1962). Classification into Two Multivariate Normal Distributions with Different Covariance Matrices. *Annals of Mathematical Statistics*, **33**, 420–431.

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

Johnson, R. A. and Wichern, D. W. (2012). *Applied Multivariate Statistical Analysis*. 6th Edition, Prentice Hall India Learning Private Limited.

Krzanowski, W. J., and Hand, D. J. (2009). *ROC Curves for Continuous Data*. CRC Press.

Liu, A., Schisterman, E. F. and Zhu, Y. (2005). On linear combinations of biomarkers to improve diagnostic accuracy. *Statistics in Medicine*, **24(1)**, 37-47.

Provost, F. J., Fawcett, T. and Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. *In Proceedings of International Conference of Machine Learning*, **98**, 445-453.

Reiser, B. and Faraggi, D. (1997). Confidence intervals for the generalized ROC criterion. *Biometrics*, **53(2)**, 644-652.

Schisterman, E. F., Faraggi, D. and Reiser, B. (2004). Adjusting the generalized ROC curve for covariates. *Statistics in Medicine*, **23(21)**, 3319-3331.

Su, J. Q. and Liu, J. S. (1993). Linear combinations of multiple diagnostic markers. *Journal of the American Statistical Association*, **88(424)**, 1350-1355.

Yu, W., Chang, Y. I. and Park, E. (2014). A modified area under the ROC curve and its application to marker selection and classification. *Journal of the Korean Statistical Society*, **43**, 161 – 175